REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	11/30/2002	Final Progress Report 11/01/96 - 10/31/02

4. TITLE AND SUBTITLE
Multidisciplinary Research for Demining

5. FUNDING NUMBERS
DAAH04-96-1-0448

6. AUTHOR(S)
Lawrence Carin

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Duke University
Department of Electrical and Computer Engineering
Durham, NC 27708-0291

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
36278.21-EV-MUR

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report summarizes research progress on the Duke University led demining MURI (Multidisciplinary University Research Initiative), encompassing researchers from Duke, Caltech, Georgia Tech and Ohio State University. Sensors examined by this team include radar, electromagnetic induction, acoustic, olfactory and MEMS. In addition to sensor development, significant effort has been directed toward development of optimal signal processing algorithms.

14. SUBJECT TERMS
land mines

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	[redacted]

Ultra-Wideband, Short-Pulse Ground-Penetrating Radar: Simulation and Measurement

Stanislav Vitebskiy, *Student Member, IEEE*, Lawrence Carin, *Senior Member, IEEE*, Marc A. Ressler, *Member, IEEE*, and Francis H. Le

***Abstract*— Ultra-wideband (UWB), short-pulse (SP) radar is investigated theoretically and experimentally for the detection and identification of targets buried in and placed atop soil. The calculations are performed using a rigorous, three-dimensional (3-D) Method of Moments algorithm for perfectly conducting bodies of revolution. Particular targets investigated theoretically include anti-personnel mines, anti-tank mines, and a 55-gallon drum, for which we model the time-domain scattered fields and the buried-target late-time resonant frequencies. With regard to the latter, the computed resonant frequencies are utilized to assess the feasibility of resonance-based buried-target identification for this class of targets. The measurements are performed using a novel UWB, SP synthetic aperture radar (SAR) implemented on a mobile boom. Experimental and theoretical results are compared.**

I. Introduction

THE INSIDIOUS nature of mines has stimulated significant research—spanning over more than half a century—on techniques for buried-mine detection and identification. A significant tool for such is ground penetrating radar [1]–[18]. The frequencies used in GPR are chosen such that good ground penetration can be achieved, which necessitates relatively low operating frequencies ($<$1 GHz) for the losses common in typical soils [19]–[21]. Further, the bandwidth must be as large as possible to achieve sufficient temporal (spatial) resolution. With these motivations, ultra-wideband (UWB), short-pulse (SP) systems have received significant interest recently [22], [23]. Many of these systems operate in the time domain and have an instantaneous frequency spectrum of approximately 0.1–1.0 GHz ($>$100% bandwidth).

We have developed new three-dimensional (3-D) numerical algorithms for the modeling of UWB-SP scattering from realistic targets placed atop and buried in lossy, dispersive soils. Our Method of Moments (MoM) algorithm is implemented in the frequency domain, thereby allowing the modeling of dispersive soils [15]–[18]. Although the targets must be bodies of revolution, most practical targets of interest (buried mines, canisters, etc.) can be modeled. Our current code [16]–[18] assumes that the target is perfectly conducting, and the soil is modeled as a lossy, dispersive half space.

To use the MoM, one requires the accurate and efficient computation of the half-space dyadic Green's function over ultra-wide bandwidths . Unfortunately, this Green's function cannot be expressed in closed form, although each component of the dyadic can be represented as a Sommerfeld integral [15], [24]–[26]. Various numerical and asymptotic algorithms have been developed for the efficient analysis of the Sommerfeld integrals [15], [27]–[29], however none of these are sufficiently general and efficient to be useful for UWB applications. Recently, Chow *et al.* [30]–[32] have refined the method of complex images to efficiently compute the half-space Green's function, making possible the efficient and accurate computation of UWB-SP scattering from general buried targets [16], [17] (and the complex resonances thereof [18]). Previously we have concentrated on the algorithm's mathematical derivation and numerical implementation [16]–[18], while here we investigate the UWB-SP resonances and scattered fields from realistic targets; in particular, we consider conducting anti-personnel mines, anti-tank mines, and a 55-gallon drum.

Concerning target resonances, Baum [33] introduced the singularity expansion method (SEM) to demonstrate that the late-time transient fields scattered from a target can be expressed compactly in terms of the target's resonant modes. Further, the complex resonant frequencies represent unique discriminants that can be exploited in the context of time-domain ATR. Over the last two decades, researchers have been pursuing SEM-based ATR, through the use of such schemes as Prony 's method [34] and the matrix-pencil method [35], [36]. Such techniques have also been pursued for GPR, with several impressive results reported [5]. However, one would expect that resonance-based identification of buried targets will be very difficult for low-Q targets and/or targets buried in high-loss soils. We address this issue by theoretically investigating the late-time resonant frequencies of a realistic buried mine.

Where appropriate, we compare our theoretical calculations with measurements performed using a novel UWB-SP synthetic aperture radar (SAR) which is implemented by placing a bipolar radar on a mobile boom. The fundamental properties of the SAR are described in [37]. The system measures VV, HH, HV, and VH polarizations and operates in the time domain with an instantaneous bandwidth of approximately 0.1–1.5 GHz, and has an approximate peak radiated power of 150 kW. We have recently used this system to perform UWB-SP SAR measurements at the Yuma Proving Grounds in Yuma, AZ. All our calculations are performed using soil parameters measured from typical Yuma soil, and some of our theoretical results are compared with measured data.

Manuscript received April 29, 1996; revised October 15, 1996.

S. Vitebskiy and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA (e-mail: lcarin@ee.duke.edu).

M. A. Ressler and F. H. Le are with the Army Research Laboratory, AMSRL-SE-RU, Adelphi, MD 20783 USA.

Publisher Item Identifier S 0196-2892(97)03457-8.

0196–2892/97$10.00 © 1997 IEEE

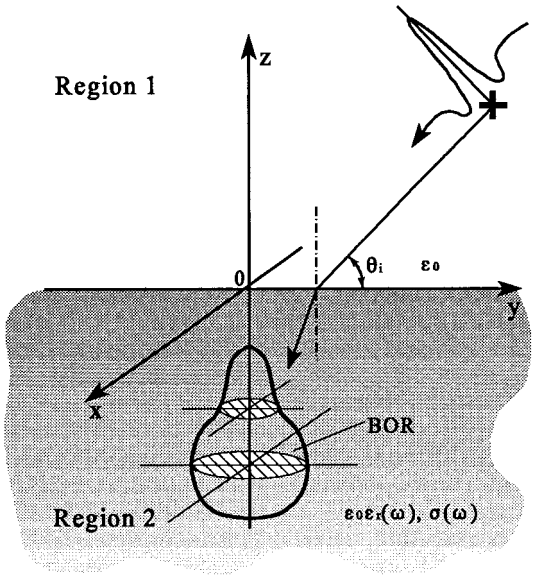


Fig. 1. Schematic of short-pulse plane-wave scattering from a buried perfectly conducting body of revolution.

The remainder of the text is organized as follows. In Section II a brief summary of our numerical algorithm is presented, followed in Section III by calculated results for the scattered fields and resonances of anti-personnel and anti-tank mines, as well as buried canisters. In Section IV we describe the properties of our UWB-SP SAR, and compare measured data with the results of some of our computations. Finally, conclusions are addressed in Section V.

II. Summary of Numerical Algorithm

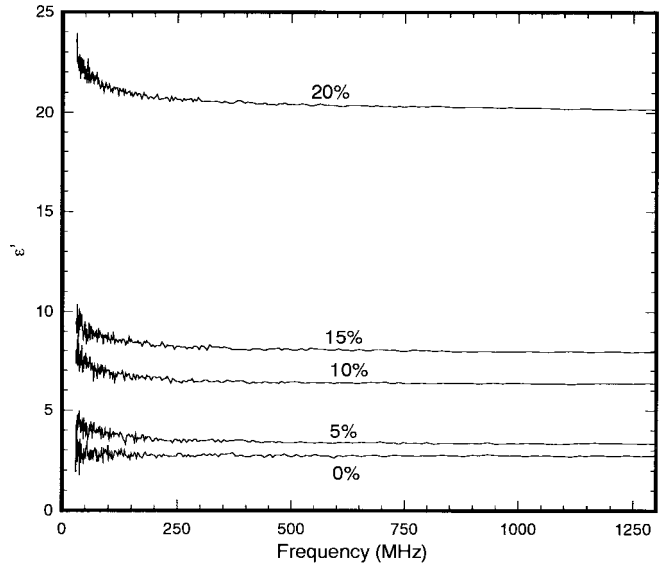
A general perfectly conducting body of revolution (BOR) buried in a lossy, dispersive half space is schematized in Fig. 1. In our algorithm, the target can be placed in (as in Fig. 1) or above the model soil. The latter case is important for military applications, since anti-personnel mines are often placed on the earth's surface. By restricting ourselves to bodies of revolution, the azimuthal variation of the unknown currents on the surface of the target can be expanded in terms of a discrete Fourier basis [17], [18], [38], with subsectional basis functions used to model variation along the generating arc. In the MoM analysis, testing functions [17], [18], [38] with properties similar to the basis functions are used to enforce the boundary condition of vanishing tangential electric field on the target surface, which decouples the problem into an infinite number of matrix equations for the unknown currents—one equation for each azimuthal Fourier component. Only a finite number of Fourier components are necessary to achieve convergence, and therefore a relatively small number of equations are obtained for the two-dimensional currents along the generating arc (the third dimension is accounted for subsequently by summing the various Fourier harmonics [17], [18], [38]). For the scattering problem, inhomogeneous matrix equations are derived for each Fourier component, the driving function of each represented by the azimuthal Fourier transform of the incident tangential electric field on the target surface, evaluated for the discrete Fourier component under consideration [17], [18], [38]. The time-domain scattered fields are synthesized by weighting the frequency-domain scattered fields by the spectrum of the desired incident pulse shape, which is converted to the time domain via Fourier transform. On the other hand, the target resonant frequencies corresponding to a particular azimuthal Fourier component are solved for by setting the driving function (incident field) equal to zero, which yields a homogeneous matrix equation for the modal currents, with nontrivial solutions at the complex frequencies for which the matrix determinant vanishes; thus, by setting the determinant of the matrix to zero, one can solve for the complex resonant frequencies, from which the modal currents can then be computed [18].

The above procedures for computing the scattered fields and complex resonant frequencies of bodies of revolution have been investigated extensively for targets in free space [38]. Our present problem is more complicated because the half-space dyadic Green's function must be expressed in the form of highly oscillatory, slowly convergent Sommerfeld integrals [16]–[18]. Recently Chow [30]–[32] has developed the method of complex images, which represents the spectral-domain reflection coefficient in terms of a sum of exponentials, with parameters found via Prony [34] or matrix-pencil [35], [36] fitting to the exact spectral reflection coefficient. The subsequent Sommerfeld integrals for each term in the exponential sum can be evaluated in closed form via the Weyl identity [39]. The restriction on such a scheme is that the BOR's axis of rotation must be perpendicular to the air-ground interface so that the (half space)-target composite preserves BOR symmetry (see Fig. 1). While this may be a significant restriction for some applications, for military problems (e.g., mines) such a target orientation is often found in practice.

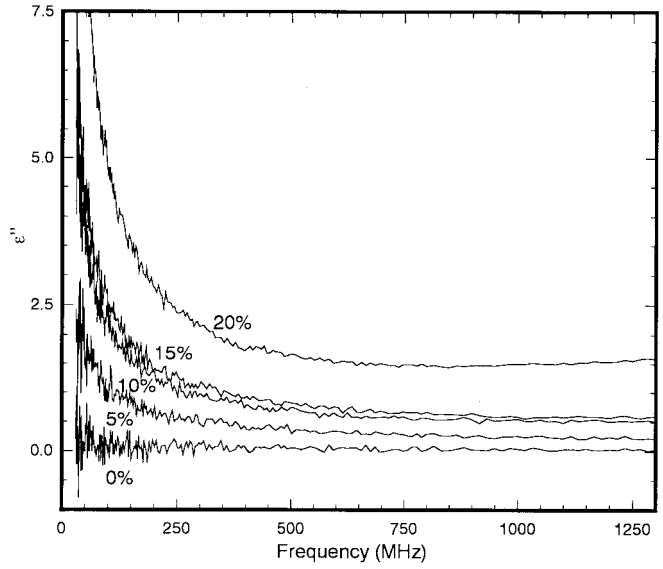
III. Short-Pulse Scattering from Three-Dimensional Targets

We consider below the short-pulse scattered fields and complex resonances of anti-personnel mines, anti-tank mines, and 55-gallon drums placed atop or buried within soil. Such that some of our calculations can be compared with measurements, the electrical parameters of the soil used in our model were measured from soil samples taken from Yuma, AZ, where experiments were performed using an UWB-SP SAR. The frequency-dependent (dispersive) soil parameters were measured with a network analyzer, and consisted of measuring the reflection coefficient from an open-circuit coaxial probe. Measurements were performed as a function of water content, by percentage weight, and are shown in Fig. 2 for the soil sample used in all subsequent scattering and resonance calculations, with results plotted in the form of complex relative permittivity $\epsilon_r = \epsilon_r' - j\epsilon_r''$.

There are several inherent difficulties in performing material measurements of the type reported in Fig. 2. In particular, the measured complex dielectric constant is a function of the pressure with which the open-circuit coaxial probe is pressed against the soil; in our measurements the probe pressure was increased until the measured complex permittivity stabilized,

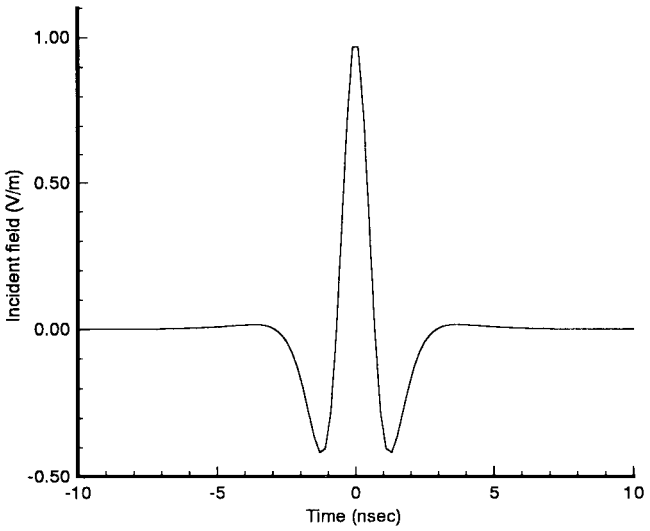


(a)

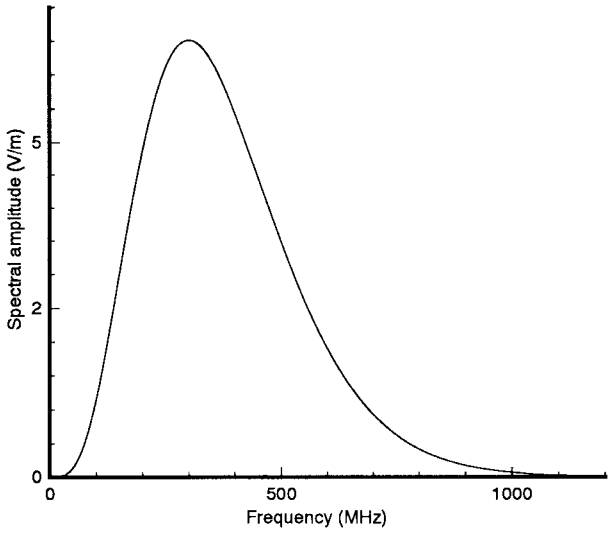


(b)

Fig. 2. Frequency-dependent complex permittivity $\epsilon_r = \epsilon_r' - j\epsilon_r''$ measured from a soil sample taken from Yuma, AZ. Results are plotted as a function of water content, by weight. (a) ϵ_r' (b) ϵ_r''.



(a)



(b)

Fig. 3. Time-domain waveform, and its spectrum, used to describe the shape of the pulsed plane wave considered in subsequent computations. (a) time-domain signal and (b) spectrum.

and the measurements were repeated five times and averaged to produce the data in Fig. 2. However, the applied pressure may influence the local percentage of water content in the vicinity of the probe, affecting the measured permittivity. In all subsequent scattering and resonances calculations, the curves corresponding to 5% water were utilized, in attempt to be consistent with what is typically found in the soils of Yuma, AZ.

A. Buried 55-Gallon Drum

First, we consider the UWB-SP fields scattered from a buried perfectly conducting 55-gallon drum with axis oriented perpendicular to the air-ground interface. The diameter of the drum is 60 cm, its length is 90 cm, and we consider a case for which its top surface is buried 30.48 cm from the air-ground interface. The time-domain shape of the incident pulsed plane wave is shown in Fig. 3, along with its Fourier spectrum. This waveform is consistent with many current pulsed sources [22], [23], and the low frequencies in the spectrum (<1.0 GHz) provide significant ground penetration. In many SAR applications, the incident waveform is near grazing, so we consider an incidence angle of 20° with respect to the ground, for both VV and HH polarization (the cross-polarized scattered fields for bodies of revolution is zero).

The relative power transmitted into the ground for vertically and horizontally polarized plane waves incident at 20° is shown in Fig. 4. For comparison, in Fig. 4 we also plot the relative transmitted power for a plane wave incident at the approximate Brewster angle at the center frequency ($\theta_B = 28.25°$). Note that the Brewster angle is frequency dependent, so Brewster-angle excitation is not possible for all

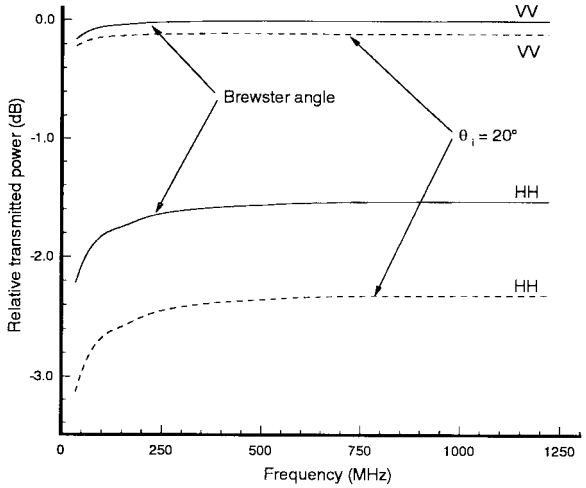


Fig. 4. Relative power transmitted into a lossy halfspace, as a function of frequency, for two angles of incidence (see Fig. 1) and polarization. The halfspace electrical parameters are described in Fig. 2, for the case of 5% water content.

frequencies in the incident pulse; moreover, for the case of lossy materials, no real Brewster angle exists at any frequency. Fig. 4 shows that the two-way transmission loss for 20° incidence is approximately 0.1 dB at the central frequency for VV polarization, and 2.4 dB for HH polarization. Further, from the 5%-water curve in Fig. 2, at the center frequency, there is approximately 4.35 dB of loss associated with propagation from the air–ground interface to and from the center of the drum. Our time-domain results simultaneously account for loss due to interface effects, propagation loss through the soil, and scattering from the target, and therefore these effects cannot be considered separately; however, from the above discussion, at the pulse's center frequency, the loss from interface and soil-propagation effects combined is approximately 4.45 dB for VV polarization and 6.75 dB for HH polarization.

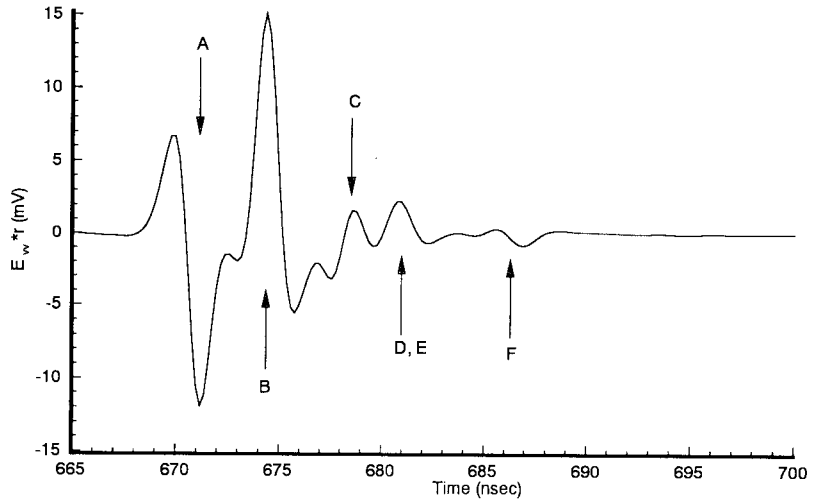
The normalized VV and HH time-domain backscattered fields are shown in Fig. 5. Several wavefronts are identified by arrows explained in Fig. 6. Included in Fig. 6 are diffractions from the the top front edge (A), diffraction from the top back edge (B), a reverberation between the front edge and the air-ground interface (C), creeping-wave circumnavigation of the cylinder (D), diffraction from the bottom front edge (E), and multiple diffraction between the top and bottom surface of the drum (F). The VV and HH signals differ in several important ways. For example, for VV polarization both edge diffractions "A" and "B" are strong, while for HH polarization wavefront "B" is substantially weaker than "A." The reason for these differences is unclear, due to the complexity of this problem. For example, in reality there is no "front" or "back" edge, but rather a single edge which encircles the entire top surface. Further comparing these results, we see that phenomena "D" and "E" are excited strongly for VV polarization while they are nearly absent for HH polarization. Phenomena "D" and "E" occur too closely in time to be individually identified. Therefore, we cannot be sure as to whether this strong wavefront for VV polarization is due to a creeping wave, edge diffraction, or both. However, for the 20° angle of incidence, the refracted wave travels at an angle of approximately 30° relative to the cylinder axis. If the incident wave were to propagate along the axis, no creeping wave is excited at all; therefore we speculate that the creeping wave will be weakly excited for this example in which the incident wave travels *almost* along the cylinder axis. Interestingly, the multiple edge diffraction, "F," is very similar for both polarizations.

B. Buried Anti-Tank Mine

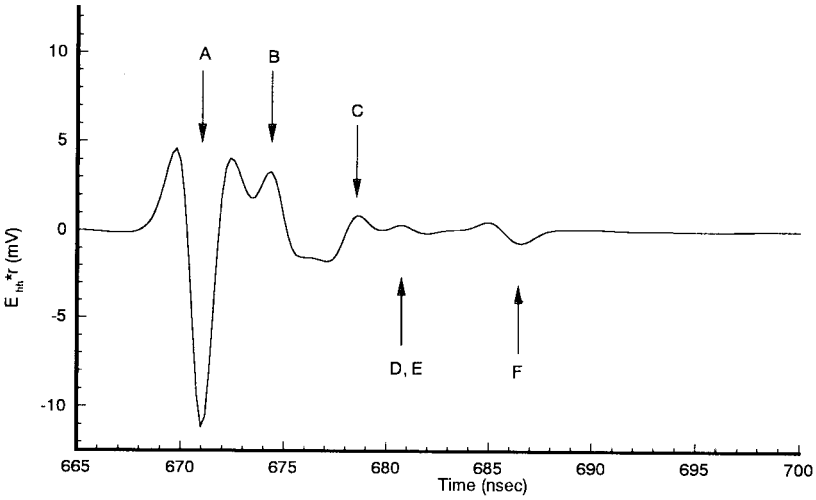
There are various anti-personnel and anti-tank mines: some are metal and others dielectric. Their shapes vary widely. In many cases, mines are conducting and disc-like in shape. As examples, we consider two metal mine-like prototypes which simulate a class of anti-personnel and anti-tank mines. Both mines are modeled as perfectly conducting cylinders with axis of rotation perpendicular to the air-ground interface; the anti-personnel mine has a 6.35-cm diameter and 5.08-cm height, while the respective anti-tank mine dimensions are 38.1 cm and 6.35 cm. We first consider short-pulse scattering from the anti-tank mine, which is usually buried at shallow depths. The incident pulsed plane wave is the same as in Fig. 3. However, we consider incidence at the Brewster angle ($\theta_B = 28.25°$), with soil transmission properties described in Fig. 4.

Results are shown in Fig. 7 for the VV and HH transient fields scattered from the anti-tank mine buried at depths of 2.54 cm, 7.62 cm, 12.7 cm, and 17.8 cm (1, 3, 5, and 7 in, respectively) from the top of the target. The most striking characteristic of these results is the similarity between the VV and HH fields. Apparently the incident pulse does not have sufficient temporal (spatial) resolution to resolve features on the target, so we do not see the distinctive signatures found in Fig. 5 for the much larger 55-gallon drum. Interestingly, initially the peak scattered signal increases as the target depth increases, in contrast with our anticipation. We attribute this phenomenon to reverberations between the top of the target and the air-ground interface that constructively interfere with wavefronts scattered from the target. As expected, as the target depth is further increased, the peak scattered waveform starts to diminish (the reverberated waveform is attenuated and the temporal overlap of the primary and reverberated wavefronts diminishes).

Fig. 7 does not show an obvious late-time resonant signature associated with the buried anti-tank mine. To explain, we plot in Fig. 8 the complex resonant frequencies of the lowest-order SEM mode for the anti-tank mine considered in Fig. 7. For comparison, the resonant frequency of the lowest-order mode is also plotted for the same target situated in a homogeneous medium of the same electrical properties. Notice that the resonant frequency changes as a function of target depth. In Fig. 9 are plotted the normalized modal currents associated with the mode at each depth. There is a noticeable change in the modal current shape as the target depth is adjusted, with the currents concentrating under the target, nearer the high-dielectric soil, as the mine depth decreases. This latter phenomenon is consistent with the well-known concentration of fields in regions of relatively high dielectric constant.



(a)



(b)

Fig. 5. Normalized backscattered fields from a buried perfectly conducting cylinder of 60-cm diameter, 90-cm length, and top surface 30.48 cm from the air-ground interface; the normalization parameter r is the distance from the surface of the halfspace (on the target axis) to the observer. The incident pulsed plane wave has a temporal shape as in Fig. 3, 20° angle of incidence, and the soil parameters are for the 5% case in Fig. 2. (a) VV and (b) HH.

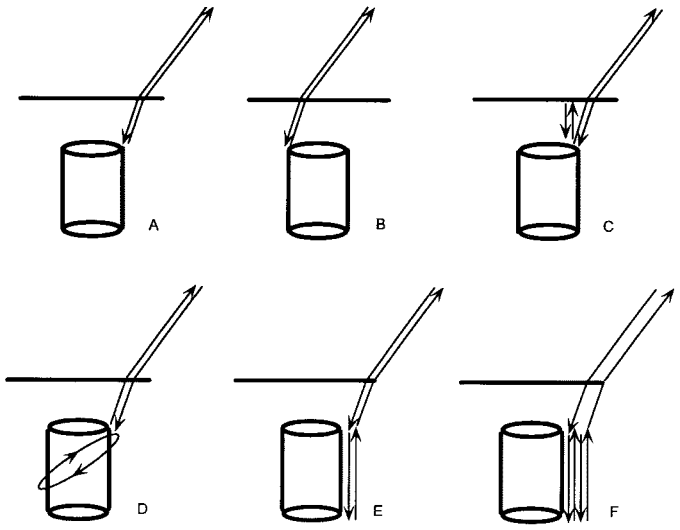
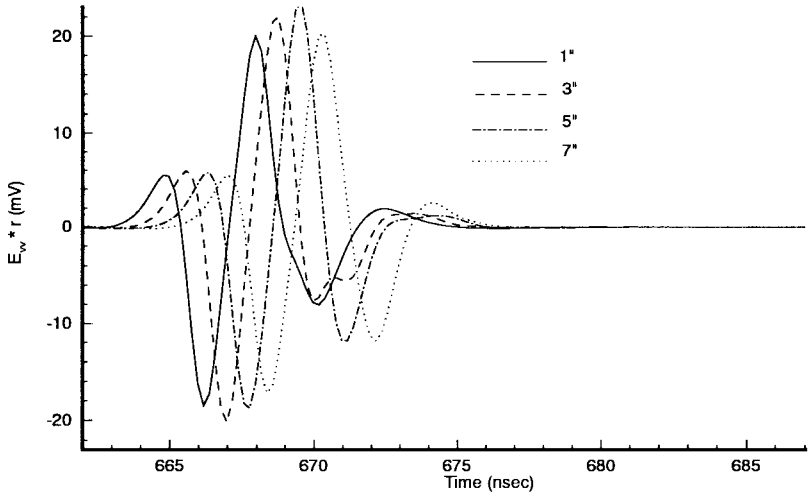


Fig. 6. Schematization of phenomenology characteristic of scattering from a buried cylinder, as considered in Fig. 5. Each scattering mechanism is summarized in the text.

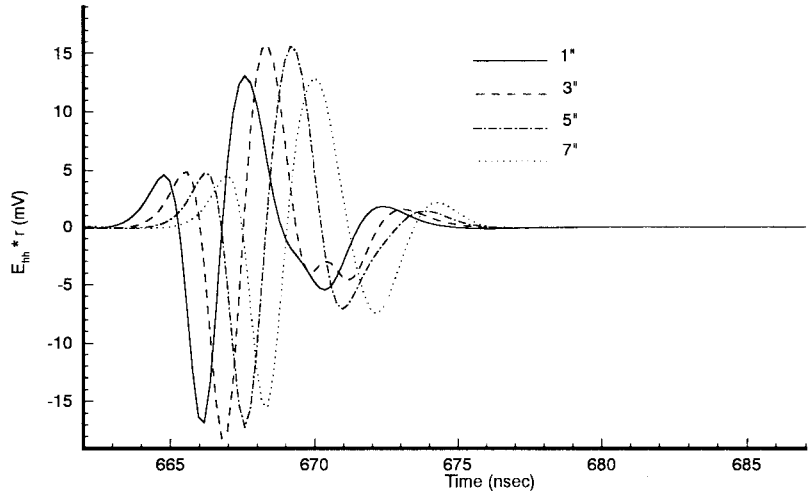
The late-time resonant modes decay with time t as $\exp(-\omega_i t)$, for complex resonant frequency $\omega_r + j\omega_i$. After n periods the resonant fields decay by $\exp(-2\pi n\omega_i/\omega_r)$. Using the computed resonant frequencies from Fig. 8, after only one oscillation ($n = 1$), the resonant signatures corresponding to depths of 2.54 cm, 7.62 cm, 12.7 cm, and 17.8 cm decay by 0.0038, 0.0042, 0.0041, and 0.0026, respectively. These results explain the absence of a discernable resonant signature in Fig. 7, and indicate the extreme difficulty of resonance-based identification for buried targets of the type in Figs. 7–9.

C. Anti-Personnel Mines

Typical anti-personnel mines are significantly smaller than anti-tank mines. Additionally, in many cases such mines are placed on the surface of the ground. Obviously, for such situations no ground penetration is required, and we therefore consider higher operating frequencies, such that enhanced down-range resolution can be achieved. In particular, we consider the same waveform in Fig. 3, except its center frequency



(a)



(b)

Fig. 7. Normalized fields backscattered from a model anti-tank mine buried at depths of 2.54 cm, 7.62 cm, 12.7 cm, and 17.8 cm from the top of the target. The anti-tank mine is modeled as a perfectly conducting cylinder of diameter 38.1 cm and height 6.35 cm, with axis perpendicular to the air-ground interface. The pulsed plane wave is incident at the Brewster-angle and incident-pulse shape, soil properties, and normalization are as in Fig. 5. (a) VV and (b) HH.

is moved to 3 GHz (for a total spectrum of approximately 1–10 GHz).

We consider a pulsed plane wave incident at 20° relative to the air-ground interface, and the perfectly conducting anti-personnel mine is placed atop the surface of the Yuma soil (5% water content). The VV and HH polarized scattered fields are shown in Fig. 10, each characterized by two wavefronts. The time interval between the two wavefronts in Fig. 10 corresponds approximately to round-trip propagation over a distance equal to the mine diameter, with account taken for the oblique angle of incidence. Finally, a major difference between the VV and HH scattered waveforms involves the amplitude of the second wavefront; this issue is addressed further in Section IV, in the context of measured data.

On a numerical note, the electric-field integral equation we use to solve the scattering problem is generally corrupted by artifacts at the internal resonances of a closed target [40], such as that considered in Fig. 10. Due to the high-frequency nature of the incident fields considered in Fig. 10, there are several internal resonances in our spectrum, which affect the scattered fields. This problem has been ameliorated by explicitly enforcing the condition of vanishing electric fields inside the perfectly conducting model mine [40].

IV. UWB-SP SAR Measurements

A. UWB-SP BoomSAR

The Army Research Laboratory (ARL) has been pursuing the use of UWB-SP SAR's for a number of years to help understand foliage penetrating and ground penetrating radar and develop and qualify suitable wide-band models for predicting performance of proposed systems. These radars provide 1 GHz of instantaneous bandwidth and the full polarization matrix to accomplish this task. The current implementation, hereafter referred to as the BoomSAR, is a mobile extension of the original design that allows data collection over a wide range of clutter and target-in-clutter scenarios. The majority of the hardware is mounted to the basket of a telescoping boom lift

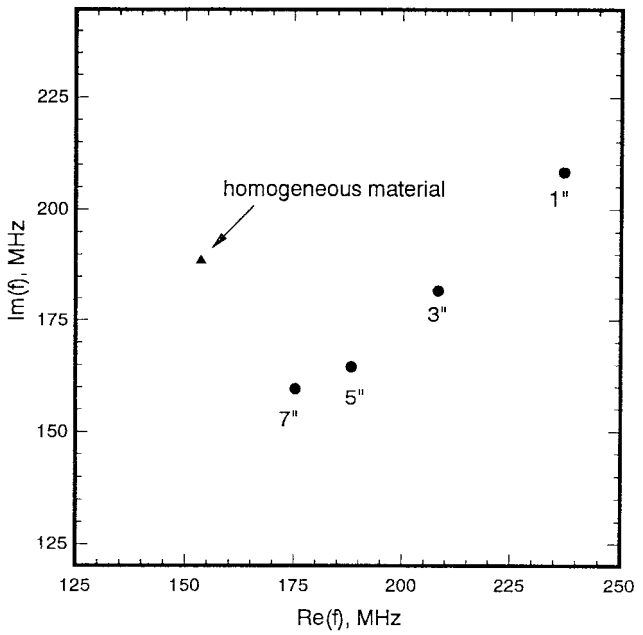


Fig. 8. Complex resonant frequencies of the lowest-order resonant mode for the anti-tank mine considered in Fig. 7.

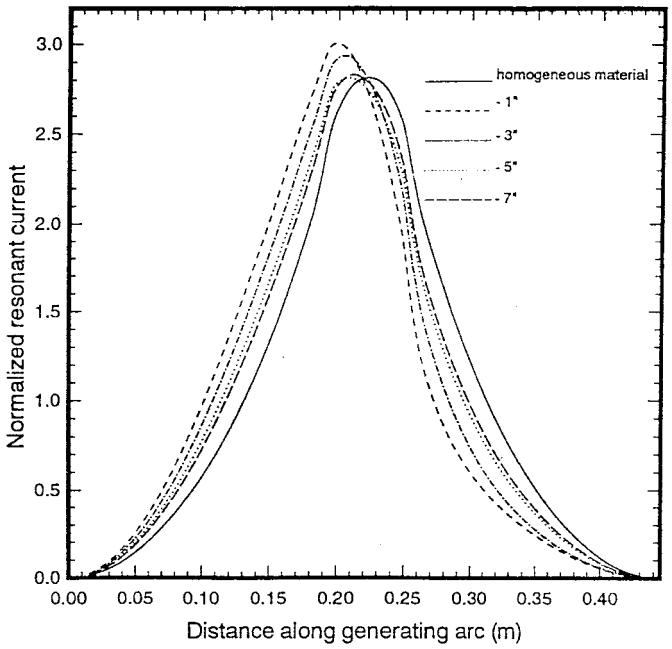


Fig. 9. Normalized modal currents, as a function of target depth, corresponding to the resonant modes in Fig. 8. The currents are normalized such that they integrate to unity when integration is performed along the generating arc of the body of revolution.

capable of moving at approximately 1 km/h while the basket is elevated to 45 m. For typical collection geometries, down-look angles to the target vary from 45° to approximately 10°, depending on the range to the target and the height of the boom.

The radar consists of several major modules. The transmitter is based on a gallium arsenide bulk avalanche semiconductor switch (GaAs BASS). A pair of Power Spectra BASS 02X, pseudo-exponential waveform impulse generators drive the

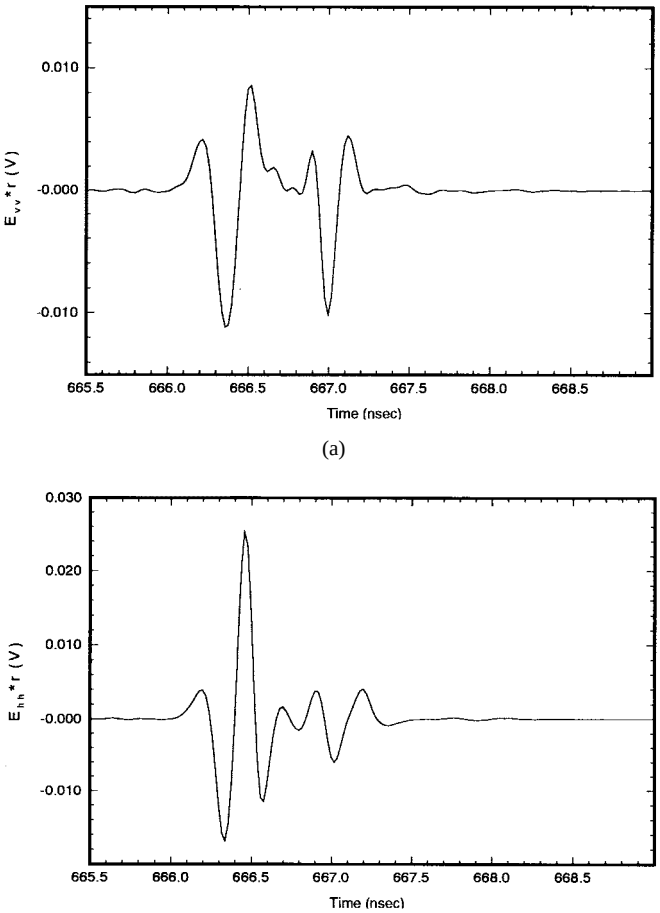
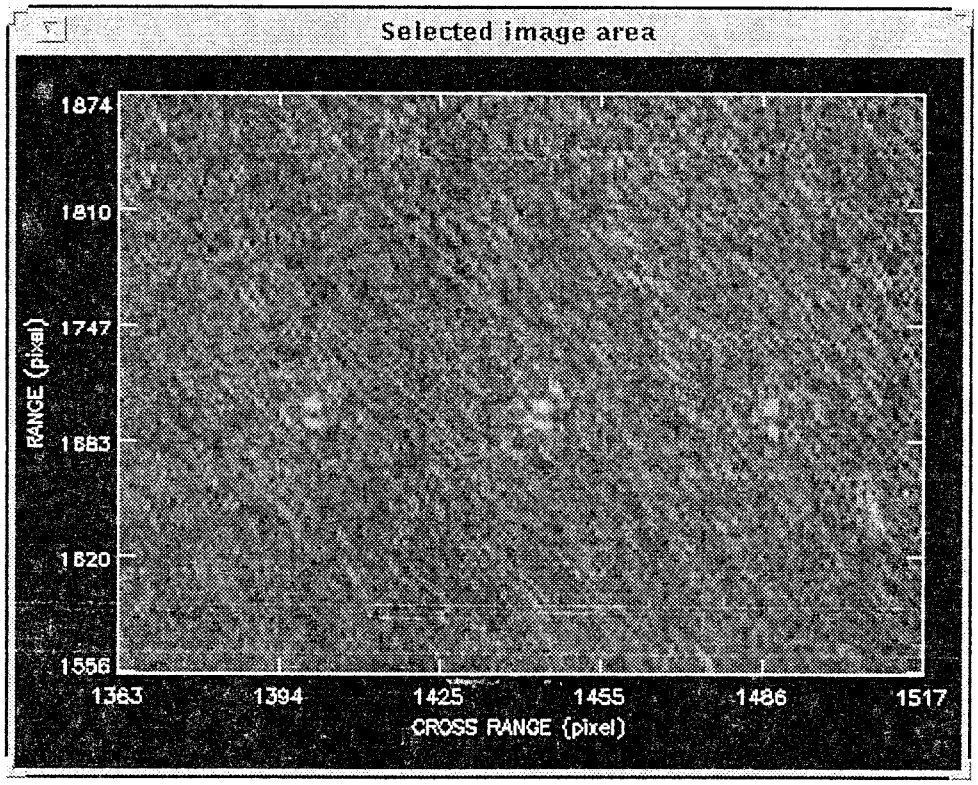


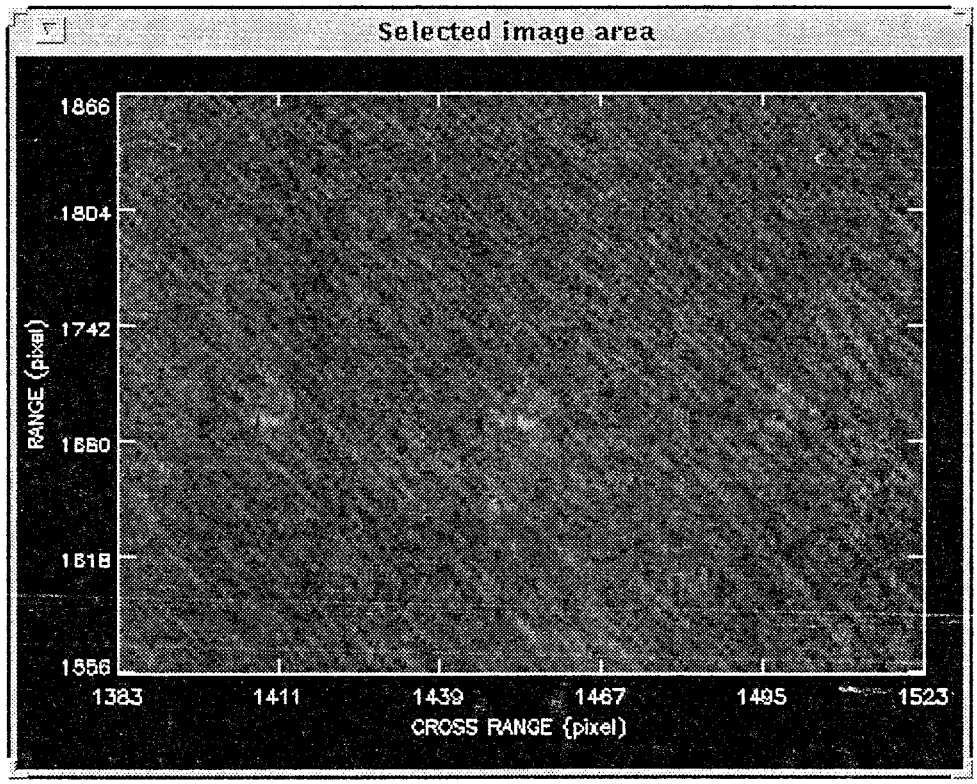
Fig. 10. Normalized fields backscattered from an anti-personnel mine placed on the surface of Yuma soil (see Fig. 2). The model anti-personnel mine is a perfectly conducting cylinder of 6.35-cm diameter and 5.08-cm height, with axis perpendicular to the air-ground interface. The angle of incidence is 20°, the incident pulse is as in Fig. 3 (with its center frequency shifted to 3 GHz), and the time-domain fields are normalized as in Fig. 5. (a) VV and (b) HH.

transmit antennas, allowing fast polarization switching by having the processor select which transmitter to enable. The transmit waveform has a 150-ps risetime and approximately a 2-ns falltime with a peak power of approximately 2 MW. In the current configuration, the radar produces a burst of 128 pulses of one transmit polarization, followed by a burst of 128 pulses of the opposite polarization, and both receive channels are operated in parallel. There are four antennas: two transmit and two receive, to provide the full polarization matrix (HH, HV, VH, VV) in a quasimonostatic sense. The antennas are 200 W open sided, resistively terminated, TEM horns about 2 m in length with a 0.3-m aperture, and are fed by a wide-bandwidth balun.

With the exception of preamplifier/attenuator assembly, which sets the gain and noise figure for the system, the A/D subsystem acts as the wideband receiver for the radar. The A/D subsystem consists of a pair of Tektronix/Analytek VX2005C, 2-GHz A/D converters, and a stable reference clock. A unique feature of these A/D converters is that they provide, to 10-ps resolution, the time difference between the sample clock and trigger event. With the use of this data,



(a)



(b)

Fig. 11. Synthetic aperture radar (SAR) image of three surface mines, as measured by an ultra-wideband BoomSAR. The details of the radar and how the image was formed are described in the text. (a) VV and (b) HH.

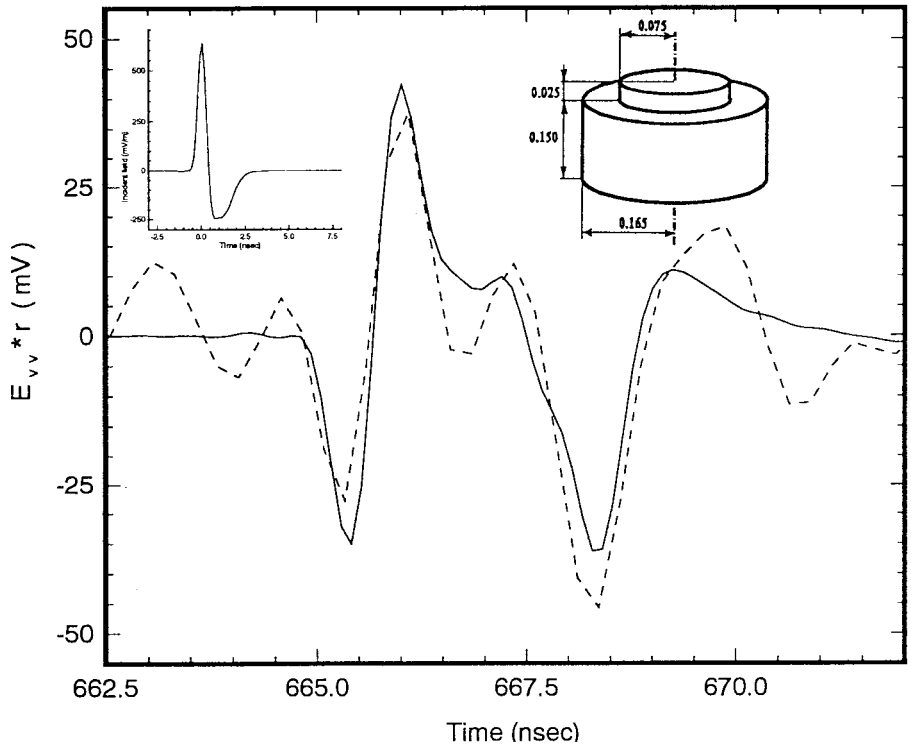


Fig. 12. Comparison between theoretical (solid) and BoomSAR (dashed) scattered fields from a surface anti-tank mine (shown inset). Results are shown for VV polarization and 20° angle of incidence; the shape of the incident pulse (inset) was measured and is characteristic of the waveform transmitted by the BoomSAR.

subsample time interpolation allows the processor to generate an interleaved record at an equivalent 64-GHz rate, much as would be provided by a sampling oscilloscope. The actual high-frequency response of the radar is defined by the 1100-MHz analog bandwidth of the A/D converters, while the low frequency response is determined by the 60-MHz response of the antenna/balun assembly.

As in conventional SAR, the BoomSAR is moved along a straight path, and for each position the scattered response is measured. In its simplest embodiment, the bipolar image amplitude $I_{m,n}$ at pixel position (m,n) is given by

$$I_{m,n} = \sum_{k=1}^{K} s_k(t = T_{k,m,n}) \tag{1}$$

where $s_k(t)$ represents the time-domain scattered field measured at the kth sensor position and $T_{k,m,n}$ is the round-trip time delay between sensor position k and the physical position in the image represented by pixel (m,n). In Section IV-C we use the results of the bipolar image, while for other applications unipolar (magnitude) SAR imagery is generated by Hilbert transforming the bipolar image [37]. The focused image has a resolution of 0.15 m in the range dimension, and 0.3 m in the cross-range dimension.

B. Example Measured Results

We have recently performed an extensive test of our BoomSAR at Yuma Proving Grounds in Yuma, AZ. We present here some relevant results. In particular, we consider surface mines, with geometries similar to that in Fig. 10. However, in those calculations, the center frequency was fixed at 3 GHz, whereas the center frequency of our BoomSAR is approximately 500 MHz. For the target sizes to be comparable, we consider a patch of (larger) anti-tank mines situated on the surface of the Yuma soil.

In Fig. 11 are shown BoomSAR measured images (magnitude) of patches of anti-tank mines, for VV and HH polarization and 20° angle of incidence. Interestingly, the signatures of the mines for VV polarization are characterized by two strong signals, while each of the mines in the HH image is characterized by only one strong response. This disparity in the VV and HH signatures motivated the calculations in Fig. 10. Considering the results in Fig. 10, we see that for VV polarization the model data is characterized by two strong returns, while for the HH data the second wavefront is substantially smaller than the first. We believe these theoretical results explain the measured contrasts in the VV and HH imagery of Fig. 11.

C. Comparison of Computed and Measured Data

The particular example considered here is for scattering from the surface anti-tank mine schematized in Fig. 12. Results are presented for VV polarization, and a 20° incidence angle is considered, consistent with the data in Fig. 10. In Fig. 12 is shown a comparison between the computed (solid) and measured (dashed) data. The measured data represents the coherent superposition of several sensor positions. Therefore, one would anticipate that the imaging process will cause some blurring in the extracted signature. Further, in the theory we assume plane-wave incidence at a fixed angle and with a particular incident-pulse shape; in practice, the incident

wave can only be approximated as planar, and the angle of incidence and incident-pulse shape can only be determined approximately. Nevertheless, we see in Fig. 12 that for VV polarization there is excellent agreement between theory and experiment. No effort was made to optimize the agreement between theory and experiment in Fig. 12; i.e., these results are typical of what we have found for several examples.

V. Conclusions

A new Method of Moments (MoM) numerical algorithm has been utilized to examine the UWB-SP fields scattered from and the resonances of several targets buried in and placed atop soil. Results have been presented particularly for soil at Yuma, AZ, with account taken for dispersion and loss. The targets considered were a 55-gallon drum, an anti-tank mine, and an anti-personnel mine. The theoretical results for an anti-tank mine compared well with measurements performed with a BoomSAR.

Results have only been presented for a small subset of targets and one soil type (albeit a relatively favorable, low-loss soil), but this study further substantiates the difficulty of radar-based detection and identification of buried and surface targets. For example, concerning the resonances of the buried anti-tank mine considered, the low-Q of such resonances, coupled with the soil loss, conspired to produce late-time resonant modes which decay extremely quickly with time, making virtually impossible resonance-based identification for such targets. However, resonant-based detection has proven useful for particular dielectric targets buried in frozen soil [5]. This dichotomy points to the need for modeling—which will yield *a priori* predictions of GPR performance—to assure that the radar is implemented under appropriate circumstances. It is highly unlikely that GPR will be an effective tool for *all* soil and target types, but when utilized properly, it can be an effective option.

Fortunately, the need for accurate modeling intersects with recent algorithmic developments, which now make possible the modeling of scattering from and the resonances of realistic (3-D) targets buried in lossy, dispersive soils. In this paper, we have been concerned with UWB-SP radar, which involves incident signals with over 100% bandwidth. Up until very recently, it was virtually impossible to model the scattering of such waveforms from realistic buried targets, due to the complexity of computing the half-space Green's function (needed for the MoM). However, the recent development of the method of complex images, which efficiently computes the Sommerfeld integrals characteristic of the half-space dyadic Green's function, has been utilized here for several realistic and complicated targets of interest. Future studies will further utilize this algorithm to quantify anticipated radar performance as a function of soil and target type. Additionally, the predicted waveforms from our model can also be utilized in the development of matched filters for target detection.

References

[1] L. Peters, Jr., and J. D. Young, "Applications of subsurface transient radars," in *Time-Domain Measurements in Electromagnetics*, E. K. Miller, Ed. New York: Van Nostrand Reinhold, 1986.

[2] D. L. Moffatt and R. J. Puskar, "A subsurface electromagnetic pulse radar," *Geophys.*, vol. 41, pp. 506–518, June 1976.

[3] M. E. Bechtel and A. V. Alogn, "Antennas and pulses for a vehicular-mounted mine detector," Calspan Corp., Buffalo, NY, Rep. MA-5366-E-1, 1974.

[4] D. J. Daniels, D. J. Gunton, and H. F. Scott, "Introduction to subsurface radar," *Proc. IEE*, vol. 135, pt. F, no. 4, pp. 278–320, Aug. 1988.

[5] L. Peters, J. J. Daniels and J. D. Young, "Ground penetrating radar as a subsurface environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.

[6] R. W. P. King and C. W. Harrison, Jr., "The transmission of electromagnetic waves and pulses into the earth," *J. Appl. Phys.*, vol. 39, pp. 4444–4452, Aug. 1968.

[7] J. A. Fuller and J. R. Wait, "Electromagnetic pulse transmission in homogeneous dispersive rock," *IEEE Trans. Antennas Propagat.*, pp. 530–533, July 1972.

[8] G. S. Smith and W. R. Scott, Jr., "A scale model for studying ground penetrating radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 27, pp. 358–363, July 1989.

[9] N. Osumi and K. Ueno, "Microwave holographic imaging of underground objects," *IEEE Trans. Antennas Propagat.*, vol. AP-33, pp. 152–159, Feb. 1985.

[10] C. Liu and C. Shen, "Numerical simulation of subsurface radar for detecting buried pipes" *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 795–798, Sept. 1991.

[11] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of a ground-penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–44, Jan. 1996.

[12] P. E. Wannamaker, G. W. Hohmann, and W. A. SanFilipo, "Electromagnetic modeing of three-dimensional bodies in layered earths using integral equations," *Geophys.*, vol. 49, pp. 60–74, Jan. 1984.

[13] H. S. Chang and K. K. Mei, "Scattering of electromagnetic waves by buried and partly buried bodies of revolution," *IEEE Trans. Geosci. Remote Sensing*, vol. 23, pp. 596–592, 1985.

[14] G. Kristensson and S. Strom, "Scattering from buried inhomogeneities—A general three dimensional formalism," in *Proc. 1977 URSI Symp. on Electromagnetic Wave Theory*, June 1977.

[15] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, Parts I and II," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–352, Mar. 1990.

[16] S. Vitebskiy and L. Carin, "Moment-method modeling of short-pulse scattering from and the resonances of a wire buried inside a lossy, dispersive half space," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 1303–1312, Nov. 1995.

[17] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 143–151, Feb. 1996.

[18] S. Vitebskiy and L. Carin, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy, dispersive half space," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.

[19] G. S. Smith and W. R. Scott, "The use of emulsions to represent dielectric materials in electromagnetic scale models," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 323–334, Mar. 1990.

[20] W. R. Scott and G. S. Smith, "Measured electrical constitutive parameters of soil as functions of frequency and moisture content," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 621–623, May 1992.

[21] J. E. Hipp, "Soil electromagnetic parameters as functions of frequency, soil density, and soil moisture," *Proc. IEEE*, vol. 62, pp. 98–103, Jan. 1974.

[22] H. L. Bertoni, L. Carin, and L. B. Felsen, Eds., *Ultra-Wideband, Short-Pulse Electromagnetics*. New York: Plenum, 1994.

[23] L. Carin and L. B. Felsen, Eds., *Ultra-Wideband, Short-Pulse Electromagnetics II*. New York: Plenum, 1995.

[24] A. Sommerfeld, "Uber die Ausbreitung der Wellen in der draht Telegraphie," *Ann. Phys.*, vol. 28, pp. 665–736, 1909.

[25] Y. Rahmat-Samii, R. Mittra, and P. Parhami, "Evaluation of Sommerfeld Integrals for lossy half-space problems," *Electromagn.*, vol. 1, no. 1, pp. 1–28, 1981.

[26] E. F. Kuester and D. C. Chang, "Evaluation of Sommerfeld integrals associated with dipole sources above the earth," Univ. Colorado Electromagn. Lab. Sci. Rep. 43, Jan. 1979.

[27] J. R. Wait, "Image theory of a quasistatic magnetic dipole over a dissipative half-space," *Electron. Lett.*, vol. 5, no. 13, pp. 281–282, June 1969.

[28] S. F. Mahmoud and A. D. Metwally, "New image representations for dipoles near a dissipative earth," *Radio Sci.*, vol. 21, pp. 605–616, Nov. 1981.

[29] I. V. Lindell and E. Alanen, "Exact image theory for the Sommerfeld half-space problem, Part III: General formulation," *IEEE Trans. Antennas Propagat.*, vol. AP-32, no. 10, pp. 1027–1032, Oct. 1984.
[30] Y. L. Chow, J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 588–562, Mar. 1991.
[31] J. J. Yang, Y. L. Chow, D. G. Fang, "Discrete complex images of a three- dimentional dipole above and within a lossy ground," *Proc. Inst. Elec. Eng.*, vol. 138, pt. H, no. 4, pp. 319–326, Aug. 1991.
[32] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.
[33] C. E. Baum, "On the singularity expansion method for the solution of electromagnetic interaction problems," Air Force Weapons Lab, Interaction Notes, Note 88, 1971.
[34] M. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. Antennas Propagat.*, vol. AP-23, pp. 777–781, Nov. 1975.
[35] Y. Hua and T. K. Sarkar, "Generalized pencil-of-function method for extracting poles of an EM system from its transient response," *IEEE Trans. Antennas Propagat.*, vol. 37, pp. 229–234, Feb. 1989.
[36] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 38, pp. 814–824, May 1990.
[37] M. A. Ressler and J. W. McCorkle, "Evolution of the Army Research Laboratory ultra-wideband test bed," in *Ultra-Wideband Short-Pulse Electromagnetics II*, L. Carin and L. B. Felsen, Eds. New York: Plenum, 1995, pp. 109–123.
[38] M. G. Andreasen, "Scattering from bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. AP-13, pp. 303–310, Mar. 1965.
[39] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. New York: IEEE Press, 1995.
[40] F. X. Canning, "Robust use of supplementary conditions for moment method solution near internal resonances," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 264–269, Mar. 1995.



Stanislav Vitebskiy (S'94) received the Dipl. Eng. (honors) degree from the Moscow Institute of Electronics Engineering, Moscow, Russia, in 1992, and the M.S. and Ph.D. degrees, both in electrical engineering, from Polytechnic University, Brooklyn, NY, in 1994 and 1996, respectively.

From 1993 to 1995, he was Graduate Research Assistant with the Department of Electrical Engineering, Polytechnic University. From 1995 to 1996, he was Graduate Research Assistant with the Department of Electrical and Computer Engineering, Duke University, Durham, NC. Currently, he is with Harris Corporation, RF Communications Division, Rochester, NY, working in the wireless communications area. His interests also include numerical modeling in electromagnetics and ground-penetrating radar.

Dr. Vitebskiy is a member of Eta Kappa Nu.

Lawrence Carin (S'85–M'86–SM'96) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department at Polytechnic University, Brooklyn, NY, as an Assistant Professor, and became an Associate Professor in 1994. In September 1995, he joined the Department of Electrical Engineering, Duke University, Durham, NC, where he is an Associate Professor. His current research interests include quasiplanar transmission lines, short-pulse scattering and propagation, and signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.



Marc A. Ressler (S'66–M'69) received the B.S. degree in electrical engineering from the University of Maryland, College Park, in 1969 and the M.S. degree in electrical and computer engineering from the University of Michigan, Ann Arbor, in 1973.

He has been employed by the Army Research Laboratory (ARL) and one of its predecessors, the Harry Diamond Laboratories, since 1969, where he worked on electronic fuzes for rockets, laser simulators and instrumentation systems, adaptive antennas, low-power, stand-alone radar, acoustic, and seismic sensor processors, multi-target tracking for the Installation Security Radar, and a microprocessor-based metering system for the electric power industry (for which a patent was awarded). More recently, he has been involved in battlefield data processing and fusion, and a microwave modular radar system that had applications to both airborne and ground-based systems, including the multi-sensor Elevated Target Acquisition System. He is currently assigned to the Sensors and Electron Devices Directorate of ARL where his area of interest is low-frequency, ultra-wideband radar for foliage and ground penetration applications.



Francis H. Le [redacted] He received the B.S.E.E. degree from the University of South Florida, Tampa, in 1986, and the M.S.E.E. degree from the John Hopkins University, Baltimore, MD, in 1991.

From 1987 to 1993, he was a part of the design and development team of the Missile Systems Branch, Army Research Laboratory (ARL), to develop analog and digital proximity fuzes for the Patriot missile system. In 1993, he joined the Microwave Branch at ARL to focus on research and development of the UWB (ultra-wideband) SAR (synthetic aperture radar) for FOPEN (foliage penetration) and GPEN (ground penetration). His work covers UWB SAR data reduction, SAR image algorithms development, signal processing, and GUI tools development.

Time-Domain Sensing of Targets Buried Under a Rough Air–Ground Interface

Traian Dogaru, *Student Member, IEEE*, and Lawrence Carin, *Senior Member, IEEE*

***Abstract*— We consider plane wave time-domain scattering from a fixed target in the presence of a rough (random) surface with application to ground penetrating radar. The time-domain scattering data are computed via a two-dimensional (2-D) finite-difference time-domain (FDTD) algorithm. In addition to examining the statistics of the time-domain fields scattered from such a surface, we investigate subsurface target detection by employing a (commonly used) matched-filter detector. The results of such a detector are characterized by their receiver operating characteristic (ROC), which quantifies the probability of detection and probability of false alarm. Such ROC studies allow us to investigate fundamental assumptions in the matched-filter detector: that the target response is deterministic and the clutter signal stochastic, with the two signals treated as additive and independent.**

***Index Terms*—Buried object detection, time-domain analysis.**

I. Introduction

OVER the last few decades, significant research has been undertaken on the scattering of electromagnetic and acoustic waves from rough (random) surfaces [1]–[9] with the scattered-field properties parametrized statistically. Initial work in this area was based on approximate formulations, while more recently, there has been a significant focus on the application of numerical algorithms. In most of these analytical and numerical analyses, frequency-domain operation has been considered with results presented as the mean and variance of the angular-dependent scattered field. There has been very little work done on time-domain operation (although some results were calculated using time-domain algorithms [2]) or on the characterization of the field scattered from a deterministic target in the presence of a rough (random) interface.

There are many applications for which one would be interested in detecting/identifying a target in the presence of a rough surface. In electromagnetics, one may encounter the problem of detecting/identifying a low-flying aircraft over ocean or a buried target under a rough air–ground interface. Similar problems can be posed in acoustics. In this paper, we consider the case of a deterministic (fixed) target in the presence of a rough surface, the latter being parametrized statistically. This study is motivated by electromagnetic sensing of buried targets [10]–[17] (e.g., mines, unexploded ordnance, etc.), while, as alluded to above, this general problem has wide applicability.

For electromagnetic scattering from the ocean surface, a stochastic surface parametrization is clearly required since the sea-surface characteristics generally change with time in a seemingly random fashion. However, for the sensing of an underground target, the air–ground interface is fixed (deterministic), calling into question the need for a statistical analysis. However, although a given portion of the air–ground interface is fixed, it will, in general, be different from that of another (fixed) region of the interface. Thus, while the fields scattered from any particular portion of the rough surface are deterministic, to characterize the fields scattered from such a general surface, the surface (roughness) parameters must be characterized statistically and, therefore, so must the associated scattered fields. Consequently, to characterize the fields scattered from such a rough air–ground interface, the surface properties are modeled as a random process with any particular surface constituting one realization of an ensemble, each of which is parametrized by the *same* probability density function. In turn, the scattered fields from such a surface are also treated as a random process.

As discussed above, most previous analyses of scattering from rough surfaces have been performed in the frequency domain. However, there has been significant interest recently in time-domain ground-penetrating radar (GPR) systems [10]–[19] motivating the work presented here on time-domain scattering. The time-domain field scattered from a rough (random) surface constitutes a random process, which, in general, is nonwhite. Therefore, for detecting a target in the vicinity of such a rough surface, optimal detectors [20] (under appropriate conditions) are preceded by a whitening filter [20] such that the clutter becomes a white (uncorrelated) time sequence. The whitening filter employs the clutter's correlation matrix, which for wide-sense stationary (WSS) clutter can be represented in the Fourier domain in terms of the clutter's power spectral density (PSD). As discussed below, in the cases we have examined thus far, the WSS model has been found to be quite accurate and, therefore, here we quantify the transient clutter statistics via the PSD.

Although the statistical properties of time-domain scattering from a rough (random) surface are of interest in their own right and are investigated here in the context of their PSD, the principal focus of the present paper is a detailed examination of the underlying assumptions used in classical optimum detectors with emphasis on time-domain GPR. Matched filters preceded by a whitening filter are commonly utilized for the

Manuscript received February 7, 1997; revised October 21, 1997. This work was supported in part by the Army Research Office under Grant DAAH04-96-1-0448, Demining MURI, by the Army Research Laboratory, Adelphi, MD, by the Air Force Office of Scientific Research under Grant F49620-94-1-0363, and by the Office of Naval Research under Grant N00014-96-1-0861.

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA.

Publisher Item Identifier S 0018-926X(98)02258-3.

0018–926X/98$10.00 © 1998 IEEE

detection of a known target in noise/clutter. For the case of GPR, such detectors are useful because there are generally a relatively small number of target types (e.g., mines) of interest, the signatures of which can in principle be measured or computed. In its simplest (and most common) manifestation, such a matched filter assumes that the target signature is deterministic and known exactly with the clutter/noise treated as stochastic and additive. While such properties result in a simple detector, in practice they are often not rigorously valid. To test such a detector and its underlying assumptions, our results are presented in the form of receiver operating characteristic (ROC) curves, which describe how the probability of false alarm varies with the probability of detection (as detailed below, these probabilities are varied by adjusting the detector threshold). Using classical detection theory [20], *theoretical* ROC curves are computed based on the clutter statistics. Monte Carlo simulations are then performed for a target in the presence of a random surface to get "*empirical*" detector performance. Where the theoretical and empirical ROC curves agree, we deem the underlying detector assumptions valid. Of interest is the discrepancy in the theoretical and empirical ROC curves as a function of target position, angle of incidence, and surface statistics. Such studies are important for when the simplifying assumptions are inappropriate, more sophisticated detectors, which, for example, exploit the random character of the target signature [20], may yield better performance than the simple matched filter (at the price of greater complexity).

To perform the aforementioned study, many realizations of the air–ground interface must be considered (each described by the same probability density function) and, therefore, the numerical scattering model must be versatile and highly efficient. Consequently, all results are computed using a two-dimensional (2-D) finite-difference time-domain (FDTD) algorithm [21]–[23], which incorporates a lossy half space, surface roughness, plane-wave incidence, a near-to-far-zone transformation, and a perfectly matched layer (PML) absorbing boundary condition (ABC) [24]–[29]. Acquiring statistics from a three-dimensional (3-D) surface would require prohibitive computational resources; however, it is felt that the basic issues studied here using a 2-D model are also relevant to the more realistic case of 3-D GPR systems.

The remainder of the text is organized as follows. A brief discussion of the FDTD algorithm is given in Section II with careful attention directed on the near-to-far-zone transformation. The matched-filter detector and whitening filter are discussed in Section III, wherein basic underlying assumptions are summarized. Results are presented in Section IV for several realizations of a target in the presence of a rough (random) air–ground interface, presented in the form of PSD's for the clutter and ROC curves for the matched-filter detector. This work is summarized and conclusions are given in Section V.

II. Numerical Scattering Model

A. Basic Numerical Algorithm

All numerical scattered-field data are calculated via a 2-D FDTD code, utilizing the classic Yee algorithm [21]. Results

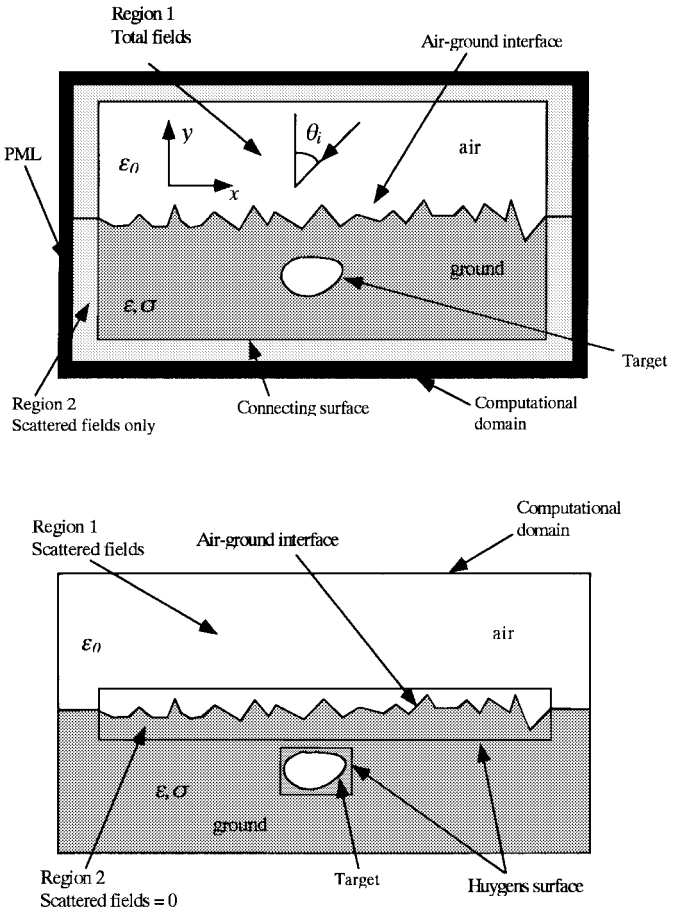


Fig. 1. Schematic of (a) FDTD computational model and (b) the Huygens surfaces used to calculate the far-zone scattered fields.

are presented for TM_z polarization $(H_z = 0)$ with similar results found (but not presented) for the TE_z $(E_z = 0)$ case (see Fig. 1). As described in [25] and [26], the plane wave incident field is modeled via the use of Huygens currents placed along a closed contour with total fields inside and scattered fields outside. We refer to "the incident field" as the plane wave that impinges on the half-space plus the fields reflected and transmitted at the planar interface in the absence of any target or surface roughness. The scattered field is produced by any perturbations to this half-space problem (target and/or surface roughness).

At the boundaries of the computational domain, we use a PML [27] to absorb the outwardly propagating scattered fields. It has been well established that the Berenger [27] PML yields a reflection coefficient at the boundary of the computational domain that is orders of magnitude smaller than absorbing boundary conditions derived from the one-way wave equation [30]. However, the PML introduced by Berenger in his original paper cannot be applied directly to lossy media and, therefore, we use a PML formulation based on [28] and [29] appropriate for lossy soils.

The far-zone scattered fields are the quantities of interest in most radar problems. Therefore, we implement a near-to-far-zone transformation in a manner similar to [26]. In particular, the perturbations (target and/or rough surface) are enclosed by Huygens surfaces and the equivalent currents thereon

(minus the incident field) are integrated with an analytic approximation [31] to the half-space Green s function. As shown in Fig. 1, separate near-to-far-zone Huygens surfaces are used for the rough surface and the target (when present). The fields are observed in backscatter for the near-grazing case applicable to many airborne subsurface radars [18].

Near-to-far-zone transformations have been used for many years to propagate FDTD computed near-zone fields into the far zone. However, much of this work has been for free-space scattering with far fewer results presented for the half-space problem [26] considered here. Therefore, we elaborate on several points germane to the half-space problem recognized in the course of this research. First, as discussed above, the far-zone fields are computed via an approximate half-space Green's function derived via an asymptotic evaluation of its exact Sommerfeld-integral representation [31]. In all such asymptotic analyses, one must first identify the highly oscillatory and relatively slowly varying portions of the integral. The Sommerfeld integrals are evaluated approximately via saddle-point integration in the vicinity of its stationary point(s) [31], the latter determined by the phase of the highly oscillatory terms [31]. For evaluation of the scattered fields above the half-space far from the interface, one usually treats the exponentials characteristic of propagation in air as the highly oscillatory terms and the remaining exponentials that describe propagation in the dielectric half-space as slowly varying. However, if the currents in the dielectric are at electrically large depths, a more accurate asymptotic representation of the Green s function may be necessary. In particular, some of the aforementioned terms may no longer be deemed slowly varying in which case they must be included when determining the stationary point(s) [32]. This issue is particularly relevant for GPR applications in which the dielectric constant of the soil is often large and, therefore, the target need not be too deep physically for the above issues to be of concern (this is especially true for the wide-band signals of interest here).

A second issue involves numerical dispersion [33], [34]. As the time-domain fields propagate through the FDTD grid, there is a inherent (nonphysical) pulse distortion caused by the difference-equation approximations to Maxwell s equations [33], [34]. This issue is exacerbated for electrically large structures, of interest for the high-dielectric-constant half-space region. Numerical dispersion is mitigated by increasing the spatial sampling in the FDTD grid, but the physical size of the problem and finite computer memory ultimately dictate the discretization limits. Therefore, to improve the accuracy of the far-zone fields, we have found it useful to place the Huygens surface as close as possible to the target to minimize the range over which FDTD computed scattered fields must propagate (of course, similar concerns are relevant for the incident-field Huygens surface as well).

For the target depths and soil properties considered in Section IV, the simple Green's function approximation has proven very accurate after performing exhaustive tests. However, similar tests also showed numerical dispersion to be a concern, requiring one to be very careful about the spatial discretization and requiring the Huygens surface to circumscribe the target (and rough surface) tightly. In the results presented below, the spatial discretization was 12 samples per wavelength (in the soil), at the highest frequency of interest in the incident fields. Petropoulos [33], [34] has demonstrated that the required sampling rate is predicated by both the dielectric constant *and* the length over which the FDTD fields must propagate. We are confident in the FDTD and near-to-far-zone results presented below, but it must be emphasized that this confidence has been obtained after a rigorous examination of numerical dispersion and its pitfalls.

B. Rough-Surface Generation

For the GPR problem, clutter can be generated from multiple sources, including surface roughness and/or subsurface inhomogeneities (e.g., rocks, roots, inhomogeneous soil, etc.). In this paper, we restrict ourselves to clutter generated by surface roughness under the assumption that the subsurface is characterized by a homogeneous lossy dielectric. Although other or additional forms of clutter may occur in practice, there has been very little work done to date on wide-band time-domain scattering from rough surfaces or from targets situated in such environments. Moreover, as discussed in Section V, the general insight from this problem may yield an appreciation for the physics inherent in more general clutter scenarios.

We assume that the surface height $f(x)$ at each transverse location x is a zero-mean Gaussian random variable with heights $f(x)$ and $f(x+\Delta x)$ related by the autocorrelation

$$E[(f(x)f(x+\Delta x)] = h^2 \exp(-\Delta x^2/l^2) \tag{1}$$

where h^2 is the variance, l is the correlation length, and $E(\cdot)$ represents the expectation operator. In actual implementation, the random process is generated in the Fourier domain [7] by passing a Gaussian white-noise process through a filter with a spatial-frequency response corresponding to the desired rough-surface power spectrum [the Fourier transform of (1)].

As described in Secion II-A, we calculate the scattered fields using a standard, 2-D Yee FDTD algorithm [21]. In such a scheme, the rough surface must be approximated in a staircase fashion [2], [21] due to the Cartesian spatial discretization. However, it has been demonstrated that if the spatial gridding is sufficiently fine, the scattered fields are in very good agreement with those calculated via a more sophisticated FDTD algorithm in which the surface roughness is modeled using a contour-integral procedure [2].

III. Time-Domain Statistics and Detection

A. Time-Domain Statistics

Frequency-domain scattering from rough (random) surfaces has been characterized in terms of the mean and variance of the scattered fields [1]–[9]. This parametrization is natural since at a given frequency the scattered field is a random variable. However, in the time domain, the scattered field is a time-dependent random process. For all the cases we have considered in the course of this work using the surface roughness in (1), the time-domain fields scattered from the rough surface are approximately wide-sense stationary (the

correlation matrix [35] computed via averaging hundreds of Monte Carlo realizations has Toeplitz symmetry with less than 5% error). Therefore, we describe the statistical properties of the rough-surface scattered field in terms of its autocorrelation [35]. One drawback of this procedure is that the autocorrelation is dependent on the incident-pulse temporal shape $g(t)$, viz., if $r_\delta(t)$ represents the autocorrelation of the scattered fields when the incident pulse shape is a delta function, then the autocorrelation of the scattered field for the input-pulse $g(t)$ is

$$r_g(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\gamma\, d\beta\, g(\gamma)g(\beta)r_\delta(t-\beta+\gamma). \quad (2)$$

The power-spectral density (PSD) of the scattered field $S_g(\omega)$ is represented by the Fourier transform of $r_g(t)$ and is expressed as

$$S_g(\omega) = |G(\omega)|^2 S_\delta(\omega) \quad (3)$$

where $S_\delta(\omega)$ and $G(\omega)$ are the Fourier transforms of $r_\delta(t)$ and $g(t)$, respectively. Therefore, although the autocorrelation and PSD of the field scattered from a rough (random) surface are dependent on the incident-pulse shape, with knowledge of $G(\omega)$, one can calculate the incident-pulse-independent PSD $S_\delta(\omega)$ characteristic of the clutter alone.

B. Optimal Detector

In addition to studying the transient fields scattered from a rough (random) surface, a major focus of this investigation involves the detection of a fixed target in the vicinity of such a surface. As discussed in Section I, this is a basic problem of interest for many applications. Detection theory has been investigated for decades [20] and, therefore, the purpose here is not to develop a new detector. However, in classical detection theory, one often makes assumptions that simplify detector design. In practice, such detectors often do not work as well as expected and it is of significant interest to understand why, such that improved detectors can be developed, if necessary. The FDTD allows us to perform controlled experiments with which underlying detector assumptions can be investigated systematically here with application to GPR.

A concise summary of simple detection theory [20] is given, such that the underlying assumptions are understood for subsequent examination. We consider a binary test for hypotheses H_0 and H_1, where H_0 states that the scattered signal consists of clutter alone and H_1 that the signal consists of the *superposition* of clutter and the *deterministic* response from a known target. Implied in this test are the assumptions that: 1) there are only two possibilities for the source of the scattered field (i.e., that the binary hypothesis test is valid); 2) the target and clutter signatures are additive; and 3) the target has a deterministic signature. Assumption 1) constitutes a simplification of the general problem of multiple hypothesis testing [20] and allows a direct examination of 2) and 3). Our numerical experiment is easily designed such that 1) is valid. Therefore, we undertake a detailed examination of 2) and 3), which are of importance for multiple hypothesis testing as well.

Under the above hypotheses, a given polarization of scattered field $e(t)$ can be expressed as $e(t) = c(t)$ under hypothesis H_0 and $e(t) = E^{1/2}s(t) + c(t)$ under H_1, where $E^{1/2}s(t)$ is the *known deterministic* target signature (with energy E) and $c(t)$ is the *stochastic* clutter characterized by a joint probability density (in the remainder of the text when E appears alone, it represents the energy of the deterministic scattered field and when it appears as $E(\cdot)$ it represents expectation). If the clutter $c(t)$ is white ($E[c(t_i)c(t_k)] = \delta_W^2\delta(t_i - t_k)$, where σ_W^2 is the variance if $c(t)$ is a zero-mean process, then the optimal detector is a matched filter [20], [36]. In general, the clutter is not white and the optimal detector consists of a matched filter preceded by a whitening filter [20], [37]. If $h_w(t)$ represents the impulse response of the linear time-invariant whitening filter, then the new noise response $c_*(t)$ and signal response $s_*(t)$ after whitening are

$$c_*(t) = \int_{-\infty}^{\infty} c(\tau)h_w(t-\tau)\, d\tau$$
$$s_*(t) = \int_{-\infty}^{\infty} s(\tau)h_w(t-\tau)\, d\tau. \quad (4)$$

If hypothesis H_0 is true, the filter output is $e_*(t) = c_*(t)$, while if H_1 is true, $e_*(t) = E^{1/2}s^*(t) + c_*(t)$. Finally, the output of the whitening filter is projected onto the function $E^{1/2}s_*(t)$ resulting in the random variable l

$$l = \int_{-\infty}^{\infty} \sqrt{E}s_*(t)e_*(t)\, dt. \quad (5)$$

If $p(l|H_0)$ is the probability of l under hypothesis H_0 and $p(l|H_1)$ is the probability of l under H_1, the decision as to whether a target is present is effected via the statistic

$$\Lambda = \frac{p(l|H_1)}{p(l|H_0)} = \frac{\exp\left[-\dfrac{(l-m)^2}{2\sigma^2}\right]}{\exp\left[\dfrac{-l^2}{2\sigma^2}\right]} \quad (6)$$

where the right side of (6) is valid if $c_*(t)$ is a zero-mean Gaussian process (it is approximately white as a consequence of the whitening filter) with $m = E^{1/2}$ and σ^2 representing the variance of $c_*(t)$. For a given threshold T, we choose H_1 if $\Lambda > T$ and H_0 if $\Lambda < T$ [20]. If $c_*(t)$ is a zero-mean white Gaussian process, it is easy to see from the right side of (6) that the statistical properties of Λ are dependent entirely on the statistics of l and, therefore, the latter is termed a sufficient statistic [20] (and l alone can be compared against a threshold).

For the surface characteristics in (1) and the examples studied here, we have found $c(t)$ to be well characterized by a Gaussian density function (this is not surprising, as dictated by the central limit theorem [35]). Therefore, we specialize the discussion to the Gaussian case. The detector performance is parametrized by computing the probability of a false alarm at threshold T

$$P_F(T) = \Pr(l > T|H_0) = \int_T^{\infty} p(l|H_0)\, dl \quad (7)$$

as well as the probability of detection

$$P_D(T) = \Pr(l > T|H_1) = \int_T^{\infty} p(l|H_1)\, dl. \quad (8)$$

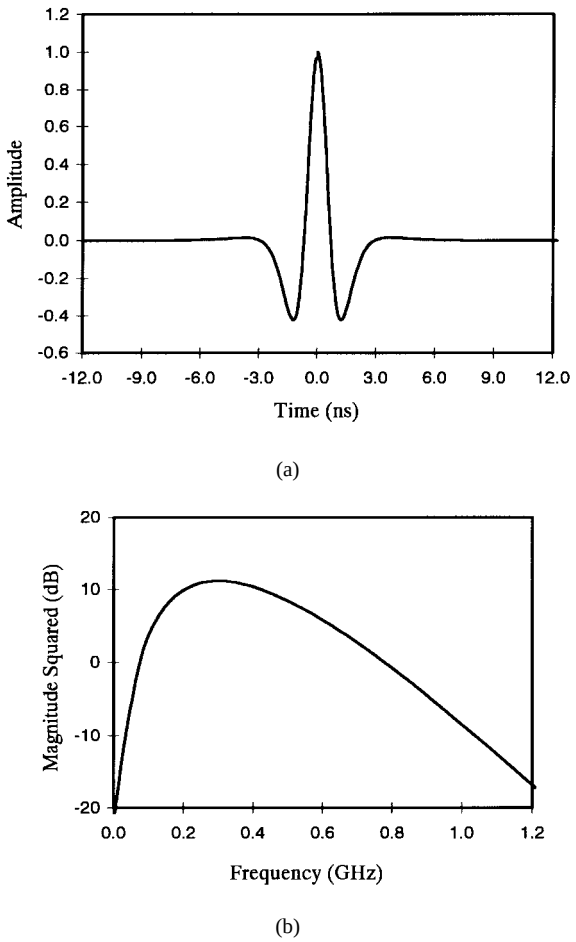
Detectors are usually characterized by plotting the P_F and P_D as a function of the threshold T and this representation is termed the receiver operating characteristic (ROC) [20]. One can adjust the parameter T to operate the detector at a P_F and P_D that is appropriate for a given application.

Summarizing, we will use the classical (whitening-filter)-(matched-filter) detector discussed above to investigate the radar detection of subsurface targets. Other more-complicated detectors could be used if the aforementioned assumptions breakdown, although this results in a significant escalation in detector complexity [20]. With regard to the results presented below, the whitening filter is implemented via a forward (linear-prediction) error filter [37], the parameters of which are computed by solving the classical Wiener–Hopf equations [37], which require the autocorrelation matrix of the clutter $c(t)$ (in the absence of the target). The autocorrelation matrix was computed by averaging the scattered fields from 300 realizations of the rough surface, each of which is described by the same rough-surface density function.

IV. Results

For all examples presented here, the excitation is in the form of an obliquely incident plane wave with incident pulse shape and spectrum shown in Fig. 2. This waveform is representative of signals generated routinely by ultra-wide-band SAR systems [17]–[19]. While we consider various statistical models for the rough-surface properties, the electrical characteristics of the soil itself are kept fixed. In particular, the soil is modeled with a frequency-independent dielectric constant $\epsilon_r = 6$ and conductivity $\sigma = 0.005$ S/m. While this model does not correspond to any particular soil, it is consistent with data measured for various soil types of interest [38]. Note that although the dielectric constant and conductivity are frequency independent in this model, the incident and scattered waveforms are dispersed as they propagate through the soil due to the frequency dependence of the index of refraction $n(\omega) = [\epsilon_r - j\sigma/\omega\epsilon_o]^{1/2}$. Thus, the equivalent (Huygens) currents used to model the incident plane wave were computed in the frequency domain before Fourier conversion to the time-domain form used in the FDTD algorithm. The 2-D target considered has a square cross section (37.5 cm $\times$ 12.5 cm) and is composed of a homogeneous lossless dielectric. Moreover, the top of the target is buried 25 cm beneath the mean position of the rough air–ground interface. The dielectric target was considered because it presents a particularly challenging case for radar systems.

In the following results, the rough-surface profile in (1) has been used, with various parameters for the standard deviation of the surface height h and the correlation length l. Moreover, two angles of incidence are considered. Finally, with regard to the target, although its shape and depth are fixed, we consider several examples for its relative dielectric constant. The purpose of this study is to explore basic target (rough surface) interactions and their impact on the performance of a common detector (the matched filter). We note, however, that there are clearly other possibilities for the surface roughness statistics, target type, and soil type. Such issues will be considered in future work.



(a)

(b)

Fig. 2. (a) The incident pulse and (b) its spectrum used in all scattered-field computations.

Before proceeding, recall that for the examples considered here, we have assumed plane wave incidence. This is motivated by the SAR problem [17]–[19] for which the sensor is quite distant from the region under interrogation. Other authors have considered a tapered excitation [2] in the FDTD computations, which may be representative of the antenna pattern for some applications; moreover, such tapers mitigate numerical difficulties at the edges of the rough surface from which fields are diffracted in a manner uncharacteristic of the surface itself. While this latter issue is important for the computation of rough-surface statistics, it is not necessary for the detection examples considered here. In particular, note from (5) that the decision test statistic l is computed by projecting the total measured waveform $e_*(t)$ onto $s_*(t)$—the latter representing the target signature for a *flat* surface. Therefore, nonzero values of $s_*(t)$ only exist over a very limited temporal support (for the targets considered here) and, therefore, the only portion of $e_*(t)$ of interest for computation of l is that for which $s_*(t)$ is nonzero. In our computations, we have carefully examined diffractions from the edges of the rough surface and these occur well outside the meaningful temporal support of $e_*(t)$. We note as well that the clutter statistics (used in the whitening filter) were also computed by considering those portions of $c_*(t)$ with support around

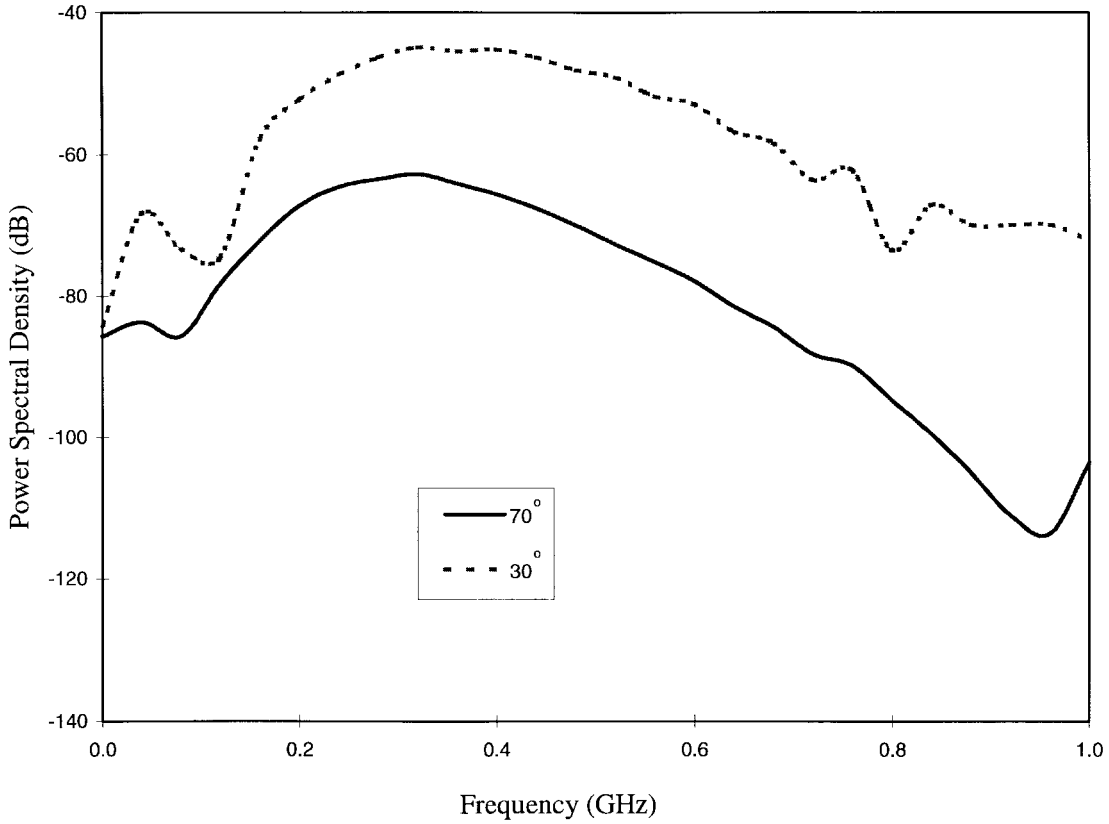


Fig. 3. Power spectral density (PSD) of the transient fields backscattered from the rough surface in (1), using the incident pulse in Fig. 2. The Gaussian rough surface has correlation length $l = 18.75$ cm and standard deviation $h = 3.95$ cm and the soil is characterized by $\epsilon_r = 6$ and $\sigma = 0.005$ S/m. Results are plotted for angles of incidence $\theta_i = 30°$ and $70°$.

the principal portion of $e_*(t)$. Therefore, edge-diffraction-induced vitiation of the detector computations was avoided by exploiting the temporal filtering of the short-pulse time-domain excitation.

A. Clutter Characteristics

As discussed in Section III-A, the clutter statistics are characterized by their power spectral density (PSD)—assuming the clutter is wide-sense-stationary, as we have found to be approximately the case for all examples considered here. In Fig. 3 we address the variation in the PSD with respect to the angles $\theta_i = 30°$ and $70°$ (see Fig. 1), for a surface roughness characterized by $h = 3.95$ cm and $l = 18.75$ cm. The most striking feature of these results is the significant decrease in the strength of the PSD as one gets closer to grazing, which is expected due to the fact that the backscattered fields excited by the rough surface diminish in strength as one approaches grazing. As a comparison with the PSD's, we consider the spectrums of waveforms backscattered from a buried target, for the same angles considered in Fig. 3. In Fig. 4, we plot the angular-dependent backscatter spectrums for a 37.5 cm $\times$ 12.5 cm target of $\epsilon_r = 2$ buried 25 cm under a *flat* air–ground interface. For the target considered here, we see from Fig. 4 that the *shape* of the backscatter spectrum is relatively insensitive to angle, while, as in Fig. 3, there is a pronounced reduction in its strength as one approaches grazing. To avoid problems inherent in deconvolving the incident pulse, the spectrums in Figs. 3 and 4 represent delta-function responses filtered by the incident pulse in Fig. 2 [see discussion concerning (2) and (3)].

The remaining results will be presented in the form of ROC curves, which depict the variation in the probability of detection P_D and probability of false alarm P_F as a function of the detector threshold T (see Section III-B). For the case of *additive* white Gaussian noise, detector performance (P_D and P_F) is dictated entirely by the parameter $d = m/\sigma$ [20], where from Section III-B m^2 is the signal energy and σ^2 is the variance of the clutter. Consequently, a reduction in the strength of the backscattered signature *does not* necessarily imply that detector performance will diminish. In fact, if the clutter variance reduces as well, then it is possible that the ratio $d = m/\sigma$ may actually increase (note in Figs. 3 and 4 the commensurate decrease in the clutter and target responses with increasing angle of incidence). As demonstrated below for the examples considered here, $d = m/\sigma$ increases with increasing angle of incidence (although individually m and σ decrease) yielding improved detector performance. However, as m and σ decrease (near grazing) other sources of noise (e.g., receiver noise, radio-frequency interference (RFI), etc.) may become dominant, which may deteriorate detector performance in practice. Such issues are not addressed here, as we concentrate on the effects of rough-surface-induced clutter on buried-target detection.

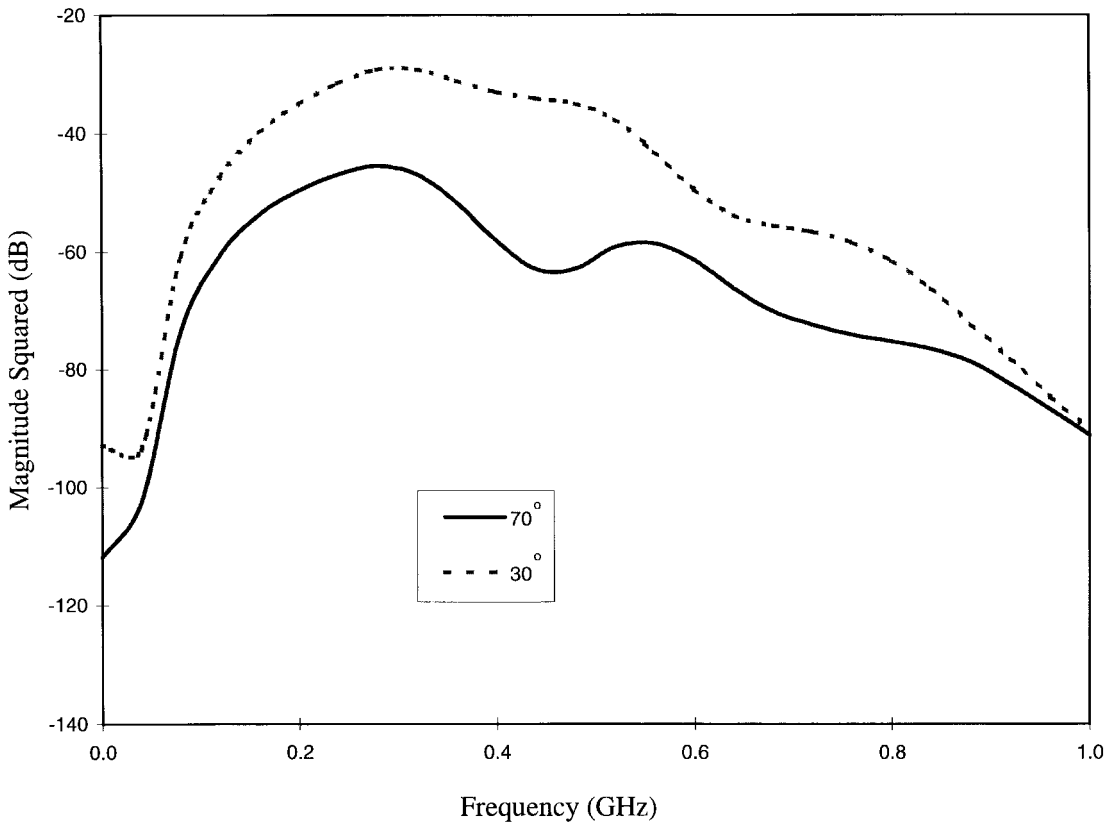


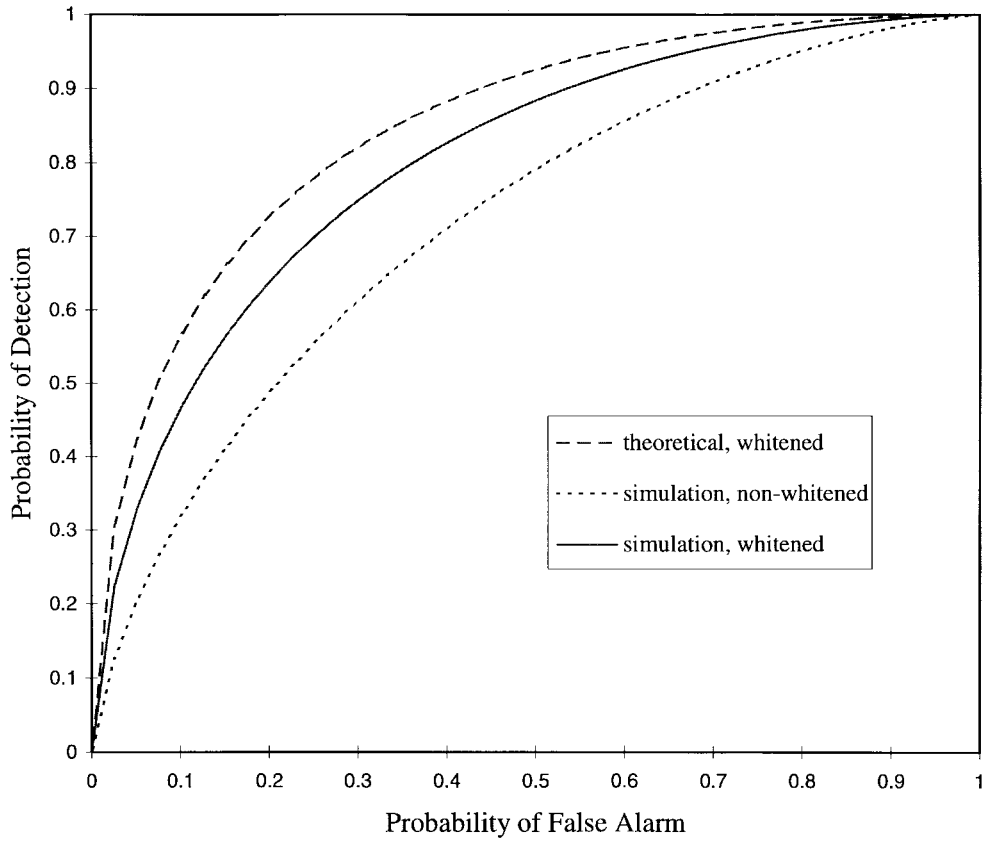
Fig. 4. Spectrum of the fields scattered from a 37.5 cm × 12.5 cm lossless dielectric target ($\epsilon_r = 2$) buried 25 cm beneath a flat air–ground interface, with the soil properties $\epsilon_r = 6$ and $\sigma = 0.005$ S/m. Results are plotted for angles of incidence $\theta_i = 30°$ and $70°$.

B. Detector Performance After Prewhitening

We have developed a (linear-predictor) error filter (ten taps) to whiten the clutter response [37]. Using the concepts in (7) and (8), we can calculate *theoretical* ROC curves, under the assumption that the detector model fits the data in question. From Section III-B, the principal detector assumptions are that the target response and the clutter are additive and that the target signature is deterministic. Under these assumptions, (7) and (8) demonstrate that the whitening filter does, in theory, improve detector performance relative to the similar computations for the case of no whitening. However, in the context of our Monte Carlo simulations from which we compute simulated (empirical) ROC performance (using the FDTD computed scattered fields in which the clutter need not be additive nor the target response deterministic), we have not seen the significant whitening-induced performance enhancement predicted by (7) and (8). The degree of improvement in Monte Carlo computed ROC performance increases as the angle of incidence decreases (as the wave approaches normal incidence), but not to the degree that (7) and (8) predict (we use 300 Monte Carlo realizations for a given surface-roughness statistics).

To demonstrate whitening-induced ROC improvement as a function of incidence angle, in Fig. 5 we reconsider a 37.5 cm × 12.5 cm target of lossless dielectric constant $\epsilon_r = 2$ buried 25 cm under the rough surface in Fig. 3. From the results in Fig. 5(a), we see that for $\theta_i = 30°$ the Monte Carlo simulations reveal a whitening-filter-induced ROC enhancement, which is relatively close to that found, theoretically, from (7) and (8). On the other hand, for the near-grazing case of $\theta_i = 70°$ [Fig. 5(b)], the Monte Carlo simulations demonstrate whitening-filter-induced enhancement that is substantially less than that expected theoretically. However, it should be pointed out that the absolute Monte Carlo predicted detector quality is better near grazing $(\theta_i = 70°)$ than it is closer to normal incidence $(\theta_i = 30°)$ despite the fact that the enhancement due to prewhitening is not as great as expected theoretically.

One can understand the discrepancy between the simulated (Monte Carlo) and theoretical [(7) and (8)] ROC performance by recognizing that the field *incident* on the target (after transmission through the rough surface) is a random process and, therefore, so are the scattered fields. Similarly, there is additional randomization induced as the scattered fields propagate out of the soil and into air through the rough surface. One can parametrize the cumulative effect of this random transmission in and out of the soil as a random process, added onto the (deterministic) response of the target (calculated when the air–ground interface is flat). This random process need not have the same statistical properties as the fields backscattered from the rough surface (in the absence of the target) with which the whitening filter was designed (in practice, there will be surface-scatter-induced clutter *and* the aforementioned random-transmission-induced clutter). The mismatch between the actual clutter and that with which the whitening filter was designed will clearly deteriorate detector performance. Alternatively, instead of viewing the random transmission



(a)

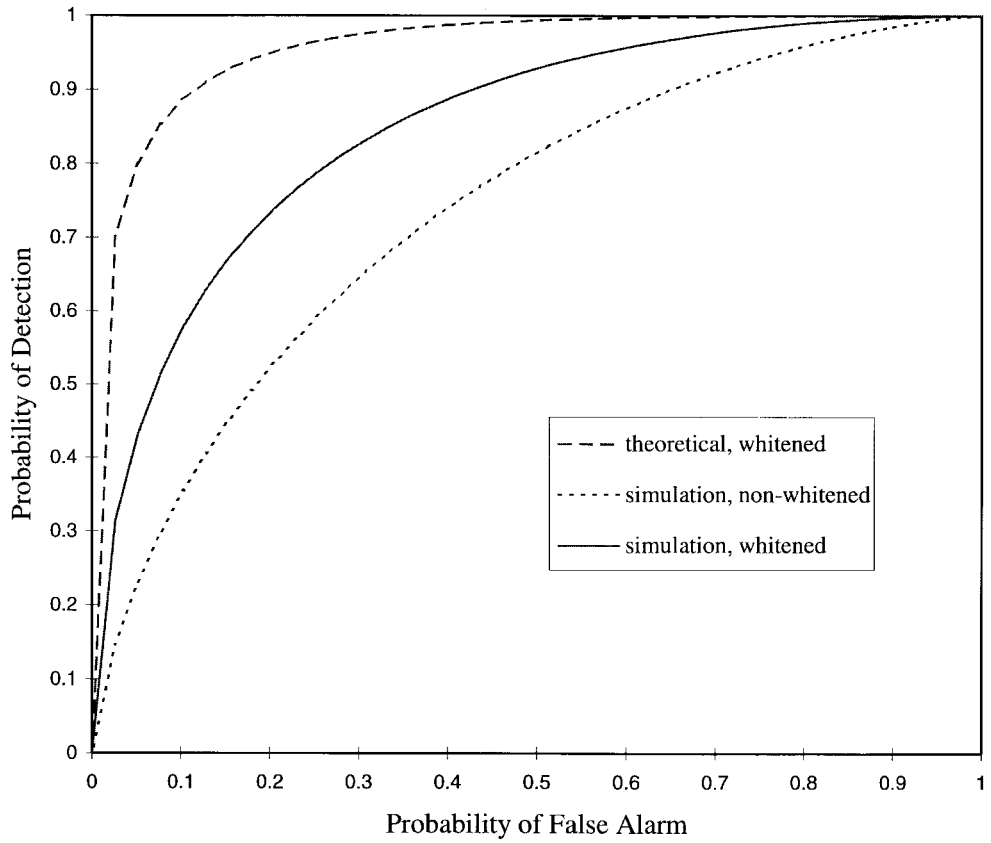
Fig. 5. Comparison of the ROC with and without prewhitening. The soil and rough air–ground interface are as in Fig. 3 and the target is a lossless dielectric ($\epsilon_r = 2$) of cross section 37.5 cm $\times$ 12.5 cm buried 25 cm beneath the mean position of the interface. The theoretical ROC [20] is calculated under the *assumption* that the scattering data fits the detector model perfectly and the simulated ROC curves (based on 300 Monte Carlo realizations) represent *actual* detector performance with and without prewhitening. (a) Angle of incidence $\theta_i = 30^\circ$.

through the air–ground interface as a random process *added* to the deterministic target signature, we can parametrize the target response itself as a random process. However, this reality runs counter to the assumptions under which our (classical) detector was designed, namely, that the target signature is deterministic (see Section III-B). In either case, the mismatch between the detector model and the physical problem is due to a randomization of the field upon entering and leaving the rough (random) air–soil interface.

For the cases considered here, the correlation length is comparable to the target width (as it often will be in practice). Therefore, the variation in the surface roughness is relatively small over the target width (the target is also relatively shallow). Thus, for near-normal incidence, the field incident upon the target and the fields backscattered from same are randomized relatively weakly, since there is modest variation (randomness) over the surface area with which these fields interact. However, as one approaches grazing the fields incident on the target interact with a wider range of the air–ground interface, sustaining more randomization (the same holds true for the backscattered fields as they are transmitted from the soil to the observer). Thus, the degree of transmission-induced randomization (upon entering and leaving the soil) increases as one approaches grazing. From the previous paragraph, it is this randomization, which is not necessarily matched to the surface clutter on which the whitening filter was designed, that causes the deterioration in detector performance. Coalescing these concepts, we can understand the reduced effectiveness of pre-whitening as the incident angle approaches grazing. However, we reiterate that in this paper, we consider one class of surface roughness, one target type, and one target depth, so the aforementioned interpretation requires further study to be stated confidently.

Finally, we reiterate that the ROC performance improves as the incidence angle approaches grazing (Fig. 5), despite the fact that the target-signature strength is smaller near grazing (Fig. 4). This implies that the variance in the clutter decreases more quickly with increasing incidence angle than does the energy in the backscattered target response. We have found this interesting result to hold for all examples considered thus far. However, in future work this phenomenon must be further tested as a function of target type and surface-roughness statistics.

In the remainder of the text, we present Monte Carlo calculated ROC curves for which the whitening filter *has* been utilized with the understanding that these results are inferior to those predicted by (7) and (8), with this disparity increasing as the incidence angle approaches grazing. However, these results are useful for they represent the type of performance one may expect from an actual system (the FDTD allows us to model



(b)

Fig. 5. (*Continued.*) Comparison of the ROC with and without prewhitening. The soil and rough air–ground interface are as in Fig. 3 and the target is a lossless dielectric ($\epsilon_r = 2$) of cross section 37.5 cm × 12.5 cm buried 25 cm beneath the mean position of the interface. The theoretical ROC [20] is calculated under the *assumption* that the scattering data fits the detector model perfectly and the simulated ROC curves (based on 300 Monte Carlo realizations) represent *actual* detector performance with and without prewhitening. (b) angle of incidence $\theta_i = 70°$.

most of the relevant physics) using the type of detector often applied in practice. The type of performance demonstrated here may be inadequate for certain applications and, therefore, in Section V, we discuss the implications of the results presented here on the development of new detectors, which properly account for the physics in question.

C. Detector Performance as a Function of Roughness Properties

We next investigate detector performance (ROC) for a fixed target and fixed angle of incidence (near grazing), when the statistical properties of the rough surface are varied. As mentioned above, the ROC's are computed using whitening filters and in these examples the change in the surface statistics implies a different whitening filter for each example studied.

A decrease in the correlation length l constitutes an increase in the variation of the rough air–ground interface and, therefore, an anticipated deterioration in detector performance. To quantify such, we consider an incident angle of $\theta_i = 70°$ a roughness standard deviation $h = 3.95$ cm and the same target as investigated in Figs. 4 and 5. From Fig. 6, we see that the correlation length $l = 31.25$ cm results in improved ROC performance relative to $l = 18.75$ cm. Additionally, we consider an example in which the correlation length is kept fixed at $l = 18.75$ cm and rough-surface standard deviations $h = 6.85$ cm and $h = 3.95$ cm are considered. The anticipated reduction in ROC performance with increased surface variation (larger h) is demonstrated in Fig. 7.

D. Target Properties

In the final example, we fix the surface-roughness properties ($l = 18.75$ cm and $h = 3.95$ cm) and consider an angle of incidence $\theta_i = 70°$. The 37.5 cm × 12.5 cm lossless target is buried 25 cm under the mean soil height (as in all previous examples) and we consider targets with dielectric constants $\epsilon_r = 2, 3$, and 4 (recall that the soil electrical properties are characterized by $\epsilon_r = 6$ and $\sigma = 0.005$ S/m). As with the results in Figs. 6 and 7, the *qualitative* ROC performance can be predicted in advance (improved detector performance with increase target-soil dielectric contrast). However, here as there, the accuracy of the FDTD computation allows us to *quantify* this phenomenon using the detector discussed above. In particular, the results in Fig. 8 demonstrate a marked improvement in detector performance with decreasing target dielectric constant (heightened target-soil contrast). For the surface roughness considered here, we see that detector performance is relatively poor for the case of a weak target-soil contrast (target with $\epsilon_r = 4$), underscoring the difficulty of detecting such targets with radar systems. Unfortunately, many buried mines have dielectric properties very similar to

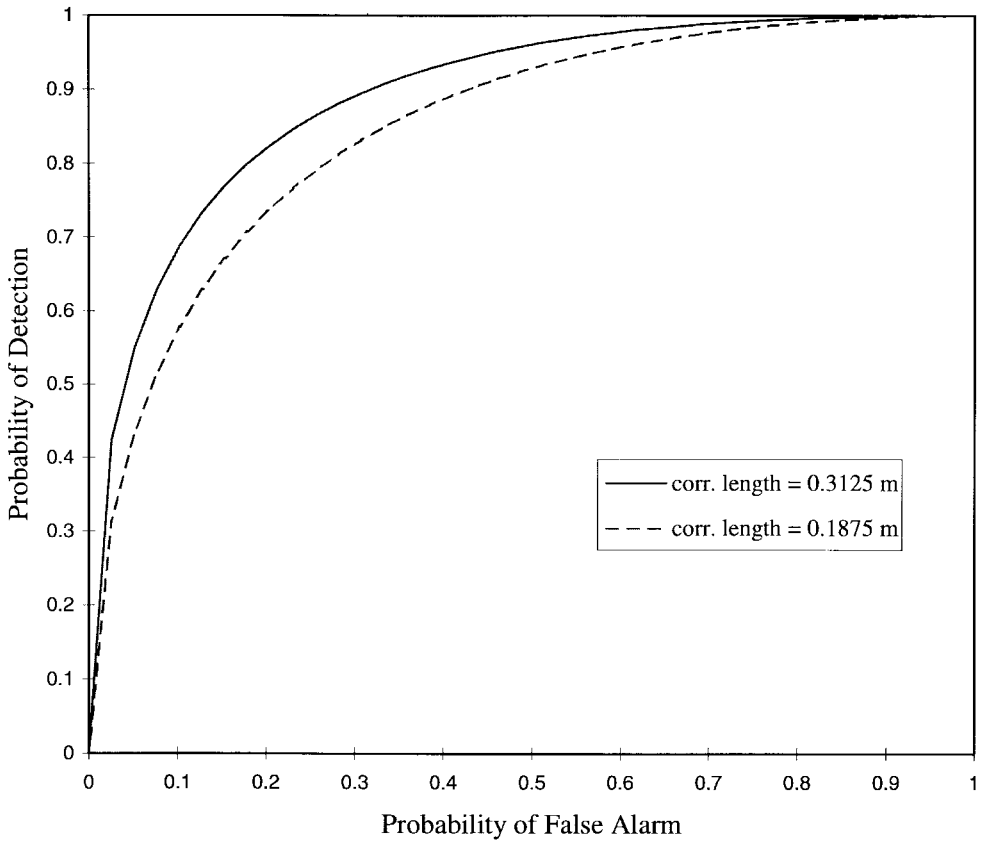


Fig. 6. Receiver operating characteristic (ROC) for a lossless dielectric target ($\epsilon_r = 2$) of cross section 37.5 cm $\times$ 12.5 cm buried 25 cm beneath the mean position of the air–ground interface (soil properties: $\epsilon_r = 6$ and $\sigma = 0.005$ S/m). The incidence angle is $\theta_i = 70^\circ$ and the standard deviation of the roughness is $h = 3.95$ cm. Results are plotted for the two correlation lengths $l = 31.25$ cm and $l = 18.75$ cm and the results were computed via 300 Monte Carlo iterations employing a whitening filter.

the surrounding soil [39], making the radar-based detection such a particularly challenging problem. In this case, the need for an improved detector, which properly accounts for the transmission randomness (Section IV-C), may be necessary to achieve useful system performance.

V. Conclusions

We have considered short-pulse electromagnetic scattering from rough (random) air–ground interfaces, with deterministic targets buried underneath (see [40] for the related problem of frequency-domain scattering; we also cite [41], which was brought to the authors attention during review). In addition to investigating the statistics of rough-surface scattering *per se*, we have addressed processing the time-domain scattered fields for the detection of buried targets. The controlled nature of the FDTD computations used to perform this study allows a systematic investigation of the assumptions underlying conventional detectors. The matched-filter detector used here, with clutter pre-whitening, invokes the assumptions that the target response is deterministic and the clutter additive. If the theoretical and calculated (via Monte Carlo simulations) agree, then we deem these assumptions valid. When such assumptions are inappropriate, more-sophisticated detectors must be designed (or one can live with the suboptimal detector performance, if such is adequate for the application in question).

After propagating through the rough air–ground interface, the fields incident upon the target are a random process, as are the fields transmitted through the interface after scattering. Therefore, the target signature is in fact *not* deterministic. At issue, therefore, is the degree to which this undermines detector performance. For the examples investigated here, we have found that such a mismatch between the physics and the detector does, in fact, result in a deterioration of detector performance relative to theoretical expectations based on a perfect fit between the data and the detector. Further, the deterioration is more noticeable as one approaches grazing. An explanation of this phenomenon was given in Section IV-C, but future research is required to investigate detector performance for a wider range of surface-roughness statistics and target types.

The results of this research indicate that, under certain conditions, classical detectors may be inadequate for the important problem of detecting targets under a rough (random) interface. To ameliorate this problem, more sophisticated detectors must be developed. Detectors in which the target response is random have been developed [20]. However, such detectors incorporate conditional probability density functions, which imply a statistical analysis of the target signature. Using the FDTD, we can parametrize the statistics associated with wave propagation into and out of the soil (for the case of a random surface), the results of which can then be

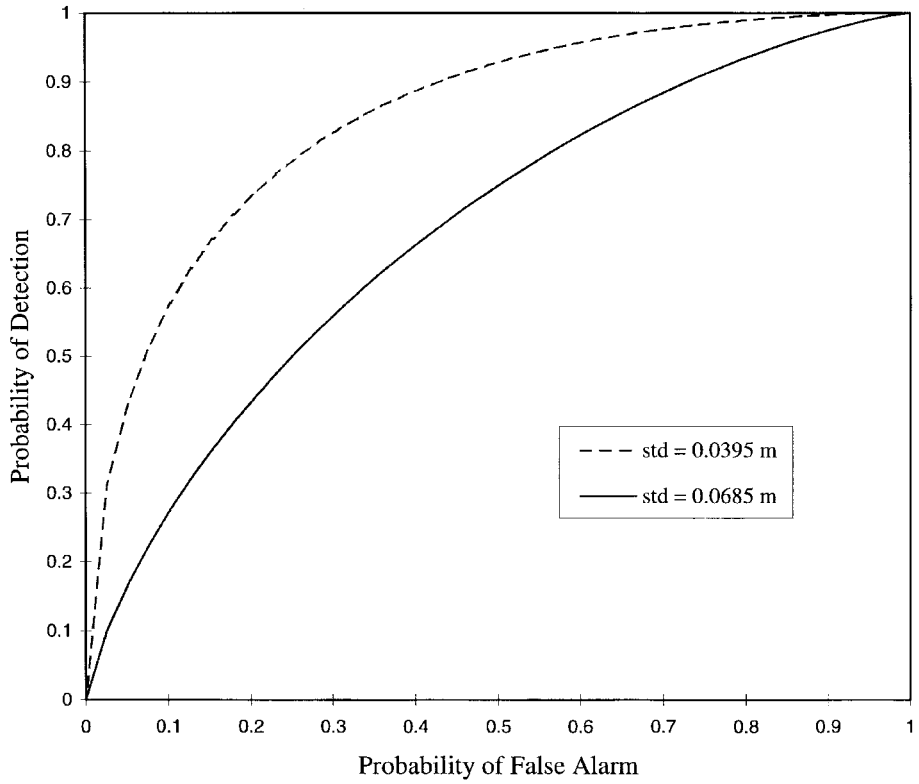


Fig. 7. Receiver operating characteristic (ROC) for the parameters in Fig. 6 except here the correlation length is fixed at $l = 18.75$ cm and the two standard deviations $h = 6.85$ cm and $h = 3.95$ cm are considered. The results were computed via 300 Monte Carlo iterations employing a whitening filter.

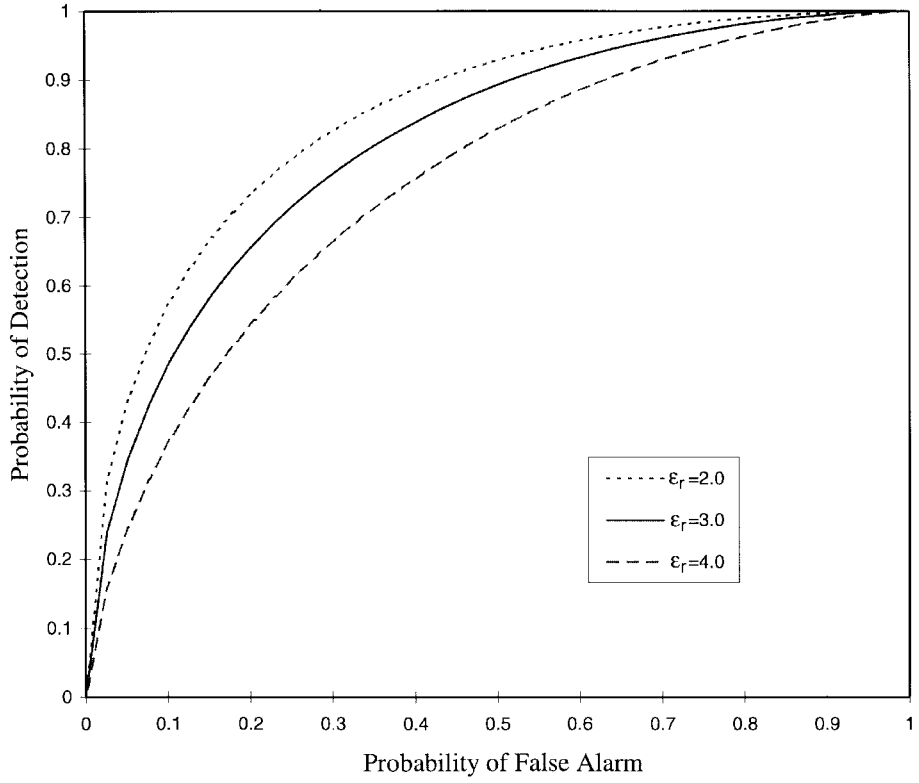


Fig. 8. Receiver operating characteristic (ROC) for a surface-roughness defined by the parameters $l = 18.75$ cm and $h = 3.95$ cm (soil properties: $\epsilon_r = 6$ and $\sigma = 0.005$ S/m). The backscattered fields are observed for incidence angle $\theta_i = 70°$, the target shape and position are as in Figs. 6 and 7, and we consider target dielectric constants $\epsilon_r = 2, 3,$ and 4. The results were computed via 300 Monte Carlo iterations employing a whitening filter.

utilized in the context of an improved detector. This is an example wherein the statistical results from forward modeling (e.g., FDTD) can be used to improve detector performance, which is of particular importance because the measurement of such transmission statistics may be difficult (in contrast to the backscatter statistics that are measured routinely). Thus, one can measure the statistical topography of a region in question from which the transmission statistics can be modeled numerically, subsequently effecting an improved detector.

The accuracy of the FDTD allows a controlled investigation of expected detector performance for various scenarios of interest—the rough-surface scattering investigated here representing one example. Such forward modeling yields access to the fully polarimetric scattered fields, while here our detector was based on only a single polarization (a time-dependent *scalar* function). A fully polarimetric system, however, is far more complicated than the case considered here. Thus, it is of interest to investigate theoretically the expected improvement in ROC performance one can ideally attain if a fully polarimetric detector is effected. Thus, another important topic of future research involves polarimetric processing with the goal of *quantifying* ultimate detector improvements with polarimetric systems. Vector detection theory has been developed [20] and will be exploited in this context.

Finally, this paper has focused on investigating and utilizing the statistics of scattering from randomness introduced by a rough (random) air–ground interface. Inhomogeneities in the soil may also necessitate a stochastic treatment. Recall from Section I that in the case of scattering from a rough surface, the scattering itself is completely deterministic. However, the details of the rough surface change as one considers different regions of the same surface. Thus, to characterize scattering or detection in such an environment, the results must be described statistically for an ensemble of surface-roughness profiles, each characterized by the same probability distribution. By extension, rocks, roots, shrapnel, and other soil inhomogeneities also induce a deterministic scattered waveform when interrogated by electromagnetic excitation. However, the location and characteristics of such inhomogeneities can only be parametrized stochastically, resulting as well in a statistical model for the fields scattered from such environments. Thus, soil inhomogeneities give rise to another important class of clutter, which may impact the radar detection of underground targets. Moreover, as the incident wave propagates through such inhomogeneities, it will experience distortion that must be treated statistically. Therefore, the same issues addressed here with regard to surface roughness are relevant for the problem of statistical soil inhomogenities. Following the basic constructs in this paper, this will be a subject future research.

ACKNOWLEDGMENT

The authors would like to thank Duke colleagues Prof. L. Collins, Prof. J. Krolik, and Prof. L. Nolte for their helpful discussions.

REFERENCES

[1] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. New York: Academic, 1978.
[2] F. D. Hastings, J. B. Schneider, and S. L. Broschat, "A Monte-Carlo FDTD technique for rough surface scattering," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 1183–1191, Nov. 1995.
[3] G. S. Brown, "The validity of shadowing corrections in rough surface scattering," *Radio Sci.*, vol. 19, pp. 1461–1468, 1984.
[4] ——, "A new approach to the analysis of rough surface scattering," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 943–948, July 1991.
[5] E. I. Thorso and D. R. Jackson, "The validity of the perturbation approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Amer.*, vol. 86, pp. 261–277, 1989.
[6] L. Tsang, S. H. Lou, C. H. Chan, and A. Ishimaru, "Application of the finite element method to Monte Carlo simulations of scattering of waves by random rough surfaces with the periodic boundary condition," *J. Electron. Waves Appl.*, vol. 5, pp. 835–855, Aug. 1991.
[7] E. I. Thorsos, "The validity of the Kirchoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Am.*, vol. 83, pp. 78–92, 1988.
[8] R. Devayya and D. H. Wingham, "The numerical calculation of rough surface scattering by the conjugate gradient method," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 645–648, Mar. 1992.
[9] R. Collin, "Electromagnetic scattering from perfectly conducting rough surfaces using a new full-wave method," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 1466–1477, Dec. 1992.
[10] S. Vitebskiy and L. Carin, "Moment-method modeling of short-pulse scattering from and the resonances of a wire buried inside a lossy, dispersive half space," *IEEE Trans. Antennas Propag.*, vol. 43, pp. 1303–1312, Nov. 1995.
[11] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 143–151, Feb. 1996.
[12] S. Vitebskiy and L. Carin, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy, dispersive half space," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.
[13] L. Peters, J. J. Daniels, and J. D. Young, "Ground penetrating radar as a subsurface environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.
[14] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of a ground-penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–44, Jan. 1996.
[15] P. E. Wannamaker, G. W. Hohmann, and W. A. San Filipo, "Electromagnetic modeing of three-dimensional bodies in layered earths using integral equations," *Geophys.*, vol. 49, pp. 60–74, Jan. 1984.
[16] D. L. Moffatt and R. J. Puskar, "A subsurface electromagnetic pulse radar," *Geophys.*, vol. 41, pp. 506–518, June 1976.
[17] S. Vitebskiy, L. Carin, M. Ressler, and F. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, to be published.
[18] M. A. Ressler and J. W. McCorkle, "Evolution of the Army Research Laboratory ultra-wideband test bed," in *Ultra-Wideband Short-Pulse Electromagnetics—2*, L. Carin and L. B. Felsen, Eds. New York: Plenum, 1995, pp. 109–123.
[19] S. L. Earp, E. S. Hughes, T. J. Elkins, and R. Vickers, "Ultra-wideband ground-penetrating radar for the detection of buried metallic mines," *IEEE Aerosp. Electron. Syst. Soc. Mag.*, pp. 30–34, 1996.
[20] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.
[21] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell s equations in isotropic media," *IEEE Trans. Antennas Propagat.*, vol. AP-14, pp. 302–307, May 1966.
[22] K. S. Kunz and R. J. Luebbers, *The Finite-Difference Time-Domain Method for Electromagnetics*. Boca Raton, FL: CRC, 1993.
[23] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 1995.
[24] D. E. Merewether, R. Fisher, and F. W. Smith, "On implementing a numerical Huygens surface in a finite difference program to illuminate scattering bodies," *IEEE Trans. Nucl. Sci.*, vol. NS-27, pp. 1829–1833, Dec. 1980.
[25] T.-T. Hsu and L. Carin, "FDTD analysis of plane-wave diffraction from microwave devices on an infinite dielectric slab," *IEEE Microwave Guided Wave Lett.*, vol. 6, pp. 16–18, Jan. 1996.
[26] K. Demarest, Z. Huang, and R. Plumb, "An FDTD near-to-far-zone transformation for scatterers buried in stratified grounds," *IEEE Trans. Antennas Propagat.*, vol. AP-44, pp. 1150–1157, Aug. 1996.

[27] J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, vol. 114, pp. 185–200, Oct. 1994.
[28] J. Fang and Z. Wu, "Generalized perfectly matched layer an extension of Berengers perfectly matched layer boundary condition," *IEEE Microwave Guided Wave Lett.*, vol. 5, pp. 451–453, Dec. 1995.
[29] Z. Wu and J. Fang, "Numerical implementation and performance of perfectly matched layer boundary condition for waveguide structures," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 2676–2683, Dec. 1995.
[30] R. L. Higdon, "Absorbing boundary conditions for difference approximations to the multi-dimensional wave equations," *Math. Comput.*, vol. 47, no. 176, pp. 437–459, Oct. 1986.
[31] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
[32] L. B. Felsen and F. Niu, "Spectral analysis and synthesis options for short-pulse radiation from a point dipole in a grounded dielectric layer," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 747–754, June 1993.
[33] P. G. Petropoulos, "Stability and phase error analysis of FDTD in dispersive dielectrics," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 62–69, Jan. 1994.
[34] P. G. Petropoulos, "Phase error control for FDTD methods of second and fourth order accuracy," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 859–862, June 1994.
[35] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.
[36] L. Franks, *Signal Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1969.
[37] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[38] J. E. Hipp, "Soil electromagnetic parameters as functions of frequency, soil density, and soil moisture," *Proc. IEEE*, vol. 62, pp. 98–103, Jan. 1974.
[39] A. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, Eds., "Detection technologies for mines and minelike targets," in *SPIE Proc.*, Orlando, FL, Apr. 1995, vol. 2496.
[40] K. O'Neil, R. F. Lussky, Jr., and K. D. Paulsen, "Scattering from a metallic object embedded near the randomly rough surface of a lossy dielectric," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 367–376, Mar. 1996.
[41] Y. Miyazaki, "Statistical refelection properties of an electromagnetic pulse by buried objects in random media using the FDTD," in *Proc. Int. Symp. Antennas Propagat.*, Chiba, Japan, Sept. 1996.



Traian Dogaru (S'96) was born in Bucharest, [redacted] He received the Engineering degree from the Polytechnic University of Bucharest, Romania, in 1990, and the M.S. degree in electrical engineering from Duke University, Durham, NC. He is currently working toward the Ph.D. degree at the same university.

Between 1992–1995, he held different engineering positions in the magnetic recording industry. His main research interests are in electromagnetic wave theory, computational electromagnetics, and radar-related signal processing.

Lawrence Carin (SM'96) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering at the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department at the Polytechnic University, Brooklyn, NY, as an Assistant Professor and became an Associate Professor there in 1994. In September 1995 he joined the Electrical Engineering Department at Duke University, Durham, NC, where he is an Associate Professor. His current research interests include quasi-planar transmission lines, short-pulse scattering and propagation, and signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.

canopy, the forward scattering of the trunk layer shows great influence on the scattering properties of the forest. Applying the approximate analytical solutions for the first-order and second-order scattering fields obtained above, the variations of the forward scattering cross sections of two adjacent conducting cylinders and dielectric cylinders via the incidence angles are calculated in this paper. To verify the validation, the scattering results are compared with those obtained by using the Numerical Electromagnetic Code (NEC), which is a computational package based on the moment method, and only those of conducting cylinders are compared because the version of NEC we have only provides for metallic objects.

Suppose the length of cylinders is 18 cm and the radius is 0.2 cm. The conducting cylinder is perfect, and the relative dielectric constant of dielectric cylinder is (9.6, j4.0). The frequency of the excitation wave is 9.6 GHz, and the incident plane-wave is arranged in x–z plane, i.e., the azimuthal incidence angle is 180°.

Figs. 2 and 3 give the VV-polarized azimuthal patterns of forward RCS for two conducting cylinders when the elevation incidence angle is 35°. In Fig. 2, the position of the second cylinder is $(\tilde{\rho}, \tilde{\phi}) = (18 \text{ cm}, 180^\circ)$. In this case, the two cylinders are located far enough apart that the secondary scattering shows very little effect. The first- and second-order scattering results are all in agreement with the moment-method computation results. In Fig. 3, the two cylinders are relatively close and the position of the second cylinder is $(\tilde{\rho}, \tilde{\phi}) = (2.5 \text{ cm}, 180^\circ)$. It may be seen that the second-order scattering result is in agreement with the moment method data and provides a reasonable approximation.

Figs. 4 and 5 give the VV-polarized elevation patterns of forward RCS for two conducting cylinders when the azimuthal scattering angle is $\phi_s = 0^\circ$. The position of the second cylinder is $(\tilde{\rho}, \tilde{\phi}) = (4 \text{ cm}, 180^\circ)$ in Fig. 4 and 8 cm, 180° in Fig. 5. The peak in the first-order scattering results is due to the constructive interference of scattering fields from two cylinders and may be decreased significantly by the inclusion of the secondary scattering terms. From these two figures, we can see that the difference between the first- and second-order results increases with the incidence angle except in the interference peak. This is because the angle of the primary forward scattering cone becomes larger when the incidence angle increases, and this leads to the increment of the secondary scattering length $L1$. The HH-polarized results are not given for the reason that the values of HH-polarized RCS of conducting cylinders are small, and, therefore, the effect of the secondary scattering is small.

For two dielectric cylinders, the VV- and HH-polarized elevation patterns of forward RCS, which also show the differences between the first- and second-order scattering results, are given in Figs. 6 and 7. In Fig. 6, the position of the second cylinder is $(\tilde{\rho}, \tilde{\phi}) = (4 \text{ cm}, 180^\circ)$, and in Fig. 7 it is $(4 \text{ cm}, 90^\circ)$. From these figures, we can see that the secondary scattering has a greater effect on VV-polarized response than HH-polarized response.

IV. Conclusions

The closed-form solution of the scattered field up to second order from two adjacent finite length cylinders has been obtained in this paper based on the reciprocity theorem. Its validity was verified by comparing analytical results with method of moments computations, and agreement was obtained. The results show that the secondary scattering due to interaction between the two cylinders has some effects on the scattered field. The present work should provide a theoretical basis for the simulations of forest canopy scattering considering the effect of multiple scattering between trunks.

References

[1] M. A. Karam and A. K. Fung, "Electromagnetic scattering from a layer of finite-length, randomly oriented dielectric circular over a rough interface with application to vegetation," *Int. J. Remote Sensing*, vol. 9, no. 6, pp. 1109–1134, 1988.

[2] M. K. Karam, A. K. Fung, R. H. Lang, and N. S. Chauhan, "A microwave scattering model for layered vegetation," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 767–784, July 1992.

[3] K. Sarabandi and P. F. Polatin, "Electromagnetic scattering from two adjacent objects," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 510–516, Apr. 1994.

[4] K. Sarabandi, P. F. Polatin, and F. T. Ulaby, "Monte Carlo simulation of scattering from a layer of vertical cylinders," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 465–474, Apr. 1993.

[5] G. T. Ruck, D. E. Barrick, W. D. Stuart, and C. K. Krichbaum, *Radar Cross-Section Handbook.* New York: Plenum, 1970, pp. 272–274.

Polarimetric SAR Imaging of Buried Landmines

Lawrence Carin, Ravinder Kapoor, and Carl E. Baum

***Abstract*— If the fields incident on a buried body of revolution are polarized vertically or horizontally (relative to the ground), the backscattered fields are exclusively copolarized (i.e., there are no cross-polarized backscattered fields). After substantiating this theoretically, *measured* ultrawideband (UWB) synthetic aperture radar (SAR) data are used for corroboration, considering real, buried landmines that approximate bodies of revolution.**

***Index Terms*—Ground-penetrating radar (GPR), polarization, synthetic aperture radar (SAR).**

I. Introduction

Over the last several decades, ground-penetrating radar (GPR) has been an important tool for buried-target detection and identification [1]–[3]. In most GPR systems, the antenna is placed on or near the ground, such that the buried target is generally in the near zone of the often complicated antenna pattern. Since the fields incident on the target are usually not polarized simply (especially as the target-sensor orientation varies), polarimetric processing is difficult. Alternatively, there has been recent interest in synthetic aperture radar (SAR) for buried-mine detection [4], [5]. Such systems offer the advantage of a significant standoff distance; moreover, since the targets are in the far zone of the source, relatively simple antenna designs can realize linearly polarized incident fields. While a SAR system may have difficulty detecting all individual mines, it can be an effective tool for mine-field detection, after which conventional GPR, electromagnetic

Manuscript received October 21, 1997; revised December 8, 1997.

L. Carin is with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA (e-mail: lcarin@ee.duke.edu).

R. Kapoor is with the Army Research Laboratory, AMSRL-SE-RU, Adelphi, MD 20783 USA.

C. E. Baum is with Phillips Laboratory, Kirtland AFB, Albuquerque, NM 87117-5776 USA.

Publisher Item Identifier S 0196-2892(98)05625-3.

0196–2892/98$10.00 © 1998 IEEE

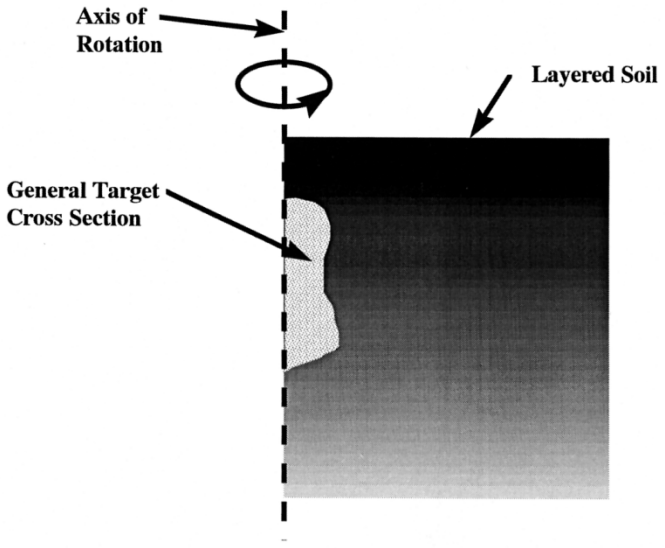


Fig. 1. Schematization of a general body of revolution.

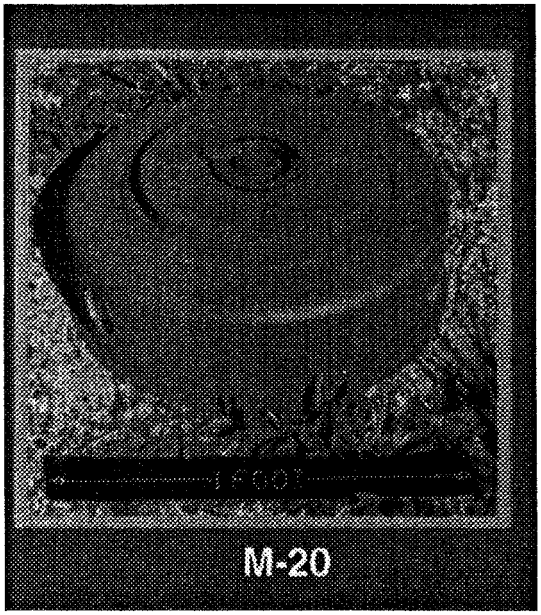


Fig. 2. The M20 anti-tank mine used in the measurements (the ruler indicates 30.48 cm).

induction (EMI), magnetometers, or other local sensors [6], [7] can be employed.

Given that a SAR system affords the potential for polarimetric detection of buried mines, in this letter, we explore this issue by considering a special but important class of land mines. In particular, we assume that the mine and the background media can be modeled as a body of revolution (BOR), implying that the buried mine and its surrounding environment can be modeled approximately by rotating a general two-dimensional (2-D) profile about an axis (see Fig. 1). A special case of such is a BOR target buried in a half space, with the BOR axis perpendicular to the air–ground interface [8]–[10]; more generally, the soil can be layered, discretely or continuously. Finally, note that the BOR mine can be metal, dielectric, or a combination of both, although the measured data considered here are restricted to the case of metal mines.

We assume that the target-sensor distance is large, such that the fields incident on the target are characterized by either a vertically or horizontally polarized plane wave, incident obliquely. For this special class of BOR targets, a symmetry plane is defined by the rotation axis and the line connecting the target center and the sensor. Since the incident plane wave satisfies symmetry properties about this plane, and because the target is symmetric about the same plane, the scattered fields will satisfy symmetry properties [11], [12]. In particular, if the incident fields are vertically polarized, the incident electric fields on either side of the symmetry plane are identical and, therefore, the same is true of the scattered electric fields; this constraint is satisfied if and only if the backscattered fields (in the plane of symmetry) are vertically polarized. There are therefore no cross-polarized backscattered fields. However, note that such symmetry does not imply that the *bistatic* fields are vertically polarized, only that at angles $\pm\theta$ from the symmetry plane the bistatic fields are polarized identically, with respect to this plane. Using the same logic, if the incident fields are horizontally polarized, the backscattered fields are as well. Therefore, in summary, for this special but important class of buried targets, the cross-polarized (VH and HV) fields are zero in the backscatter direction. The above polarimetric properties apply to any discretely or continuously layered soil profile, and they can be confirmed through consideration of the Green's functions for special cases (e.g., the half-space problem [8]–[10]).

In most SAR systems, the fields are measured in backscatter for a particular sensor position, along what subsequently becomes a synthetic aperture. Therefore, we expect the above polarimetric properties to manifest themselves in actual SAR imagery. This therefore provides an important and, to our knowledge, new tool for discriminating a ubiquitous class of mines from natural clutter (e.g., rocks, roots, etc.) that generally do not satisfy the BOR model. To demonstrate the potential of such a discrimination strategy, we consider polarimetric SAR imagery measured for an anti-tank mine, the "M20" shown in Fig. 2. The M20 mine was placed on the soil surface and buried at a depth of 6 in, although space constraints restrict the results presented here to the buried case. Upon conclusion, we discuss our experience with the surface mines and a class of anti-personnel mines.

The measurements were performed at Yuma Proving Grounds, Yuma, AZ (USA), with the data collected using an ultrawideband (UWB), time-domain radar [4]. In short, the radar transmits a pulse with bandwidth from approximately 50 to 1200 MHz. The radiation is effected via flared, parallel-plate horn antennas, which, over a relatively large beamwidth, emit nearly pure linearly polarized radiation. Four antennas are employed: vertically and horizontally oriented transmit antennas as well as vertical and horizontal receive antennas. Although the measurements are therefore not strictly backscatter, the distance between the four antennas is inconsequential relative to the target-sensor distance. The four antennas are placed atop a boom lift, 150 ft above the ground, and the entire unit is driven to effect a synthetic aperture. The mines are placed in the ground using military doctrine, and the region under consideration is contaminated by rocks, desert vegetation, and animal burrows. Every attempt is made to make the field under test as close as possible to what would be expected in an actual minefield (i.e., these are not sandbox-type experiments).

The measurements are performed in the time domain, and the SAR image is formed via a simple "delay-and-sum" procedure [4], utilizing the speed of light. We produce four images, one for each polarization (VV, HH, HV, and VH). We have considered many buried and surface mines, with the results presented here representative. The SAR imagery (Fig. 3) clearly demonstrates that the copolarized (VV and HH) target images are large relative to their cross-polarized (HV) counterparts. In the HV image, the sequence of mines (identified by arrows in the HH and VV images) are virtually invisible. While such

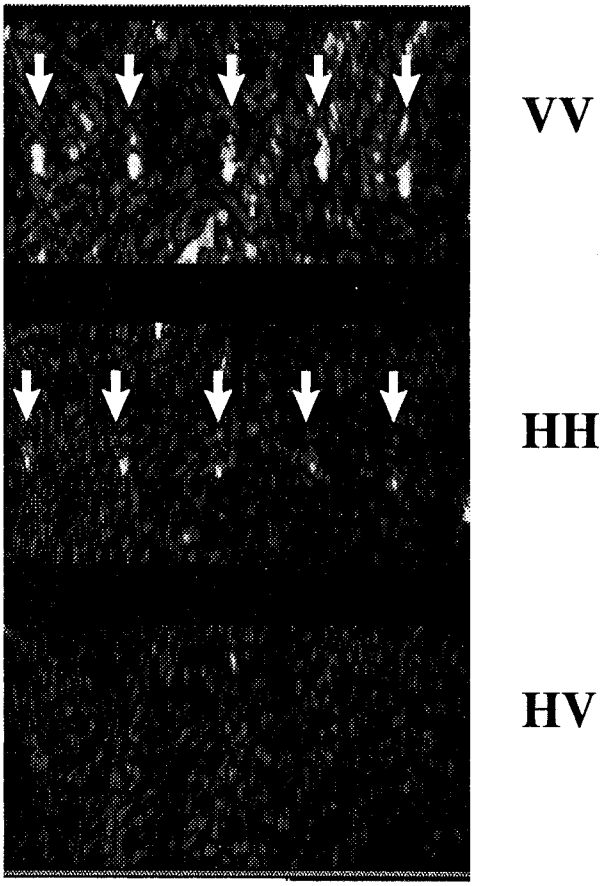


Fig. 3. SAR imagery for a row of five M20 mines buried 6 in (15.24 cm) deep, as measured by a UWB radar system. The three images correspond to VV-, HH-, and HV-polarizations, from top to bottom, and the mine locations are indicated by the arrows.

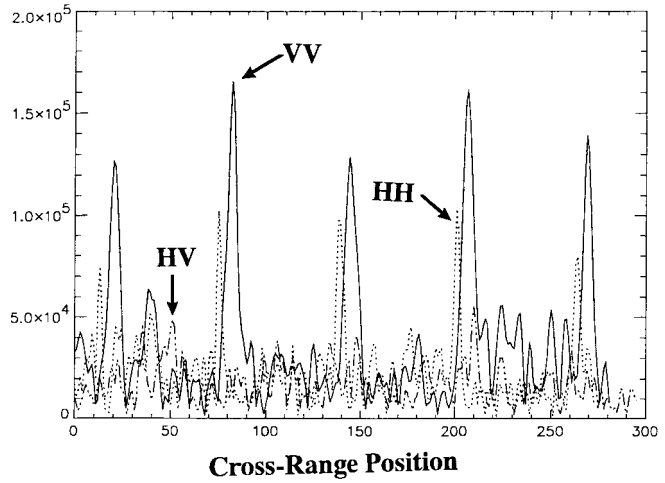


Fig. 4. Waveforms extracted by taking a "cut" out of the SAR imagery, in the cross-range direction, through the target centers. Results are shown for VV- (solid), HH- (dashed), and HV-polarizations (solid-dashed). The five peaks in the VV- and HH-polarizations correspond to the arrows in Fig. 3.

imagery gives a qualitative feel for the polarization dependence of the SAR imagery, we quantify such as follows. In Fig. 4 are plotted waveforms extracted from the SAR imagery, taking a cross-range "cut" in the image through the center of the mines (using knowledge of the actual target position). These data correspond to a 60° aperture integration, over which the polarization of the transmitted radiation is approximately linearly polarized (±30° about a line perpendicular to the synthetic aperture, along which the target center resides). We can see from these results—which were repeated faithfully for all M20 mines considered (as well as other BOR mines not shown)—that the copolarized component is significantly larger than its cross-polarized counterpart. Note from Fig. 2 that the M20 mine has small features that break the rigorous BOR symmetry, but appear to be unimportant for the wavelengths considered here. Note also that the VV and HH characteristics of the M20 mine are themselves different (although each is significantly stronger than the cross-polarized component); this issue has been explored in detail in a previous publication [10] and could be exploited in polarimetric mine detection.

II. Conclusion

In conclusion, we have explained theoretically that the backscattered fields from a BOR excited by a vertically or horizontally polarized plane wave are characterized by a zero cross-polarization component. These expectations have been confirmed experimentally using a UWB SAR system, considering anti-tank and anti-personnel mines that approximate BOR's (in addition to the M20, we considered the Valmara anti-personnel mine, results for which were not shown, for brevity). These polarimetric characteristics are valid, independent of the mine's electrical properties, as long as it satisfies the BOR model (i.e., they are valid for plastic and metal mines, although the measurements considered here were restricted to metal mines). Moreover, the soil can in general be continuously or discretely layered, as typically found in practice. Most natural and anthropic clutter do not satisfy the BOR model, and therefore, polarimetric SAR appears to be an effective tool for detection of an important and ubiquitous class of land mines. The principal limitation of this approach is the ability of electromagnetic waves to penetrate lossy soil; however, this is mitigated to some extent by the large integration afforded by the synthetic aperture (in contrast with conventional, local GPR systems [1]–[3], for which such integration gain is difficult to attain).

References

[1] L. Peters, J. J. Daniels, and J. D. Young, "Ground penetrating radar as a subsurface environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.

[2] G. S. Smith, "Directive properties of antennas for transmission into a material half-space," *IEEE Trans. Antennas Propagat.*, vol. AP-32, pp. 232–247, Feb. 1984.

[3] P. J. B. Clarricoats, "Portable radar for the detection of buried objects," in *Proc. Radar 77 Inst. Elect. Eng. Conf.*, London, U.K., 1977.

[4] M. A. Ressler and J. W. McCorkle, "Evolution of the Army Research Laboratory ultra-wideband test bed," in *Ultra-Wideband Short-Pulse Electromagnetics 2*, L. Carin and L. B. Felsen, Eds. New York: Plenum, 1995, pp. 109–123.

[5] S. L. Earp, E. S. Hughes, T. J. Elkins, and R. Vickers, "Ultra-wideband ground-penetrating radar for the detection of buried metallic mines," *IEEE Aerosp. Eng. Syst. Mag.*, pp. 30–34, Sept. 1996.

[6] A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, Eds., *Detection Technologies for Mines and Minelike Targets.* Orlando, FL: SPIE, 1995, vol. 2496.

[7] A. C. Dubey and R. L. Barnard, Eds., *Detection and Remediation Technologies for Mines and Minelike Targets.* Orlando, FL: SPIE, 1997, vol. 3079.

[8] S. Vitebskiy and L. Carin, "Short-pulse plane wave scattering from a buried perfectly conducting body of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 112–120, Feb. 1996.

[9] S. Vitebskiy, R. Sturgess, and L. Carin, "Late-time resonant frequencies of buried bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.

[10] S. Vitebskiy, L. Carin, M. Ressler, and F. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Theory and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.
[11] C. E. Baum and H. N. Kritikos, Eds., *Electromagnetic Symmetry.* New York Taylor and Francis, 1995.
[12] C. E. Baum, "Symmetry and electromagnetic scattering as a target discriminant," Phillips Lab., Interaction Note 523, Oct. 1996.

A Low-Frequency Radar Experiment for Measuring Vegetation Biomass

Marc L. Imhoff, Steven Carson, and Patrick Johnson

***Abstract*— Optical depth and backscatter measures were made at 18 frequencies between 50 and 450 MHz on two forest stands having dry biomass densities of 323 and 87 tons/ha to identify radar frequencies capable of penetrating heavy stands of vegetation. Good separation between stands was only achieved below 120 MHz.**

***Index Terms*— Biomass measurement, forest stands, low-frequency radar, vegetation, VHF radar.**

I. Introduction

Measuring the amount of carbon stored in the earth's dense tropical forests, in the form of biomass, has been an important part of efforts to quantify the exchange of carbon between the biosphere and atmosphere [1]. The amount of biomass stored in moist to wet tropical forest stands is considerable, with above-ground dry biomass densities in mature stands ranging from about 200 to 700 tons/ha [2]. Knowledge about both the biomass density and extent of the stands is important for estimating the amount of photosynthetically captured carbon stored in these stands on a regional and global basis.

A considerable investment has been made exploring the use of radar sensors to remotely measure vegetative biomass because of their ability to penetrate cloud cover and the way the transmitted electromagnetic radiation interacts with the physical structure of targets. Of particular interest has been the use of synthetic aperture radar (SAR) [3]–[5]. Many studies have reported correlations between radar backscatter and forest stand physical structural parameters that relate to biomass, such as height, basal area, and stocking density, and not surprisingly to stand biomass itself [6]–[10]. While these results are encouraging and SAR systems can be used to map biomass in some stands, saturation points, or the biomass level at which radar backscatter no longer increases with biomass for P-, L-, and C-band radar systems, have been reported as occurring at fairly low biomass loads. Saturation limits for a collection of coniferous forest stands using P-band (0.44 GHz), L-band (1.25 GHz), and C-band (5.3 GHz) SAR was reported to be at about 200 tons/ha for P-band and 100 tons/ha for L-band. No saturation limit was determined for C-band [9]. Other work examining the saturation limits of some tropical forest stands and temperate coniferous stands, using the same SAR system, showed signal saturation limits with respect to biomass to be approximately 100 tons/ha for P-band, 40 tons/ha for L-band, and 20 tons/ha for C-band [11].

The exact location of these saturation limits and their effect on making biomass inventories has varied depending on the forest type and structure and whether polarimetric data were used. Smaller statured forests, and some larger coniferous forests, can be successfully measured, as has been demonstrated by some (see [10]–[13] for temperate forests and [14]–[17] for tropical forests). Ultimately, however, the biomass saturation could still have a profound effect on global biomass inventories. An estimated 81% of the total terrestrial phytomass resides in forest stands with biomass densities above 100 tons/ha and nearly 41% in stands with densities above 200 tons/ha [11]. The growing awareness of the limitations presented by the saturation phenomenon has renewed interest in developing lower frequency radar systems to measure heavy biomass stands.

This paper describes initial results of an effort to explore the development of a low-frequency radar sensor to map biomass in heavy forest stands. The overall goal was to develop a radar sensor capable of operating at low frequencies and provide an inexpensive alternative to imaging radar systems. The work reported here is from the first phase. The work was carried out under NASA's Small Business Innovation Research Program (SBIR) with Zimmerman Associates, Inc./American Electronics, Inc. (ZAI/AMELEX), Vienna, VA, as the prime contractor for NASA's Goddard Space Flight Center, Greenbelt, MD. While the overall goal of this project (ongoing) is to develop an inexpensive system for measuring biomass in heavy stands, Phase I was primarily concerned with proof-of-concept and defining the limits of biomass saturation as a function of frequency from 450 (approximately P-band, which is the lower limit of NASA's current capability) to 50 MHz.

II. Methodology

In this first phase, a horizontal measurement geometry was used to conduct the tests (i.e., the direction of radar signal propagation was parallel with the ground). This geometry provided a convenient way to make low-cost, ground-based measurements as part of a feasibility study. However, this geometry creates interpretation problems when considering a synoptic perspective of canopy architecture. The dominant scatterers for this geometry are the tree trunks rather than the branches and leaves of the canopy, and this geometry represents a worst case for signal attenuation as a function of aspect angle at high-stocking densities. As such, these results could represent a lower bound on the saturation phenomenon as a function of frequency.

The horizontal viewing geometry used in this experiment requires the modeling of the backscatter from the forest as volumetric scatter because of the target's extent in range, as well as azimuth and elevation. Most of the radar data for biomass estimation in the literature are reported as surface scatter, in units of m^2 per m^2, not volumetric scatter. The results reported here are reported in m^2 per m^3 because each radar range gate measures the radar reflectivity of the volume of trees, due to the measurement geometry. This prevents the direct comparison to data from previous studies.

The first objective of this work was to establish the frequency requirements for penetrating stands of heavy vegetation for this measurement geometry. The Phase I measurements were taken with vertically polarized radiating elements for practical measurement

Manuscript received May 12, 1997; revised December 17, 1997. This work was supported by NASA under Contract NAS5-32 735.

M. L. Imhoff is with the Biospheric Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: mimhoff@ltpmail.gsfc.nasa.gov).
S. Carson is with SAIC, Arlington, VA 22203 USA.
P. Johnson is with Zimmerman Associates, Inc., American Electronics, Inc., Vienna, VA 22182-2623 USA.

Publisher Item Identifier S 0196-2892(98)08278-3.

0196–2892/98$10.00 © 1998 IEEE

Letters

Analysis and Processing of Ultra Wide-Band SAR Imagery for Buried Landmine Detection

David Wong and Lawrence Carin

***Abstract*—Experimental and theoretical results are presented for ultra wide-band (UWB) synthetic aperture radar (SAR) signatures of buried anti-tank and anti-personnel mines. Such are characterized by resonance-like peaks as well as valleys, across the 50–1200 MHz bandwidth considered. Consequently, frequency subbanding is used to highlight one target over another, of application to discriminating targets (mines) from clutter.**

***Index Terms*—Buried object detection, synthetic aperture radar.**

Ultra wide-band (UWB) synthetic aperture radar (SAR) is a new tool that has shown effectiveness for the detection of targets buried underground [1]. To examine the expected UWB SAR frequency-dependent signatures of buried landmines, we have developed a method-of-moments (MoM) algorithm, which models the target as a conducting body of revolution (BOR) embedded in a lossy dielectric half-space. The details of the MoM algorithm have been discussed elsewhere [3], [4]. The focus of this work is to demonstrate that: 1) close agreement is achievable between theoretical and experimental landmine signatures despite the complexity of the mines and the surrounding environment and 2) results from the models can be used to dictate which frequency bands are optimal for the detection of particular target classes.

The theoretical and experimental data correspond to landmines buried at Yuma Proving Grounds, Yuma, AZ. Soil samples were taken from this site from which the complex dielectric constant was measured. The results of the frequency-dependent soil characterization is shown in Fig. 1, with the measurements performed via a conventional coaxial probe [4]. These data correspond to a moisture content of 5% by weight, which is felt representative of the characteristics of this generally dry soil.

The two landmines considered here (both primarily metal) are an M20 anti-tank mine, which we consider buried to a depth of 15.24 cm, and a Valmara anti-personnel mine, buried just under the surface. Both targets approximate BOR's. In the computations, the target axis must be perpendicular to the air–ground interface such that the target half-space composite preserves the BOR symmetry [3], [4]. While this is clearly a simplification, in practice, landmines are generally so buried. Space limitations preclude a detailed accounting of the dimensions of these mines, but we note that the M20 anti-tank mine has an outer diameter of 33 cm and a total height of 17.5 cm, while the Valmara anti-personnel mine has an outer diameter of 10.16 cm and a total height of 20.32 cm (pictures of the mines are inserted in Fig. 2). With regard to the Valmara, note that it has several protruding metal spikes, which are not included in the BOR model. These prongs

Manuscript received January 18, 1998.

D. Wong is with the Army Research Laboratory, AMSRL-SE-RU, Adelphi, MD 20783 USA.

L. Carin is with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA.

Publisher Item Identifier S 0018-926X(98)08899-1.

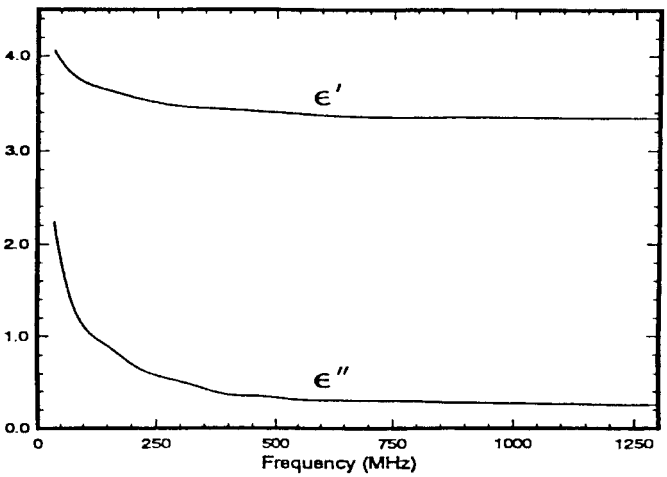


Fig. 1. Real (ϵ') and imaginary (ϵ'') part of the complex dielectric constant for Yuma soil.

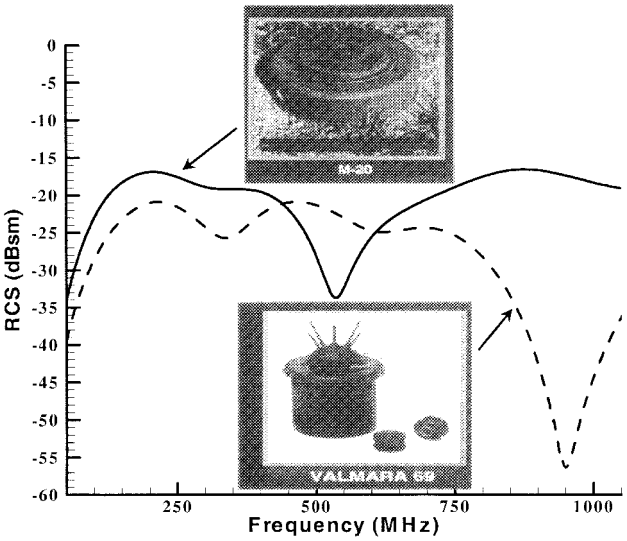


Fig. 2. Theoretical radar cross section (RCS) of the M20 and Valmara mines as a function of frequency for an incidence angle of 30° with respect to grazing.

are each approximately 10 cm in length and are likely to become most important at higher frequencies, when one approaches their resonant regime (∼1 GHz for the soils considered). Although not considered here, the BOR model can be modified in principle to more accurately model the Valmara's prongs [5]. With regard to the UWB SAR system, details can be found in [1] and [2]. We note that the system operates over the 50–1200 MHz bandwidth, with measurements performed directly in the time domain.

In Fig. 2, we consider the theoretical RCS for the M20 and Valmara mines (VV (vertical–vertical) polarization), with results presented here for $\theta_i = 30^\circ$. Note that the RCS of the M20 mine

0018–926X/98$10.00 © 1998 IEEE

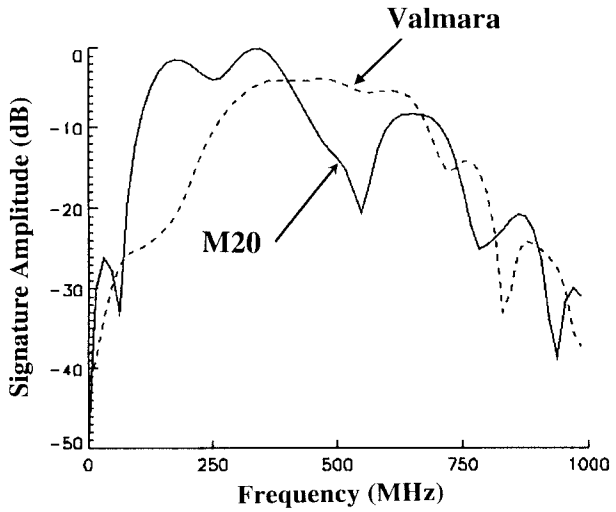


Fig. 3. Spectrum of the scattered signals from M20 and Valmara mines, as extracted from measured SAR imagery.

is characterized by two subbands over which the scattered fields are relatively large (approximately 160–400 MHz and 680–920 MHz), as well as an intervening subband over which the RCS is quite small (approximately 440–600 MHz). The two frequency subbands over which the RCS is large are representative of target resonances. Considering the first resonant band, a complex resonant frequency [4] with real part near 300 MHz (the center of this band) corresponds to a half wavelength of approximately 30 cm for the soil in Fig. 1, consistent with the general size (diameter) of the M20 mine. The broad frequency extent of this resonance-like feature is also consistent with that expected for a very low Q target [4].

The Valmara anti-personnel mine is considerably smaller than the M20 anti-tank mine and, therefore, one would expect its RCS to be smaller as well. However, from Fig. 2 we see that the RCS of the Valmara is, in fact, larger than that of the M20 in the frequency subband for which the M20 has a null (480–600 MHz). To examine whether these theoretical observations are manifested in actual SAR imagery, we consider the *experimental* spectrum of these two mines. Several rows of M20 and Valmara mines were buried in the Yuma soil, with burial depths as discussed above. From the measured (bipolar) SAR imagery [1], [2], one-dimensional down-range cuts (of 5-m length) were taken through ground-truth-dictated target positions. By taking into account the speed of light, these down-range cuts can be viewed as time-domain waveforms and are Fourier transformed to produce the spectrums in Fig. 3. Although we considered numerous examples of each mine, the general features of the results in Fig. 3 were faithfully reproduced. We see from Fig. 3 that the experimental spectrum of the M20 mine is characterized as in Fig. 2 by two resonant-like broad peaks separated by a sharp null. Moreover, over much of the spectrum the response of the M20 mine is larger than that of the Valmara, except in the region of the null in the M20 response.

The agreement between Figs. 2 and 3 is good, especially considering the following issues. While the RCS calculations in Fig. 2 have calibrated out the effects of the incident waveform, in practice, it is difficult to know the antenna response exactly. Therefore, the results in Fig. 3 are contaminated by the response of the antenna and, therefore, do *not* explicitly represent RCS (although the antenna spectrum is believed to be relatively flat over the bandwidth and imaging angle considered [1]). Moreover, since the results in Fig. 3 were extracted from a SAR image, they are not characterized by a single angle of incidence (angles from $\theta_i = 20^\circ$ to $\theta_i = 40^\circ$ are superposed when forming a single SAR image). The imaging process itself introduces artifacts [1]. Additionally, while every effort was made to make the experimental parameters the same as those considered in the model, the local soil properties are unlikely to be exactly like those in Fig. 1 and there was likely variability in the exact mine placement (depth and orientation). Finally, recall that the BOR model did not account for the metal prongs atop the Valmara, which may explain why the very sharp dip in the theoretical Valmara RCS for frequencies near 1 GHz is not as obvious in the measured data (i.e., wire resonances with such not accounted for in the theory, may explain this discrepancy). Nevertheless, at frequencies less than approximately 850 MHz, the general agreement between theory and experiment is encouraging.

Based on the above theoretical and experimental evidence, instead of utilizing the full 50–1200 MHz bandwidth to form an SAR image, we perform frequency-domain filtering of the measured data from which subbanded SAR images are formed. A tapered finite-impulse response (FIR) filter was used to process the following subbands: 100–200 MHz, 200–500 MHz, 480–580 MHz, 700–1000 MHz. From Figs. 2 and 3, the RCS of the M20 is larger than that of the Valmara for all but the 480–580 MHz subband. Space limitations prohibit presentation of each of the subband images, but we summarize the results as follows. We calculated the energy associated with each mine, as extracted from the particular subband SAR images. On average, for the 100–200 MHz, 200–500 MHz, and 700–1000 MHz subbands, the energy in the M20 mine was, respectively, 10.2, 2.6 and 3.2 dB larger than that of the Valmara mines. For the 480–580 MHz subband, approximately representative of the null in the M20 response (Figs. 2 and 3), the Valmara mine had an energy on average 1-dB larger than that of the M20 mine. The relative energies summarized above are as expected from the data in Figs. 2 and 3, except for the 480–580 MHz subband for which the discrepancy in the Valmara and M20 energies may have been expected to be larger. We attribute this to the fact that the width of the measured null (Fig. 3) is smaller than that of the theory (Fig. 2) with which the subband filters were designed. Nevertheless, the general expectations have been confirmed, despite the complexity of the theory and measurements.

REFERENCES

[1] M. A. Ressler and J. W. McCorkle, "Evolution of the army research laboratory ultra-wideband test bed," in *Ultra-Wideband Short-Pulse Electromagnetics 2*, L. Carin and L. B. Felsen, Eds. New York: Plenum, 1995, pp. 109–123.

[2] S. Vitebskiy, L. Carin, M. Ressler, and F. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Theory and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.

[3] S. Vitebskiy and L. Carin, "Short-pulse plane wave scattering from a buried perfectly conducting body of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 112–120, Feb. 1996.

[4] S. Vitebskiy, K. Sturgess, and L. Carin, "Late-time resonant frequencies of buried bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.

[5] T. E. Durham and C. G. Christodoulou, "A method for treating junctions between bodies of revolution and arbitrary surfaces," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 213–219, Feb. 1994.

On the Low-Frequency Natural Response of Conducting and Permeable Targets

Norbert Geng, *Member, IEEE,* Carl E. Baum, *Fellow, IEEE,* and Lawrence Carin, *Senior Member, IEEE*

***Abstract*— The low-frequency natural response of conducting, permeable targets is investigated. We demonstrate that the source-free response is characterized by a sum of nearly purely damped exponentials, with the damping constants strongly dependent on the target shape, conductivity, and permeability, thereby representing a potential tool for pulsed electromagnetic induction (EMI) identification (discrimination) of conducting and permeable targets. This general concept is then specialized to the particular case of a body of revolution (BOR), for which the Method-of-Moments (MoM)-computed natural damping constants from several targets are compared with measurements. Moreover, theoretical natural (equivalent) surface currents and damping coefficients are shown for other targets of interest. Finally, we investigate the practical use of such natural signatures in the context of identification, wherein Cramer–Rao bound (CRB) studies address signal-to-noise ratio (SNR) considerations.**

I. Introduction

THE NATURAL (source-free) modes of a linear system are independent of the driving function inducing their excitation [1]–[6] and therefore constitute useful signatures for identification (however, the excitation strengths of the various modes are generally strongly dependent on the incident fields). In electromagnetics, interest in such natural modes motivated development of the singularity expansion method (SEM) [7]–[15], which demonstrated that multiple interactions between scattering centers—after the incident fields have interrogated the entire target (at what are termed "late times")—can be repackaged compactly in terms of a sum of oscillatory, exponentially damped resonances (i.e., in terms of the SEM resonant modes). While the amplitudes of the SEM modes are dependent on the excitation waveform, the complex resonant frequencies themselves are termed "aspect independent" (the same for all target orientations and incident fields). After the original SEM framework was formalized [13], numerous researchers applied it to modeling [7]–[8], measurements [9]–[10], and signal processing (termed electromagnetic singularity identification, EMSI) [11]–[12], motivated largely by its potential for aspect-independent target identification.

The resonant frequencies of a particular target are determined by the spatial separation between scattering centers (i.e., by characteristic target lengths). For example, considering a conducting wire of length d, we would expect resonant frequencies at approximately $f_n = cn/(2d)$, where c is the speed of light and $n > 0$ is an integer. However, if we excite a target at frequencies below its lowest SEM resonance, the fields and currents on the target will still exist after departure of the incident fields, except the late-time fields will not oscillate in the manner characteristic of resonances. Considering excitation of the wire at frequencies $f \ll f_1$, we would expect the late-time fields to decay approximately as $\exp[-(R/L)t]$, where R and L are the wire resistance and inductance, respectively (the inductance can be readily computed if we consider a wire *loop* [3]). In the former, resonant regime, the late-time fields are characterized by oscillations at nearly real frequencies (oscillatory natural-mode poles near the real-frequency axis), while the latter case is characterized approximately by exponential damping (evanescent natural-mode poles near the imaginary-frequency axis). These two extremes are characteristic as well of the modes of a resonant cavity, at frequencies above and below the lowest cavity resonance, respectively.

In this paper, we demonstrate that the low-frequency exponential damping characteristic of the simple wire example above can be extended to general target shapes and materials. This is done by employing 1) low-frequency circuit analogies (analogous to the wire example above) as well as through 2) development of a full-wave Method-of-Moments (MoM) analysis, with particular application to a conducting, permeable body of revolution (BOR). For such targets, it is demonstrated that the damping is strongly dependent on the target shape and material parameters. Therefore, if we can measure such exponentially damped waveforms accurately, the damping coefficient can be used for possible target identification. Of particular interest here are frequencies less than 1 MHz, for which the conductivity- and permeability-dependent skin depth can vary significantly (manifested in wider decay-constant variability with target properties, thereby aiding in discrimination). This should be contrasted with microwave and millimeter-wave frequencies, for which most metals can be considered perfect conductors (and therefore only the shape of the metal target influences its scattering properties).

An additional benefit of the low-frequency, pulsed-decay discrimination scheme discussed above is that, for the good conductors and low frequencies of interest here, the conduction current induced in most good metals is often considerably stronger than the displacement current, which can therefore often be neglected. Thus, the dielectric properties of the surrounding media are often relatively unimportant. This is particularly important for the low-frequency (termed elec-

Manuscript received July 21, 1997; revised March 16, 1998.

N. Geng and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA (e-mail: lcarin@ee.duke.edu).

C. E. Baum is with Phillips Laboratory/WSQW, Kirtland Air Force Base, Albuquerque, NM 87117-5776 USA.

Publisher Item Identifier S 0196-2892(99)00119-9.

0196–2892/99$10.00 © 1999 IEEE

tromagnetic induction, EMI) identification of buried targets (e.g., mines), for example, for which the soil properties can often be ignored since the conductivity of the soil is generally many orders of magnitude less than that of good metals. Analogous to the EMSI referred above, we term this magnetic singularity identification (MSI). By contrast, ground-penetrating radar (GPR) [16]–[20], and the associated EMSI, is strongly influenced and limited by the soil properties. As a caveat, however, we should note that, if the metal target is very small, the conduction currents induced in the expansive, low-loss soil may become comparable to or exceed the currents induced in a small, localized, high-conductivity target, at which point the background medium does become important (this is particularly relevant for low-metal-content mines, for example); this will be the subject of future research.

We restrict ourselves here to the special case of conducting, permeable targets in free space (as mentioned, the soil conductivity and permittivity not being significant at the relevant low frequencies [3]). After using circuit theory to explain the basic wave phenomenology, numerical damping coefficients and currents are computed for the special case of conducting, permeable BOR's. While such analyses have been applied to the case of scattering from perfectly conducting and low-loss dielectric BOR's [21]–[24], [25]–[30], its application to calculation of natural modes for highly (but not perfectly) conducting, permeable targets is new. To validate the accuracy of the numerical data, comparisons are made with measured data [31]–[32] for several variations in material properties. Additionally, we present numerical results for the decay coefficients and natural-mode currents of other conducting and permeable targets.

Finally, in the context of target identification, such decay constants must be extracted from measured data, which are generally noisy. For such situations, a model-based estimation algorithm is often applied [14]–[15]. Each noisy measured waveform constitutes one realization of a random process, and therefore the estimated decay coefficients are random variables. Assuming the estimation algorithm is unbiased (the estimates for the decay constants are correct in the mean), we are interested in their variance, from which we quantify the expected accuracy of the decay-constant estimates for a given measured waveform. In this context, we investigate the Cramer–Rao bound (CRB) [33]–[35] for estimation of the damping coefficient of a noisy, exponentially damped waveform. The CRB constitutes the lowest variance any unbiased estimator can achieve for such parameter estimation. The CRB is strongly dependent on the signal-to-noise ratio (SNR). Therefore, by employing the CRB, we quantify the SNR requirements for unambiguous (low-variance) estimation of pulsed-EMI decay constants. Such an analysis represents an example in which wave modeling and basic signal processing coalesce to yield fundamental system constraints (SNR requirements in this case).

The remainder of the text is organized as follows. In Section II, we present a circuits-based explanation of the damped low-frequency EMI response from conducting, permeable targets, from which fundamental relationships are derived. We also formulate an appropriate boundary integral equation for the special case of conducting, permeable BOR's, which we solve via the MoM. The latter numerical formulation is discussed only briefly since similar analyses have been performed previously for different but related problems. In Section III, we address the issue of extracting exponentially damped waveforms from noisy measured data, through consideration of the CRB. The results of the numerical code are compared with measurements in Section IV, and several additional sets of numerical results are presented as well. Finally, conclusions are addressed in Section V.

II. Theory

A. Circuit-Based Explanation of Damped Low-Frequency EMI Response

We first consider an approximate circuit-based analysis of low-frequency scattering from conducting targets. This simple analysis demonstrates the basic phenomenology of low-frequency, time-domain scattering from such targets, motivating the rigorous MoM analysis discussed subsequently.

An incident low-frequency magnetic field $\boldsymbol{H}_{\rm inc}$ induces currents $\boldsymbol{J}(\boldsymbol{r}')$ in a conducting object. The magnetic dipole moment

$$\boldsymbol{m} = \tfrac{1}{2} \iiint_{V'} [\boldsymbol{r}' \times \boldsymbol{J}(\boldsymbol{r}')]\, dV' \tag{1}$$

associated with these currents generates the scattered field $\boldsymbol{H}^{\rm scat}$. For observation distances large relative to characteristic target dimensions, the scattered quasimagnetostatic field can be expressed as [3], [6]

$$\begin{aligned} \boldsymbol{H}^{\rm scat}(\boldsymbol{r}) &= \frac{1}{4\pi r^3}\,[3\hat{\boldsymbol{r}}\hat{\boldsymbol{r}} - \overleftrightarrow{\boldsymbol{I}}] \cdot \boldsymbol{m} \\ &= \frac{1}{4\pi r^3}\,[3\hat{\boldsymbol{r}}\hat{\boldsymbol{r}} - \overleftrightarrow{\boldsymbol{I}}] \cdot \overleftrightarrow{\boldsymbol{M}} \cdot \boldsymbol{H}^{\rm inc}(\boldsymbol{r}=0) \end{aligned} \tag{2}$$

where $\overleftrightarrow{\boldsymbol{I}}$ is the identity dyadic and $\overleftrightarrow{\boldsymbol{M}}$ is the magnetic polarizability dyadic, which relates the incident magnetic field to the induced magnetic dipole moment.

Let the scattering object have a simple shape so that it can have a magnetic moment in only a single direction, e.g., thin conducting disks or loops that support induced currents only in circular paths around the axis of symmetry (the magnetic moment of such an object is parallel to this axis, defined here as z). For simplification and because natural frequencies are independent of the excitation and observation point, we only consider a homogeneous z-directed incident magnetic field and the scattered field on the z-axis, expressed as

$$H_z^{\rm scat}(z) = \frac{m_z}{2\pi r^3} = \frac{1}{2\pi r^3}\, M_{zz} H_z^{\rm inc} = \frac{1}{2\pi r^3}\, M_{zz} H_{z0}. \tag{3}$$

Note that, if these assumptions are violated, only the amplitudes of the various modes will change, when the target is excited by a particular incident field.

The most simple target, a thin wire loop of area A, is characterized entirely by its resistance R and inductance L, assuming that displacement currents can be neglected compared to the conduction currents (zero capacitance). By

Faraday's law, an incident magnetic field parallel to the loop axis induces a voltage $V(t) = -\mu_0 A \partial H_{z0}/\partial t$ around the loop, which produces a current according to the impedance given by R and L. Using the Laplace-transform variable $s = j\omega$ (ω complex), the loop current in the Laplace domain is given by [3], [19], [32]

$$\begin{aligned} I(s) &= \frac{V(s)}{Z(s)} = -\mu_0 A \, \frac{s}{R+sL} \, H_{z0}(s) \\ &= \frac{-\mu_0 A}{L} \, \frac{s}{s+R/L} \, H_{z0}(s) \end{aligned} \tag{4}$$

where we have assumed that the imposed field is initially zero. The current defines a magnetic moment $\boldsymbol{m} = IA\hat{\boldsymbol{z}}$, and therefore, the z-directed scattered magnetic field can be written using (3) as

$$H_z^{\rm scat}(z,s) = \frac{-\mu_0 A^2}{2\pi r^3 L} \, \frac{s}{s+R/L} \, H_{z0}(s) \tag{5}$$

resulting in a purely exponentially damped step response (i.e., for an excitation of the form $H_{z0}(t) = H_{z0}u(t)$) in the time domain

$$\begin{aligned} H_z^{\rm scat}(z,t) &= \frac{-\mu_0 A^2}{2\pi r^3 L} \, H_{z0} e^{-t} \, u(t) \\ &= \frac{-\mu_0 A^2}{2\pi r^3 L} \, H_{z0} e^{-t/\tau_{\rm wire}} \, u(t) \end{aligned} \tag{6}$$

where $u(t)$ is the unit step function.

Next, the effects of the mutual coupling of multiple filament loops within a solid target are illustrated considering the simple case of two circular coaxial wire loops. The impedance matrix of the coupled system of the two loops may be written as

$$[\boldsymbol{Z}(s)] = \begin{bmatrix} Z_{11}(s) & Z_{12}(s) \\ Z_{21}(s) & Z_{22}(s) \end{bmatrix} = \begin{bmatrix} R_1 + sL_1 & sM \\ sM & R_2 + sL_2 \end{bmatrix} \tag{7}$$

where M stands for the mutual inductance of the two loops. A voltage is induced onto each loop from the incident magnetic field, which is again assumed to be homogeneous and parallel to the loop axes. The currents in the two loops are then

$$\begin{aligned} \begin{pmatrix} I_1(s) \\ I_2(s) \end{pmatrix} &= [\boldsymbol{Z}(s)]^{-1} \begin{pmatrix} V_1(s) \\ V_2(s) \end{pmatrix} \\ &= \frac{-\mu_0 s H_{z0}(s)}{(R_1+sL_1)(R_2+sL_2) - (sM)^2} \\ &\quad \cdot \begin{pmatrix} (R_2+sL_2)A_1 - sMA_2 \\ -sMA_1 + (R_1+sL_1)A_2 \end{pmatrix} \end{aligned} \tag{8}$$

and therefore using (3) the scattered magnetic field (on the z-axis) can be written as

$$\begin{aligned} &H_z^{\rm scat}(z,s) \\ &\quad = \frac{A_1 I_1(s)}{2\pi r_1^3} + \frac{A_2 I_2(s)}{2\pi r_2^3} \\ &\quad \approx \frac{1}{2\pi r^3} \left[A_1 I_1(s) + A_2 I_2(s) \right] \\ &\quad \approx \frac{-\mu_0}{2\pi r^3} \, \frac{s(R_2+sL_2)A_1^2 - 2s^2 M A_1 A_2 + s(R_1+sL_1)A_2^2}{(R_1+sL_1)(R_2+sL_2) - (sM)^2} \\ &\qquad \cdot H_{z0}(s) \end{aligned} \tag{9}$$

where we assumed an observation distance much larger than the loop separation. The resonances are given by the singularities of (9). If we write $s_{01} = -R_1/L_1 = -1/\tau_{\rm wire1}$ and $s_{02} = -R_2/L_2 = -1/\tau_{\rm wire2}$ for the negative real poles of each loop in the absence of the other loop, and $0 \le k^2 = M^2/L_1 L_2 \le 1$ for the mutual coupling coefficient, the resonances are

$$\begin{aligned} s_{1,2} &= \frac{(s_{01}+s_{02}) + \sqrt{(s_{01}-s_{02})^2 + 4k^2 s_{01} s_{02}}}{2(1-k^2)} \\ &= \frac{(s_{01}+s_{02}) \pm \sqrt{(s_{01}+s_{02})^2 - 4(1-k^2) s_{01} s_{02}}}{2(1-k^2)}. \end{aligned} \tag{10}$$

Again the poles of the system are negative real, corresponding to pure exponential damping. As expected, the poles reduce to those of the isolated loops if the mutual coupling vanishes ($k^2 = 0$). For nonzero coupling, the step response of the system has the form

$$\begin{aligned} &H_z^{\rm scat}(z,t) = \\ &\frac{-\mu_0}{2\pi r^3} \, H_{z0} [f_1(A_{1,2}, R_{1,2}, L_{1,2}, k) e^{-t/\tau_1} \\ &\qquad + f_2(A_{1,2}, R_{1,2}, L_{1,2}, k) e^{-t/\tau_2}] u(t). \end{aligned} \tag{11}$$

For zero coupling between the two loops, it reduces to a superposition of the two step responses for the individual loops according to (6)

$$\begin{aligned} &H_z^{\rm scat}(z,t) \\ &\quad = \frac{-\mu_0}{2\pi r^3} \, H_{z0} \left[\frac{A_1^2}{L_1} \, e^{-t/\tau_{\rm wire1}} + \frac{A_2^2}{L_2} \, e^{-t/\tau_{\rm wire2}} \right] u(t). \end{aligned} \tag{12}$$

Considering the R and L of simple nonpermeable shapes (loops), we can show that the time constants are in general proportional to the square of some characteristic object dimension times the conductivity. In the context of our rigorous MoM studies discussed below, we will demonstrate that such properties have been found to hold for more complicated targets than the simple loops for which the circuit theory is directly applicable. Finally, the variation of the poles or decay time constants as a function of the permeability is more complicated and therefore cannot be derived easily using the simple circuit-based theory. A solution for a conducting and permeable sphere is given in [3], [19], and [32], but there is no analytical method available for arbitrary three-dimensional (3-D) objects. This can only be accomplished using numerical methods like the MoM solution for BOR's described next.

B. Surface Integral Equation and MoM

We consider electromagnetic scattering from highly (but not perfectly) conducting and permeable targets situated in a homogeneous environment, e.g., free space. Moreover, we specialize the solution to the case of a BOR. Because the surface integral equation approach for treating problems involving perfectly conducting as well as low-loss dielectric BOR's has

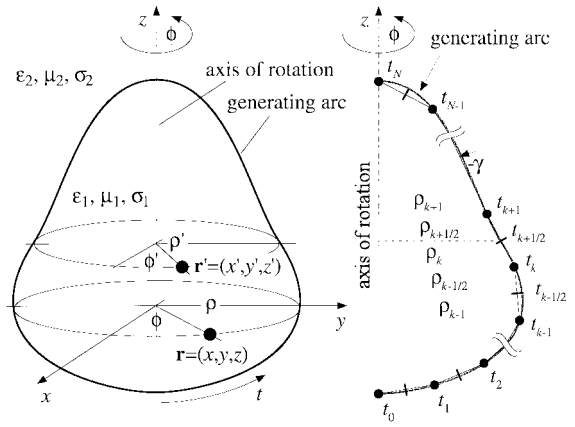


Fig. 1. General geometry for a body of revolution (BOR) and discretization scheme for MoM solution using a staggered pulse-basis along the generating arc [23], [24], [27].

been studied by various authors [21]–[23], [24]–[30], only a brief summary of the basic theory will be given here.

Considering Fig. 1, regions 1 and 2 are characterized by (possibly frequency dependent) medium parameters $(\varepsilon_1, \mu_1, \sigma_1)$ and $(\varepsilon, \mu_2, \sigma_2)$, respectively. According to surface equivalence principles [36]–[37], exterior scattered fields $(\boldsymbol{E}^{\text{scat}}, \boldsymbol{H}^{\text{scat}})$ and interior total fields $(\boldsymbol{E}, \boldsymbol{H})$ may be determined from a set of equivalent electric and magnetic surface currents. These equivalent surface currents are related to the total tangential magnetic and electric field by $\boldsymbol{J}_s = \pm\hat{\boldsymbol{n}} \times \boldsymbol{H}$ and $\boldsymbol{M}_s = \pm\boldsymbol{E} \times \hat{\boldsymbol{n}}$, respectively, where the upper sign holds for the exterior and the lower sign for the interior region. Coupled integral equations

$$\boldsymbol{E}^{\text{inc}}_{\tan}(\boldsymbol{r}) = \left[\sum_{i=1}^{2} \left(j\omega \boldsymbol{A}_i(\boldsymbol{r}) + \nabla\Phi_{ei}(\boldsymbol{r})\right)\right]_{\tan} + \left[\nabla \times \sum_{i=1}^{2} \frac{\boldsymbol{F}_i(\boldsymbol{r})}{\sigma_i + j\omega\varepsilon_i}\right]_{\tan} \tag{13a}$$

$$\boldsymbol{H}^{\text{inc}}_{\tan}(\boldsymbol{r}) = \left[-\nabla \times \sum_{i=1}^{2} \frac{\boldsymbol{A}_i(\boldsymbol{r})}{\mu_i}\right]_{\tan} + \left[\sum_{i=1}^{2} \left(j\omega \boldsymbol{F}_i(\boldsymbol{r}) + \nabla\Phi_{mi}(\boldsymbol{r})\right)\right]_{\tan} \tag{13b}$$

for the surface currents are obtained by satisfying the boundary conditions at the interface [27]–[30], where $(\boldsymbol{E}^{\text{inc}}_{\tan}, \boldsymbol{H}^{\text{inc}}_{\tan})$ represent the tangential components of the incident field. Herein the electric and magnetic vector and scalar potentials are defined as

$$\boldsymbol{A}_i(\boldsymbol{r}) = \mu_i \oiint_{S'} G_i(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{J}_s(\boldsymbol{r}')\, dS' \tag{14a}$$

$$\boldsymbol{F}_i(\boldsymbol{r}) = \frac{\sigma_i + j\omega\varepsilon_i}{j\omega} \oiint_{S'} G_i(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{M}_s(\boldsymbol{r}')\, dS' \tag{14b}$$

$$\Phi_{ei}(\boldsymbol{r}) = \frac{j\omega}{\sigma_i + j\omega\varepsilon_i} \oiint_{S'} G_i(\boldsymbol{r}, \boldsymbol{r}')q_{es}(\boldsymbol{r}')\, dS' = \frac{-1}{\sigma_i + j\omega\varepsilon_i} \oiint_{S'} G_i(\boldsymbol{r}, \boldsymbol{r}')\nabla' \cdot \boldsymbol{J}_s(\boldsymbol{r}')\, dS' \tag{15a}$$

$$\Phi_{mi}(\boldsymbol{r}) = \frac{1}{\mu_i} \oiint_{S'} G_i(\boldsymbol{r}, \boldsymbol{r}')q_{ms}(\boldsymbol{r}')\, dS' = \frac{-1}{j\omega\mu_i} \oiint_{S'} G_i(\boldsymbol{r}, \boldsymbol{r}')\nabla' \cdot \boldsymbol{M}_s(\boldsymbol{r}')\, dS' \tag{15b}$$

with the Green's function

$$G_i(\boldsymbol{r}, \boldsymbol{r}') = \frac{e^{-jk_i|\boldsymbol{r}-\boldsymbol{r}'|}}{4\pi|\boldsymbol{r}-\boldsymbol{r}'|} \quad \text{with} \quad k_i = \sqrt{\omega^2\varepsilon_i\mu_i - j\omega\mu_i\sigma_i} \quad \text{and} \quad i = 1, 2 \tag{16}$$

for the homogeneous interior $(i = 1)$ and homogeneous exterior $(i = 2)$ region, respectively. The complex wavenumbers of the interior and exterior region are given by k_1 and k_2, and the vectors $\boldsymbol{r}$ and $\boldsymbol{r}'$ represent source and observation point, respectively.

Equations (13)–(16) are valid for an arbitrarily shaped homogeneous target in a homogeneous environment, for which an MoM solution could be applied [36]. However, we only consider the special case of a BOR (Fig. 1), which is formed by rotating a generating arc about an axis that is chosen to be the z-axis of a Cartesian coordinate system. For numerical simulations, the generating arc is approximated by a sequence of linear segments. Coordinates (ϕ, t) are introduced, where the angle ϕ is equal to the one used in cylindrical coordinates (ρ, ϕ, z) and t is the length variable along the curve generating the arc. To take advantage of the rotational symmetry, the incident field, all currents, and scalar Green's functions are expanded into discrete Fourier series by taking a Fourier transform in the azimuthal (ϕ) direction, from which the general solution reduces to an infinite number of distinct MoM problems for each of the Fourier modes (each with azimuthal variation $\exp(jm\phi)$). The resulting sets of simultaneous equations may be represented in matrix form as

$$[\boldsymbol{Z}^m]\boldsymbol{I}^m = \boldsymbol{V}^m \quad \text{for } m = 0, \pm1, \pm2, \pm3, \cdots \tag{17}$$

where $[\boldsymbol{Z}^m]$ is the moment matrix, $\boldsymbol{I}^m$ is a column vector containing the unknown basis function coefficients, and $\boldsymbol{V}^m$ is the driving vector for the mth Fourier mode [21]–[23], [25]–[26]. Details regarding the calculation of the impedance matrix or the driving vector can be found in the literature.

If we are interested in the natural resonances of the target, the driving vector $\boldsymbol{V}^m$ is set to zero. Then the system (17) of linear equations has a nontrivial solution only if the determinant of the MoM impedance matrix is zero

$$\det[\boldsymbol{Z}^m(s_{m,\nu} = j\omega_{m,\nu}] = 0 \quad \text{with} \quad s_{m,\nu} = j\omega_{m,\nu} = j[\omega'_{m,\nu} + j\omega''_{m,\nu}] = -\omega''_{m,\nu} + j\omega'_{m,\nu}. \tag{18}$$

The roots of (18) in the complex frequency plane are the resonant frequencies of the modes (m, ν). Searching for these complex roots is relatively easy because, for highly conducting and permeable targets, the negative real part of $s_{m,\nu}$ is much larger than the imaginary part (almost pure exponential damping, as expected from Section II-A). Thus, it is sufficient to search first along the negative real s axis for an approximate solution, after which Mueller's method is performed to yield an

accurate solution for the complex resonant frequency (where we have found five–ten iterative steps generally sufficient).

For mode identification, it is in principle necessary to compute the detailed field distribution inside and outside the target. But in a limited sense, this may be accomplished by studying the resonant surface current distribution only. For each complex resonant frequency, the surface currents or, equivalently, the tangential fields along the surface can be readily calculated.

In contrast to MoM calculations for perfectly conducting objects, for which subsectional basis-function discretization only depends on the wavelength in the outside region (often free space), the maximum subsection length here has to be chosen with respect to the wavenumber inside the highly conducting and permeable object. Neglecting displacement currents and assuming a real permeability μ_1, this wavenumber is given by

$$k_1 = k_1' - jk_1'' = \sqrt{\frac{\sigma_1\mu_1}{2}}\,(\sqrt{|\omega| + \omega''} - j\sqrt{|\omega| - \omega''}). \quad (19)$$

Reasonable results are obtained if the subsection length Δt (along the generating arc) satisfies the conditions $k_1'\Delta t \le 2\pi/10$ and $k_1''\Delta t \le 1/2$, which means at least ten subsections per wavelength and at least two per skin depth. Although the required number of subdomains for a target that is much larger than the skin depth is impractical, this does not cause any problems because, for these cases (high frequency, high conductivity, and/or large target), the problem can be easily solved using a perfectly electric conducting (PEC) assumption.

Considering the wavenumber in (19), we discuss an alternative means of deriving the resonant frequency scaling relationship discussed in Section II-A (i.e., the relationship that the decay constant is proportional to the product of conductivity times the square of a characteristic target dimension). For purely imaginary frequencies, the wavenumber is real. Assuming real permeabilities $\mu_1 \gg \mu_2$ and a highly conducting target ($|\sigma_1 + j\omega\varepsilon_1| \approx \sigma_1 \gg |\sigma_2 + j\omega\varepsilon_2|$), from the theory of cavities or dielectric resonators [37], [28]–[30], the product of the real resonant wavenumber and a characteristic dimension D of the object does not change for a given shape. This results in

$$D\sqrt{\omega_{m,\nu}''\sigma_1\mu_1} = \text{const} \Rightarrow \omega_{m,\nu}'' = \frac{1}{\tau_{m,\nu}} \propto \frac{1}{D^2\sigma_1\mu_1}. \quad (20)$$

III. Cramer–Rao Lower Bound

The previous section demonstrated that the low-frequency transient response from a conducting target is characterized by exponential damping. Moreover, the decay constant is a function of the target shape, conductivity, and permeability. Therefore, in principle, it is possible to use the decay constants for target identification by estimating them from measured data [9], [14]–[15]. However, for noisy data, the extracted decay parameters must be characterized as random variables, with particular mean and variance. Assuming the mean is correct (constituting an unbiased estimate), it is of interest to quantify the variance, from which we can assess the practicality of achieving unambiguous target identification via decay-constant discrimination. To this end, we consider the Cramer-Rao lower bound (CRLB) [33]–[34], which represents the lowest possible variance for estimation of particular parameters from noisy data. The CRLB is a function of the SNR. Therefore, if we can quantify the maximum tolerable decay-parameter variance, for unambiguous identification (based on the range of decay constants anticipated), the CRLB will then quantify what SNR is required for this variance to be achieved, from which critical system requirements can be assessed. It should be noted that the CRLB does not suggest that an estimator exists that can achieve this minimum variance (termed an efficient estimator), only that no estimator can do better; however, several nearly efficient estimators are available for the estimation of exponential waveforms [14]–[15].

Let the vector $\boldsymbol{r}$ represent a discretized version of the measured time-domain waveform (of N samples)

$$\boldsymbol{r} = [r_0 \quad r_1 \quad r_2 \quad \cdots r_{N-1}]^T = \boldsymbol{g}(\boldsymbol{a}) + \boldsymbol{n} \quad (21)$$

where $\boldsymbol{g}(\boldsymbol{a})$ is the N-dimensional signal vector and $\boldsymbol{n}$ represents additive noise (here, assumed white Gaussian noise with variance σ^2). The signal $\boldsymbol{g}$ is assumed to be a function of the P-dimensional vector $\boldsymbol{a}$, which is to be estimated.

The CRLB states that the covariance matrix $[\boldsymbol{C}]$ of the estimated vector satisfies [33]–[34]

$$[\boldsymbol{C}] - [\boldsymbol{I}(\boldsymbol{a}]^{-1} = E\{(\hat{\boldsymbol{a}} - \boldsymbol{a})(\hat{\boldsymbol{a}} - \boldsymbol{a})^T\} - [\boldsymbol{I}(\boldsymbol{a})]^{-1} \ge 0 \quad (22)$$

where ≥ 0 means that the matrix is positive semidefinite and the Fisher information matrix [33]–[34] is expressed as

$$\begin{aligned}[\boldsymbol{I}(\boldsymbol{a})]_{ij} &= -E\left\{\frac{\partial^2 \ln\, p_{\boldsymbol{r}|\boldsymbol{a}}(\boldsymbol{r}|\boldsymbol{a})}{\partial a_i \partial a_j}\right\} \\ &= E\left\{\frac{\partial\, \ln\, p_{\boldsymbol{r}|\boldsymbol{a}}(\boldsymbol{r}|\boldsymbol{a})}{\partial a_i} \cdot \frac{\partial\, \ln\, p_{\boldsymbol{r}|\boldsymbol{a}}(\boldsymbol{r}|\boldsymbol{a})}{\partial a_j}\right\} \\ &\qquad \text{with} \quad i,j = 1,\cdots,P. \end{aligned} \quad (23)$$

Therefore, the lower bounds for the variances of the individual estimates ($\hat{a}_p$ $(p = 1,2,\cdots,P)$ are given by the pth diagonal element in the inverse of the Fisher matrix. The partial derivatives in (23) are evaluated at the true values, and the expectation is taken with respect to the conditional probability density function $p_{\boldsymbol{r}|\boldsymbol{a}}(\boldsymbol{r}|\boldsymbol{a})$.

We are interested in the parameters describing the natural response of a system. Assuming M poles s_m of single order, which is true for the conducting and permeable targets of interest here [3], [19], the system function in the Laplace domain and the corresponding impulse response are given by

$$\begin{aligned} H(s) &= \sum_{m=1}^{M} \frac{B_m}{s - s_m} \Leftrightarrow h(t) = \sum_{m=1}^{M} B_m e^{s_m t} u(t) \\ &= \sum_{m=1}^{M} b_m e^{j\beta_m} e^{j2\pi f_m' t} e^{-2\pi f_m'' t} u(t) \end{aligned} \quad (24)$$

where the excitation coefficients for pairs of conjugate complex poles $s_i = s_j^*$ satisfy $B_i = B_j^*$. Therefore the components

of the signal vector $\boldsymbol{g}(\boldsymbol{a})$ can be written as

$$g_n(\boldsymbol{a}) = \sum_{m=1}^{M} b_m e^{j\beta_m} e^{j2\pi f'_m n\Delta t} e^{-2\pi f''_m n\Delta t} \tag{25}$$

with $n = 0, 1, 2, \cdots, N-1$, where Δt represents the time increment for the sampling and, hence, $N\Delta t$ the total observation time. The vector $\boldsymbol{a}$ of the $P = 4M$ unknown parameters (excluding the noise variance σ^2) is given by

$$\begin{aligned}\boldsymbol{a} &= [a_1 \quad a_2 \quad \cdots \quad a_p]^T \\ &= [b_1 \quad \cdots \quad b_M \quad \beta_1 \quad \cdots \quad \beta_M \quad f'_1 \\ &\qquad \cdots \quad f'_M \quad f''_1 \quad \cdots \quad f''_M]^T\end{aligned} \tag{26}$$

and the likelihood function $p_{\boldsymbol{r}|\boldsymbol{a}}(\boldsymbol{r}|\boldsymbol{a})$ is [33]–[34]

$$\begin{aligned}p_{\boldsymbol{r}|\boldsymbol{a}}(\boldsymbol{r}|\boldsymbol{a}) &= p_{\boldsymbol{n}}(\boldsymbol{r} - \boldsymbol{s}(\boldsymbol{a})) \\ &= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{(\boldsymbol{r} - g(\boldsymbol{a}))^T \cdot (\boldsymbol{r} - g(\boldsymbol{a}))}{2\sigma^2}\right].\end{aligned} \tag{27}$$

Using this probability density function, the signal vector $\boldsymbol{g}$ in (25), and the parameter vector $\boldsymbol{a}$ in (26), the Fisher information matrix $[\boldsymbol{I}(\boldsymbol{a})]$, and the CRLB matrix $[\boldsymbol{C}]$ can be derived [35] for all $4M+1$ unknowns, including the noise variance σ^2. No further simplifications are necessary here, but for the actual computation of the CRLB matrix, the excitation strengths of the individual damped exponentials have to be known.

A simple approximate solution for the CRLB as a function of the SNR that does not require any knowledge of the excitation strengths can be derived, assuming known noise variance σ^2 and excitation of only the resonant mode of interest. This is a "best-case" study that provides a "lower bound" for the CRLB. Taking into account only the exponential decay term associated with the lowest purely imaginary resonant frequency (fundamental mode), i.e., neglecting all other poles, the time series simplifies to

$$\begin{aligned}g_n(B, f''_{\text{res}}) &= Be^{-2\pi f''_{\text{res}} n\Delta t} \\ &= Be^{-n\Delta t/\tau} \quad \text{with} \quad n = 0, 1, \cdots, N-1\end{aligned} \tag{28}$$

where we have renamed the amplitude and the resonant frequency for simplification. After some algebra, and assuming a sampling time much smaller and a total observation time much larger than the decay time constant associated with this lowest resonant frequency, i.e.,

$$\begin{aligned}\Delta t/\tau &= 2\pi f''_{\text{res}} \Delta t \ll 1 \quad \text{and} \\ N\Delta t/\tau &= 2\pi f''_{\text{res}} N\Delta t \gg 1\end{aligned} \tag{29}$$

the CRLB for the normalized resonant frequency or decay time, respectively, simplifies to

$$\text{var}\left\{\frac{\hat{f}''_{\text{res}}}{f''_{\text{res}}}\right\} = \text{var}\left\{\frac{\hat{\tau}}{\tau}\right\} = 4\sigma^2 \Big/ \sum_{n=0}^{N-1} g_n^2 = \frac{4}{\text{SNR}}. \tag{30}$$

As expected, the variance is reduced with increasing SNR. We have found that the error incurred by using (30) instead of the general CRLB in (22) is less than 5%, as long as the sampling time and the number of sampling points satisfy $\Delta t/\tau < 0.55$ and $N\Delta t/\tau > 3$.

IV. Results

The MoM code for BOR-type targets is used to calculate natural complex frequencies and resonant surface currents for different shapes and material parameters. For a sphere, thin disk, and cylinder, these results are compared with measurements. In addition, numerical results for a more general BOR are presented. Only results for $m = 0$ will be shown here, these rotationally symmetric modes often being most important in practice. Since from now on we assume free space for medium 2, we drop the subscript 1 to indicate material parameters of the target.

A. Sphere

To illustrate the frequency dependence of the MoM impedance matrix determinant, Fig. 2 shows its magnitude along the imaginary frequency axis for a sphere of radius $a = 1.27$ cm and azimuthal mode index $m = 0$. Although the resonances are not exactly located on the imaginary axis, such a plot gives a good idea as to the number and location of the resonances. The imaginary frequencies at which the minima of the magnitude occur are used as starting points for a Mueller method to search for the complex roots. Fig. 2 also shows a comparison between different material properties. The curves for nonpermeable targets $(\mu_r = 1)$ with conductivities $\sigma = 5 \cdot 10^6$ S/m and $\sigma = 20 \cdot 10^6$ S/m lie on top of each other (note the scaling of the frequency axis by a factor of four), which is consistent with the resonant frequency scaling relationship (20). The curves for $\sigma = 5 \cdot 10^6$ S/m and permeabilities $\mu_r = 1$ and $\mu_r = 100$, respectively, also use different frequency axes, here scaled by a factor of 100. TM^z-resonances with $m = 0$, i.e., modes with the magnetic field only in the azimuth direction (and therefore not excited by an incident z-directed magnetic field), scale exactly according to (20). For TE^z-resonances, this is only true if the permeability satisfies $\mu \gg \mu_0$. A simple explanation for this will be given in Section IV-C.

In [31]–[32], an impulse detector was used to measure the step response of certain canonical objects, in which the impulse response of the detector itself has been deconvolved from the measured data. Using a Helmholtz coil, the generated incident magnetic field was nearly uniform, as assumed in Section II-A. The scattered field was measured with a quadrupole receiver coil. Table I presents a direct comparison of the measured and MoM-computed decay times for a 1-in diameter sphere made of aluminum, brass, lead, and 304 stainless steel, respectively [32]. Taking into account that in the literature the conductivities given for different metals are highly variable, the agreement between measurement and numerical results is excellent. In fact, assuming that the measurements were performed with a high degree of accuracy, we can use our theory to extract from data the exact conductivities for the targets considered in the experiments.

B. Thin Disk

Next we consider aluminum disks of diameter $2a = 5$ cm and nominal thickness 3.175, 1.588, 0.794, and 0.508, respectively (1/8, 1/16, 1/32, and 1/50 in), for which measured decay

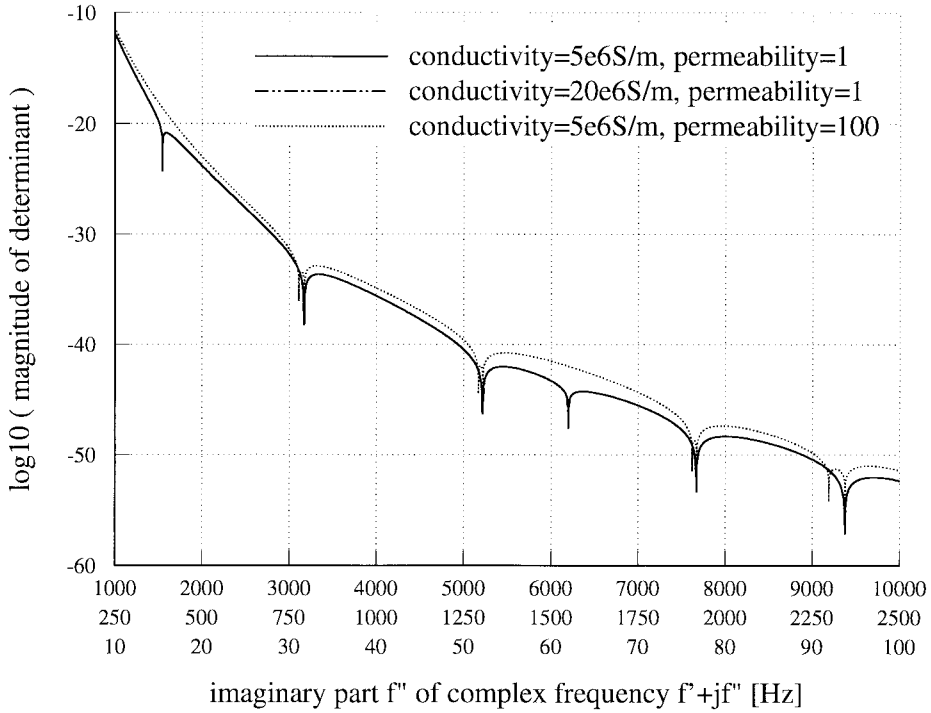


Fig. 2. Magnitude of MoM impedance matrix determinant ($m = 0$) along the imaginary axis of the complex frequency plane for spheres ($a = 1.27$ cm) of various material properties.

TABLE I
COMPARISON BETWEEN MEASURED DECAY TIME $\tau_{\rm meas} = 1/(2\pi \cdot {\rm Im}\{f_{\rm meas}\})$ [31]–[32] AND MoM SOLUTION FOR CONDUCTING SPHERES OF RADIUS $a = 1.27$ cm

metal type	τ_{meas} [msec]	Im{f_{meas}} [Hz]	Im{$f_{0,1}$} [Hz]	Re{$f_{0,1}$} [Hz]
aluminum ($\sigma \approx 30 \cdot 10^6$S/m)	≈ 0.477	≈ 333	260.5	0.016
brass ($\sigma \approx 12 \cdot 10^6$S/m)	≈ 0.255	≈ 624	651.2	0.041
lead ($\sigma \approx 5 \cdot 10^6$S/m)	≈ 0.095	≈ 1675	1562.9	0.097
stainless steel 304SS ($\sigma \approx 1.3 \cdot 10^6$S/m)	≈ 0.026	≈ 6121	6011.1	0.374

TABLE II
COMPARISON BETWEEN MEASURED DECAY TIME $\tau_{\rm meas} = 1/(2\pi \cdot {\rm Im}\{f_{\rm meas}\})$ [31]–[32] AND MoM SOLUTION FOR ALUMINUM DISKS OF RADIUS $a = 2.5$ cm AND DIFFERENT THICKNESSES

thickness of aluminum disk (radius a=2.5cm, $\sigma \approx 21 \cdot 10^6$S/m)	τ_{meas} [msec]	Im{f_{meas}} [Hz]	Im{$f_{0,1}$} [Hz]	Re{$f_{0,1}$} [Hz]
h≈3.175mm (1/8 inch)	≈ 0.363	≈ 438	466.1	0.169
h≈1.588mm (1/16 inch)	≈ 0.164	≈ 970	883.4	0.490
h≈0.794mm (1/32 inch)	≈ 0.089	≈ 1788	1718.3	1.213
h≈0.508mm (1/50 inch)	≈ 0.058	≈ 2744	2658.7	1.483

times are also available in [31]–[32] (Table II). Although the disks were machined from different sheets of aluminum of different alloys and unknown conductivities [32], we have assumed the same conductivity of $\sigma = 21 \cdot 10^6$ S/m for all disks. The fact that the decay time scales approximately with thickness does not violate (20) because (20) was derived assuming that the target shape does not change. Here only the thickness varies while the radius a is constant.

The resonant surface currents, i.e., the nontrivial solution of (17) for zero driving vector $\boldsymbol{V}^m = 0$ at a given resonant frequency, can be used for mode identification. For the aluminum disk with $h \approx 3.175$ mm (1/8 in.) Fig. 3 shows the resonant surface current distribution for the two lowest modes. Only the ϕ-component of the electric and the t-component of the magnetic surface currents are nonzero (TE^z-modes with $m = 0$). That is, because at very low frequencies, a thin disk can only support currents in circular paths around the axis of symmetry.

C. Cylinder

Despite the fact that a circular cylinder is very simple in shape, there is no closed form solution available for its natural resonances or scattering behavior. We discuss an approximate solution for modes with $m = 0$ later, but let us first compare numerical results with measured decay times for a solid cylinder of radius $a = 1.27$ cm, height $h = 2.54$ cm, and different conductivities. The decay times derived from the measured step responses in [31]–[32] and converted into imaginary resonant frequencies are compared to the calculated ones (Table III). The slight discrepancies could be due to the fact that the true conductivities of the targets are unknown.

Fig. 4 shows the frequency-dependent magnitude of the MoM determinant for different material properties, using different scales on the imaginary frequency axes. Again, all resonant frequencies scale almost exactly with $1/\sigma$ (the two curves for nonpermeable cylinders lie on top of each other and so only one curve is visible), but only a few resonances scale with $1/\mu$. Note also that the resonant frequencies for highly permeable objects occur in pairs with only slightly shifted resonant frequencies. We now give a simple explanation for

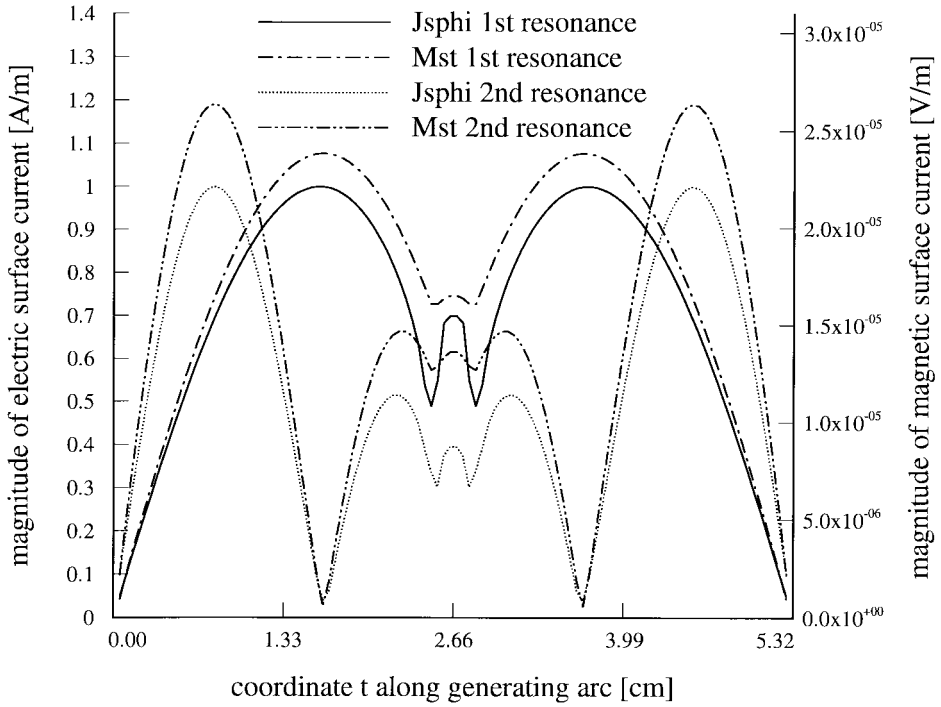


Fig. 3. Electric and magnetic resonant surface currents for the lowest two resonances ($m = 0$, TE^z-modes) of a 5-cm diameter aluminum disk with thickness $h \approx 3.175$ mm (1/8 in.).

TABLE III
COMPARISON BETWEEN MEASURED DECAY TIME $\tau_{\text{meas}} = 1/(2\pi \cdot \text{Im}\{f_{\text{meas}}\})$ [31]–[32] AND MoM SOLUTION FOR ALUMINUM DISKS OF RADIUS $a = 1.27$ cm AND HEIGHT $h = 2.54$ cm

metal type	τ_{meas} [msec]	Im$\{f_{meas}\}$ [Hz]	Im$\{f_{0,1}\}$ [Hz]	Re$\{f_{0,1}\}$ [Hz]
aluminum ($\sigma \approx 30 \cdot 10^6$S/m)	≈ 0.620	≈ 257	208.6	0.006
brass ($\sigma \approx 12 \cdot 10^6$S/m)	≈ 0.295	≈ 539	521.7	0.014
lead ($\sigma \approx 5 \cdot 10^6$S/m)	≈ 0.096	≈ 1658	1251.9	0.034
stainless steel 304SS ($\sigma \approx 1.3 \cdot 10^6$S/m)	≈ 0.036	≈ 4421	4815.4	0.130

this behavior. Neglecting displacement currents ($\sigma \gg \omega\varepsilon$) and assuming a real permeability, the wavenumber inside the target is real for purely imaginary frequencies [see (19)]. Therefore, we can apply the theory for circular cross-section cavities [37], if we modify the boundary conditions to derive the radial wavenumber k_p. For TM$^z_{0np}$-modes (azimuthal mode index $m = 0$, radial mode index n, axial mode index p), i.e., modes with the magnetic field only in the ϕ-direction, the boundary condition for the normal component of the electric field at the surface can be satisfied only if $E_p(\rho = a) \approx 0$ and $E_z(z = \pm h/2) \approx 0$ because $\sigma + \omega\varepsilon \approx \sigma \gg \omega\varepsilon_0$. The resonant frequencies are therefore given by

$$f''_{0np} \approx \frac{1}{2\pi\sigma\mu}\left[\left(\frac{\chi'_{0n}}{a}\right)^2 + \left(\frac{p\pi}{h}\right)^2\right] \quad \text{with} \quad n = 1, 2, 3, \cdots \quad \text{and} \quad p = 1, 2, 3, \cdots \tag{31}$$

where χ'_{0n} represents the nth zero of the derivative of the Bessel function $J_0(x)$. For nonpermeable targets, the corresponding TE^z-modes have lower resonant frequencies because the magnetic field components H_ρ and H_z inside the target have finite values at the boundary. For increasing permeability, however, the resonant frequencies of the TE^z-modes converge to those of the TM$^z_{0np}$-modes given in (31) (pairs of resonances in Fig. 4) because, dual to the TM^z-case, the boundary conditions $H_\rho(\rho = a) \approx 0$ and $H_z(z = \pm h/2) \approx 0$ have to be satisfied for $\mu \gg \mu_0$. The numerically calculated resonant frequencies for TM$^z_{0np}$-modes and the theoretical ones in (31) agree within less than 0.02% for all modes included in Fig. 4 ($mnp = 011$, 012, 013, 014, 021, 022), providing an excellent verification of our MoM code for BOR type objects.

To illustrate the explanation above, Figs. 5 and 6 show the resonant surface currents of corresponding TE^z- and TM^z-modes (for $m = 0$) for a nonpermeable cylinder with conductivity $\sigma = 5 \cdot 10^6$ S/m and the same dimensions as in Table III. The principal behavior of the electric/magnetic surface currents for the TE^z-modes is about the same as the magnetic/electric surface currents for the TM^z-modes. However, while the $M_{s\phi}$-currents (corresponding to E_t at the surface) for the TM^z-modes vanish at the cylinder edges, this is not true for the $J_{s\phi}$-currents (corresponding to H_t at the surface). Therefore, for TE^z-modes, the normal magnetic field does not vanish at the interface between target and free space. Increasing the permeability to $\mu_r = 100$, the resonant frequencies of these two TE^z-modes ($j26.24$ and $j57.52$ Hz) approach those of the TM^z-modes ($j26.94$ and $j57.94$ Hz), where the latter ones almost exactly scale with $1/\mu$. The TM^z resonant surface currents for the permeable cylinder look exactly like those in Fig. 6 (the electric surface currents only scaled by a factor $1/\mu_r$) and the TE^z surface currents show an almost dual behavior with no obvious differences ($J_{s\phi}$ looks like $M_{s\phi}$ in Fig. 6, M_{st} only scaled by a factor of approximately 1400 compared to J_{st} in Fig. 6).

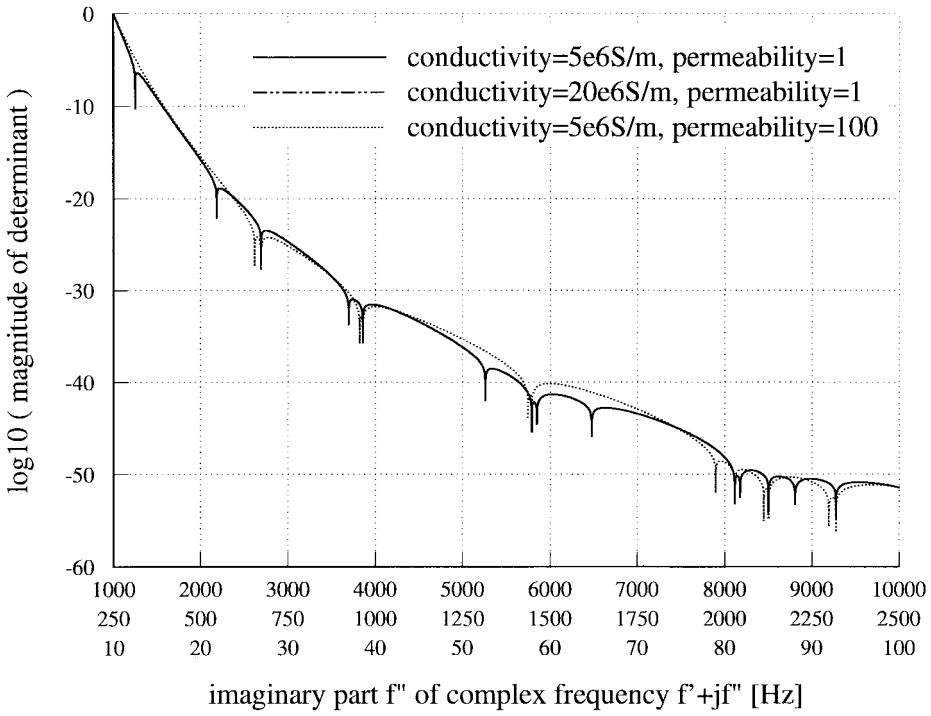


Fig. 4. Magnitude of MoM impedance matrix determinant ($m = 0$) along the imaginary axis of the complex frequency plane for cylinders ($a = 1.27$ cm, $h = 2.54$ cm) of various material properties.

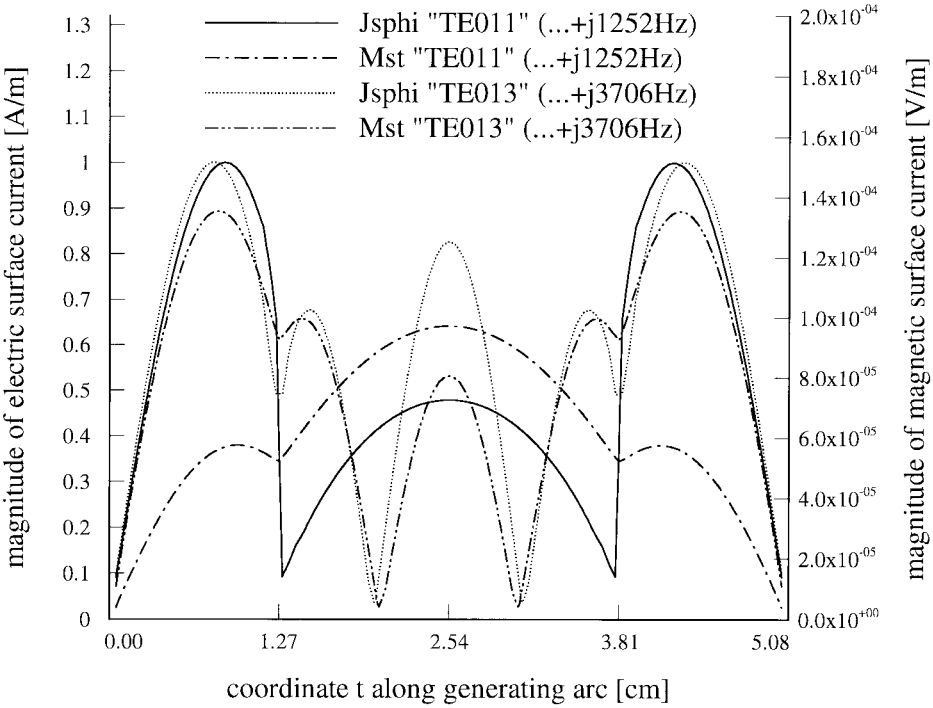


Fig. 5. Electric and magnetic resonant surface currents for two different TE^z-modes ($m = 0$) of a conducting and nonpermeable cylinder ($\sigma = 5 \cdot 10^6$ S/m, $a = 1.27$ cm, $h = 2.54$ cm).

D. BOR of Arbitrary Shape

The MoM solution in Section II-B is applicable not only to canonical type targets (sphere, disk, cylinder) but to arbitrary BOR's. As an example, we now look at the natural resonances of the target shown as inset in Fig. 7, which we believe cannot be calculated analytically. The characteristic dimension of the target is about the same as for the sphere and the cylinder in Section IV-A and C, respectively, and therefore we expect resonant frequencies of the same order. For the first three TE^z- and TM^z-modes with $m = 0$ the calculated resonant frequencies are given in Table IV for the nonpermeable conducting case ($\sigma = 5 \cdot 10^6$ S/m) as well as for conductivity

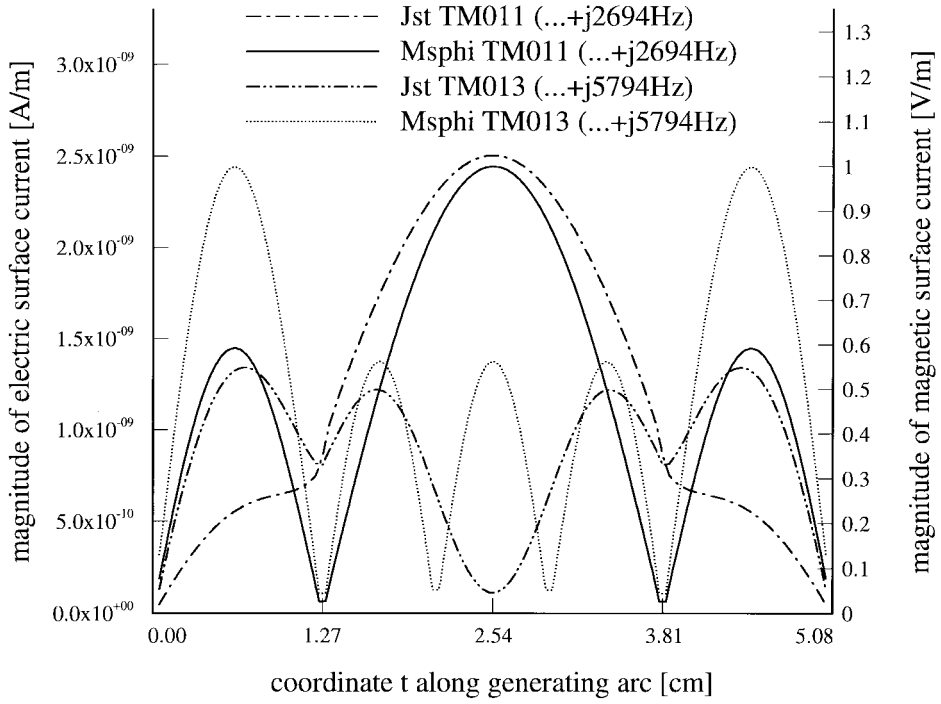


Fig. 6. Electric and magnetic resonant surface currents for two different modes of a conducting and nonpermeable cylinder ($\sigma = 5 \cdot 10^6$ S/m, $a = 1.27$ cm, $h = 2.54$ cm).

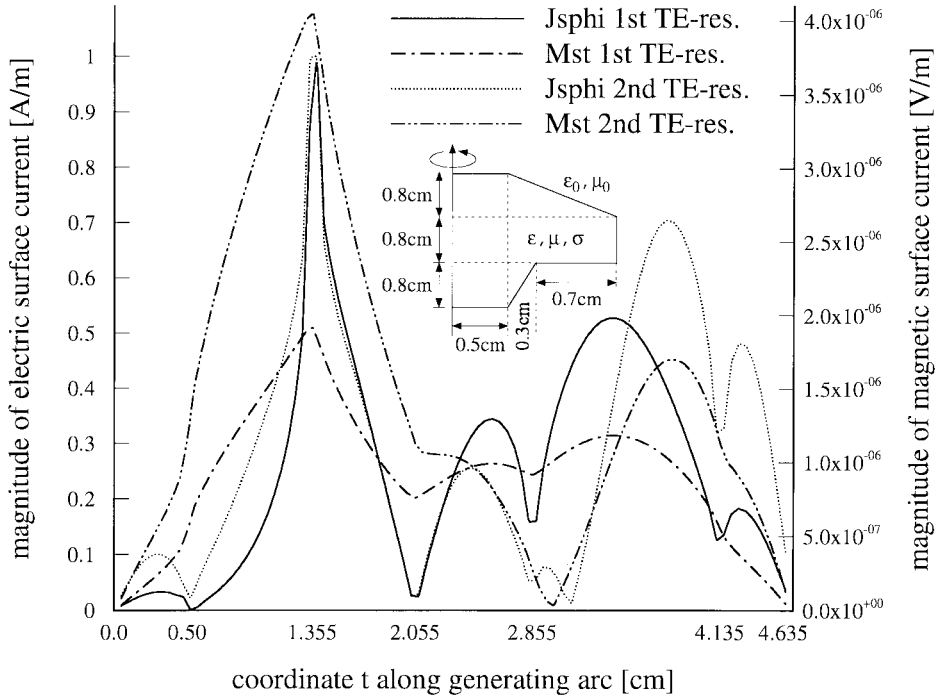


Fig. 7. Electric and magnetic resonant surface currents for the first two TE^z modes ($m = 0$) of a target (shown as inset) with conductivity $\sigma = 5 \cdot 10^6$ S/m and relative permeability $\mu_r = 100$.

$\sigma = 5 \cdot 10^6$ S/m and relative permeability $\mu_r = 100$. As for the canonical objects before, the resonant frequencies for the TM^z-modes exactly scale with $1/\mu$ (not the case for the TE^z-modes), and the TE^z resonant frequencies approach those for the TM^z-modes for increasing permeabilities. Finally, Fig. 7 shows the resonant surface currents for the lowest two TE^z-modes for the permeable case in Table IV ($\sigma = 5 \cdot 10^6$ S/m, $\mu_r = 100$).

E. Example for CRLB

Let us consider as an example the CRLB (Section III) for the cylinder of radius 1.27 cm and height 2.54 cm in

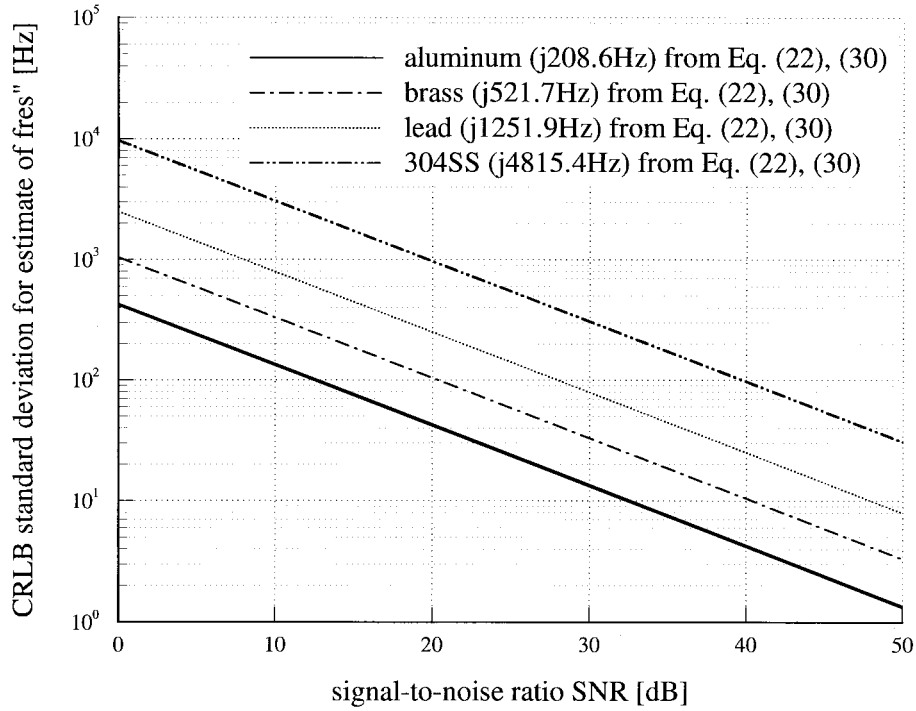


Fig. 8. Minimum standard deviation (Cramer–Rao) for the estimate of the imaginary part of the complex resonant frequency of conducting cylinders made of metals with different conductivities ($a = 1.27$ cm, $h = 2.54$ cm) for sampling with $N = 256$ and $\Delta t = 10$ μs.

TABLE IV
RESONANT FREQUENCIES OF THE TARGET SHOWN IN THE INSET OF FIG. 7 FOR $m = 0$ AND DIFFERENT PERMEABILITIES

σ=5·10^{6}S/m and μ_r=1		σ=5·10^{6}S/m and μ_r=100	
$f_{0,\nu}$ [Hz] TE^z-modes	$f_{0,\nu}$ [Hz] TM^z-modes	$f_{0,\nu}$ [Hz] TE^z-modes	$f_{0,\nu}$ [Hz] TM^z-modes
0.124+j·1370.8	0.035+j·3170.9	0.00036+j·31.007	0.00035+j·31.709
0.262+j·3638.4	0.132+j·6518.3	0.00124+j·64.083	0.00132+j·65.183
0.714+j·4489.3	0.145+j·7008.5	0.00170+j·69.336	0.00145+j·70.085

Section IV-C. The lowest resonant frequency, corresponding to the longest decay time, is $j208.6$ Hz for aluminum, $j521.7$ Hz for brass, $j1251.9$ Hz for lead, and $j4815.4$ Hz for stainless steel, respectively (see Table III). Taking into account only this lowest resonance, the standard deviation for the imaginary part of the complex resonant frequency as a function of the SNR is given in Fig. 8. Since we used $\Delta t = 10$ μs and $N = 256$ here (conditions in (29) are satisfied), the CRLB results from (30) differ only slightly from the general solution in (22) (the curves lie on top of each other in Fig. 8). Note that (30) gives the CRLB for the *relative* variance, but the standard deviation is plotted in Fig. 8. Moreover, if we assume Δ represents the minimum difference between the resonant frequencies considered above, and that a standard deviation of 0.25Δ is required for unambiguous identification of the four targets, an SNR of about 23 dB is needed in this example. If we relax this condition to 0.5Δ, an SNR of 17 dB is sufficient.

The example above considered a fixed target shape with variable conductivities. Now we vary the shape of a target while the conductivity and permeability are fixed. We are interested in the minimum SNR required to distinguish between 5-cm diameter aluminum disks of different thickness h ($\sigma = 21 \cdot 10^6$ S/m, $\mu = \mu_0$) (see Section IV-B). The MoM results for the lowest resonant frequencies (longest decay times) for $h \approx 3.175$, 1.588, 0.794, and 0.508 mm (1/8, 1/16, 1/32, and 1/50 in) can be found in Table II. The minimum standard deviation for the imaginary part of the complex resonant frequency (CRLB) as a function of SNR is given in Fig. 9 for the four different thicknesses. Although the coarse sampling assumed here ($\Delta t = 50$ μs, $N = 16$) does not satisfy the conditions in (29) very well, the error using (30) is still small [see Fig. 9, the true standard deviations from (22) are slightly larger than those from (30)]. If we again assume that a standard deviation less than 0.25Δ (definition of Δ analogue to first example) is required for unambiguous identification of the disks, an SNR of about 28 dB is needed (for 0.5Δ, an SNR of 22 dB is sufficient). In like manner, SNR requirements can be computed for any set of targets of interest.

V. CONCLUSIONS

A circuit-based analysis of low-frequency, electromagnetic scattering from conducting targets has been presented, from which it was demonstrated that the late-time (natural) response from such targets is represented by an exponentially damped response. A rigorous, MoM analysis was then applied to the case of conducting, permeable BOR's in free space, for which the damped natural response was demonstrated as still appropriate. Moreover, the MoM-computed decay constants were found to be in good agreement with measured data.

Several important relationships were derived and demonstrated, as follows.

1) For a conducting, nonpermeable target, the modal decay constants were shown to scale with conductivity, and

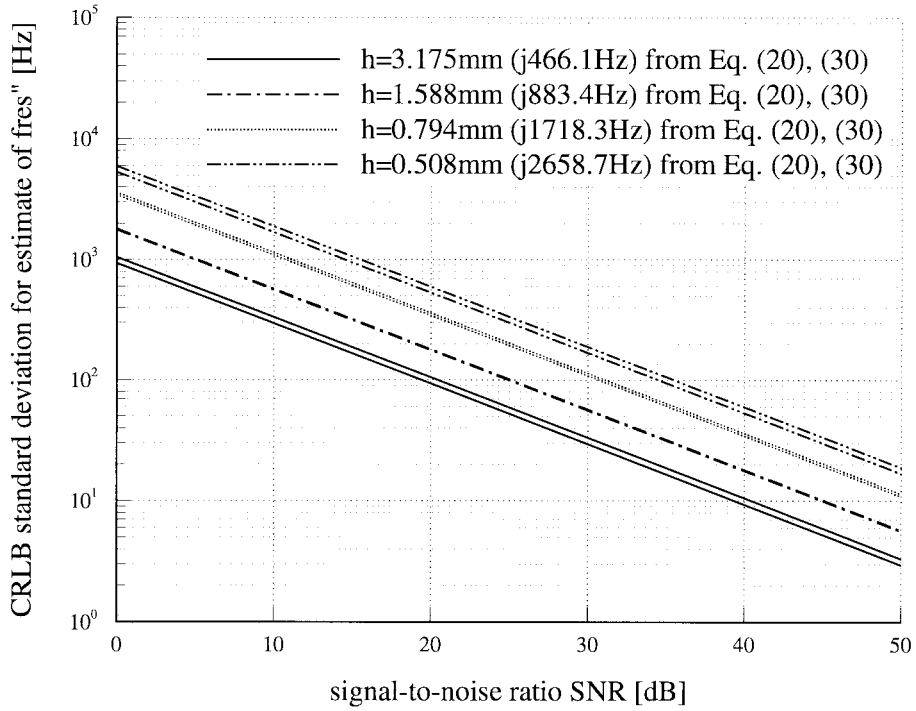


Fig. 9. Minimum standard deviation (Cramer–Rao) for the estimate of the imaginary part of the complex resonant frequency of 5-cm diameter aluminum disks of different thickness h ($\sigma = 21 \cdot 10^6$ S/m, $\mu = \mu_0$) for sampling with $N = 16$ and $\Delta t = 50$ μs [the results from (22) are slightly larger than those using the approximation (30)].

therefore results need only be computed for a representative conductivity, assuming the target shape remains constant.

2) Decay constants (with $m = 0$) for highly permeable and conducting targets can be computed very accurately ($< 1\%$ error) using a simple resonant-cavity analysis (particularly useful for simple shapes such as cylinders and discs).
3) CRLB for estimation of the decay parameters can often be simplified to a simple relationship that only depends on the SNR (computed from the noise variance and the decay coefficient).

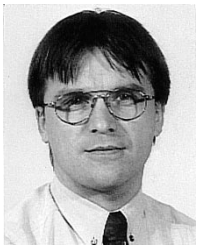
These relationships allow us to perform quick "back-of-the-envelope" computations of the system requirements for unambiguous target discrimination via decay-parameter estimation. In this context, for the examples considered here, we have found that an SNR of roughly 20 dB is required.

The low-frequency, decay-constant discrimination of conducting, permeable targets has many applications. For example, we can use it for concealed-weapon *identification* [31] as well as for identification of buried metal mines (in which case discrimination mitigates the need to dig up each piece of anthropic metal clutter). With regard to the latter example, the effects of the lossy soil become important as the size of the metal target diminishes (this is especially relevant for plastic mines that possess very small metal content). Therefore, in future work, we will consider the BOR targets buried in a lossy half space. This problem constitutes a significant escalation in complexity, due to the need to compute the half-space dyadic Green's function.

REFERENCES

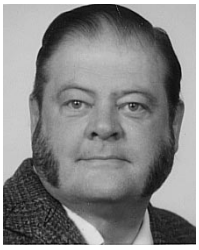
[1] B. L. Merchant, P. J. Moser, A. Nagl, and H. Überall, "Complex pole patterns of the scattering amplitude for conducting spheroids and finite-length cylinders," *IEEE Trans. Antennas Propagat.*, vol. 36, pp. 1769–1778, Dec. 1988.
[2] S. R. Vechinski and T. H. Shumpert, "Natural resonances of conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 1133–1136, July 1990.
[3] C. E. Baum, "Low-frequency near-field magnetic scattering from highly, but not perfectly, conducting bodies," Phillips Lab., Interaction Note 499, Nov. 1993.
[4] S. Vitebskiy and L. Carin, "Moment-method modeling of short-pulse scattering from and the resonances of a wire buried inside a lossy, dispersive half-space," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 1303–1312, Nov. 1995.
[5] ——, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy dispersive half-space," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.
[6] C. E. Baum, N. Geng, and L. Carin, "Integral equations and polarizability for magnetic singularity identification," Phillips Lab., Interaction Note 524, Mar. 1997.
[7] F. M. Tesche, "On the analysis of scattering and antenna problems using the singularity expansion technique," *IEEE Trans. Antennas Propagat.*, vol. AP-21, pp. 53–62, Jan. 1973.
[8] L. Marin, "Natural-mode representation of transient scattered fields," *IEEE Trans. Antennas Propagat.*, vol. 21, pp. 809–818, Nov. 1973.
[9] M. L. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. Antennas Propagat.*, vol. AP-23, pp. 777–781, Nov. 1975.
[10] A. J. Poggio, M. L. Van Blaricum, E. K. Miller, and R. Mittra, "Evaluation of a processing technique for transient data," *IEEE Trans. Antennas Propagat.*, vol. AP-26, pp. 165–173, Jan. 1978.
[11] D. G. Dudley, "Parametric modeling of transient electromagnetic systems," *Radio Sci.*, vol. 14, pp. 387–396, May–June 1979.
[12] C. E. Baum, R. J. Rothwell, K.-M. Chen, and D. P. Nyquist, "The singularity expansion method and its application to target identification," *Proc. IEEE*, vol. 79, pp. 1481–1492, Sept. 1991.
[13] C. E. Baum, "On the singularity expansion method for the solution of electromagnetic interaction problems," Phillips Lab., Interaction Note 88, Dec. 1971.

[14] Y. Hua and T. K. Sarkar, "Generalized pencil-of function method for extracting poles of an EM system from its transient response," *IEEE Trans. Antennas Propagat.*, vol. 37, pp. 229–234, Feb. 1989.
[15] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 814–824, May 1990.
[16] R. W. P. King and C. W. Harrison, "The transmission of electromagnetic waves and pulses into the earth," *J. Appl. Phys.*, vol. 39, pp. 4444–4452, Aug. 1968.
[17] D. L. Moffatt and R. J. Puskar, "A subsurface electromagnetic pulse radar," *Geophysics*, vol. 41, pp. 506–518, June 1976.
[18] G. S. Smith and W. R. Scott, "A scale model for studying ground penetrating radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 27, pp. 358–363, July 1989.
[19] C. E. Baum, Ed., *Detection and Identification of Visually Obscured Targets*. New York: Taylor & Francis, 1998.
[20] C. Liu and L. C. Shen, "Numerical simulation of subsurface radar for detecting buried pipes," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 795–798, Sept. 1991.
[21] M. G. Andreasen, "Scattering from bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 13, pp. 303–310, Mar. 1965.
[22] J. R. Mautz and R. F. Harrington, "Radiation and scattering from bodies of revolution," *Appl. Sci. Res.*, vol. 20, pp. 405–435, June 1969.
[23] S. D. Gedney and R. Mittra, "The use of the FFT for the efficient solution of the problem of electromagnetic scattering by a body of revolution," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 313–322, Mar. 1990.
[24] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse plane-wave scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 143–151, Feb. 1996.
[25] T. Wu and L. L. Tsai, "Scattering from arbitrarily-shaped lossy dielectric bodies of revolution," *Radio Sci.*, vol. 12, pp. 709–718, Sept.–Oct. 1977.
[26] J. R. Mautz and R. F. Harrington, "Electromagnetic scattering from a homogeneous material body of revolution," *AEÜ*, vol. 33, pp. 71–80, Feb. 1979.
[27] A. W. Glisson and D. R. Wilton, "Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces," *IEEE Trans. Antennas Propagat.*, vol. 28, pp. 593–603, Sept. 1980.
[28] A. W. Glisson, D. Kajfez, and J. James, "Evaluation of modes in dielectric resonators using a surface integral equation formulation," *IEEE Trans. Microwave Theory Tech.*, vol. 31, pp. 1023–1029, Dec. 1983.
[29] A. W. Glisson, "Integral equation techniques," in *Dielectric Resonators*, D. Kajfez and P. Guillon, Eds. Dedham, MA: Artech House, 1986, pp. 259–325.
[30] D. Kajfez, A. W. Glisson, and J. James, "Computed modal field distributions for isolated dielectric resonators," *IEEE Trans. Microwave Theory Tech.*, vol. 32, pp. 1609–1616, Dec. 1984.
[31] G. D. Sower and S. P. Cave, "Detection and identification of mines from natural magnetic and electromagnetic resonances," in *Proc. SPIE*, Orlando, FL, vol. 2496, pp. 1015–1024, 1995.
[32] G. D. Sower, "Eddy current resonances of canonical metallic targets—Theory and measurements," EG&G MSI, Interaction Note, Feb. 1997.
[33] K. Kroschel, *Statistische Nachrichtentheorie—Erster Teil— Signalerkennung und Parameterschätzung*. Berlin, Germany: Springer-Verlag, 1986 (in German).
[34] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[35] W. M. Steedly and R. L. Moses, "The Cramer-Rao bound for pole and amplitude coefficient estimates of damped exponential signals in noise," *IEEE Trans. Signal Processing*, vol. 41, pp. 1305–1318, Mar. 1993.
[36] R. F. Harrington, *Field Computation by Moment Methods*. New York: Macmillan, 1968.
[37] C. A. Balanis, *Advanced Engineering Electromagnetics*. New York: Wiley, 1989.



Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Karlsruhe, Germany, in 1991 and 1996, respectively.

He was with the Institut für Höchstfrequenztechnik und Elektronik (IHE), University of Karlsruhe, from 1991 to 1996, working on full-wave propagation modeling for radio communication systems. In January 1997, he joined the Department of Electrical and Computer Engineering, Duke University, Durham, NC, as a Visiting Postdoctorate for 18 months. His current research interests include computational methods in electromagnetics, with emphasis on short-pulse scattering and wave propagation modeling.



Carl E. Baum (S'62–M'63–SM'78–F'84) [redacted] He received the B.S. (with honosr), M.S., and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1962, 1963, and 1969, respectively.

He was commissioned in the United States Air Force in 1962 and was stationed at the Air Force Research Laboratory, Directed Energy Directorate (formerly, Phillips Laboratory, formerly, Air Force Weapons Laboratory), from 1963 to 1967 and from 1968 to 1971. Since 1971, he has served as a civil servant with a position as a Senior Scientist at Phillips Laboratory, Kirtland Air Force Base, Albuquerque, NM. He has published three books: *Transient Lens Synthesis: Differential Geometry in Electromagnetic Theory*, *Electromagnetic Symmetry*, and *Ultra-Wideband, Short-Pulse Electromagnetics 3*. He is Editor of several interagency note series on electromagnetic pulse (EMP) and related subjects. He has led an EMP short course and HPE workshops at numerous locations around the globe.

Dr. Baum received the Air Force Research and Development Award in 1970, the Richard R. Stoddard Award from the IEEE EMC Society in 1984, the IEEE Harry Diamond Memorial Award in 1987 with the citation "for outstanding contributions to the knowledge of transient phenomena in electromagnetics," the AFSC Harold Brown Award in 1990, and Phillips Laboratory Fellow in 1996. He is a member of Commissions A, B, and E of the United States National Committee of the International Union of Radio Science (URSI). He is Founder and President of the SUMMA Foundation, which sponsors various electromagnetics-related activities, including scientific conferences, publications, short courses, fellowships, and awards.

Lawrence Carin (S'85–M'89–SM'96) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

He was with the Electrical Engineering Department at Polytechnic University, Brooklyn, NY, as an Assistant Professor in 1989, and he became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, Durham, NC, where he is an Associate Professor. His current research interests include quasiplanar transmission lines, short-pulse scattering and propagation, and signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.

An Improved Bayesian Decision Theoretic Approach for Land Mine Detection

Leslie Collins, *Member, IEEE*, Ping Gao, *Student Member, IEEE*, and Lawrence Carin, *Senior Member, IEEE*

***Abstract*— A rigorous signal detection theoretic analysis is used to improve detectability of land mines. The development is performed for sensors that integrate time-domain information to provide a single data point (standard metal detector), those that provide a sampled portion of the time-domain waveform, and those that operate at several discrete frequencies. This approach is compared to standard thresholding techniques, and it is shown to provide substantial improvements when evaluated on measured data.**

***Index Terms*—Bayes procedures, data processing, electromagnetic induction, signal detection.**

I. Introduction

THE GOAL of any detection system is to achieve a high detection rate while minimizing the number of false alarms. While many mine detectors achieve the first of these goals, it is often at the cost of a prohibitively large false alarm rate. This often results from the fact that mine detectors simply seek anomalies, which may occur as a result of mines or clutter, and do not incorporate any knowledge of the statistics of the sensor response to mines and clutter. In this paper, a Bayesian decision-theoretic approach to the detection of mines is presented that incorporates the statistical properties of both the target and clutter. Results are presented for three different classes of electromagnetic induction (EMI) devices: integrated time-domain, time-domain, and frequency-domain. Detectors for each class of device are derived theoretically and evaluated based on data obtained in the field. In particular, we consider the Geonics EM61, an integrated time-domain, pulsed-induction device; the Geonics EM61-3D, a prototype, multichannel time-domain, pulsed-induction device; and Geophex's GEM-3, a prototype, broadband, frequency-domain device. Our first goal is to determine whether, at a fixed detection rate, a Bayesian decision-theoretic approach substantially reduces false alarm rates relative to the standard thresholding techniques commonly used on such data. In addition to determining whether such an approach is useful for mitigating false alarms, we investigate the degree to which additional knowledge of the time- or frequency-domain signatures (EM61-3D and GEM-3, respectively) could be used within the same Bayesian detection framework to improve performance over that obtained with the more standard integrated time- or frequency-domain measurements (EM61).

Using signal detection theory, the physics of each of the EMI systems is incorporated into the detector explicitly. The signal detection-theoretic approach prescribes a method for optimally processing the signals for detection and identification in an uncertain or random environment. When applied correctly, this approach guarantees optimal performance in the sense of minimum probability of error, or maximum detection rate for a given false alarm rate. Additionally, signal detection theory can be used to incorporate *a priori* knowledge concerning the set of values that parameters of the environment or the signal can take on. This approach provides both an improved detection algorithm and quantitative performance evaluation measures in the form of a functional relationship between probability of detection (P_d) and probability of false alarm (P_{fa}), termed the receiver operating characteristic (ROC). We present results in which the statistical, Bayesian approach significantly outperforms (lower P_{fa} for the same P_d) the standard threshold detector on the data obtained from the DARPA backgrounds clutter data collection experiment [1], which was sponsored by the Defense Advanced Research Projects Agency.

In the remainder of this paper, the data and the sensors that were used for this analysis are described. The underlying physics of the EMI sensors is reviewed briefly and followed by a discussion of the detection strategies. In addition to describing the Bayesian detection strategies, the standard techniques commonly applied to the data are described and their relationship to the Bayesian techniques are addressed. Finally, the results of the Bayesian and standard detection techniques for each of the three sensors are discussed.

II. Review of Underlying Physics

EMI systems generally operate at very low frequencies (<1 MHz) at which the displacement-current component to the near-field loop-induced fields can often be ignored. Therefore, such systems are relatively insensitive to dielectric discontinuities and primarily sense changes in conductivity and permeability. Thus, EMI systems are of primary use for the detection of mines with substantial metal content, or for unexploded ordnance (UXO), which generally contain significant metal. Traditional EMI systems are of little or no use for purely dielectric targets (mines).

Nevertheless, most mines have at least some metal content, and EMI systems are currently the only fieldable sensors for detecting such targets. Moreover, when operated in the time- or

Manuscript received September 29, 1997; revised June 8, 1998. This work was supported by the Army Research Office under Grant DAAH04-96-1-0448 (Demining MURI).

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA (e-mail: lcollins@ee.duke.edu).

Publisher Item Identifier S 0196-2892(99)00841-4.

0196–2892/99$10.00 © 1999 IEEE

wideband frequency-domain, EMI systems afford the potential for target discrimination, which mitigates the need to dig up each piece of metal detected. In particular, it has been shown [2]–[5] that the pulsed-EMI response $s(t)$ from a general conducting, permeable target can be expressed as

$$s(t) = \sum_{n=1}^{M} A_n \exp(-\alpha_n t). \tag{1}$$

Thus, the target response is characterized as a sum of M-weighted, exponentially damped modes. In practice, the lowest mode dominates the response, and only one exponential term (α_1) need be considered. Moreover, theoretical and experimental data demonstrate that α_1 is unique to each metal type and is a function of conductivity and permeability. It is this phenomenon we wish to exploit to distinguish targets from clutter. The amplitude of the response A_1 corresponds to the excitation level, and it is a complicated function of, among other things, the distance between the target and sensor and the orientation of the target. In this initial work, we chose to normalize the amplitude of the response to unity to focus on the discrimination capability inherent in the decay rate, as well as to avoid uncertainty in the target to sensor distance. Thus, hereafter, $A_1 = 1$.

III. Detector Design

A. Integrated Time-Domain Sensors

In order to improve detection performance, we consider utilizing signal detection theory to incorporate the statistics of the target and clutter into the detector design. The signal model that is adopted assumes that the sensor integrates the voltage induced in the receive coil from an initial time t_0 to some final time t_1 by sampling the received signal and summing the samples to a scalar value s, i.e.,

$$s = \sum_{k=0}^{N-1} \exp(-\alpha(t_0 + k\Delta t)) \tag{2}$$

where $t_0 + (N-1)\Delta t = t_1$. Finally, it is assumed that the sensor is subject to additive noise $n(k)$ at each sample point so that the received data r can be modeled as

$$r = \sum_{k=0}^{N-1} [\exp(-\alpha(t_0 + k\Delta t)) + n(k)]. \tag{3}$$

1) Known Parameter Case: The detector will be formulated in terms of a binary hypothesis test assuming equally likely hypotheses. H_1 denotes the target-present hypothesis, and H_0 denotes the null, or no-target-present, hypothesis. The likelihood ratio test (LRT) to be performed on the received data r is defined as

$$\Lambda(r) = \frac{p(r/H_1)}{p(r/H_0)} \tag{4}$$

where $p(r/H_1)$ and $p(r/H_0)$ are the probability density functions describing the received data under H_1 and H_0, respectively [6]. If $\Lambda(r) > \delta$, where δ is a threshold set to achieve a certain false alarm rate, H_1, is chosen, whereas, if $\Lambda(r) < \delta$, H_0 is chosen. If $p(r/H_1)$ and $p(r/H_0)$ are Gaussian distributions with identical variances, but different means, the likelihood ratio test reduces to the standard threshold test. However, when these assumptions are not met, a threshold test is not, in general, optimal.

In order to effect a likelihood ratio detector, it is necessary to obtain a valid estimate of the probability density functions of the data under H_1 and H_0. Under ideal conditions, in which both the clutter and targets are completely homogeneous, $s(t) = \exp(-\alpha_1 t)$ under H_1 and $s(t) = \exp(-\alpha_0 t)$ under H_0 (α_1 is thus representative of the target, while α_0 is induced by the background, e.g., soil). In this case, α_1 and α_0 are assumed to be constant and known. Under these assumptions and assuming that the sensor is subject to additive Gaussian noise $n(k)$, with zero mean and known variance σ^2, then under H_1 the received signal, r, is

$$r = \sum_{k=0}^{N-1} [\exp(-\alpha_1(t_0 + k\Delta t)) + n(k)] \tag{5}$$

and under H_0 the received signal is

$$r = \sum_{k=0}^{N-1} [\exp(-\alpha_0(t_0 + k\Delta t)) + n(k)]. \tag{6}$$

(The variance σ^2 was computed from the data described in Section IV-A by calculating the variance in the responses obtained when the measurements were obtained at a single location multiple times prior to the beginning of the data collection.) Under these assumptions, the density functions follow a Gaussian distribution, or $p(r/H_1) \sim N(K_1, N\sigma^2)$ and $p(r/H_0) \sim N(K_0 \sim N\sigma^2)$, where $K_1 = \sum_{k=0}^{N-1} \exp(-\alpha_1(t_0 + k\Delta t))$ and $K_0 = \sum_{k=0}^{N-1} \exp(-\alpha_0(t_0 + k\Delta t))$. When $K_1 > K_0$, the threshold detector is in fact optimum.

2) Unknown Parameter Case: When the parameters α_1 and α_0 are not known exactly, or are known to vary spatially, the probability density distributions under the two hypotheses will not be as described above. Since targets and clutter are rarely homogeneous, we expect that these assumptions are not met in practice; this was confirmed upon inspection of the EM61 data from the DARPA study. In this case, the LRT becomes [6]

$$\Lambda(r) = \frac{\int p(r/H_1, \alpha) p(\alpha_1)\, d\alpha_1}{\int p(r/H_0, \alpha_0) p(\alpha_0)\, d\alpha_0} \tag{7}$$

where $p(\alpha_1)$ and $p(\alpha_0)$ are the probability density functions describing α_1 and α_0, respectively.

In the mine detection problem, it is usually possible to develop estimates of these probability density functions prior to initiating the detection process either by gathering data and estimating the density functions, or by employing *a priori* information regarding the metal content, and thus decay rates, of the targets in question. In this instance, the former approach was taken, both since a "calibration area" was provided in the data set (see Section IV-A for a description) and the statistical nature of the clutter process was expected to vary across the data collection sites. In order not to overfit the measured data, the parameters of either a Gaussian model (i.e., mean and

variance) or a uniform model (i.e., upper and lower limits) were estimated from the data. This approach was followed for all of the sensors. In all cases, the performance did not vary significantly as a function of the statistical form of the model that was used.

To simplify the notation, let $s(\alpha_1) = \sum_{k=0}^{N-1} \exp(-\alpha_1(t_0 + k\Delta t))$ and $s(\alpha_0) = \sum_{k=0}^{N-1} \exp(-\alpha_0(t_0 + k\Delta t))$. For the case of additive white Gaussian noise, the LRT can be expressed as

$$\Lambda(r) = \frac{\displaystyle\int_{\alpha_1^L}^{\alpha_1^U} \exp\left(-\frac{(r - s(\alpha_1))^2}{2N\sigma^2}\right) p(\alpha_1)\, d\alpha_1}{\displaystyle\int_{\alpha_0^L}^{\alpha_0^U} \exp\left(\frac{(r - s(\alpha_0))^2}{2N\sigma^2}\right) p(\alpha_0)\, d\alpha_0} \tag{8}$$

where the limits of integration α_1^U, α_1^L, α_0^U and α_0^L are determined from the appropriate density functions and may be plus or minus infinity in some cases.

Although this integral can be calculated via Monte Carlo integration [7], we will prove that when the mean of the density function of α_1 is less than the mean of the density function of α_0 (i.e., responses to targets decay more slowly on the average than responses to the background or clutter), this likelihood ratio is a monotonic function of r, and therefore r is a sufficient statistic [6]. Thus, if the likelihood ratio is a monotonic function of r, a threshold test of the received data is optimal, even when the parameters of the data are uncertain. To prove this relationship, we take the derivative of the likelihood ratio with respect to the received data r and show that it is always greater than zero. Writing $\Lambda(r)$ as $\Lambda(r) = f(r)/g(r)$, we need to demonstrate that

$$h(r) = \left(\frac{\partial}{\partial r} f(r)\right) g(r) - f(r)\left(\frac{\partial}{\partial r} g(r)\right) > 0. \tag{9}$$

Taking the appropriate derivatives

$$\begin{aligned} h(r) = &\int_{\alpha_1^L}^{\alpha_1^U} \frac{-2(r - s(\alpha_1))}{2N\sigma^2} \exp \\ &\left(-\frac{(r - s(\alpha_1))^2}{2N\sigma^2}\right) p(\alpha_1)\, d\alpha_1 \int_{\alpha_0^L}^{\alpha_0^U} \exp \\ &\left(-\frac{(r - s(\alpha_0))^2}{2N\sigma^2}\right) p(\alpha_0)\, d\alpha_0 \\ &- \int_{\alpha_0^L}^{\alpha_0^U} \frac{-2(r - s(\alpha_0))}{2N\sigma^2} \exp \\ &\left(-\frac{(r - s(\alpha_0))^2}{2N\sigma^2}\right) p(\alpha_0)\, d\alpha_0 \int_{\alpha_1^L}^{\alpha_1^U} \exp \\ &\left(-\frac{(r - s(\alpha_1))^2}{2N\sigma^2}\right) p(\alpha_1)\, d\alpha_1. \end{aligned} \tag{10}$$

Simplifying, and combining the integrals

$$\begin{aligned} h(r) = &\frac{1}{N\sigma^2} \int_{\alpha_1^L}^{\alpha_1^U} \int_{\alpha_0^L}^{\alpha_0^U} (s(\alpha_1) - s(\alpha_0)) \exp \\ &\left(-\frac{(r - s(\alpha_1))^2}{2N\sigma^2}\right) \exp\left(-\frac{(r - s(\alpha_0))^2}{2N\sigma^2}\right) \\ &\cdot p(\alpha_1) p(\alpha_0)\, d\alpha_1\, d\alpha_0. \end{aligned} \tag{11}$$

Because probability density functions are always greater than or equal to zero, $h(r)$ is always greater than or equal to zero if the first term of the above equation $s(\alpha_1) - s(\alpha_0)$ is greater than zero. When $\alpha_1 < \alpha_0$, the first term is positive and $h(r) > 0$. If α_1 is not always less than α_0, but the mean of $p(\alpha_1)$ is less than the mean of $p(\alpha_0)$, the double integral can be divided into four areas of integration as

$$\begin{aligned} &\int_{\alpha_0^L}^{\alpha_0^U} \left(\int_{\alpha_1^L}^{\alpha_0^L} (\,)\, d\alpha_1 + \int_{\alpha_0^L}^{\alpha_1^U} (\,)\, d\alpha_1\right) d\alpha_0 \\ &\quad + \int_{\alpha_1^U}^{\alpha_0^U} \left(\int_{\alpha_1^L}^{\alpha_0^L} (\,)\, d\alpha_1 + \int_{\alpha_0^L}^{\alpha_1^U} (\,)\, d\alpha_1\right) d\alpha_0. \end{aligned} \tag{12}$$

In the first, third, and fourth integrals, $\alpha_1 < \alpha_0$, so these contribute only positive area to $h(r)$. For example, in the first integral, $\alpha_1^L < \alpha_1 < \alpha_0^L$ and $\alpha_0^L < \alpha_0 < \alpha_1^U$; thus, $\alpha_1 < \alpha_0$. In the second integral, the fact that α_1 and α_0 are integrated over identical regions, coupled with the additive nature of and similarity in form of $s(\alpha_1)$ and $s(\alpha_0)$ results in a contribution of zero from this term since

$$\begin{aligned} &\int_{\alpha_0^L}^{\alpha_1^U} \int_{\alpha_0^L}^{\alpha_1^U} \left[\sum_{k=0}^{N-1} \exp(-\alpha_1(t_0 + k\Delta t)) \right. \\ &\qquad \left. - \sum_{k=0}^{N-1} \exp(-\alpha_0(t_0 + k\Delta t))\right] d\alpha_1\, d\alpha_0 \\ &= \int_{\alpha_0^L}^{\alpha_1^U} \left[\sum_{k=0}^{N-1} \int_{\alpha_0^L}^{\alpha_1^U} \exp(-\alpha_1(t_0 + k\Delta t))\, d\alpha_0 \right. \\ &\qquad \left. - \sum_{k=0}^{N-1} \int_{\alpha_0^L}^{\alpha_1^U} \exp(-\alpha_0(t_0 + k\Delta t))\, d\alpha_0\right] d\alpha_1 \\ &= 0. \end{aligned} \tag{13}$$

Therefore, $h(r) > 0$ for the previously noted assumptions on the parameters involved.

This analysis has shown that, for integrated time-domain sensors, such as the EM61 operating under the assumption of Gaussian sensor noise with known statistics, the optimal processor for the detection of targets is a simple threshold statistic. This result follows directly from the likelihood ratio when the target and clutter responses are known exactly. However, even under conditions of uncertainty, as long as the rate of decay of target signatures is slower on average than that of clutter, the likelihood ratio is a monotonic function of the integrated response, and thus the optimal processor is still a threshold test. It will be shown in the next section, however, that, if the entire signal is available for processing, the optimal processor is not a threshold test performed on the integrated data values.

B. Time-Domain Sensors

In addition to incorporating a model of the received waveform into the detector formulation, we exploit the fact that an EMI device is effectively a "change detector," as long as the size of the objects of interest are small with respect to the spacing of the response measurements. This approach implicitly incorporates the assumption that the background

clutter is spatially correlated. For simplicity, it is assumed that, when no target is present, the statistics of the responses at two adjacent points are the same. However, if a small metal and/or permeable target is present, the response at the two locations will necessarily be different. The target must be small with respect to the spacing of the response measurements to ensure that target responses are not observed at sequential measurement locations.

By taking this approach, it is not necessary to specify the target response exactly, as would be the case for a matched filter [6]. This approach also avoids limiting the detector to searching for a particular decay rate, which corresponds to looking for a particular type of metal. Instead, this approach allows the uncertainty in the ground surface, composition of the mine, and its placement in the surrounding environment to be incorporated into the design of the processor.

Let $r_{x,y}(k)$ be the normalized, sampled, received signal at location $\{x, y\}$ on a site, where

$$r_{x,y}(k) = s_{x,y}(k) + n_{x,y}(k). \tag{14}$$

Again, $n_{x,y}(k)$ is assumed to be a white Gaussian noise process with known statistics associated with the sensor, and as before

$$s_{x,y}(k) = \exp(-\alpha_{x,y}(t_0 + k\Delta t)). \tag{15}$$

Note that under the no-target, or null hypothesis (H_0), $\alpha_{x,y}$ is associated with the response of the ground, or a clutter object, whereas under the target hypothesis, H_1, $\alpha_{x,y}$ is determined by the target response. If $\alpha_{x,y} = \alpha_1$ is a single, known constant for targets, and a different, known constant for clutter (α_0), the optimum test is a matched filter [6], or correlation receiver, in which the received signal is correlated with the difference between a decaying exponential due to a target and that due to clutter, or

$$\Lambda(\mathbf{r}_{x,y}) = \sum_{k=0}^{N-1} r_{x,y}(k)[\exp(-\alpha_1(t_0 + k\Delta t)) - \exp(-\alpha_0(t_0 + k\Delta t))]. \tag{16}$$

It is important to note that, under these assumptions, the optimal processing approach is a weighted sum of the received data, and not an unweighted sum, such as that performed by integrated EMI sensors. However, modeling and experimental results, such as those obtained in the DARPA experiment, indicate that the decay rates across targets and clutter are quite variable. Therefore, the matched filter approach is not necessarily optimal.

An alternate approach can be obtained by assuming that the operator travels linearly in the y direction, targets are small with respect to the data sampling rate, and the terrain characteristics are slowly varying in space. Note that, although this approach is not appropriate without modification for targets that extend across multiple spatial locations, it is useful to help screen out clutter items that are extended in space. Given these assumptions, under H_0

$$s_{x,y}(k) = s_{x,y-1}(k) \tag{17}$$

or equivalently

$$\alpha_{x,y}(k) = \alpha_{x,y-1}(k). \tag{18}$$

Under the target hypothesis, H_1, $s_{x,y-1}(k)$ and $s_{x,y-1}(k)$ are not identical, such that

$$\alpha_{x,y}(k) \neq \alpha_{x,y-1}(k). \tag{19}$$

Based on these preliminaries, the nature of the received signal $r_{x,y}(k)$ at locations $\{x, y\}$ and $\{x, y-1\}$ may be determined. Under H_0

$$\begin{aligned} r_{x,y}(k) &= \exp(-\alpha_0(t_0 + k\Delta t)) + n_{x,y} \\ r_{x,y-1}(k) &= \exp(-\alpha_0(t_0 + k\Delta t)) + n_{x,y-1} \end{aligned} \tag{20}$$

where $n_{x,y}$ is assumed to be an independent, identically distributed Gaussian random process with zero mean and known variance σ^2 (estimated as described in the previous section). Under H_1

$$\begin{aligned} r_{x,y}(k) &= \exp(-\alpha_1(t_0 + k\Delta t)) + n_{x,y} \\ r_{x,y-1}(k) &= \exp(-\alpha_0(t_0 + k\Delta t)) + n_{x,y-1}. \end{aligned} \tag{21}$$

Again, the LRT for the received data vector $\mathbf{r}$ is defined as

$$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r}/H_1)}{p(\mathbf{r}/H_0)} \tag{22}$$

where $p(\mathbf{r}/H_1)$ and $p(\mathbf{r}/H_0)$ are the probability density functions of the received data vector under H_1 and H_0, respectively. As before, when $\Lambda(\mathbf{r}) > \delta$, H_1 is chosen, whereas when $\Lambda(\mathbf{r}) < \delta$, H_0 is chosen.

Since the distribution of the received signal $\mathbf{r}$ is dependent on the parameters α_1 and α_0 under H_1, and on α_0 under H_0, and these parameters are unknown, if we let $s_{\alpha_1}(k) = \exp(-\alpha_1(t_0 + k\Delta t))$ and $s_{\alpha_0}(k) = \exp(-\alpha_0(t_0 + k\Delta t))$, the LRT becomes as in (23), shown at the bottom of the page, where $p(\alpha_0)$ and $p(\alpha_1)$ are the probability density functions associated with the decay times α_1 and α_0. In this case, Monte Carlo integration [7] is utilized to calculate the decision statistic specified above.

This decision statistic can be computationally expensive since Monte Carlo integration of the integrals can, in some cases, take many iterations to converge. Therefore, the suboptimal generalized likelihood ratio test (GLRT) has also been formulated [6]. This formulation avoids the evaluation

$$\frac{\iint \left(\prod_k \exp\left(-\frac{(r_{x,y-1}(k) - s_{\alpha_0}(k))^2}{2\sigma^2} - \frac{(r_{x,y}(k) - s_{\alpha_1}(k))^2}{2\sigma^2} \right) \right) p(\alpha_0)p(\alpha_1)\, d\alpha_0\, d\alpha_1}{\int \left(\prod_k \exp\left(-\frac{(r_{x,y-1}(k) - s_{\alpha_0}(k))^2}{2\sigma^2} - \frac{(r_{x,y}(k) - s_{\alpha_0}(k))^2}{2\sigma^2} \right) \right) p(\alpha_0)\, d\alpha_0} \tag{23}$$

of the multidimensional integration and results in a simplified detector. In the GLRT formulation, the maximum likelihood estimate for each of the unknown parameters (in this case, α_1 and α_0) is used in the formulation of the likelihood ratio. For this problem, the GLRT has the form

$$\Lambda(\mathbf{r}) = \frac{\prod_k \exp\left(-\dfrac{(r_{x,y}(k) - \hat{s}_{x,y}(k,\,\alpha_1))^2}{2\sigma^2}\right)}{\prod_k \exp\left(-\dfrac{(r_{x,y}(k) - \hat{s}_{x,y}(k,\,\alpha_0))^2}{2\sigma^2}\right)}. \tag{24}$$

Here, $\hat{s}_{x,y}(k,\,\alpha_1)$ and $\hat{s}_{x,y}(k,\,\alpha_0)$ are the maximum likelihood estimators [6] for $s_{x,y}(k,\,\alpha_1)$ and $s_{x,y}(k,\,\alpha_0)$, and they can be estimated from the appropriate $r_{x,y}(k)$. In this case, the likelihood estimator for the mean is simply the sample mean [6] and this estimator in (24) is applied to the data obtained in the calibration area (see Section IV-A).

In order to determine whether the change detector formulation provides an improvement in performance over a pixel-by-pixel Bayesian approach, the performance using the likelihood ratio for the pixel-by-pixel case was also evaluated. In this case, Monte Carlo integration was used to evaluate

$$\Lambda(r) = \frac{\int \prod_k \exp\left(-\dfrac{(r_{x,y}(k) - s_{x,y}(k,\,\alpha_1))^2}{2\sigma^2}\right) p(\alpha_1)\,d\alpha_1}{\int \prod_k \exp\left(-\dfrac{(r_{x,y}(k) - s_{x,y}(k,\,\alpha_0))^2}{2\sigma^2}\right) p(\alpha_0)\,d\alpha_0}. \tag{25}$$

The GLRT formulation for the pixel-by-pixel analysis is equivalent to that shown in (24). In all cases, parameters of the *a priori* density functions, such as $p(\alpha_1)$, were estimated from the data measured in the calibration area (see Sections IV-A and III-A).

C. Wideband Frequency-Domain Sensors

In order to effect a likelihood ratio detector for each of the components of a frequency-domain sensor (in-phase and quadrature measured for each frequency), as well as for the entire data set, it is again necessary to obtain a valid estimate of the probability density functions of the data under H_1 and H_0. Based on the data that were gathered with the GEM-3 system during the DARPA experiment, we will restrict the development to the case in which data are measured at two frequencies, resulting in four components. If each of the recorded components is a deterministic, known quantity for all targets and types of clutter, and is subject to additive, zero-mean Gaussian noise, denoted n, under H_1, the ith received component can be modeled as

$$r_i = A_{1,i} + n. \tag{26}$$

In the case of the GEM-3, $i = 1$ denotes the in-phase component at 4050 Hz, $i = 2$ denotes the quadrature component at the same frequency, etc. Under H_0, the ith received component can be modeled as

$$r_i = A_{0,i} + n \tag{27}$$

where $A_{1,i} > A_{0,i}$. In this idealized case, the threshold detector is in fact optimal.

However, in the case in which the parameters $A_{1,i}$ and $A_{0,i}$ are not known exactly, or are known to vary spatially, the probability density distributions will be more complicated in general than described above. Since targets and clutter are rarely homogeneous, we expect that these assumptions are in general not valid (confirmed through data collected with the GEM-3 during the DARPA experiment). To mitigate this problem, two approaches have been taken. In the first, the density functions are estimated from the collected data, using both parametric [8] and nonparametric [9] approaches (similar to the approach used for the time-domain and integrated time-domain sensors). These density functions are then used to form the likelihood ratio test. A second approach, which does not rely so directly on the data, invokes the Gaussian assumption, with parameters assumed unknown. In this case, the likelihood ratio test becomes

$$\Lambda(r_i) = \frac{\int p(r_i/(H_1,\,A_{1,i}))p(A_{1,i})\,dA_{1,i}}{\int p(r_i/(H_0,\,A_{0,i}))p(A_{0,i})\,dA_{0,i}} \tag{28}$$

where $p(A_{1,i})$ and $p(A_{0,i})$ are the probability density functions for $A_{1,i}$ and $A_{0,i}$, respectively. Again, the parameters of these density functions were estimated from data taken in the calibration area for each site (see Section IV-A). More specifically

$$\Lambda(r_i) = \frac{\int \exp\left(-\dfrac{(r_i - A_{1,i})^2}{2\sigma^2}\right) p(A_{1,i})\,dA_{1,i}}{\int \exp\left(-\dfrac{(r_i - A_{0,i})^2}{2\sigma^2}\right) p(A_{0,i})\,dA_{0,i}}. \tag{29}$$

Monte Carlo integration is used to evaluate the integrals in the above equation.

The same two approaches were used to develop a detector for the full data set (in-phase and quadrature at 4050 and 12 270 Hz). In the first approach, the multidimensional density functions are estimated from the collected data and these density functions are then used to form the likelihood ratio test. In the second approach, the Gaussian assumption is used, but the parameters of the distribution are assumed to be unknown.

D. Threshold Detectors and Other Suboptimal Detectors

In order to address whether the complexity of the optimal receiver results in significant performance improvements, and to compare the results of the model-based processor to those of other processors, several other simplified detection strategies are investigated. In these approaches, the physics of the target and ground responses are utilized to realize *ad hoc* detection strategies [10]–[12].

When the clutter and target data follow Gaussian distributions with identical variances, but different means, the LRT reduces to the standard threshold test for unidimensional data, or to the matched filter in the multidimensional case. A threshold test is commonly applied to EMI data, and it thus was evaluated for the EM61 data, along with each of the individual

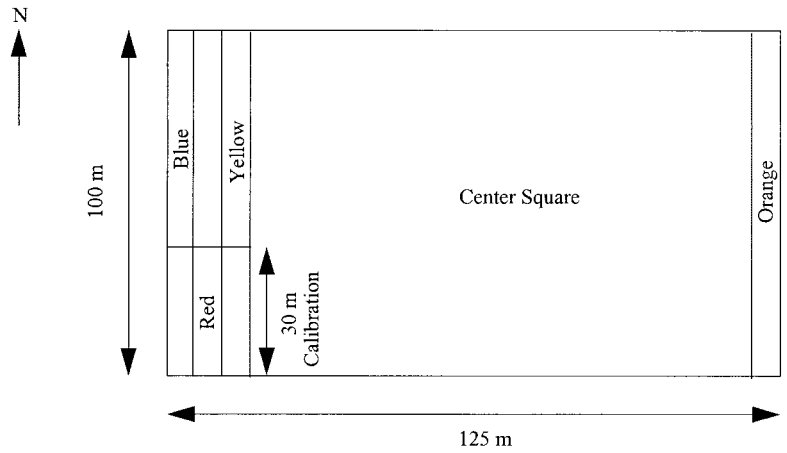


Fig. 1. Typical site layout from DARPA backgrounds clutter data collection experiment.

components of the GEM-3. The detection performance of a matched filter was also analyzed for the EM61-3D. For the matched filter, the received signal was correlated with the difference between a "known" signal (the "known" signal was extracted from the data of targets that were emplaced in the site) and the average clutter signal. Similar results to those described below were obtained for all "known" target signals.

Under the hypothesis that the ground is less conductive than potential targets, the rate of decay of the clutter signatures will be faster than that of the targets (this neglects anthropic clutter with significant metal content). Several detectors may be formulated to exploit this observation. One such detector is an energy detector, which was applied to the EM61-3D and the GEM-3 data. In addition, for the EM61-3D, a decay rate estimator was formulated. In the results presented here, a least-squares estimate of the rate of decay applied [7].

IV. Experimental Details

A. Data

The objective of the background clutter data collection experiment [1] was to collect data to aid in the understanding of the effects of clutter on system performance. During the course of the experiment, data were collected using four types of sensors: ground penetrating radar (GPR), EMI, magnetometer, and infrared (IR). Data were collected at four sites, distributed over two locations (Fort Carson, CO, and Fort A. P. Hill, VA). The locations represented three different soil types, along with substantial variability in man-made contamination (anthropic clutter). Three of the four sites had a small amount of anthropic clutter, while at one of the sites (Fort A. P. Hill) the amount of such was substantial. In addition to the indigenous anthropic clutter, a few synthetic calibration targets, unexploded ordnance (UXO) items, and land mines were emplaced on each site.

A typical layout for a test site is illustrated in Fig. 1. The site measured 125 $\times$ 100 m and consisted of three areas: a center square, side bars, and a calibration area. The center square, measuring 100 $\times$ 100 m, was intended primarily for clutter characterization, although 15 calibration targets were placed there to aid equipment calibration and location registration. Replicas of the calibration targets, along with synthetic UXO items and land mines were emplaced in the red, blue, yellow, and orange side bars on the site. A complete description of the targets and their deployment can be found in [1]. In general, the red lane consisted of copies of the registration targets, other targets for calibration, and system-stressing targets. The blue side bar contained UXO and demagnetized spheres, the yellow side bar contained mines, IR calibration sources, and dielectric targets. The orange side bar consisted of a combination of these and other targets. The total area of the blue yellow and orange side bars was 1300 m^2. The calibration area consisted of a 30 $\times$ 15 m portion of the red, blue, and yellow side bars. The locations and descriptions of the targets in the calibration area (14 of which could be seen by EMI systems) and the center square (ten of which could be seen by EMI systems) were released to the public, the remaining locations (blue, yellow, and orange side bars) were not disclosed. In the analyses that are described below, only data from the calibration areas and the center square are processed to evaluate performance of the various detectors.

B. Sensors

In this paper, we consider data from three classes of EMI sensors: sensors that integrate time-domain information to provide a single data point (standard metal detector), those that provide a sampled portion of the time-domain waveform, and those that operate at several discrete frequencies.

The first system, the Geonics EM61, consists of a single-channel pulsed induction system with a 0.5-m transmitter coil positioned approximately 0.3 m above the ground. Data are received in both the transmitter coil and a second receiver coil that is located 0.4 m above the transmitter. The system operates at a center frequency of 75 Hz; the received signal is integrated from 0.18 to 0.87 ms after each transmit pulse, and the result, a single data point, is stored for later processing. Data were collected along survey lines spaced 0.5 m in the easterly direction and 0.2 m on the average in the northerly direction.

The second sensor, the Geonics EM61-3D, is a prototype three-component time-domain induction system having a 1-

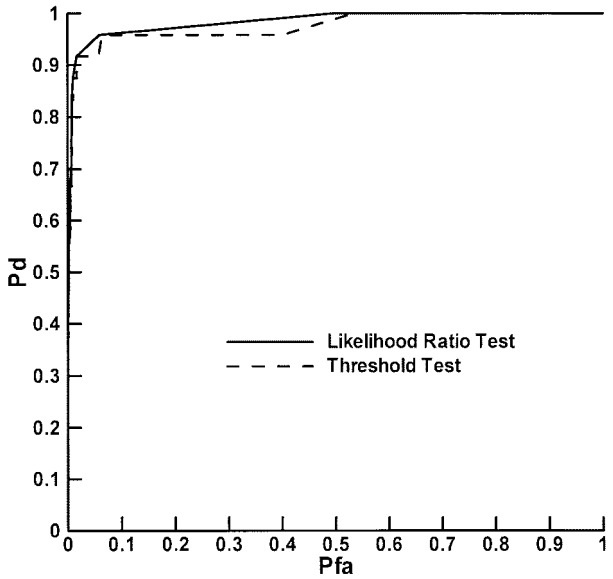


Fig. 2. Comparison of threshold (dashed line) and likelihood ratio test (solid line) performance—EM61, Turkey Creek site.

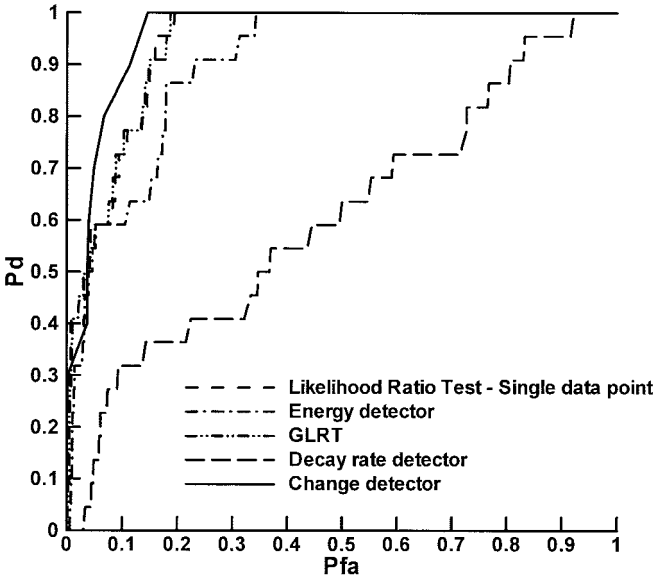


Fig. 3. Comparison of detector performance—EM61-3D, firing point 20 site. Solid line = Bayesian change detector performance, dashed line = Bayesian pixel-by-pixel detector, dotted line = GLRT performance, dashed-dotted line = energy detector, long-dashed line = decay rate detector.

m square transmitter coil and three orthogonal 0.5-m receiver coils positioned approximately 0.3 m above the ground. The system operates at a center frequency of 7.5 Hz. Sensor output is measured and recorded at 20 geometrically spaced time gates, spanning a time range from 32 to 320 μs. Data were collected along survey lines spaced 1 m apart in the easterly direction and at a rate of 3 samples per second, or approximately 0.2 m in the northerly direction.

The third system that was analyzed was Geophex's prototype GEM-3 sensor [13]. The GEM-3 is a monostatic, multifrequency EMI instrument. The GEM-3 can function as a time- or frequency-domain sensor; however, for the DARPA data collection project, the frequency-domain capabilities were used exclusively. Data acquisition parameters for all four test sites included a 10-Hz sampling rate and two transmit frequencies, 4050 and 12 270 Hz. The data used for detection in the subsequent analysis consist of in-phase and quadrature readings at each transmit frequency. Data were collected along survey lines spaced 0.5 m apart in the easterly direction and 0.17 m on the average in the northerly direction.

V. Results

In this section, the results of the different detection strategies are discussed. The raw data were analyzed by dividing each site into an appropriate grid based on the sampling rate of each of the sensors, and then averaging all signals recorded within each grid point. In the interest of space, data from only one site is presented in the figures for each sensor, although variability in performance across sites is discussed.

A. Integrated Time-Domain Sensor: EM61

As was expected based on the analysis presented in Section III-A, Fig. 2 indicates no difference between the performance achieved using the Bayesian processor and a standard thresholding technique. This results because the likelihood ratio is a monotonic function of the data (Section II-A) when the data are the integrated time-domain signals. Again, it should be emphasized that this is not the case when samples of the raw time-domain signal are available for processing.

B. Time-Domain Sensor: EM61-3D

The superiority of the single pixel Bayesian processor, as compared to standard thresholding and various *ad hoc* techniques, is demonstrated for the EM61-3D in Fig. 3. The performance of the GLRT processor is approximately equivalent to that of the single-pixel Bayesian formulation. It is interesting to note that the energy detector is superior to the *ad hoc* processors based on the decay rate and the matched filter (not shown since the performance was almost identical to that of the decay-rate processor); this is likely a result of the significant variability in target and clutter signatures, not taken into account by either of the *ad hoc* processors. In addition, the Bayesian change detector formulation provides improved performance over the pixel-by-pixel formulation. Absolute performance of the different approaches varies across sites, and it is poorest at Firing Point 20, Ft. A. P. Hill, VA, known to be the most cluttered site. The detection rate for the Bayesian change detector formulation is larger than that of more standard approaches by a factor of between 1.4 and 2 at a fixed false alarm rate of 5%, depending on the site.

C. Wideband Frequency-Domain Sensor: GEM-3

Fig. 4 illustrates the improvement in performance of the Bayesian approach relative to standard threshold techniques for the individual components measured by the GEM-3. Substantial improvements are observed for all components except the 12 270 Hz in-phase component, for which the performance of the two approaches is equivalent. In some cases, detection performance at a false alarm rate of 5% can be improved by

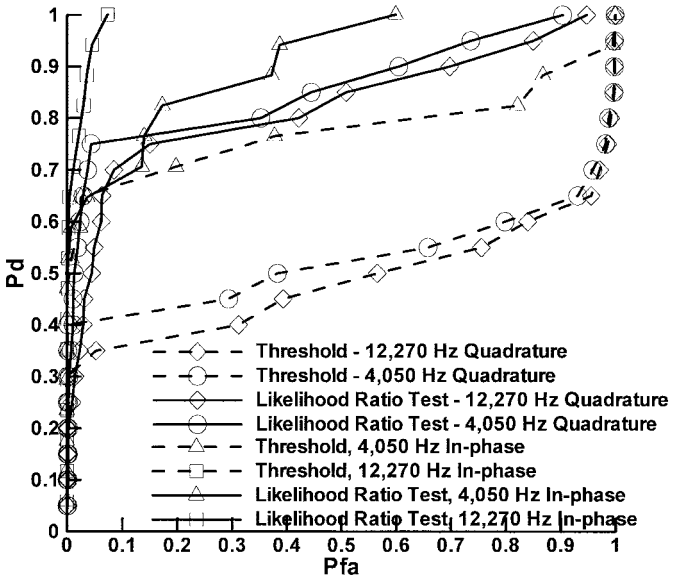


Fig. 4. Comparison of detection performance—single components of GEM-3, Firing Point 22 site. Dashed lines indicate detection performance for threshold tests, solid lines indicate Bayesian detector performance. Triangles = 4050 Hz in-phase data, circles = 4050 Hz quadrature data, squares = 12 270 Hz in-phase data, diamonds = 12 270 Hz quadrature data.

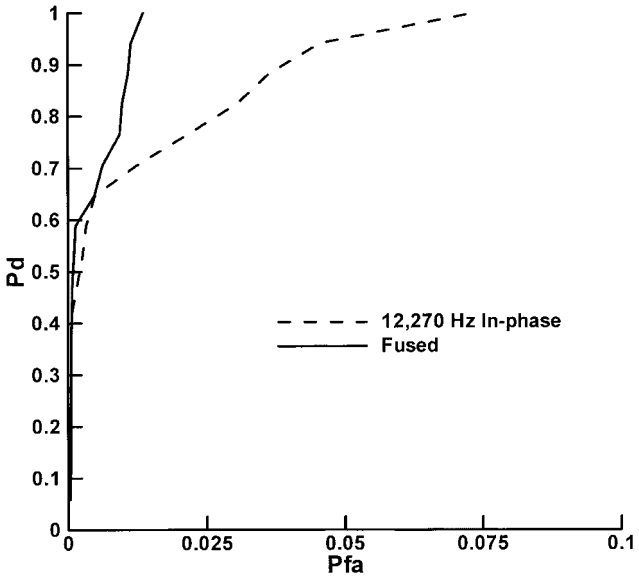


Fig. 5. Comparison of detector performance-fused (solid line) versus single component (dashed line, 12 270 Hz in-phase) of GEM-3, Firing Point 22 site.

a factor of two to three. Similar performance improvements were observed across all four sites, with some variability in absolute level of performance. In all cases, the 12 270 Hz data results in the best detection performance, regardless of the site. It is possible that the improved detection at 12 270 Hz (over that at 4050 Hz) is a result of differences in skin-depth effects between target and clutter that are only observable at higher frequencies. This result suggests that careful selection of the transmitting frequencies [14], based on the physics of the underlying process, may allow additional improvements in detection performance.

Fig. 5 illustrates the improvement in performance obtained when the entire data set is processed using the Bayesian approach, as compared to the best performance achievable with a single component (dashed line, 12 270 Hz, in-phase). The fused data set provides better detection performance across all sites than that achievable with a single component.

VI. Conclusions

We have presented results which indicate that a rigorous, Bayesian signal detection theoretic analysis can be used to improve detectability of a set of targets, including land mines and some small UXO objects using EMI sensors over standard thresholding techniques. A suite of EMI sensors was evaluated, including those that integrate time-domain information to provide a single data point (standard metal detector), those that provide a sampled portion of the time-domain waveform, and those that operate at several discrete frequencies.

Theoretical analyses of the statistical nature of the signal obtained from standard metal detectors indicated that this signal was in fact monotonic with the likelihood ratio. Thus, a threshold detector is the optimal processor for a sensor that integrates the time-domain waveform to generate a single data point. Using data obtained from the DARPA experiment, the performance of the likelihood ratio processor and a threshold detector were evaluated and shown to be equivalent. This indicated that the assumptions used to derive the LRT and its equivalence to a threshold test were valid, at least for the data in question.

When a sampled version of the induced time-domain response from the target is available for processing, further theoretical analysis indicated that an integration of the time-domain waveform is not the optimal form of processing. In this case, the optimal processor was implemented and its performance was compared not only to the performance of the threshold test, but to several *ad hoc* processors that have been suggested in the literature [11], [12]. The results of this analysis indicated that a Bayesian detector operating on the data exceeded that of either the threshold detector or the *ad hoc* strategies tested. In addition, a Bayesian approach developed using the notion that such sensors are effectively change detectors provided an additional improvement in detection performance.

Similar results were obtained for a Bayesian detector designed to process the measured frequency response from a target. The form of the detector did not reduce to a threshold test, and in all but one case, in which the performance of the two detectors was equivalent, the performance of the Bayesian detector far exceeded that of the threshold test. In addition, a detector designed to process all of the frequency data concurrently performed significantly better than the detectors designed to operate on the single-frequency data. The results from this analysis indicated that detection performance is a function of frequency and that careful selection of operating frequencies may result in further improvements in detection performance [14].

In the development of the likelihood ratios, probability density functions for the data, and for the parameters upon which the signals depended, were required. In general, these

density functions were derived from the data that was acquired by each sensor in a calibration area, in which the location of the targets was known and the background, or clutter, was typical to that contained in the remainder of the test site. To avoid "overfitting" the data, parameters of two general classes of density functions, a Gaussian and a uniform, were estimated and the performance achieved using the different density functions was compared to evaluate the robustness of the methodology. The performance of these systems was only slightly affected by the selection of the density function, which is similar to results obtained in other studies [14]. It was necessary, however, to estimate these density functions separately for each site, particularly for the clutter statistics, since these statistics varied substantially across sites. Using an estimate of the density functions based on one of the sites at a different site usually degraded the performance substantially.

Although target responses to EMI systems have been studied both experimentally and theoretically [2]–[5], [10], it is only recently that multichannel time- and frequency-domain sensors have become available so that the Bayesian approach to detector design can be exploited. The DARPA experiment enabled a more complete characterization of statistics, thus aiding in the design of such detectors. We have shown that the approach provides improved detection performance over standard detection technique. In addition, the results presented here indicate that detection performance is improved if the entire time-domain signal, or multifrequency data are utilized since this allows exploitation of the phenomenology embedded in the signatures. The theoretical development of the statistically based algorithm, along with the proof of concept application of the detector formulation to the DARPA data, combined to yield encouraging progress toward the problem of mine detection.

ACKNOWLEDGMENT

The authors would like to thank Dr. T. Altshuler, V. George, and Dr. R. Dugan for many helpful discussions pertaining to this work. They would also like to thank the four anonymous reviewers for their helpful comments.

REFERENCES

[1] V. George, T. Altshuler, A. Andrews, J. Nicoll, E. Cespedes, D. Butler, T. Broach, and R. Mehta, "Background data collection plan," DARPA/Defense Science Office, Dec. 1996.

[2] C. E. Baum, "Low-frequency near-field magnetic scattering from highly, but not perfectly, conducting bodies," Phillips Lab. Interaction Note 499, Nov. 1993.

[3] G. D. Sower and Steven P. Cave, "Detection and identification of mines from natural magnetic and electromagnetic resonances," in *Proc. SPIE,* Orlando, FL, Apr. 1995, vol. 2496.

[4] Y. Das, J. E. McFee, and R. H. Cherry, "Time-domain response of a sphere in the field of a coil: Theory and experiment," *IEEE Trans. Geosci. Remote Sensing,* vol. GE-22, pp. 360–367, July 1984.

[5] Y. Das and J. E. McFee, "A simple analysis of the electromagnetic response of buried conducting objects," *IEEE Trans. Geosci. Remote Sensing,* vol. 28, pp. 278–288, May 1990.

[6] H. L. Van Trees, *Detection, Estimation, and Modulation Theory.* New York: Wiley, 1968.

[7] J. O. Berger, *Statistical Decision Theory.* Berlin, Germany: Springer-Verlag, 1980

[8] B. W. Silverman, *Density Estimation for Statistics and Data Analysis.* New York: Chapman & Hall, 1986.

[9] G. Roussas, *Nonparametric Functional Estimation and Related Topics.* New York: Kluwer, 1990.

[10] Y. Das, J. E. McFee, J. Toew, and G. C. Stuart, "Analysis of an electromagnetic induction detector for real-time location of buried objects," *IEEE Trans. Geosci. Remote Sensing,* vol. 29, pp. 343–344, Mar. 1991.

[11] A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, Eds., "Detection technologies for mines and minelike targets," in *Proc. SPIE,* Orlando, FL, Apr. 17–21, 1995, vol. 2496.

[12] A. C. Dubey and R. L. Barnard, Eds., "Detection and remediation technologies for mines and minelike targets H," in *Proc. SPIE,* Orlando, FL, Apr. 21–24, 1997, vol. 3079.

[13] I. J. Won, D. A. Keiswetter, and D. R. Hanson, "GEM-3: A monostatic broadband electromagnetic induction sensor," *J. Environ. Eng. Geophys.,* vol. 2, pp. 53–64, Mar. 1997.

[14] N. Geng, P. Garber, L. Collins, L. Carin, D. Hansen, D. Keiswetter, and I. J. Won, "Wideband electromagnetic induction for metal-target identification: Theory, measurement and signal processing," Tech. Rep., Duke Univ., Durham, NC, Sept. 1997.

Leslie Collins (M'96) [redacted] She received the B.S.E.E. degree in 1985 from the University of Kentucky, Lexington, and the M.S.E.E. degree in 1986 and the Ph.D. degree in electrical engineering in 1995, both from The University of Michigan, Ann Arbor.

She was a Senior Engineer at the Westinghouse Research and Development Center, Pittsburgh, PA, from 1986 to 1990. In 1995, she became an Assistant Professor in the Electrical and Computer Engineering Department at Duke University, Durham, NC. Her current research interests include incorporating physics-based models into statistical signal processing algorithms, and she is pursuing applications in demining as well as enhancing speech understanding by hearing impaired individuals.

Dr. Collins is a member of the Tau Beta Pi, Eta Kappa Nu, and Sigma Xi honor societies.



Ping Gao (S'97) received the B.S. and M.E. degrees (honors) in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1993 and 1996, respectively, and the M.S.E.E. degree from Duke University, Durham, NC, in 1997. She is currently pursuing the Ph.D. degree in the Electrical and Computer Engineering Department at Duke University.

She has worked as a Research Assistant at Duke University since 1997. Her main research interests include signal detection and estimation theory, statistical signal modeling and processing, and applied electromagnetics.

Lawrence Carin (S'85–M'89–SM'96), for a photograph and biography, see p. 359 of the January 1999 issue of this TRANSACTIONS.

Ultra-Wide-Band Synthetic-Aperture Radar for Mine-Field Detection

Lawrence Carin[1], Norbert Geng[1], Mark McClure[1], Jeffrey Sichina[2], and Lam Nguyen[2]

[1]Department of Electrical and Computer Engineering
Duke University
Durham, NC 27708-0291
Tel: +1 (919) 660-5252
Fax: +1 (919) 660-5293
E-mail: lcarin@ee.duke.edu

[2]Jeffrey Sichina and Lam Nguyen
Army Research Laboratory
AMSRL-SE-RU
Adelphi, Maryland 20783

Keywords: Radar, radar detection, radar applications, synthetic aperture radar, ultra-wide-band radar, mine detection, polarimetric radar

1. Abstract

A full-wave model is developed for electromagnetic scattering from buried and surface land mines (both conducting and plastic), taking rigorous account of the lossy, dispersive, and potentially layered properties of soil. The (polarimetric) theoretical results are confirmed via synthetic-aperture radar (SAR) measurements, performed using the US Army Research Laboratory's BoomSAR, with which fully polarimetric ultra-wide-band (50-1200 MHz) SAR imagery is produced. The SAR system is used to acquire a large database of imagery, including a significant distribution of naturally occurring clutter. Several techniques are used for mine detection with such data, including several detectors that are based on target features gleaned from the modeling, as well as a matched-filter-like detector that directly incorporates the target signatures themselves. In addition, the theoretical model is used to predict wave phenomenology in various environments (beyond the limited range of parameters that can be examined experimentally). Since the efficacy of radar-based subsurface sensing depends strongly on the soil properties, we perform a parametric study of the dependence of such on the target RCS, and on possible land-mine resonances.

2. Introduction

Ground-penetrating radar (GPR) constitutes one of the oldest technologies for subsurface sensing [1-9]. Most such systems are placed in direct or near-direct contact with the earth surface [3, 4, 8, 9]. A significant drawback of this approach is the lack of "standoff," a particular problem for the detection of buried ordnance (e.g., mines [10]). Moreover, it is time consuming to use such hand-held systems for large-area interrogation over variable terrain and foliage. The US Army Research Laboratory (ARL) has therefore undertaken the development of a synthetic-aperture-radar (SAR) system, with which significant standoff can be achieved [11]. This system provides surveillance of large areas, the results from which can dictate the smaller regions over which conventional hand-held systems (e.g., electromagnetic induction [10], conventional GPR [1-9], and magnetometers [10]) should be deployed. Therefore, while ideally we would like to use the SAR system to detect and distinguish *each* mine, we are particularly interested in mine-*field* detection, over very large areas. Consequently, we need not detect each mine, but rather mine *clusters*. This simplified problem statement still constitutes a significant technological challenge, particularly in highly cluttered (natural and anthropic) environments.

To enhance discrimination between targets and clutter, the excitation should be as wide-band as possible. However, it is also essential to operate at frequencies for which adequate soil (and foliage) penetration can be effected, particularly for deeply buried targets, or regions in which there has been significant foliage growth (for humanitarian de-mining, one often must deal with mine fields that have been in the ground for decades [10]). The ARL ultra-wide-band (UWB) SAR therefore transmits and receives waveforms with usable bandwidth from 50-1200 MHz. Moreover, this system is fully polarimetric, providing access to HH, HV, VH, and VV imagery.

While radar-based subsurface sensing is an old technology, until recently, there has been very little modeling done to characterize expected system performance as a function of mine type and environment. In practice, there can be wide variability in the mines deployed and in environmental (background) conditions. Therefore, system performance has historically been anecdotal, with excellent performance under certain conditions and unsatisfactory performance elsewhere. Recently, however, we have developed rigorous electromagnetic-scattering models [12-16] for conducting and plastic mines buried in lossy, layered soil (including, potentially, a top layer of snow [15]). Our models are based on a full-wave formulation of Maxwell's equations, with solution via the Method of Moments (MoM). The principal challenge to making such an analysis tractable is computation of the half-space or layered-medium Green's function [17], which we have performed efficiently via the method of complex images [18]. Moreover, we have restricted ourselves to targets that can be simulated as a body of revolution (BOR) [11-16], simplifying computational complexity significantly. While not all mines can be modeled as a body of

 1045-9243/99/$10.00©1999 IEEE

revolution, the vast majority of land mines (both conducting and plastic) are accurately so modeled [10]. The details of these models are described elsewhere, while here we concentrate on their application to measured SAR imagery.

The phenomenological insight garnered from such models can guide system design, and automatic target detection and discrimination. In particular, we examine the potential of radar-based sensing as a function of mine type and environment, in the context of target RCS and possible target resonances. These results quantify the dramatic variability in radar performance that can be achieved, depending on the detailed mine and soil characteristics. We also document that system performance for plastic mines can be enhanced significantly by modifying the properties of the surrounding environment (e.g., enhancement of the dielectric mismatch between a plastic mine and soil, induced by increasing the soil water content). Other issues that we explore include system performance as a function of bandwidth, in which the benefits of enhanced bandwidth (resolution) are examined in the context of the attendant loss in soil penetration. Finally, we address the potential of polarimetric sensing of buried and surface land mines. It is demonstrated theoretically and experimentally that the special symmetry properties of land mines (and the lack of such for clutter) make polarimetric imaging attractive.

In addition to using wave modeling to elucidate phenomenology and guide system design, the target models play an important role in detection and identification algorithms. In particular, optimal detectors require an accurate model for the signature of the target in question [19]. For the case of buried mines, until recently, one was required to use *measured* signatures for such applications. However, empirical target models for buried mines are of limited value, due to the strong dependence of the target signature on the mine type, target depth, and soil properties. We have therefore utilized our simulated target models as integral parts of a detector, with results presented here for data measured by the BoomSAR at Yuma Proving Ground (Yuma, AZ, USA). Results are presented in the form of receiver-operating-characteristic (ROC) curves. We also address wavelet-based compression techniques, which reduce algorithm complexity.

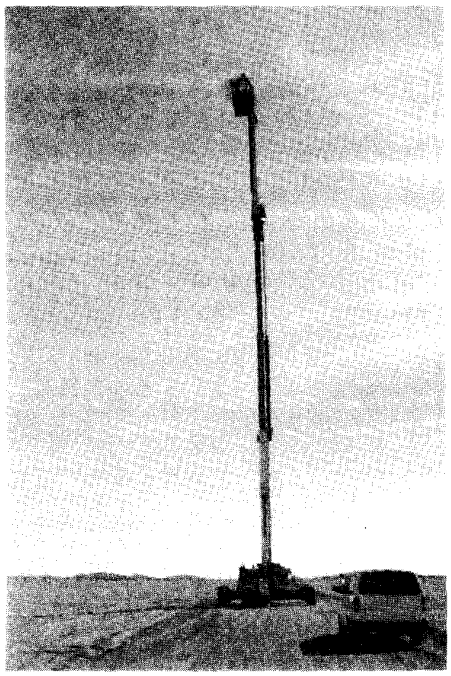


Figure 1. A photograph of the ARL BoomSAR.

The remainder of the text is organized as follows. In Section 3, we present a detailed description of the UWB BoomSAR, used to perform the measurements. Moreover, we give a brief summary of the modeling algorithm, the details of which are discussed elsewhere. Section 4 considers a comparison of measured and computed scattering data, for surface and buried land mines, in which we quantify agreement between theoretical and measured data. In addition to looking at a single polarization, we experimentally examine predicted polarimetric properties. Having calibrated the model, in Section 5 we discuss the design of a land-mine detector, which incorporates the computed target models (accounting for uncertainty in the target depth and soil properties), with results presented in the form of receiver-operating-characteristic curves. For comparison, the results of simpler detectors are also discussed. In Section 6, we use the modeling algorithm to perform parametric studies of the RCS and resonances of a buried plastic model land mine, to address issues of optimal bandwidth and soil modification for signature enhancement. Conclusions and future directions are discussed in Section 7.

3. Measurement system and theoretical model

3.1 BoomSAR system

The Army Research Laboratory (ARL) has been pursuing the use of UWB SARs for a number of years, to help understand foliage-penetrating and ground-penetrating radar, and to develop and qualify suitable wide-band models for predicting performance of proposed systems. These radars provide 1 GHz of instantaneous bandwidth and the full polarization matrix. The current implementation–hereafter referred to as the BoomSAR–is a mobile extension of the original design, allowing data collection over a wide range of clutter and target-in-clutter scenarios. The majority of the hardware is mounted to the basket of a telescoping boom lift, capable of moving at approximately 1 km/hr while the basket is elevated to 45 m (see Figure 1). For typical collection geometries, down-look angles to the target vary from 45° to approximately 10° (from grazing), depending on the range to the target and the height of the boom.

The radar consists of several major modules. The transmitter is based on a gallium-arsenide bulk-avalanche semiconductor switch (GaAs BASS). The transmitted waveform has a 150 ps rise time, and, approximately, a 2 ns fall time, with a peak power of approximately 2 MW. In the current configuration, the radar produces a burst of 128 pulses of one transmitted polarization, followed by a burst of 128 pulses of the opposite polarization, and both receiver channels are operated in parallel (for both polarizations). There are four antennas: two transmitting and two receiving, to provide the full polarization matrix (HH, HV, VH, VV) in a quasi-monostatic sense. The antennas are 200 W open-sided, resistively-terminated TEM horns, about 2 m in length, with a 0.3-m aperture, and are fed by a wide-bandwidth balun [20].

With the exception of the preamplifier/attenuator assembly, which sets the gain and noise figure for the system, the A/D subsystem acts as the wide-band receiver for the radar. The A/D subsystem consists of a pair of Tektronix/Analytek VX2005C, 2 Gsamples/sec A/D converters, and a stable reference clock. A

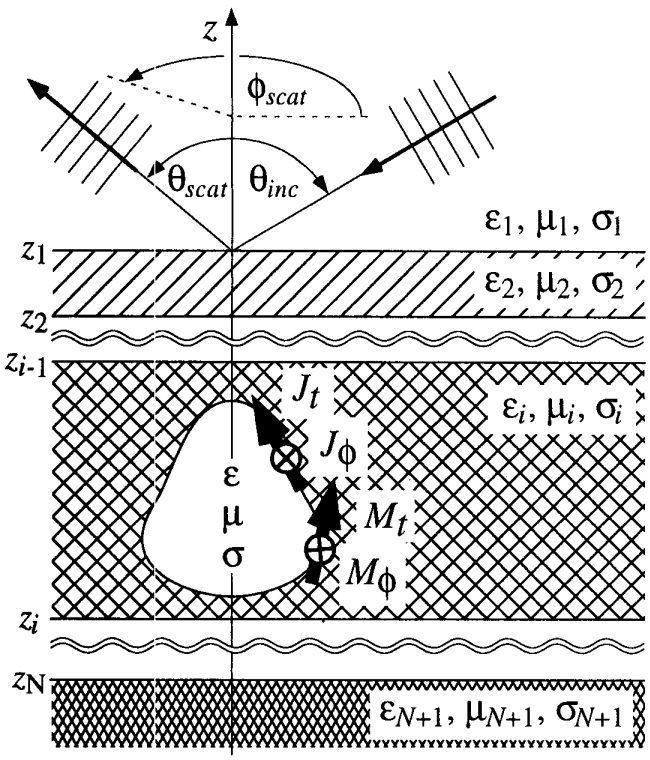


Figure 2. A schematic diagram of a body of revolution (BOR) buried in a layered medium. The BOR can be perfectly conducting or a dielectric.

unique feature of these A/D converters is that they provide, to 10 ps resolution, the time difference between the sample clock and trigger event. With the use of this data, sub-sample time interpolation allows the processor to generate an interleaved record at an equivalent 64 GHz rate, much as would be provided by a sampling oscilloscope.

As in a conventional SAR, the BoomSAR is moved along a straight path, and for each position the scattered response is measured. In its simplest embodiment, the bipolar image amplitude, $I_{m,n}$, at pixel position (m, n), is given by

$$I_{m,n} = \sum_{k=1}^{K} s_k\left(t = T_{k,m,n}\right), \tag{1}$$

where $s_k(t)$ represents the time-domain scattered field, measured at the kth sensor position, and $T_{k,m,n}$ is the round-trip time delay between sensor position k and the physical position in the image represented by pixel (m, n). The focused image has a resolution of 0.15 m in the range dimension, and 0.3 m in the cross-range dimension.

3.2 Numerical model

The phenomenology associated with UWB scattering from plastic and conducting land mines has been investigated through development of a full-wave scattering model for a body of revolution (BOR) buried in a lossy, dispersive, layered medium (see Figure 2). In addition to studying UWB scattering from such targets [15], the algorithm has been modified to model their natural resonances [16], as well. The body-of-revolution model permits a two-and-one-half-dimensional Method of Moments (MoM) analysis, in which the unknown currents exist along a two-dimensional generating arc [21-23], with the third dimension handled via a Fourier-series expansion in the azimuthal direction [21-23]. While the body-of-revolution assumption cedes some loss of generality, the vast majority of land mines are accurately so modeled [10]. Although a three-dimensional finite-difference time-domain (FDTD) algorithm [9] can be used to model electromagnetic scattering from such targets, such three-dimensional algorithms are computationally more expensive than the two-and-one-half-dimensional MoM analysis (although the FDTD algorithm is more general). Moreover, the FDTD and similar time-domain algorithms require an excitation field, and, therefore, are inappropriate for the computation of the (source-free) target resonances [24, 25]. While target resonances could be extracted from the computed time-domain scattered fields–via a parametric technique, such as Prony's method [26] or the matrix-pencil method [27]–the resonant-mode excitation is strongly dependent on the incident fields (e.g., on their bandwidth and incidence angle). By contrast, a frequency-domain algorithm, such as the MoM, is naturally applicable to source-free resonant-frequency computation [16, 23].

We note that the MoM analysis used here is not the first frequency-domain scheme to model buried targets as a body of revolution. Chang and Mei [28] used such a model several years ago, with solution via a hybrid (finite-element)-(integral-equation) formulation. Their FEM analysis is more general than the surface-MoM formulation used here, in that it can handle general body-of-revolution heterogeneity. However, this gain in generality has significant computational costs. The MoM analysis is sufficient for conducting (impenetrable) land mines, while, for plastic mines, the homogeneous or discretely inhomogeneous capabilities of the surface-MoM formulation are generally adequate.

For both the MoM and FEM formulations, the principal numerical hurdle involves efficient and accurate computation of the background-medium Green's function. While Chang and Mei

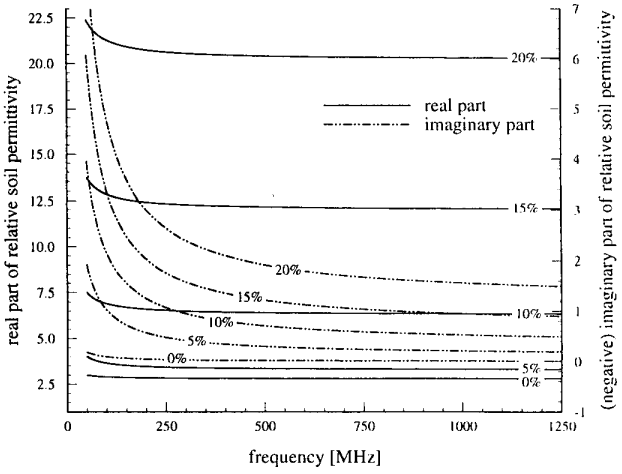


Figure 3. The measured real (ε') and imaginary (ε'') parts of the frequency-dependent complex dielectric constant of soil from Yuma Proving Ground, Yuma, Arizona, USA. The results are plotted as a function of water content, by percentage weight.

[28] considered a half-space background, as did our earlier work [11-14], our recent work considers the more-general case of a layered-medium environment (see Figure 2): representative of general layered soil, and/or soil with a covering snow layer. The efficient computation of the layered-medium Green's function is effected via the method of complex images [18], with careful attention for the surface-wave and leaky-wave pole positions encountered in the Sommerfeld integral [29]. The latter issue is of particular concern for the complex frequencies characteristic of the natural target resonances [16].

To our knowledge, this work represents the first rigorous analysis of scattering from and the resonances of *dielectric* bodies of revolution, buried in a general lossy, dispersive, layered medium. The details of the algorithm are described in two recent papers [15, 16]. Here, we concentrate on the application of the associated results to actual SAR imagery, as well as on the investigation of associated wave phenomenology.

4. Comparison of theory and measurement

We consider targets buried in the soil characterized in Figure 3, measured using soil samples taken from Yuma Proving Ground, Yuma, Arizona, USA. The soil characterization was performed with a Hewlett-Packard (HP) coaxial-probe test fixture, in conjunction with an HP network analyzer, with measurements performed as a function of water content (by weight). The initial set of scattering examples, below, are for soil with approximately 5% water, representing relatively low loss (these were the conditions encountered when we performed our SAR measurements). However, system performance depends strongly on the soil properties, with the dependence of such addressed in Section 6.

4.1 Target scattering

Equation (1) gives the connection between the measured (time-domain) fields and the subsequent SAR image. For a comparison of theoretical and measured SAR imagery, it is essential to use (in the theory) an incident waveform consistent with that of the actual UWB SAR system. A through measurement was therefore performed, from which an approximation to the incident pulse was determined (shown in Figure 4). It should be noted, however, that, in general, the antennas have frequency-dependent properties that have a strong angular dependence [20]. Moreover, each of the waveforms $s_k(t)$, measured along the SAR aperture, constitutes a different angle between the sensor and the target. Therefore, in actuality, there is *not* a single pulse shape incident on the target, from all aperture positions. Similar issues hold with regard to the polarization purity: an antenna may be vertically polarized at boresight, but the polarization can change markedly with angle. To account for this, only a limited number of aperture positions are used to image a given pixel in the SAR image (such that the properties of the incident field on that pixel are similar for all waveforms used in image formation). In this work, the image is segmented into a collection of contiguous tiles (sub-images), with 60° of aperture used for each tile, ±30° with respect to the normal between the SAR aperture and the tile center.

We compare the scattering model with measurements, by forming theoretical SAR images. In computing $s_k(t)$, we account for the k-dependent angle of incidence to a given pixel, assuming plane-wave excitation, pure polarization properties, and a fixed

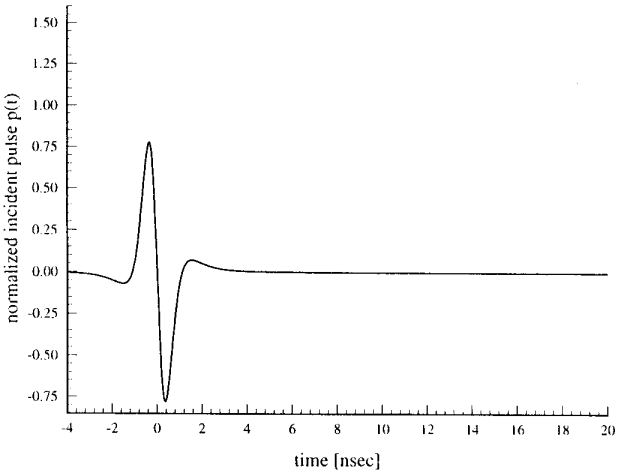


Figure 4a. The incident pulse waveform used in the scattering computations.

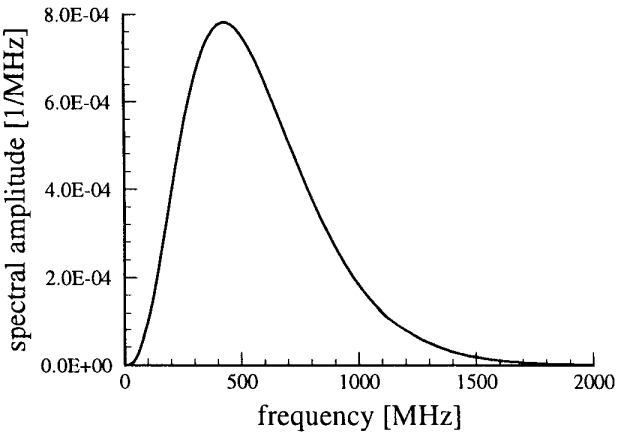


Figure 4b. The spectrum of the incident pulse shown in Figure 4a.

incident-pulse shape. The latter two conditions are dictated by our incomplete knowledge of the antenna properties over UWB frequencies, while the plane-wave-incidence approximation is appropriate for the SAR applications of interest. Finally, the waveforms, $s_k(t)$, used to form the theoretical image are weighted as a function of aperture position, to reflect the properties of the antenna pattern. A Gaussian weighting was used. In particular, for aperture position y, we use the weighting $w(y) = \exp\left[-(y - y_i)^2 / 2\sigma^2\right]$, where y_i represents the aperture point at which a normal from the pixel in question intersects the aperture, and $\sigma = 21.34\,\text{m}$ was found to yield theoretical cross-range performance similar to that found from the measured SAR imagery.

We initially consider the M20 anti-tank mine, a relatively large conducting target, shown in Figure 5. The measured and computed (bipolar) SAR imagery are compared in Figures 6a and 6b, for the M20 on the surface of the Yuma soil; in Figures 6c and 6d, for the mine flush-buried, just under the surface; and in Figures 6e and 6f, for the target buried six inches (15.24 cm). In all cases, VV polarization is considered (vertically incident polariza-

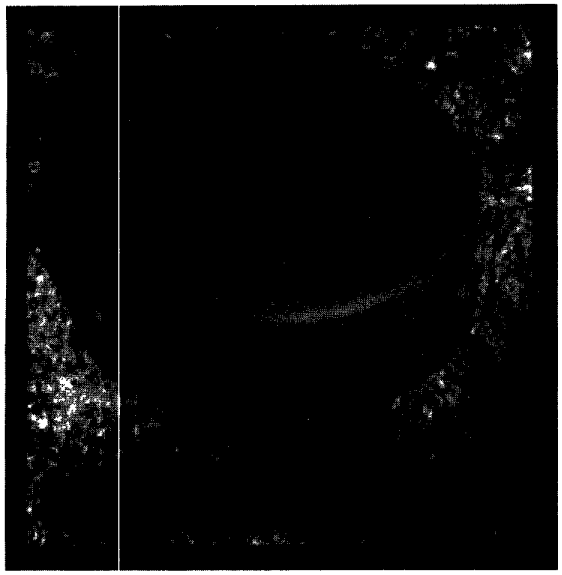


Figure 5a. An M20 anti-tank mine, used in the measurements and calculations.

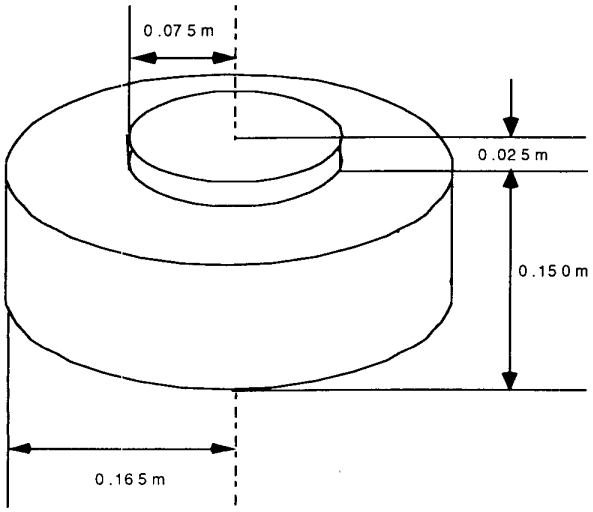


Figure 5b. The dimensions of the M20 anti-tank mine.

tion, vertically received polarization). From these results we see that the agreement between the measured and computed imagery is good (particularly so for the surface mine) but, for the case of the *buried* targets, the background clutter corrupts the measured signal. Nevertheless, these results bolster our confidence that scattering from the three-dimensional mine is modeled accurately. Note that the imagery is characterized by significant structure (particularly for the surface and flush-buried mines), a result of the UWB bandwidth and the cross-range resolution provided by the SAR aperture. Such structure is needed for accurate target discrimination, and will be diminished as the target size is reduced (for a fixed bandwidth), a particular issue for smaller anti-personnel mines. This issue is addressed further in Section 6.

4.2 Polarization properties

As mentioned in Section 3.2, the body-of-revolution model is appropriate for most land mines, and modeling the targets as such yields important insight into the underlying wave phenomenology. In particular, considering a vertically (V) or horizontally (H) polarized plane wave, incident along the direction β, a plane of symmetry P_s exists, as defined by β and the body-of-revolution axis of rotation. Since the incident V- or H-polarized plane wave satisfies symmetry properties about P_s, and the body-of-revolution target is also symmetric about P_s, the scattered fields will likewise satisfy symmetry properties. With respect to the plane P_s, the incident V-polarized fields are identical at (r,θ,ϕ) and $(r,\theta,-\phi)$, where the azimuthal angle, ϕ, is referenced to P_s. Consequently, for V-polarized excitation, the *scattered* fields at (r,θ,ϕ) are identical to those at $(r,\theta,-\phi)$, with respect to P_s. In the limit $\phi \to 0$ (i.e., in the plane P_s, representative of backscattering), this implies that the fields are vertically polarized. Therefore, if the incident plane wave is purely V polarized, the backscattered fields will also be V polarized. Using similar logic, if the incident fields are H polarized, the backscattered fields are also H polarized. Thus, for a buried body of revolution (such as in Figure 2), the *backscattered* fields have no cross-polarization component (the HV and VH fields are zero); note that the *bistatic* fields need not be purely V or H polarized.

This polarimetric property is potentially very useful, in that the body-of-revolution symmetry of most land mines is uncharacteristic of most natural clutter (rocks, roots, etc.). Moreover, note that each waveform $s_k(t)$, used to constitute the SAR image, is measured in backscatter, so the aforementioned polarimetric properties will be manifested in SAR imagery, as well. As an example of such, in Figure 7 we plot *measured* polarimetric SAR imagery for the M20 mine, buried 15.24 cm in Yuma soil. It is clear that the mines are virtually invisible in the cross-polarized (HV) image, while they appear prominently in the VV and HH images. Such issues have been addressed in detail in a recent letter [30].

5. Mine and mine-field detection

The results in Section 4 indicate that the models discussed in Section 3.2 accurately represent the phenomenology characteristic of UWB scattering from buried and surface land mines. The examples of that section were for conducting targets, as are the detection results of this section. The more challenging problem of plastic mines is addressed in Section 6. The detection results presented here do not exploit the potential of polarimetry, outlined in Section 4.2. This is due to the polarization properties of the horn antennas used for the measurements, for which pure V or H polarization can only be achieved over a limited range of angles with respect to boresight (the results in Figure 7 used a truncated range of angles when forming the image, thereby undermining cross-range resolution). Therefore, the results presented here are sub-optimal, in that we do not exploit the potential of polarimetry; this will require a better understanding of, and correction for, the antenna-polarization properties, constituting the subject of ongoing research.

5.1 Model-based template matching: image domain

If the clutter was white, Gaussian, and additive, and the target signature known exactly, the optimal detector would be a matched

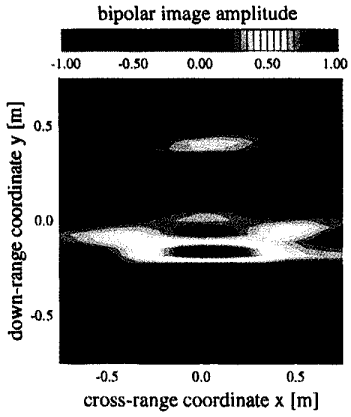


Figure 6a. The measured synthetic-aperture-radar image for the mine in Figure 5, on the surface of the soil in Figure 3.

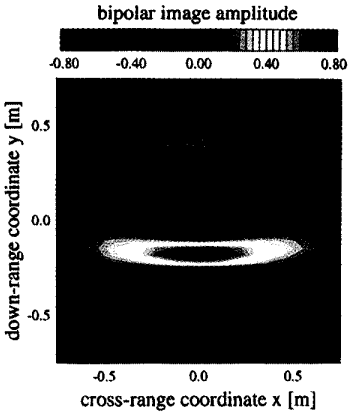


Figure 6b. The computed synthetic-aperture-radar image for the mine in Figure 5, on the surface of the soil in Figure 3. This should be compared with the measurements of Figure 6a.

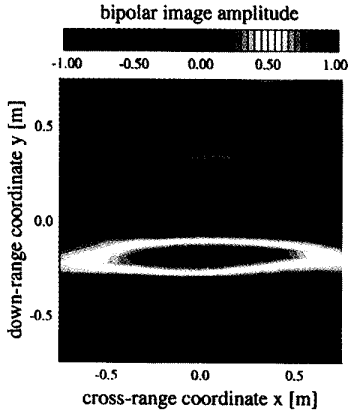


Figure 6c. The measured synthetic-aperture-radar image for the mine in Figure 5, flush-buried just under the surface of the soil in Figure 3.

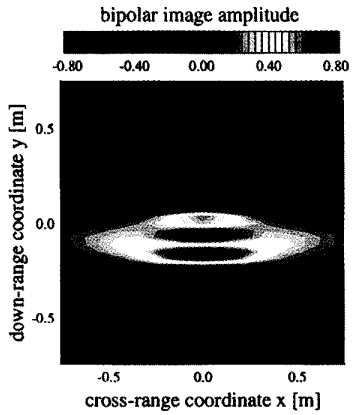


Figure 6d. The computed synthetic-aperture-radar image for the mine in Figure 5, flush-buried just under the surface of the soil in Figure 3. This should be compared with the measurements of Figure 6c.

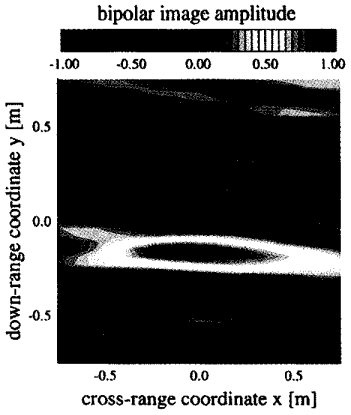


Figure 6e. The measured synthetic-aperture-radar image for the mine in Figure 5, buried 15.24 cm beneath the surface of the soil in Figure 3.

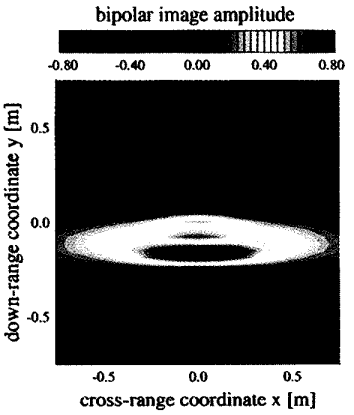


Figure 6f. The computed synthetic-aperture-radar image for the mine in Figure 5, buried 15.24 cm beneath the surface of the soil in Figure 3. This should be compared with the measurements of Figure 6e.

filter [19]. In reality, for the data we have measured, the clutter is generally non-stationary, characterized by discrete scatterings from localized naturally occurring obstacles (bushes, rocks, petrified wood, animal burrows, etc.). Nevertheless, the results of Section 4.1 indicate that the scattering model is in accord with the measurements, suggesting the development of a matched-filter-type detection scheme. It should be emphasized that this approach makes good use of our model, but is sub-optimal in that it does not exploit the statistics of the clutter (i.e., it improperly assumes that the clutter is white and Gaussian). The strong non-stationary quality of the clutter makes determination of the clutter statistics difficult, and therefore we do not attempt such here. However, we note that an alpha-stable statistical model is appropriate for random scattering from discrete entities [31], and the exploitation of such for the clutter is being explored.

Even if the aforementioned simplifications with regard to the clutter are utilized, the simple matched-filter detector has limitations. In particular, the target signature for the buried land mine is a function of the burial depth and the background-soil properties. Therefore, if $I_t(x,y;\gamma)$ represents the model target image, as a function of the parameter vector γ (symbolically representing the target depth and soil properties), we implement the generalized matched filter as [19]

$$l(x_0,y_0)=\int dx\int dy\int d\gamma\; M(x,y)I_t(x-x_0,y-y_0;\gamma)p_\gamma(\gamma), \quad (2)$$

where $M(x,y)$ represents the measured SAR imagery, (x_0,y_0) denotes the shifted position of the *computed* template $I_t(x,y;\gamma)$, and $p_\gamma(\gamma)$ represents the estimated density function on the parameters γ. The test statistic, $l(x_0,y_0)$, is compared to a threshold, t, and a target is declared at point (x_0,y_0) if $l(x_0,y_0)>t$, while clutter is declared (i.e., no target) if $l(x_0,y_0)<t$. By varying the threshold, t, the receiver operating characteristic (ROC) is generated [19], characterized as the probability of detection versus the probability of false alarm.

There are several issues with regard to Equation (2) that require further discussion. First, there are many examples for which the clutter response is much stronger in amplitude than that of the desired target. Consequently, even though the *shape* of $M(x,y)$ is not matched to $I_t(x,y;\gamma)$ in the vicinity of the clutter, the former's strong amplitude can yield a large l (and, hence, a false alarm). For the clutter considered here, this is a particular problem. Therefore, when computing Equation (2), we normalize $M(x,y)$ such that $\int dx\int dy M^2(x,y)=1$, where the integration is performed over the principal support of $I_t(x,y;\gamma)$. A second issue with regard to Equation (2) is the possibility of obtaining multiple "hits" on the same target (or clutter), in which several closely situated *discrete* positions (x_0,y_0) exceed the threshold. Therefore, after running the detector, we distinguish localized pixel clusters that exceed the threshold, and declare the total cluster as a detection (correct or otherwise). This cluster therefore defines a "halo," and if the halo intersects the position of a target (known from "ground truth"), then a correct detection is declared; otherwise, it is scored as a false alarm. The size of the clusters, and hence the halo, is defined to be the putative size of the buried target.

The parameter density function, $p_\gamma(\gamma)$, is used to account for environmental uncertainty. In particular, while the target types that may be encountered are often known a priori in de-mining, the exact soil properties and target depth are generally unknown. However, $p_\gamma(\gamma)$ accounts for the anticipated variability in these parameters. In the examples considered here, we sought, for example, M20 mines buried to a depth of 15.24 cm. For those results, $p_\gamma(\gamma)$ was designed such that the depth varied normally between 12.70 to 17.78 cm (to account for real-world variability in the burial depth). Moreover, we considered soil moisture content (Figure 3) varying normally between 0 to 6%. We also considered uniform distributions for these parameters, and little change was witnessed in the results.

The procedure in Equation (2) exploits the accurate models discussed in Section 3.2, while accounting for realistic environmental variability. However, the requisite computations are intensive for a large swath of SAR imagery (two-dimensional spatial integration, or summation, for each realization of γ, with the number of γ s required dictated by the density, $p_\gamma(\gamma)$). However, there are techniques available to make such two-dimensional spatial integrations tractable. In particular, we perform the two-dimensional integration, for *all* shifts (x_0,y_0) *simultaneously*, through exploitation of the properties of the fast Fourier transform (FFT). Both the templates $I_t(x,y;\gamma)$ and the measured imagery $M(x,y)$ are converted to the two-dimensional wavenumber domain, via a two-dimensional FFT in the coordinates x and y. The product of the two spectral-domain functions is performed, followed by the two-dimensional inverse FFT, from which all spatial shifts are acquired simultaneously. This procedure is performed separately for each γ required in Equation (2), as dictated by $p_\gamma(\gamma)$.

While the two-dimensional FFT results in a significant reduction in computational complexity, matters can be simplified further through use of a multi-resolution search strategy. In particular, we initially consider the templates $I_t(x,y;\gamma)$ and images $M(x,y)$ at low resolution, by filtering their spectral content, thereby allowing down-sampling. The test statistic of Equation (2) is computed at low resolution, and only those pixels for which a given threshold is exceeded are processed at the next, higher-resolution level (for which greater spectral content is used in both $I_t(x,y;\gamma)$ and $M(x,y)$). At each stage of the multi-resolution hierarchy, the regions in $M(x,y)$ at which mines are anticipated is pruned, and through this pruning process, the highest-resolution (full-band) computation of Equation (2) is only performed over a relatively limited domain of $M(x,y)$. We have performed this multi-resolution scheme via use of the Daubecheis perfectly reconstructing wavelet [32] (in two dimensions). For the imagery of interest here, we have found four levels useful (full band, 1/4, 1/16, and 1/64 full band). Therefore, through this procedure, Equation (2) is only computed for *all* shifts (x_0,y_0) at 1/64 the resolution, followed by pruning. This process is repeated at 1/16 full band, then at 1/4 full band, and, finally, at full band. Therefore, the full-band, most-computationally-expensive computations in Equation (2) are performed over a relatively limited portion of $M(x,y)$. In this process, one must carefully account for the shift-dependent properties of the wavelet transform [33, 33].

5.2 Template matching: down-range "cuts"

In the above section, we have discussed model-based template matching in the image domain. While this exploits all the

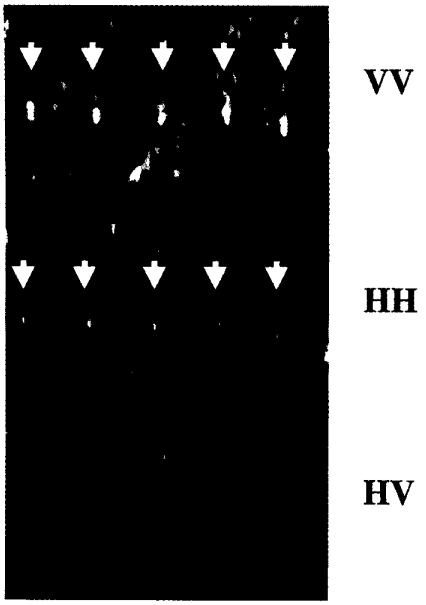


Figure 7. Polarimetric synthetic-aperture-radar imagery for an M20 mine, buried 15.24 cm in Yuma soil (Figure 3). The mine locations are indicated with the arrows.

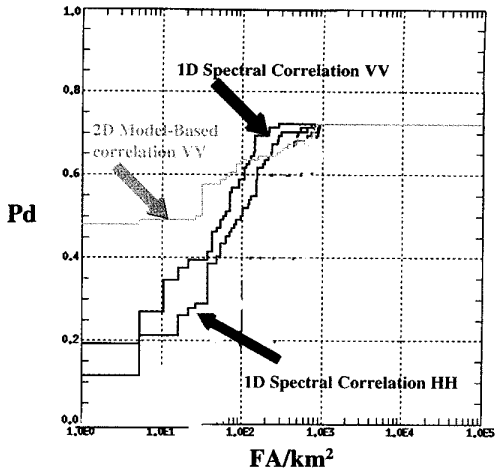
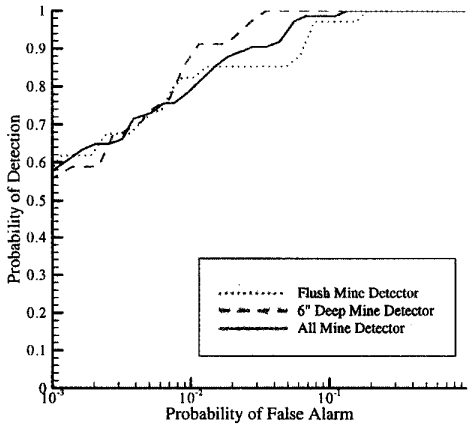


Figure 8. The receiver operating characteristic for the detectors discussed in Sections 5.1 and 5.2. Results for the detector in Section 5.1 are shown for VV polarization, while both VV and HH polarizations are considered for the detector in Section 5.2.

Figure 9. The receiver operating characteristic for a matched-filter detector, using a single template. Results are plotted for a detector utilizing (i) a template for a flush-buried target, and (ii) a template for the 15.24-cm-deep target, both in soil of 3% water content (Figure 3). Also shown are the results from Equation (2), using 27 templates, with parameters selected so as to span the range of target depths and moisture contents anticipated.



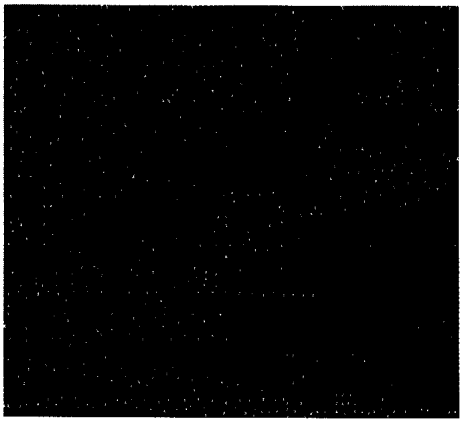


Figure 10a. Synthetic-aperture-radar imagery for a 0.46 km² section of the total 1.9 km² area imaged (Yuma Proving Ground, Yuma, Arizona). The locations of detections are marked with an "X". This shows the result when targets were declared at all points for which the SAR amplitude was consistent with that of a buried mine.

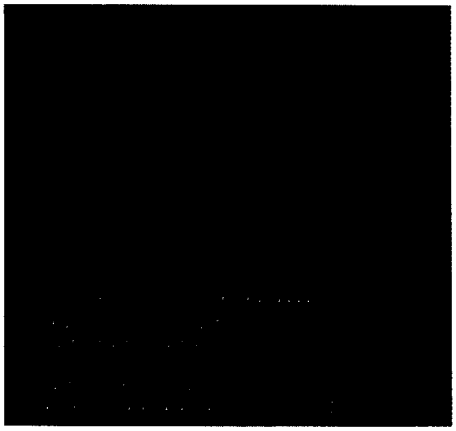


Figure 10b. The same type of SAR imagery as in Figure 10a, but showing the result of the detector in Equation (2), with the threshold t selected for a probability of detection of 0.7, and a false alarm rate of 1/km².

information available, it is computationally expensive. We have also considered a less-computationally-intensive approach, in which we perform correlations using one-dimensional down-range "cuts" in the image (these are formed by taking a cut through images like those in Figure 6, in the down-range direction, through the target center). For the UWB time-domain system employed here, such down-range cuts are analogous to time-domain waveforms. Similar down-range cuts are formed in regions of interest in the image $M(x,y)$, with which the one-dimensional templates are correlated (the correlations are performed in the spectral–Fourier–domain, where results are less sensitive to misalignment). While such an approach is sub-optimal–in that it does not exploit the cross-range properties of the target signature (as does the former algorithm)–it is computationally quite efficient. Moreover, this procedure exploits differences in the down-range properties of targets and natural clutter, which we have consistently found over much of the imagery considered. To further improve computational speed, instead of integrating over the parameter γ, as done in Equation (2) for the two-dimensional templates, here, we *train* the algorithm to determine that (single) down-range template that is most correlated with the data, and least correlated with the clutter. Therefore, this approach has two significant differences from Equation (2): the aforementioned reduction to a one-dimensional template, and the need for training data to determine the best down-range template for the target of interest. The advantage of this approach, after training, is that it is computationally quite efficient.

5.3 Detector results

We present here results for detection of M20 mines. There were several other mine types placed in the test zone (e.g., several Valmara 69 anti-personnel mines), which were deemed false alarms if detected by the algorithm: i.e., these additional mines were considered (anthropic) clutter to the detector designed for the M20.

The detector performance is indicated in Figure 8. The detector in Equation (2) used full resolution (not exploiting the wavelet, multi-resolution scheme), and results are shown for VV polarization. This detector used, in Equation (2), the templates for the M20 on the surface, flush-buried, and buried to a depth of 15.24 cm (constituting all M20 mines in the image). Results from the detector in Section 5.2 are shown as well, for both VV and HH polarizations. A total of 75 M20 mines were placed in the scene, and therefore the receiver-operating-characteristic (ROC) curves are somewhat discontinuous in form (due to the small number of targets available for generating the probability of detection). Nevertheless, these results demonstrated significant promise. For example, the detector discussed in Section 5.1, which exploits two-dimensional model-based templates (with no training), achieved a probability of detection of nearly 0.5 at one false alarm per square kilometer. While this is clearly inadequate for detection of *each* land mine, it is quite encouraging for mine-*field* detection. Additionally, the simpler detectors of Section 5.2 demonstrated respectable performance, as well. Note that these results had a peak detection probability of approximately 0.7; this is attributed to the fact that we initially scanned the imagery with a "pre-screener," essentially an energy detector, that determined regions of interest for the algorithms in Sections 5.1 and 5.2. In this example, the pre-screening process also eliminated several of the targets, and these targets cannot be recovered subsequently by the more-sophisticated algorithms discussed above. Therefore, a probability of detection of 0.7 in Figure 8 implied that the algorithms in Sections 5.1 and 5.2 have properly detected *all* targets remaining after pre-screening. The number of targets lost is dictated by the threshold used after pre-screening, and the results in Figure 8 are representative, but by no means optimal.

The detector in Section 5.1 appears to be particularly promising, so it is discussed in further detail. In Equation (2), integration of the parameter vector γ yields increased algorithmic complexity, relative to performing detection with a single template (a single γ). To assess the attendant performance gain, in Figure 9 we plot detector performance using a single template, for two template choices: a template for the 15.24-cm-deep target, and a template for a flush-buried target, both in soil of 3% water content (Figure 3). For comparison, we also show detector results computed by integrating (summing) over 27 templates, representative of the detector in Equation (2), with these templates spanning the range of parameters (target depth and moisture content) anticipated. In all results presented in Figure 9, scoring was assessed relative to *all* M20 mines in the image (i.e., all M20 mines were deemed targets, regardless of position). Moreover, in this example we have used Equation (2) over the entire image, without a pre-screener, and therefore a detection probability of 1.0 is achievable (c.f. Figure 8). Finally, while Figure 8 quantified false alarms through false-alarms per square kilometer, in Figure 9 we plot the associated *probability* of false alarm (for this data, a probability of false alarm of 10^{-3} corresponds to approximately one false alarm per square kilometer).

Considering the results in Figure 9, as anticipated, the detector in Equation (2) out-performed each of the detectors formed with a single template. However, the single-template detectors still performed reasonably well, with results that may be satisfactory for some applications, especially with regard to the significantly reduced algorithmic complexity relative to Equation (2). The reasonable performance of a single template can be tracked to the relatively slight variation in the templates for the surface, flush-buried, and 15.24-cm-deep mine (Figure 6). The utility of Equation (2) will become more apparent for terrain and mine properties of greater diversity.

Space limitations prohibit presentation of all the results from our investigations of detector performance, but other observations are summarized. With regard to the wavelet-based pruning strategy for the detector of Section 5.1, we did not notice any significant reduction in detector performance relative to a full-band implementation of Equation (2), despite the reduced algorithmic complexity. In these studies, we considered two-, three-, and four-wavelet stages (scales). The threshold settings at the various scales determines the rate of pruning (the extent of area, at a given scale, that is deemed clear of mines, and hence is not searched at higher wavelet resolutions). There is therefore a tradeoff between algorithmic complexity and accuracy (as for the pre-screener discussed above, once a target is lost at a low wavelet resolution, it cannot be recovered when one moves to higher resolution). We found the final results relatively insensitive to the pruning procedure, as long as it was not too dramatic at a given scale. While this was a positive feature of the multi-resolution scheme, the lack of shift invariance in the wavelet coefficients [33, 34] constitutes a serious problem (reducing the total gains in computational efficiency). Relative to performing Equation (2) at full resolution, the two-level wavelet algorithm required 90% the computational complexity; three levels, 80%; and four levels, 73%. While these constitute significant savings, especially given the robust final detector perform-

ance, a shift-invariant multi-resolution scheme would provide far more savings.

Using the detector in Section 5.1, designed for both flush-buried and 15.24-cm-deep mines, we plot, in Figure 10, the locations in the image that exceed a prescribed threshold, t (with t selected for a probability of detection of 0.7, and a false-alarm rate of 1/km^2). The detections are marked with an "X". We see, in Figure 10, that most of the mines in the minefield at the bottom left were properly identified. There are a few stray detections at which ground truth indicated no mines, and these detections, therefore, constituted false alarms. Nevertheless, these results indicate that the mine field has been located properly. Note that each mine need not be detected to constitute a mine field, and stray detections can be discarded as false alarms. As a comparison of the quality of the detector in Equation (2), relative to a simpler approach, in Figure 10 we also identify portions of the SAR image that have signature amplitudes consistent with that expected for the mine in question. From Figure 10, we see that such an energy-based detection criterion (which does not exploit the detailed target signatures, in Figure 6) results in an unacceptably high false-alarm rate.

6. Plastic mines

The detection results in Section 5 are encouraging. However, one could argue that the examples in that section were relegated to the relatively simple case of *conducting* anti-tank mines (M20s). In fact, plastic mines *were* placed in the region imaged, but the signatures of these targets were corrupted by the clutter response, and could not be convincingly seen in the data. The detection of such targets constitutes the principal challenge in mine detection, because such low-metal-content mines are virtually invisible to metal detectors (electromagnetic-induction [10] and magnetometer [10] sensors). In this section, we exploit the scattering model in Section 3.2 to address this problem in detail, with application to UWB radar.

While the properties of the plastic mine are not easily changed, once in the ground, it is possible to change the electrical properties of the soil. For example, the plastic mines considered in our SAR measurements had a dielectric constant near $\varepsilon_{rB} = 3$, while the soil (Figure 3), for the approximately 5% water content, also had a real part of the dielectric constant near 3. With such a minute electrical contrast between the target and background, their near-invisibility is expected in the SAR image. However, the results in Figure 3 indicate that the contrast can be enhanced by increasing the soil water content. The enhanced contrast will also be accompanied by increased attenuation in the soil. The tradeoffs between increased contrast and loss are explored here in detail.

6.1 Target scattering

We consider UWB plane-wave scattering from a "PMN2" plastic anti-personnel (AP) mine (shown in the inset in Figure 13), with the incident waveform described in Figure 4. The mine is characterized by a dielectric constant $\varepsilon_{rB} = 2.9 - j0.01$. The results in Figure 11 present the time-domain scattered fields from the PMN2 AP mine, buried to a depth of 2 cm in Yuma soil (Figure 3), considering 0%, 5%, 10%, 15%, and 20% water by weight, and an incidence angle of $\theta_i = 60°$ with respect to the normal. As a comparison, scattering results are also shown for the 15.24-cm-deep M20 mine, which constituted the weakest scatterer considered in Section 5. We see, in Figure 11, for VV and HH polarization, that the waveform scattered from the mine in 5% soil was dramatically smaller than that of the buried M20 mine, explaining why such mines were virtually impossible to see in the SAR imagery.

The results in Figure 11 demonstrate that the increased water content dramatically increases the signatures of plastic anti-personnel (AP) mines. This phenomenon is aided, in large measure, by the fact that AP mines are usually buried at quite shallow depths. With soil penetration constituting less of an issue for such situations, it suggests consideration of systems that operate over an even wider bandwidth than the ARL system. By so doing, we will generate scattered waveforms with more structure, especially for the small AP mines, aiding detection algorithms of the form of Equation (2). While the bandwidth of the ARL system is adequate for the relatively large anti-tank mine considered in Sections 4 and

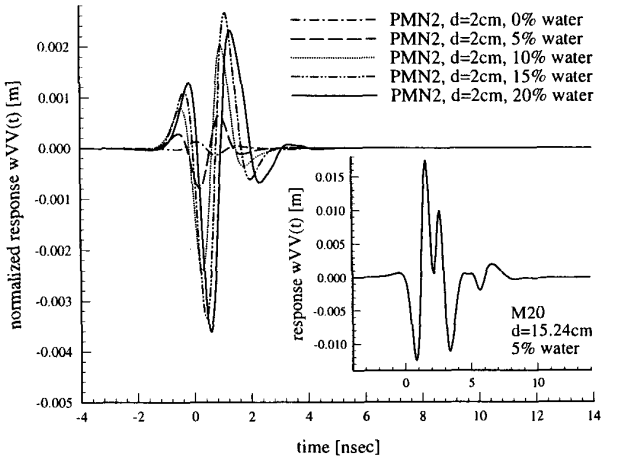


Figure 11a. Ultra-wide-band VV plane-wave scattering from a "PMN2" ($\varepsilon_{rB} = 2.9 - j0.01$) plastic mine (shown in the inset in Figure 13a), with the incident waveform described in Figure 4. The mine is buried at a depth of 2 cm in Yuma soil (Figure 3). Results are shown for 0%, 5%, 10%, 15%, and 20% water by weight, and an incidence angle of $\theta_i = 60°$. As a comparison, scattering results are also shown for the 15.24-cm-deep M20 mine (5% water content).

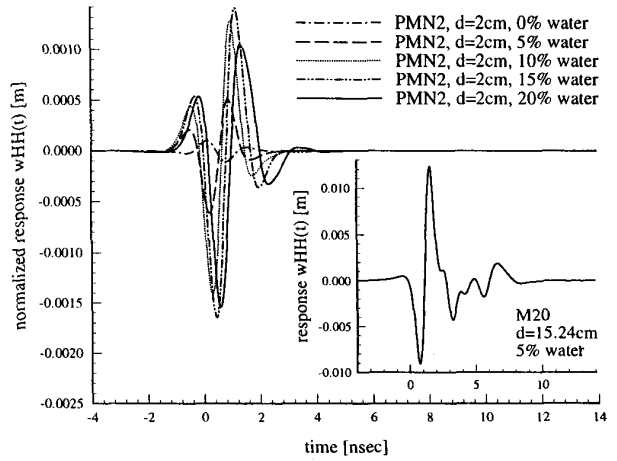


Figure 11b. The same results as in Figure 11a, but for HH polarization.

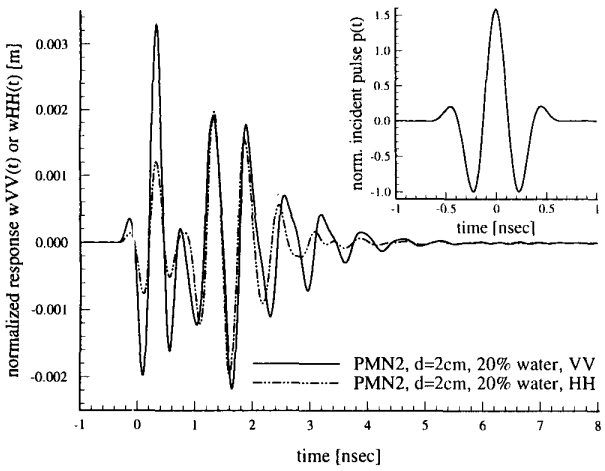


Figure 12. The scattered field for the PMN2 plastic mine (see the inset in Figure 13a), buried 2 cm beneath the Yuma soil, with 20% water content (Figure 3). The incident pulse (inset) has a center frequency of 2 GHz (instead of the 430 MHz center frequency considered in Figure 11).

5, further bandwidth is required to resolve the features on the PMN2, shown in Figure 11 (and similar AP mines). However, the development of such systems is expensive, and such development must be justified by *quantifiable* performance enhancement. The algorithm discussed in Section 3.2 affords us this opportunity.

In Figure 12, we plot scattering results for the PMN2 plastic mine, situated 2 cm beneath the Yuma soil, with 20% water content. However, in these examples, we consider an incident pulse with a center frequency of 2 GHz (instead of the 420 MHz center frequency considered in Figure 11). The corresponding incident pulse is shown inset in Figure 12. It is important to note that the incident pulses in Figures 11 and 12 are designed such that they contain the same energy (despite the differing bandwidths). For this case of relatively high water content–which the results in Figure 11 demonstrated was salutary for the detection of plastic AP mines–we see that the use of higher frequencies did not result in appreciable additional attenuation (for these realistic, shallow targets). Moreover, the increased system bandwidth resulted in dramatically increased structure in the scattered waveform, relative to scattering within the bandwidth of the ARL BoomSAR. Additionally, note that the scattered waveforms for VV and HH were markedly different, aiding polarimetric processing (recall, from Section 4.2, that the cross-polarized fields were zero).

6.2 Target resonances

It has been demonstrated previously [11] that the resonances of conducting mines, such as the M20, are characterized by very low-Q resonances that would be virtually impossible to extract from measured data. However, since energy can penetrate, reverberate inside, and subsequently radiate from a *plastic* mine, the potential of measuring resonances exists for such targets. In fact, measured results have been presented [4, 8] for which multiple resonances were extracted for a buried plastic mine. However, one would anticipate that the utility of resonance-based discrimination will depend strongly on the contrast between the plastic mine and the background soil. Our model gives us the requisite tools to address this issue, in detail.

In Figures 13a-13c, we plot the complex resonant frequency of the fundamental resonant mode for the PMN2 AP mine, for Yuma soil of 10%, 15%, and 20% water (Figure 3). The results are plotted as a function of target depth. In addition to plotting the depth-dependent resonant frequencies, we also plot the fundamental resonant frequency for the same target, embedded in a homogeneous medium characterized by the same electrical parameters as the soil. Several phenomena can be observed from these results. First, note that the resonant frequencies tend to spiral with changing depth. Similar behavior has been seen for a wire over a ground plane [35, 36], and for a wire buried in a half-space [14]. This phenomenon has been explained in detail in those papers, and analogous effects are occurring here (the reader is referred to [14, 35, 36] for a detailed explanation). For resonant frequency $\omega = \omega' + j\omega''$, the Q is defined as $\omega'/2\omega''$, and we note from Figure 13 that the Q increases with increasing water content. This is expected, for the increased target-soil contrast, which gave rise to the increased scattered signal in Figure 11, will also increase the contrast between the resonant structure and its background. As for the scattering results, this process is aided by the fact that the plastic AP mines are generally very close to the surface, obviating the enhanced soil attenuation with increased water content.

The benefits of an enhanced Q can be quantified through consideration of the Cramer-Rao lower bound [37]. In particular, resonance-based discrimination [4, 8, 24] is usually performed by subjecting the late-time scattered fields to parametric estimation for the resonant frequency, via Prony's method [26] or the matrix-pencil method [27] (or similar algorithms [38]). If the measurements are contaminated by noise and/or clutter, the scattered fields constitute a random process, with the results of any particular measurement constituting one realization of an ensemble. Therefore, the resonant frequencies extracted from such data must be parametrized as random variables. While algorithmic performance in the mean is important (and a correct mean constitutes an "unbiased" estimator [37]), of greater significance are the *variances* of the estimated resonant frequencies. The degree of variation (variance) in the estimated frequency will dictate whether the resonant frequencies can provide *unambiguous* discrimination. The Cramer-Rao lower bound (CRLB) quantifies the lowest variance that can be achieved by *any* estimator. If such variation is too large for unambiguous estimation, *no* algorithm will be effective for resonance-based discrimination. Moreover, the Cramer-Rao lower bound is a function of the signal-to-noise ratio (SNR). Therefore, by quantifying what variance is desired for unambiguous discrimination, the Cramer-Rao lower bound dictates what SNR is required to achieve such (assuming an algorithm that meets the Cramer-Rao lower bound exists, termed an "efficient" algorithm [37]). In Figure 14, we use the Cramer-Rao lower bound to quantify the normalized standard deviation for ω', considering the three examples in Figure 13. In these calculations, we utilize the resonant frequency for the target in a homogeneous medium (which, from Figure 13, is close to that of the resonant frequencies at the various depths considered). We see that, to achieve a standard deviation of less than 10% of the resonant frequency (real part), SNRs of 22 dB, 16 dB, 12 dB, and 9 dB are required for soils with, respectively, 5%, 10%, 15%, and 20% water. For 5% water, results are shown for the PMN2 in a homogeneous-soil background: the very slight target-soil contrast for this case made computation of the half-space results intractable (and, therefore, results are not shown in Figure 13 for 5% water). The Cramer-Rao lower-bound computations assume additive white Gaussian noise [40] (sampled above the Nyquist rate, with the data length at least three times the time constant), but similar results are derivable for arbitrary noise/clutter statistics. Figure 14 demonstrates that the properties

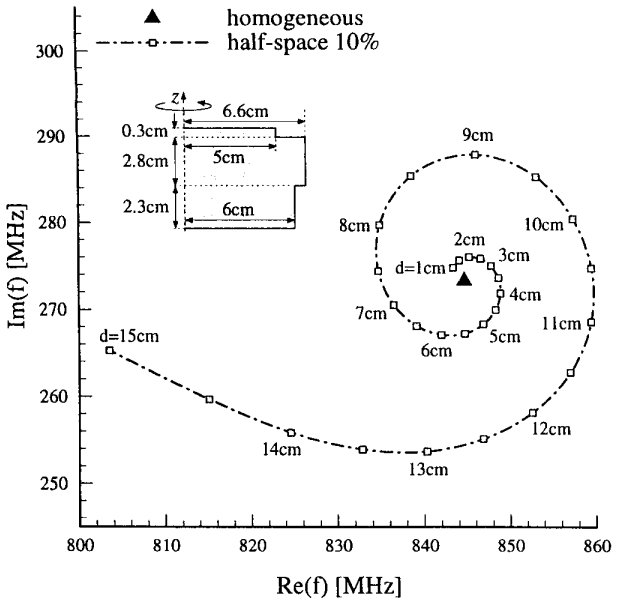


Figure 13a. The complex resonant frequency of the fundamental resonant mode for the PMN2 mine (inset), for Yuma soil of 10% water (Figure 3). The results are plotted as a function of target depth, *d*.

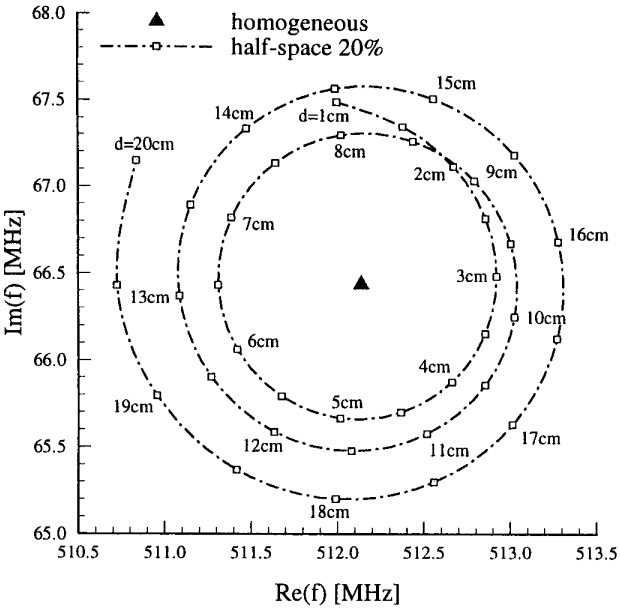


Figure 13c. The complex resonant frequency of the fundamental resonant mode for the PMN2 mine, for Yuma soil of 20% water (Figure 3). The results are plotted as a function of target depth, *d*.

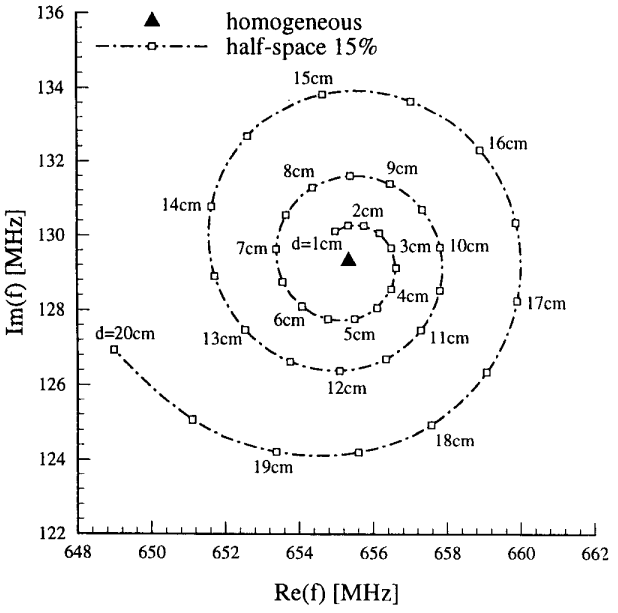


Figure 13b. The complex resonant frequency of the fundamental resonant mode for the PMN2 mine, for Yuma soil of 15% water (Figure 3). The results are plotted as a function of target depth, *d*.

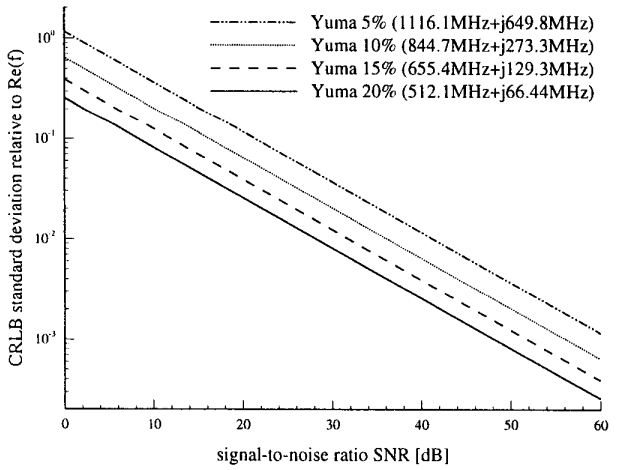


Figure 14. The Cramer-Rao lower bound for the normalized standard deviation for f', where f' is the real part of the complex resonant frequency. Results are shown for the PMN2 mine (inset in Figure 13a), embedded in a homogeneous medium characterized by the soil in Figure 3, considering water contents of 5%, 10%, 15%, and 20% by weight.

of the soil dramatically influence the efficacy of resonance-based discrimination, and that, for the examples presented here, simply increasing the soil water content can significantly enhance the utility of resonance-based discrimination for plastic mines.

7. Conclusions

Radar is one of the oldest tools for subsurface sensing. Remarkably, until very recently, there has been little rigorous modeling of the fields scattered from or the resonances supported by buried targets. This is, in large part, due to the complexity of such analyses, which have been aided greatly of late by revolutionary increases in computational power. Therefore, for much of its existence, the efficacy of radar-based subsurface sensing has been assessed based on a limited set of measured data. This has yielded results, in some cases, that suggested radar is a panacea and, at the other extreme, results for cases in which radar failed entirely. It is now understood that the utility of radar-based sensing for buried targets is largely dictated by the details of the problem (target and soil) in question, with the advances in modeling providing the key to appropriate radar deployment.

In de-mining applications, one generally must interrogate a large swath of terrain. Moreover, standoff is of obvious benefit, when attempting to locate these insidious targets. Therefore, the Army Research Laboratory (ARL) has focused on development of a SAR-based system that affords both of these desired attributes. Moreover, the ARL BoomSAR operates over ultra-wide-band frequencies (50-1200 MHz), simultaneously providing soil penetration and resolution. Synthetic-aperture radar results have been presented for data measured at Yuma Proving Ground, Yuma, Arizona. These data were used to confirm our MoM-based scattering models, as well as predictions with regard to the polarimetric properties of most land mines. In particular, we confirmed that mines that approximate a body of revolution (BOR) excite no cross-polarized backscattered fields.

After substantiating the model accuracy, we used the computed target signatures to effect a detector. This detector accounted for the real-world situation in which the expected mines are generally known, but the exact (spatially varying) soil properties and target depths are often not known exactly. Therefore, these latter properties have been treated within a statistical framework, in which the soil properties and target depth are characterized via statistical distributions. Results have been presented in the form of the receiver operating characteristic (ROC), and indicate that such an approach has promise. We accurately detected a large percentage of the mines, at a relatively low false-alarm rate. Moreover, the false alarms tended to be spread sporadically throughout the SAR image (based on a posteriori ground-truth knowledge), while the correct mine detections tended to be clustered, as expected of a mine field. Hence, SAR appears to be an attractive option for mine-*field* detection. Finally, a wavelet-based optimization scheme was investigated, to improve detector efficiency. Results from this preliminary multi-resolution study appear attractive, but significantly greater computational savings can be achieved by exploiting a multi-resolution scheme that is shift invariant (unlike the spatial wavelet transform). A detailed discussion of the ARL mine-detection paradigm can be found in [40].

While the SAR detector results are encouraging, there were some disconcerting issues learned as a consequence of the measurements. In particular, the buried *plastic* mines were virtually invisible to the radar and were *not* detected. Since many mines have only trace metal content, this failure is particularly troubling. The near-invisibility of the plastic mines to radar may be traced to the fact that the electrical contrast between the plastic target and soil was insignificant, for the examples considered here. However, while the properties of the plastic mine cannot be changed, the electrical characteristics of the soil can. In particular, we have examined the utility of applying water to a suspected mine field, thereby increasing the electrical contrast between the target and soil. This issue was addressed in the context of the scattered-field amplitude, as well as with regard to the natural resonances of such targets. For the soil and (representative) plastic target considered, it was demonstrated that this technique potentially has significant utility. This matter is aided by the fact that most anti-personnel mines are buried near the surface, mitigating the increased soil attenuation.

In addition to the weak target-background contrast, making difficult the detection of plastic mines, the SAR image is contaminated by artifacts due to imperfections in the imaging itself (e.g., angle-dependent effects in the transmitted waveform and in the polarization, which do not currently allow image calibration), as well as radio-frequency interference (RFI). It is felt that significant improvements can still be accrued in these areas (for example, through better understanding of the antennas), and that the subsequently improved SAR imagery will play an important role in detecting low-contrast (plastic) targets. Hence, while soil modification will play an important role in detecting such targets, it is felt that significant systems-level improvements are also important. While the current manifestation of the BoomSAR represents a notable engineering achievement, it is felt that significant improvements are still possible.

The M20 mines, buried to a depth of 15.24 cm, were the weakest scatterers detected from measured data in this study. To quantify the complexity of the problem for plastic mines, our theoretical results indicate that the plastic mine had a scattered amplitude roughly 4% that of the buried M20 mine (with the plastic mine placed in the same soil as considered in the measurements, *buried to a depth of only 2 cm*). This scattered amplitude is well beneath the noise (random system noise, imaging artifacts, and RFI) and clutter floor, and explains why the plastic mines were invisible in the imagery. After increasing the water content to 20%, the plastic mine had a scattered amplitude approximately 20% that of the 15.24-cm-deep M20. This is still a very stressing target, but with better imaging techniques and better removal of radio-frequency interference, it may yet be possible to detect plastic mines.

Finally, the matched-filter-like detector, described in Equation (2), which was so effective for the measured data considered here, implicitly assumes that the air-soil interface is flat. In humanitarian de-mining, for which the mines may have been in the ground for decades, this interface may be rough (or foliated by randomizing overgrowth). In this case, electromagnetic transmission through and scattering from the interface must be parametrized as a random process. While this has not been a concern for the data considered here, there are many applications for which such a formulation is essential. We have addressed this issue in detail [41, 42], and future SAR-based investigations will investigate the utility of such a framework, through consideration of measured data.

8. References

1. R. W. P. King and C. W. Harrison, Jr., "The Transmission of Electromagnetic Waves and Pulses into the Earth," *Journal of Applied Physics*, **39**, August 1968, pp. 4444-4452.

2. J. A. Fuller and J. R. Wait, "Electromagnetic Pulse Transmission in Homogeneous Dispersive Rock," *IEEE Transactions on Antennas and Propagation,* **AP-20**, July 1972, pp. 530-533.

3. D. L. Moffat and R. J. Puskar, "A Subsurface Electromagnetic Pulse Radar," *Geophysics*, **41**, June 1976, pp. 506-518.

4. L. Peters, Jr. and J. D. Young, "Applications of Subsurface Transient Radars," in E. K. Miller (ed.), *Time-Domain Measurements in Electromagnetics*, New York, Van Nostrand Reinhold, 1986.

5. G. S. Smith and W. R. Scott, Jr., "A Scale Model for Studying Ground Penetrating Radars," *IEEE Transactions on Geoscience and Remote Sensing*, **GRS-27**, July 1989, pp. 358-363.

6. C. Liu and C. Shen, "Numerical Simulation of Subsurface Radar for Detecting Buried Pipes," *IEEE Transactions on Gcoscience and Remote Sensing*, **GRS-29**, September 1991, pp. 795-798.

7. N. Osumi and K. Ueno, "Microwave Holographic Imaging of Underground Objects," *IEEE Transactions on Antennas and Propagation*, **AP-33**, February 1985, pp. 152-159.

8. L. Peters Jr., J. J. Daniels and J. D. Young, "Ground Penetrating Radar as an Environmental Sensing Tool," *Proceedings of the IEEE*, **82**, December 1994, pp. 1802-1822.

9. J. M. Bourgeois and G. S. Smith, "A Fully Three-Dimensional Simulation of Ground Penetrating Radar: FDTD Theory Compared with Experiment," *IEEE Transactions on Geoscience and Remote Sensing*, **GRS-34**, January 1996, pp. 36-28.

10. A. C. Dubey and R. L. Barnard (eds.), "Detection and Remediation Technologies for Mines and Minelike Targets," *SPIE Proceedings*, **3079**, Bellingham, Washington, SPIE, 1997

11. S. Vitebskiy, L. Carin, M. A. Ressler and F. H. Le, "Ultra-Wideband, Short-Pulse Ground-Penetrating Radar: Simulation and Measurement," *IEEE Transactions on Geoscience and Remote Sensing*, **GRS-35**, May 1997, pp. 762-772.

12. S. Vitebskiy and L. Carin, "Moment-Method Modeling of Short-Pulse Scattering from and the Resonances of a Wire Buried Inside a Lossy, Dispersive Half-Space," *IEEE Transactions on Antennas and Propagation*, **AP-43**, November 1995, pp. 1303-1312.

13. S. Vitebskiy, K. Sturgess and L. Carin, "Short-Pulse Plane-Wave Scattering from Buried Perfectly Conducting Bodies of Revolution," *IEEE Transactions on Antennas and Propagation*, **AP-44**, February 1996, pp. 143-151.

14. S. Vitebskiy and L. Carin, "Resonances of Perfectly Conducting Wires and Bodies of Revolution Buried in a Lossy Dispersive Half-Space," *IEEE Transactions on Antennas and Propagation*, **AP-44**, December 1996, pp. 1575-1583.

15. N. Geng and L. Carin, "Wideband Electromagnetic Scattering from a Dielectric BOR Buried in a Layered Lossy, Dispersive Medium," submitted to *IEEE Transactions on Antennas and Propagation.*

16. N. Geng, D. R. Jackson and L. Carin, "On the Resonances of a Dielectric BOR Buried in a Dispersive Layered Medium," submitted to *IEEE Transactions on Antennas and Propagation.*

17. K. A. Michalski and D. Zheng, "Electromagnetic Scattering and Radiation by Surfaces of Arbitrary Shape in Layered Media, Parts I and II," *IEEE Transactions on Antennas and Propagation*, **AP-38**, March 1990, pp. 335-352.

18. R. M. Shubair and Y. L. Chow, "A Simple and Accurate Complex Image Interpretation of Vertical Antennas Present in Contiguous Dielectric Half-Spaces," *IEEE Transactions on Antennas and Propagation*, **AP-41**, June 1993, pp. 806-812.

19. H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part 1*, New York, NY, Wiley, 1968, pp. 261-264.

20. M. A. Ressler and J. W. McCorkle, "Evolution of the Army Research Laboratory Ultra-Wideband Test Bed," in L. Carin and L. B. Felsen (eds.), *Ultra-Wideband Short-Pulse Electromagnetics 2*, New York, Plenum Press, 1995, pp. 109-123.

21. T. Wu and L. L. Tsai, "Scattering from Arbitrarily-Shaped Lossy Dielectric Bodies of Revolution," *Radio Science*, **12**, September-October 1977, pp. 709-718.

22. J. R. Mautz and R. F. Harrington, "Electromagnetic Scattering from a Homogeneous Material Body Of Revolution," *AEÜ*, **33**, February 1979, pp. 71-80.

23. A. W. Glisson, D. Kajfez, and J. James, "Evaluation of Modes in Dielectric Resonators using a Surface Integral Equation Formulation," *IEEE Transactions on Microwave Theory and Techniques*, **MTT-31**, December 1983, pp. 1023-1029.

24. C. E. Baum, "On the Singularity Expansion Method for the Solution of Electromagnetic Interaction Problems," *Air Force Weapons Laboratory Interaction Notes*, Note 88, 1971.

25. E. Heyman and L. B. Felsen, "A Wavefront Interpretation of the Singularity Expansion Method," *IEEE Transactions on Antennas and Propagation*, **AP-33**, 1985, pp. 706-718.

26. M. L. Van Blaricum and R. Mittra, "A Technique for Extracting the Poles and Residues of a System Directly from its Transient Response," *IEEE Transactions on Antennas and Propagation*, **AP-23**, November 1975, pp. 777-781.

27. Y. Hua and T. K. Sarkar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped/Undamped Sinusoids in Noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-38**, May 1990, pp. 814-824.

28. H. S. Chang and K. K. Mei, "Scattering of Electromagnetic Waves by Buried and Partly Buried Bodies of Revolution," *IEEE Transactions on Geoscience and Remote Sensing*, **GRS-23**, 1985, pp. 596-592.

29. T. Tamir and A. A. Oliner, "Guided Complex Waves. I. Fields at an Interface. II. Relation to Radiation Patterns," *Proceedings of the IEE (London)*, **110**, 1963, pp. 310-334.

30. L. Carin, R. Kapoor, C. E. Baum, "Polarimetric SAR Imaging of Buried Land Mines," *IEEE Transactions on Geoscience and Remote Sensing*, May 1998.

31. C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, New York, Wiley, 1995.

32. I. Daubechais, "Ten Lectures on Wavelets," Society for Industrial & Applied Mathematics, May 1992.

33. H. Sari-Sarraf and D. Brzakovic, "A Shift-Invariant Discrete Wavelet Transform," *IEEE Transactions on Signal Processing*, **ASSP-45**, October 1997, pp. 2621-2626.

34. I. Cohen, S. Raz and D. Malah, "Orthonormal Shift-Invariant Wavelet Packet Decomposition and Representation," *Signal Processing*, **57**, 1997, pp. 251-270.

35. L. S. Riggs and T. H. Shumpert, "Trajectories of Singularities of a Thin Wire Scatterer Parallel to Lossy Ground," *IEEE Transactions on Antennas and Propagation*, **AP-27**, 1979, pp. 864-868.

36. E. J. Rothwell and M. J. Cloud, "On the Natural Frequencies of an Annular Ring above a Conducting Half Space," *Journal of Electromagnetic Waves and Applications*, **10**, February 1996, pp. 155-179.

37. S. M. Kay, *Modern Spectral Estimation*, Englewood Cliffs, NJ, Prentice-Hall, 1987.

38. S. L. Marple, Jr., *Digital Spectral Analysis with Applications*, Englewood Cliffs, NJ, Prentice-Hall, 1987.

39. W. M. Steedly and R. L. Moses, "The Cramer-Rao Bound for Pole and Amplitude Coefficient Estimates of Damped Exponential Signals in Noise," *IEEE Transactions on Signal Processing*, **41**, March 1993, pp. 1305-1318.

40. L. H. Nguyen, J. Sichina, K. Kappra, D. C. Wong, and R. Kapoor, "Minefield Detection Algorithm Utilizing Data from an Ultrawideband Wide-Area Surveillance Radar," *Proceedings of 1998 SPIE Conference*, April 13-17, 1998, Orlando, FL.

41. T. Dogaru and L. Carin, "Time-Domain Sensing of Targets Buried Under a Rough Air-Ground Interface," *IEEE Transactions on Antennas and Propagation*, **AP-46**, March 1998.

42. T. Dogaru, L. Collins, and L. Carin, "Optimal Detection of a Deterministic Target Buried Under a Randomly Rough Interface," submitted to *IEEE Transactions on Antennas and Propagation*.

Introducing Feature Article Authors

Lawrence Carin [redacted] in Washington, DC. He earned the BS, MS, and PhD degrees in electrical engineering at the University of Maryland, College Park, in 1985, 1986, and 1989, respectively. In 1989, he joined the Electrical Engineering Department at Polytechnic University (Brooklyn) as an Assistant Professor, and became an Associate Professor there in 1994. In September, 1995, he joined the Electrical Engineering University at Duke University, where he is an Associate Professor. His current research interests include quasi-planar transmission lines, short-pulse scattering and propagation, and signal processing. Dr. Carin is the principal investigator on an Army Research Office-funded Multidisciplinary University Research Initiative (MURI) on de-mining. He is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.

Norbert Geng [redacted] in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Germany, in 1991 and 1996, respectively. From 1991 to 1996, he was with the Institute for Microwaves and Electronics at the University of Karlsruhe, working on full-wave propagation modeling for radio communication systems. For his PhD thesis on this subject, he received the Mannesmann Innovation Award, in 1997. In January, 1997, he joined the Department of Electrical and Computer Engineering at Duke University, Durham, North Carolina, as a visiting Post-Doc for 18 months. Since July, 1998, he has been back with the University of Karlsruhe. His current research interests include computational methods in electromagnetics, and wave-propagation modeling.

Mark R. McClure [redacted] in Mineola, New York. He received the BS degree in 1990 from Northeastern University, Boston, MA; the MS degree in 1993 from Polytechnic University, Brooklyn, NY; and the PhD degree in 1998 from Duke University, Durham, NC, all in electrical engineering.

From 1990 to 1995, he was with the Raytheon Company, Bedford, MA. While at Raytheon, he worked on the design and development of various radar systems. Since completing his PhD, he has been with the Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, MA. His current research interests include wave-based signal processing, superresolution imaging, and ATR. Dr. McClure is a member of Tau Beta Pi and Eta Kappa Nu.

Jeffrey Sichina received the BSEE degree in 1975 from the University of Pittsburgh. He subsequently undertook graduate study at the University of Maryland, College Park.

In 1974, he joined Harry Diamond Laboratories, Adelphi, MD, a component of the US Army Materiel Command. From 1974 through 1985, he was a staff engineer, working on a variety of electronic fuzing programs, including the Patriot Air Defense System and SGT York Gun System. His primary research interests included digital signal processing, FM/CW systems, and electronic counter-counter measures. In 1986, he was named Project Officer for the Chaparral Air Defense System Target Detecting Device Product Improvement Program. In 1988, he joined the Radar Branch, and was responsible for the development of a miniature moving-target-indicator radar for unmanned air vehicles. This joint undertaking, with MIT/Lincoln Laboratory, culminated in a series of airborne demonstrations of a 100-pound, highly capable radar system. In 1992, Mr. Sichina was named Chief of the Radar Branch (of the Army Research Laboratory, an organization composed of the former Army corporate labs), and was responsible for a variety of long-term applied research undertakings within the

Army. Among these was an effort aimed at determining the effectiveness of ultra-wide-band (UWB) radar for detection of concealed objects, such as buried mines. Mr. Sichina's current research interests include algorithm development and detection theory.

Mr. Sichina is a member of Eta Kappa Nu.

Lam H. Nguyen [redacted] in Vietnam. He received the BSEE, MSEE, and MSCS degrees from Virginia Polytechnic Institute, Blacksburg, VA; George Washington University, Washington, DC; and John Hopkins University, Baltimore, MD, respectively.

He started his career with General Electric Company from 1984 to 1985. He has been with the Army Research Lab and its predecessor organization, Harry Diamond Labs, from 1986 to the present. He has been primarily engaged in the research and development of foliage penetration and ground penetration via ultra-wide-band (UWB) radars. His primary research interests include digital signal processing and algorithm development.

Mr. Nguyen is a member of Eta Kappa Nu and Tau Beta Pi.



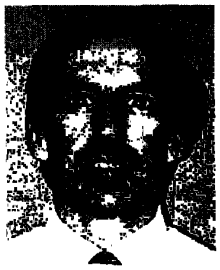
Lawrence Carin



Norbert Geng



Mark McClure



Jeffrey Sichina



Lam Nguyen

Corrections

Two errors appeared in the "Minutes of the AP-S AdCom Meeting" in the last issue (*IEEE Antennas and Propagation Magazine*, **40**, 6, December 1998). On page 54, in item 6.1.1, the 2000 International Conference on Phased Array Systems and Technology will be held at Dana Point, California. In item 6.3, on page 55, the European report was given by Gerry Crone.

Wide-Band Electromagnetic Scattering from a Dielectric BOR Buried in a Layered Lossy Dispersive Medium

Norbert Geng, *Member, IEEE,* and Lawrence Carin, *Senior Member, IEEE*

***Abstract*—A method of moments (MoM) analysis is developed for electromagnetic scattering from a dielectric body of revolution (BOR) embedded in a layered medium (the half-space problem constituting a special case). The layered-medium parameters can be lossy and dispersive, of interest for simulating soil. To make such an analysis tractable for the wide-band (short-pulse) applications of interest here, we have employed the method of complex images to evaluate the Sommerfeld integrals characteristic of the dyadic layered-medium Green's function. Example wide-band scattering results are presented, wherein fundamental wave phenomenology is elucidated. Of particular interest, we consider wide-band scattering from a model plastic mine, buried in soil, with the soil covered by a layer of snow.**

***Index Terms*— Method of moments, subserfice sensing, time-domain scattering.**

I. Introduction

ELECTROMAGNETIC scattering from bodies of revolution has been a subject of interest for more than three decades [1]–[9]. This research has been motivated by the fact that a body of revolution (BOR) can be used to simulate many man-made targets (e.g., missiles). Moreover, while a BOR is three-dimensional (3-D), the scattering of waves from such a target can be solved by using what is often a 2.5-dimensional analysis, wherein the target's azimuthal symmetry is exploited. In particular, azimuthal Fourier modes are used to represent the target's rotational variation, thereby reducing the problem to finding the unknown fields or currents along a two-dimensional (2-D) "generating arc" [1]–[9]. Such unknowns are usually solved for in the frequency domain using a method of moments (MoM) analysis [1]–[9].

Most previous such research has focused on targets in free space, representative of scattering from airborne BOR's. In such work one exploits the free-space Green's function, which can be computed in closed form and is itself rotationally symmetric; these two properties are of particular importance in simplifying the MoM numerical analysis. However, there has recently been significant interest in using electromagnetic scattering (radar) for the detection of underground targets [10]–[14]. Examples of such include buried mines, unexploded ordnance, and buried pipes [13], [14]. Each of these man-made targets has particular properties that can be exploited in the context of the scattering analysis. In the work presented here, we are interested in buried mines. It is well known that such targets often closely resemble BOR's [15]. This property has been exploited in a recent series of papers [16]–[19], in which MoM was used to analyze wide-band scattering from buried conducting BOR's, which simulated metallic mines.

The previous research on buried metal mines demonstrated the close agreement that can be obtained between theory and experiment [19], for real mines, and shed light on the underlying wave phenomenology. However, a principal challenge in mine detection involves plastic mines, with such targets generally composed of only trace metallic content (usually the tiny firing pin, often representing only a few grams of metal). In the work presented here, we therefore consider dielectric BOR's buried in a layered medium, with the lossy, dispersive layers representing the typical layered character of many soils (and/or a snow layer).

While scattering from a plastic mine can be solved via a 3-D finite-difference time-domain (FDTD) algorithm, such requires solution of the fields at all points in the computational domain [20]. Though ceding some generality, the MoM formulation only requires solution of the fields along the BOR generating arc, from which the scattered fields can be determined at any point. Moreover, a modified version of the BOR-MoM code can be used to study the properties of dielectric-mine resonances [21], something that is difficult to perform via FDTD, since mines are generally characterized by low-Q resonances. Finally, the exploitation of the mine's symmetry yields important phenomenological insight. In particular, for the case of buried-BOR targets, one can readily demonstrate [22] that there are no cross-polarized fields $(HV = VH = 0)$ in the case of backscattering, a property that can be exploited for target discrimination.

The analysis presented here is not the first to model buried dielectric discontinuities as BOR's. Chang and Mei [23] used a hybrid algorithm that combined finite-element method (FEM) modeling of the BOR with an integral equation, the latter accounting for the surrounding half-space medium. Although

Manuscript received January 13, 1998; revised Januarh 28, 1999.

The authors are with Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0018-926X(99)04766-3.

0018–926X/99$10.00 © 1999 IEEE

that work did not consider buried mines (it focused on geological discontinuities), the general framework could be applied to the case of dielectric mines. However, the FEM requires solution of the fields at all points within the BOR, while, again, the MoM analysis only requires solution for the fields on the surface of the generating arc. This of course results in some loss of generality, but yields significant computational savings, which are beneficial for the wide-band (short-pulse) applications of interest.

The BOR-MoM analysis developed here uses the mixed-potential integral-equation formulation of Michalski and Zheng [24] (what they termed "formulation C"). To effect such, one requires the space-domain dyadic Green's function for a general layered medium, here considering loss and dispersion as well. As is well known, closed-form representation of the dyadic Green's function components is only possible in the spectral domain [24], while the Sommerfeld integrals required for conversion to the space domain must be evaluated approximately [25]–[29]. Over the years various numerical and asymptotic techniques have been developed for evaluating these ubiquitous integrals [25]–[29]. Most electromagnetic mine-detection systems are of an ultra-wide-band nature [11], [14], [19] and, therefore, it is essential that the Green's function components (which are frequency dependent) be computed as efficiently as possible. To this end, we exploit here the method of complex images [27]–[29].

The method of complex images utilizes a parametric algorithm, for example Prony's method [30], [31] or the matrix-pencil method [32], [33], to express the components of the spectral-domain dyadic Green's function in terms of a sum of exponentials with complex arguments (a different such representation for each component of the dyadic). Utilizing the Weyl or Sommerfeld identity [27]–[29], each term in the summation can then be converted to the space domain (in closed form), interpreted physically as an image in complex space. Therefore, the problem of determining the space-domain Green's function for a layered medium is reduced from the laborious and computationally expensive task of evaluating Sommerfeld integrals numerically [34], to the relatively efficient task of parametric estimation [30]–[33]. As noted, the method of complex images has been used previously, primarily for conducting targets or radiators [16]–[19], [27]–[29], [35], while here we are interested in dielectric targets, thereby necessitating additional Green's function components [36].

In addition to describing the above formulation, with a focus on computational efficiency for wide-band applications, we examine the phenomenology associated with scattering from buried dielectric and conducting targets. Of particular interest are the effects of target and soil properties on the subsequent target signature, as a function of operating frequency. Moreover, the use of a layered-media Green's function, rather than the simplified half-space Green's function [16]–[19], allows us to examine several important and realistic scattering problems of interest. For example, one may be interested in detecting small mines buried in soil, with the soil covered by a snow layer. This is a problem well suited to the numerical paradigm utilized here, and the phenomenology of such is examined in detail, as a function of snow type (wetness) and snow depth.

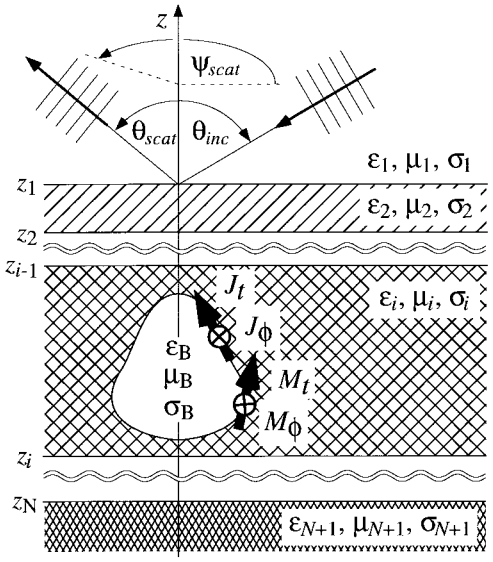


Fig. 1. Plane-wave scattering from body of revolution in planar multilayered environment.

The remainder of the text is organized as follows. The numerical formulation is discussed in Section II, wherein we detail the general MoM formulation and use of complex images. In Section III are presented several wide-band (short-pulse time-domain) numerical results, wherein basic phenomenology is examined for realistic model parameters (including lossy and dispersive soil and snow). Conclusions are addressed in Section IV.

II. Theory

We consider scattering from a (lossy) dielectric BOR situated in a planar layered environment, assuming the BOR's axis of rotation is normal to the interfaces (Fig. 1). We are interested in the scattered time domain fields for an incident ultrawide-band short-pulse plane wave. The problem is analyzed in the frequency domain using a MoM solution of the coupled surface integral equations, with the time-domain response calculated via a fast Fourier transform (FFT). For the wide-band problem of interest here, the frequency-domain analysis must be as efficient as possible. Therefore, the dyadic and scalar spatial domain Green's functions are calculated using the complex image technique introduced in [27]–[29], avoiding numerical integration of Sommerfeld type integrals.

A. Surface Integral Equation Formulation

By enforcing boundary conditions for the tangential electric and magnetic field components on the BOR surface, one obtains coupled integral equations for the electric and magnetic surface currents, $\boldsymbol{J}$ and $\boldsymbol{M}$, respectively. These integral equations can be put in a mixed potential form (MPIE, mixed

potential integral equation)

$$\begin{aligned}\boldsymbol{n}\times\boldsymbol{E}^{\text{inc}}(\boldsymbol{r}) = \boldsymbol{n}\times\Bigg[&j\omega\mu(\boldsymbol{r})\int_S \overleftrightarrow{\boldsymbol{K}}_A(\boldsymbol{r},\boldsymbol{r}')\cdot\boldsymbol{J}(\boldsymbol{r}')\,dS' \\ &-\frac{\nabla}{j\omega\varepsilon(\boldsymbol{r})}\int_S K_{\phi e}(\boldsymbol{r},\boldsymbol{r}')\nabla\cdot\boldsymbol{J}(\boldsymbol{r}')\,dS' \\ &+\nabla\times\int_S \overleftrightarrow{\boldsymbol{G}}_F(\boldsymbol{r},\boldsymbol{r}')\cdot\boldsymbol{M}(\boldsymbol{r}')\,dS'\Bigg]_{\boldsymbol{r}\in S^+} \\ &+\boldsymbol{n}\times\Bigg[j\omega\mu_B\int_S G_B(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{J}(\boldsymbol{r}')\,dS' \\ &-\frac{\nabla}{j\omega\varepsilon_B}\int_B G_B(\boldsymbol{r},\boldsymbol{r}')\nabla\cdot\boldsymbol{J}(\boldsymbol{r}')\,dS' \\ &+\nabla\times\int_S G_B(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{M}(\boldsymbol{r}')\,dS'\Bigg]_{\boldsymbol{r}\in S^-} \end{aligned} \tag{1a}$$

$$\begin{aligned}\boldsymbol{n}\times\boldsymbol{H}^{\text{inc}}(\boldsymbol{r}) = \boldsymbol{n}\times\Bigg[&j\omega\varepsilon(\boldsymbol{r})\int_S \overleftrightarrow{\boldsymbol{K}}_F(\boldsymbol{r},\boldsymbol{r}')\cdot\boldsymbol{M}(\boldsymbol{r}')\,dS' \\ &-\frac{\nabla}{j\omega\mu(\boldsymbol{r})}\int_S K_{\phi m}(\boldsymbol{r},\boldsymbol{r}')\nabla\cdot\boldsymbol{M}(\boldsymbol{r}')\,dS' \\ &-\nabla\times\int_S \overleftrightarrow{\boldsymbol{G}}_A(\boldsymbol{r},\boldsymbol{r}')\cdot\boldsymbol{J}(\boldsymbol{r}')\,dS'\Bigg]_{\boldsymbol{r}\in S^+} \\ &+\boldsymbol{n}\times\Bigg[j\omega\varepsilon_B\int_S G_B(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{M}(\boldsymbol{r}')\,dS' \\ &-\frac{\nabla}{j\omega\mu_B}\int_S G_B(\boldsymbol{r},\boldsymbol{r}')\nabla\cdot\boldsymbol{M}(\boldsymbol{r}')\,dS' \\ &-\nabla\times\int_S G_B(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{J}(\boldsymbol{r}')\,dS'\Bigg]_{\boldsymbol{r}\in S^-} \end{aligned} \tag{1b}$$

similar to those given in [17] and [24] for PEC scatterers. These integral equations have the same form as those typically used for scattering from dielectric scatterers in free space [3]–[9], for which well established numerical procedures are available [3]–[9]. The source and observation point are described by $\boldsymbol{r}'$ and $\boldsymbol{r}$, respectively, and $\boldsymbol{E}^{\text{inc}}(\boldsymbol{r})$ and $\boldsymbol{H}^{\text{inc}}(\boldsymbol{r})$ are the known incident electric and magnetic fields. The layered medium (complex) total permittivity and permeability are represented by $\varepsilon(\boldsymbol{r}) = \varepsilon'(\boldsymbol{r}) - j\sigma(\boldsymbol{r})/\omega$ and $\mu(\boldsymbol{r})$, respectively, where $\varepsilon'(\boldsymbol{r})$ is the real part of the permittivity and $\sigma(\boldsymbol{r})$ is the conductivity; $\varepsilon_B = \varepsilon'_B - j\sigma_B/\omega$ and μ_B represent these same properties for the homogeneous, lossy BOR target. Surface charge densities have been replaced by derivatives of the electric and magnetic surface current densities $\boldsymbol{J}(\boldsymbol{r}')$ and $\boldsymbol{M}(\boldsymbol{r}')$, respectively, using the continuity relation.

Explicit expressions for the spectral domain representation of the layered-medium dyadic kernel $\overleftrightarrow{\boldsymbol{K}}_A(\boldsymbol{r},\boldsymbol{r}')$, the dyadic Green's function $\overleftrightarrow{\boldsymbol{G}}_A(\boldsymbol{r},\boldsymbol{r}')$ (representing the magnetic vector potential $\boldsymbol{A}(\boldsymbol{r})$ produced by an infinitesimal electric dipole at $\boldsymbol{r}'$) and the electric scalar potential $K_{\phi e}(\boldsymbol{r},\boldsymbol{r}')$ of a point charge associated with a horizontal electric dipole have been given by Michalski and Zheng [24], where we use their "formulation C." Generalizing the formulation in [24] to (lossy) dielectric scatterers, we introduce the additional dyadics $\overleftrightarrow{\boldsymbol{K}}_F(\boldsymbol{r},\boldsymbol{r}')$ and $\overleftrightarrow{\boldsymbol{G}}_F(\boldsymbol{r},\boldsymbol{r}')$ as well as a scalar kernel function $K_{fm}(\boldsymbol{r},\boldsymbol{r}')$ associated with equivalent magnetic surface currents and surface charges, respectively; these Green's function components

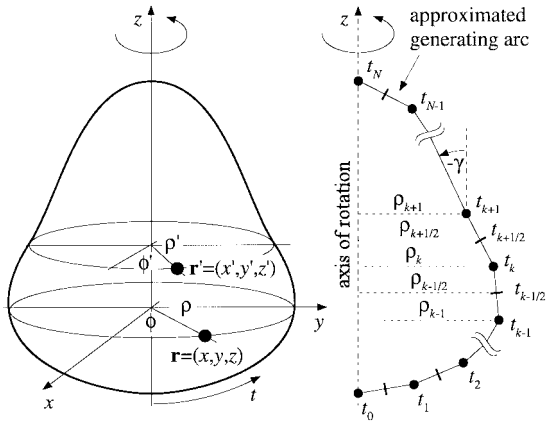


Fig. 2. Geometry for method of moments analysis of scattering from body of revolution.

can be determined from the aforementioned electric-current Green's function, via duality. For calculating the field inside the homogeneous BOR, produced by surface currents and surface charges on its surface, we utilize the homogeneous-medium Green's function $G_B(\boldsymbol{r},\boldsymbol{r}') = \exp(-jk_BR)/4\pi R$, where $R = |\boldsymbol{r}-\boldsymbol{r}'|$ represents the distance between source and observation point and k_B denotes the (complex) wave number inside the BOR.

According to "formulation C" in [24], the dyadic kernel functions $\overleftrightarrow{\boldsymbol{K}}_{A,F}(\boldsymbol{r},\boldsymbol{r}')$ and dyadic Green's functions $\overleftrightarrow{\boldsymbol{G}}_{A,F}(\boldsymbol{r},\boldsymbol{r}')$ are written as

$$\begin{aligned}\overleftrightarrow{\boldsymbol{K}}_{A,F}(\boldsymbol{r},\boldsymbol{r}') = &(\hat{\boldsymbol{x}}\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}}\hat{\boldsymbol{y}})K^{xx}_{A,F}+\hat{\boldsymbol{x}}\hat{\boldsymbol{z}}K^{xz}_{A,F}+\hat{\boldsymbol{z}}\hat{\boldsymbol{x}}K^{zx}_{A,F} \\ &+\hat{\boldsymbol{y}}\hat{\boldsymbol{z}}K^{yz}_{A,F}+\hat{\boldsymbol{z}}\hat{\boldsymbol{y}}K^{zy}_{A,F}+\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}K^{zz}_{A,F}\end{aligned} \tag{2a}$$

$$\begin{aligned}\overleftrightarrow{\boldsymbol{G}}_{A,F}(\boldsymbol{r},\boldsymbol{r}') = &(\hat{\boldsymbol{x}}\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}}\hat{\boldsymbol{y}})G^{xx}_{A,F}+\hat{\boldsymbol{z}}\hat{\boldsymbol{x}}G^{zx}_{A,F} \\ &+\hat{\boldsymbol{z}}\hat{\boldsymbol{y}}G^{zy}_{A,F}+\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}G^{zz}_{A,F}.\end{aligned} \tag{2b}$$

The coupled integral equations in (1) are valid for an arbitrarily shaped homogeneous scatterer, for which a 3-D MoM solution could be applied. However, to make an application to ultra-wide-band scattering tractable, we consider the special case of a BOR (Fig. 1), which is formed by rotating a generating arc about the z-axis. For numerical simulations, the generating arc is approximated by a sequence of linear segments (Fig. 2). Coordinates (ϕ,t) are introduced, where the angle ϕ is as used in cylindrical coordinates (ρ,ϕ,z) and t is a length variable along the arc. In addition, γ is defined as the angle between the vector $\hat{\boldsymbol{t}}$ along the arc and the z-direction (Fig. 2).

If the Green's function components in the spatial domain can be computed efficiently from their corresponding closed-form spectral domain representation [24], the coupled system of integral equations given by (1) can be solved using a modified free-space MoM code for scattering from a BOR [3]–[9].

B. Discrete Complex Image Technique

In [17], [18], [27]–[29], [35], and [36], it was shown that the discrete complex image technique is well suited for fast

numerical calculation of the spatial domain Green's functions, whereby laborious and computationally expensive integration of Sommerfeld type integrals is avoided. Here we summarize the basic steps.

Assuming the Green's function in the spectral domain is represented by $\tilde{f}(k_x, k_y, z, z')$, its spatial domain counterpart $f(x-x', y-y', z, z')$ is defined by the (shifted) Fourier transform

$$f(x-x', y-y', z, z') = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{f}(k_x, k_y, z, z') \cdot e^{-jk_x(x-x')-jk_y(y-y')}\, dk_x\, dk_y. \quad (3)$$

Introducing polar coordinates (k_ρ, α) in the spectral domain and polar coordinates (η, β) to represent $(x-x', y-y')$, and taking into account that $\tilde{f}(k_\rho, \alpha, z, z')$ is independent of the angle α, the spatial domain form (3) can be written alternatively as

$$f(\eta, \beta, z, z') = \frac{1}{2\pi} \int_0^\infty \tilde{f}(k_\rho, z, z') J_0(k_\rho \eta) k_\rho\, dk_\rho. \quad (4)$$

To avoid direct numerical integration of this highly oscillatory Sommerfeld type integral, we represent $\tilde{f}(k_\rho, z, z')$ in a form suitable for use of the Sommerfeld identity [27]–[29], [37]

$$\frac{e^{-jkR}}{4\pi R} = \frac{1}{2\pi} \int_0^\infty \frac{1}{j2k_z} e^{-jk_z u(z-z')} J_0(k_\rho \eta) k_\rho\, dk_\rho \quad (5a)$$

with

$$k_2 = k_\rho^2 + k_z^2 \quad \text{and}$$
$$R = \sqrt{\eta^2 + u(z-z')^2} = \sqrt{(x-x')^2 + (y-y')^2 + u(z-z')^2}. \quad (5b)$$

For the source and observation point both in the ith layer, this is accomplished by performing an exponential parameter fit to the spectral domain representation of the Green's function along a proper path in the k_{zi}-plane or k_ρ-plane [27]–[29], [36]

$$\tilde{f}(k_\rho, z, z') = \frac{e^{-jk_{zi}g(z,z')}}{j2k_{zi}} \tilde{h}(k_\rho) \approx \frac{e^{-jk_{zi}g(z,z')}}{j2k_{zi}} \cdot \left(\lim_{k_\rho \to \infty} \tilde{h}(k_\rho) + \sum_{\mu=1}^{M} a_\mu e^{-k_{zi}b_\mu} \right). \quad (6)$$

Here we use a least-square Prony's method [30], [31] to determine the complex coefficients a_μ and b_μ, although any such algorithm (e.g., the Matrix Pencil method [32], [33]) can be utilized. According to (6), we perform a parameter fit using a finite number of complex exponentials for the function $\tilde{h}(k_\rho) - \tilde{h}(\infty)$ (a different such for each Greens function component). Typically between $M = 6$ and $M = 15$ exponentials are sufficient to achieve an agreement between the approximate and the original spectral domain Green's function within 1% in the root mean square sense. For a detailed discussion, including the proper choice of the integration path, we refer the reader to [27]–[29] and [36].

C. MoM Solution

The coupled integral equations given by (1) are solved numerically via the well-known MoM solution for axially symmetric objects [1]–[9], [17], [18]. To take advantage of the rotational symmetry, the incident field, surface currents, and Green's functions are expanded into discrete Fourier series along the azimuthal (ϕ or ϕ') direction. The currents are represented using a subsectional pulse basis along the generating arc (coordinate t) and Fourier series expansion for the periodic azimuthal variation [5], [7]. The discretization of the generating arc and all geometrical parameters are defined in Fig. 2. Since the discrete complex image technique represents the layered-medium Green's functions as a sum of terms similar to the free-space Green's function, generalizing a free-space BOR MoM code [1]–[9] to the layered case considered here is straightforward, though somewhat laborious. Therefore, we do not repeat all the details.

Testing the integral equations corresponding to ϕ- and t-components of the incident field decouples the azimuthal Fourier modes $m = 0, \pm 1, \pm 2, \pm 3, \cdots$ [1]–[9] and leads to a set of simultaneous equations where each of them may be represented in matrix form as

$$\boldsymbol{Z}^m \cdot \boldsymbol{I}^m = \begin{bmatrix} [Z^{mlk}_{E\phi,J\phi}] & [Z^{mlk}_{E\phi,Jt}] & [Z^{mlk}_{E\phi,M\phi}] & [Z^{mlk}_{E\phi,Mt}] \\ [Z^{mlk}_{Et,J\phi}] & [Z^{mlk}_{Et,Jt}] & [Z^{mlk}_{Et,M\phi}] & [Z^{mlk}_{Et,Mt}] \\ [Z^{mlk}_{H\phi,J\phi}] & [Z^{mlk}_{H\phi,Jt}] & [Z^{mlk}_{H\phi,M\phi}] & [Z^{mlk}_{H\phi,Mt}] \\ [Z^{mlk}_{Ht,J\phi}] & [Z^{mlk}_{Ht,Jt}] & [Z^{mlk}_{Ht,M\phi}] & [Z^{mlk}_{Ht,Mt}] \end{bmatrix} \cdot \begin{bmatrix} [J^{mk}_\phi] \\ [J^{mk}_t] \\ [M^{mk}_\phi] \\ [M^{mk}_t] \end{bmatrix} = \begin{bmatrix} [V^{ml}_{E\phi}] \\ [V^{ml}_{Et}] \\ [V^{ml}_{H\phi}] \\ [V^{ml}_{Ht}] \end{bmatrix} = \boldsymbol{V}^m. \quad (7)$$

$\boldsymbol{Z}^m$ is the moment matrix, $\boldsymbol{I}^m$ is a column vector containing the unknown surface current coefficients, and $\boldsymbol{V}^m$ is the driving vector depending on the Fourier series coefficients (mth mode) of the incident electric and magnetic fields tangential to the BOR.

D. Time Domain Far-Field Plane-Wave Scattering

The MoM impedance matrices $\boldsymbol{Z}^m$ are independent of the incident field (driving vector) as well as the observation point. Therefore once the impedance matrices have been calculated, the scattering for an arbitrary number of incident fields (*near*- or *far-field* source) and/or observations points (in the *near* or *far field*) can be analyzed.

Here we are only interested in the scattered field at a point $(r, \theta_{\text{scat}}, \psi_{\text{scat}})$ in medium 1 far from the target, produced by an incident plane wave (propagating in medium 1, e.g., air) of arbitrary polarization and direction of propagation described by $(\theta_{\text{inc}}, \psi_{\text{inc}})$, where for the case of a BOR we set $\psi_{\text{inc}} = 0$ without any loss of generality, due to the rotational symmetry (Fig. 1). This is best described using the polarimetric scattering matrix $\boldsymbol{S}(\omega, \theta_{\text{inc}}, \theta_{\text{scat}}, \psi_{\text{scat}})$ (in BSA, backscatter alignment convention) defined by

$$\begin{bmatrix} E^{\text{scat}}_\theta \\ E^{\text{scat}}_\psi \end{bmatrix} = \frac{e^{jk_1 r}}{r} \begin{bmatrix} S_{\theta\theta} & S_{\theta\psi} \\ S_{\psi\theta} & S_{\psi\psi} \end{bmatrix} \cdot \begin{bmatrix} E^{\text{inc}}_\theta \\ E^{\text{inc}}_\psi \end{bmatrix} \quad (8)$$

from which the polarimetric radar cross section (RCS)

$$\mathrm{RCS}_{pq} = \lim_{r\to\infty} 4\pi r^2 \left| \frac{E_p^{\mathrm{scat}}}{E_q^{\mathrm{inc}}} \right|^2 = 4\pi |S_{pq}|^2 \quad \text{with} \quad p,q = \theta,\psi \tag{9}$$

can be easily determined. Similar to the scattering matrix $\boldsymbol{S}$, we define a normalized time domain response matrix (independent of the distance r and, therefore, only characterizing the target itself)

$$\begin{bmatrix} w_{\theta\theta}(t) & w_{\theta\psi}(t) \\ w_{\psi\theta} & w_{\psi\psi}(t) \end{bmatrix} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \begin{bmatrix} S_{\theta\theta}(\omega)\cdot P_\theta(\omega) & S_{\theta\psi}(\omega)\cdot P_\psi(\omega) \\ S_{\psi\theta}(\omega)\cdot P_\theta(\omega) & S_{\psi\psi}(\omega)\cdot P_\psi(\omega) \end{bmatrix} e^{j\omega t}\, d\omega \tag{10}$$

which only depends on the incident pulses $p_q(t)$ or the corresponding spectra $P_q(\omega)$ in the two orthogonal incident polarizations $q = \theta$ and $q = \psi$. The scattered time domain far field can then be found at an arbitrary point $(r, \theta_{\mathrm{scat}}, \psi_{\mathrm{scat}})$ as

$$\begin{bmatrix} e_\theta^{\mathrm{scat}}(t) \\ e_\psi^{\mathrm{scat}}(t) \end{bmatrix} = \frac{1}{r} \begin{bmatrix} w_{\theta\theta}(t-r/c) & w_{\theta\psi}(t-r/c) \\ w_{\psi\theta}(t-r/c) & w_{\psi\psi}(t-r/c) \end{bmatrix} \cdot \begin{bmatrix} E_{\theta 0} \\ E_{\psi 0} \end{bmatrix}, \quad \text{for} \quad \begin{bmatrix} e_\theta^{\mathrm{inc}}(t) \\ e_\psi^{\mathrm{inc}}(t) \end{bmatrix} = \begin{bmatrix} E_{\theta 0} p_\theta(t) \\ E_{\psi 0} p_\psi(t) \end{bmatrix} \tag{11}$$

assuming that medium 1 is nondispersive and lossless (e.g., air), where time domain variables are written as small letters and t here represents time.

For a calculation of the scattering matrix $\boldsymbol{S}$ (for one incident angle) we solve (7) for two different driving vectors $\boldsymbol{V}^m$, corresponding to two orthogonal polarizations. Explicit expressions for the elements of the driving vectors are given in the literature [1]–[9] for an incident plane wave in free-space (e.g., air). For the problem considered here, these expressions have to be modified, to take into account the direct (downward), reflected (upward) and/or transmitted (downward) incident field as well as the refraction at the interfaces, both depending on the layer i in which the BOR is located. The number of azimuthal modes required in the Fourier series approximation of the incident fields, $m = -m_{\max} \cdots m_{\max}$, depends on $k_\rho \rho_{\max} = k_1 \rho_{\max} \sin\theta_{\mathrm{inc}}$, using the maximum radius $\rho_{\max}$ of the BOR. In our MoM implementation $m_{\max}$ is determined adaptively to guarantee errors less than 0.1% compared to the overall incident field. If medium 1 is lossless (e.g., $k_1 \approx k_0$), this could be simplified by using the approximation

$$m_{\max} \approx 9 + 1.07 k_\rho \rho_{\max} - \frac{16}{2 + k_\rho \rho_{\max}} \quad \text{valid for } 0 \le k_\rho \rho_{\max} \le 100 \tag{12}$$

which was found to provide a close fit to $m_{\max}$ for the range indicated.

After having solved for the unknown current coefficients $\boldsymbol{I}^m$ in (7), the dyadic Green's functions needed in the calculation of the scattered far field $\boldsymbol{E}^{\mathrm{scat}}$ are evaluated asymptotically using only the saddle point contribution, neglecting surface and leaky waves [37]. Explicit equations for the far field scattered by a BOR, taking into account the Fourier series expansion in the azimuth direction and therefore calculating the ϕ'-integral in the 2-D surface current integration analytically, can be easily derived by generalizing the solution given for free space [1]–[4].

III. Wide-Band Scattering Results

The layered-medium BOR algorithm is applicable to many problems of interest in radar based land mine detection. For example, one can account for the layered manner in which soil is often distributed. Moreover, the overturned soil in which a mine is buried often has electrical properties different from those of the surrounding, undisturbed soil; the disturbed soil can be modeled approximately via a lossy dielectric BOR, in which a separate BOR (representing the model mine) is placed. Such applications are the subjects of future papers. Here we demonstrate a few examples, which show the potential and utility of the algorithm. In particular, we first examine scattering from a plastic BOR (model plastic mine) buried in a lossy, dispersive half-space (Section III-A), with half-space electrical parameters representative of data measured from real soil [38]. As a comparison, we also consider scattering from a buried perfectly conducting BOR of the same shape as the model plastic mine, with the perfectly conducting BOR modeled by use of a subset of the terms in (7) [17]. A comparison between the weak energy scattered from dielectric mines, relative to conducting mines of the same shape, underscores the difficulty of radar-based plastic mine detection. Finally, to demonstrate an example for which the layered-medium Green's function is applicable, we consider scattering from a plastic BOR buried just under the soil, with the soil covered by a layer of snow (Section III-B). In this context we examine the scattered signal as a function of snow type (snow wetness) and snow depth.

In the context of radar-based detection of buried targets, there is always a tradeoff between resolution and signal strength. To achieve good penetration in typical lossy soils, ground-penetrating radar systems are typically designed to operate at frequencies less than approximately 1.5 GHz [11], [14], [19], [20]. However, such a restriction in bandwidth reduces resolution, thereby minimizing the information content in the scattered fields. To enhance the information in the scattered fields, one must increase the system bandwidth (resolution), although this implies increased signal attenuation in lossy soils. This issue is examined by considering several incident waveforms of differing bandwidth and center frequency.

Finally, before proceeding to the results, we note that there is very little work we are aware of published on short-pulse scattering from buried dielectric targets. The computer code was checked extensively to verify its accuracy, and tests were successfully performed for a dielectric sphere in free-space and with the limited data available on short-pulse scattering from buried targets [17].

A. Dielectric and Conducting BOR Buried in a Half Space

In our first set of examples, we consider perfectly conducting (PEC) and dielectric cylinders of the same shape, buried in

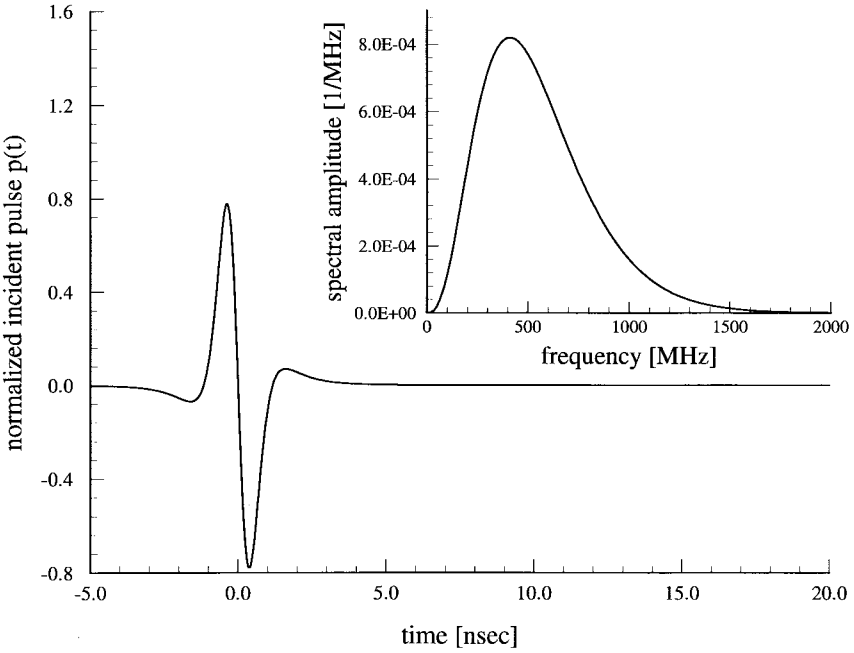


Fig. 3. Incident pulse $p(t)$ (Rayleigh pulse of order $n = 3$, center frequency $f_c = 410$ MHz [39]) and corresponding amplitude spectrum $|P(f)|$ (used in Fig. 4).

a lossy dispersive half space. The PEC and plastic targets are both cylinders of diameter 12.6 cm and height 5.3 cm, with the dielectric target characterized by $\varepsilon'_{rB} = 2.9$ and $\sigma_B = 10^{-3}$ S/m (PVC). The top of the cylinders are placed 5 cm beneath the air-ground interface, and the pulsed plane wave is incident at $\theta_{\rm inc} = 50$ deg (see Fig. 1). Finally, the half-space in which these targets are buried is described by the characteristics of Puerto Rico clay (10% water content), as reported in [38].

The first set of examples utilize the incident pulse in Fig. 3 (a Raleigh wavelet [39] of order $n = 3$ and center frequency at $f_c = 410$ MHz). This incident waveform and bandwidth are typical of many ground-penetrating systems [11], [14], [19], [20]. The far-zone backscattered fields are normalized as in (10)–(11), with results presented in Fig. 4. For a BOR target there are no cross-polarized backscattered fields, and therefore results are only plotted for VV (vertical incidence, vertical receive) and HH (horizontal incidence, horizontal receive) scattering. While there are slight differences in the VV and HH scattered fields for a given target, the most dramatic characteristic of the results in Fig. 4 is the large discrepancy in the amplitudes of the waveforms scattered from the PEC and PVC targets (despite the fact that the targets have exactly the same shape) and the different polarity of the signals, the latter resulting from a "negative" contrast in permittivity ($\varepsilon'_{rB} = 2.9$ for the PVC target compared to a soil permittivity ranging from 5 to 6).

The relatively weak backscattered fields for the plastic target are attributed to the modest contrast between the target and the background soil. Over the bandwidth considered here, the Puerto Rico clay [38] has a dielectric constant ranging from 5 to 6, as compared to $\varepsilon'_{rB} = 2.9$ for the target. It should be pointed out, however, that many dry soils have dielectric constants even closer to ε'_{rB} for which the fields scattered from the plastic target will be even smaller.

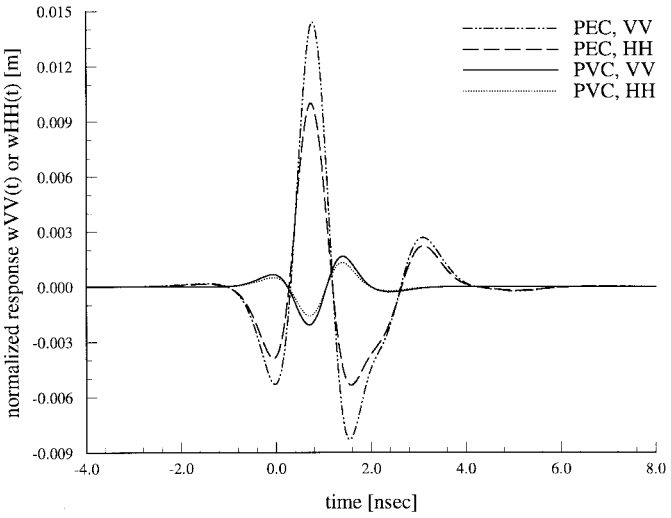


Fig. 4. Comparison of the normalized time domain response for a perfectly conducting (PEC) and a plastic (PVC: $\varepsilon'_{rB} = 2.9, \sigma_B = 10^{-3}$ S/m) cylinder of same dimensions (diameter = 12.6 cm, height = 5.3 cm). The top of the cylinder is 5 cm below the ground-air interface at $z_1 = 0$ cm (layer 1: air, layer 2: Puerto Rico clay with 10% water [38]). Results are shown for $w_{VV}(t) = w_{\theta\theta}(t)$ and $w_{HH}(t) = w_{\psi\psi}(t)$ in case of backscattering and $\theta_{\rm inc} = 50^\circ$ (no cross-polarization for backscattering from BOR [22]). The incident pulse is given in Fig. 3.

Another characteristic of the results in Fig. 4 is the fact that the time-domain scattered fields have duration similar to the incident waveform in Fig. 3 (although the wave shapes are distorted due to dispersive target scattering and propagation through the dispersive soil). Therefore, for the bandwidth in Fig. 3, the targets (both PEC and PVC) are dispersive point scatterers. Since there are many obstacles which will similarly be point scatterers (rocks, roots, etc.), target discrimination is difficult with this incident waveform. To address this issue, one can utilize a wider bandwidth incident waveform such

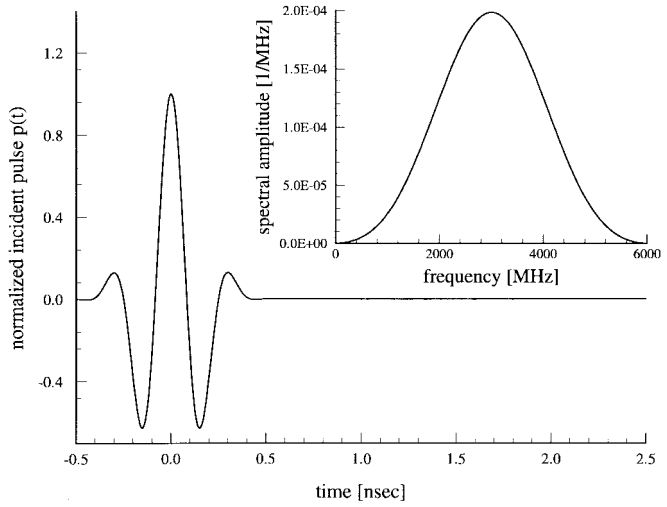


Fig. 5. Incident pulse $p(t)$ (center frequency $f_c = 3000$ MHz) and corresponding amplitude spectrum $|P(f)|$ (used in Fig. 6).

that higher resolution is achieved. We therefore consider the waveform in Fig. 5, which has a center frequency at $f_c = 3000$ MHz and considerably wider bandwidth than Fig. 3. To accommodate the increased bandwidth, however, the principal energy components are at significantly higher frequency than those in Fig. 3, resulting in increased attenuation as the wave penetrates the soil. In Fig. 6 we consider the same targets as in Fig. 4, but using the incident waveform in Fig. 5. As seen from Fig. 6(a) and (b), the increased bandwidth results in a noticeable difference in the VV and HH backscattered fields. Moreover, the shapes of the waveforms scattered from the PEC and PVC targets are different as well (note the different polarity already mentioned above), although, as in Fig. 4, the scattered fields from the PEC target are markedly stronger than those of the plastic target.

The increased bandwidth of Fig. 6 has resulted in the anticipated enhanced information content in the scattered fields (relative to Fig. 4). For the PEC target we attributed much of the additional waveforms (vis-à-vis, Fig. 4) to reverberations between the top of the target and the air-ground interface, which can be resolved using this larger bandwidth. For the plastic target there are also such reverberations, as well as energy that gets absorbed in and then reradiated from the interior of the plastic (these two phenomena are difficult to separate). The increased complexity in the scattered waveforms of Fig. 6 is accompanied by a significant reduction in the amplitude of the scattered fields (due to the aforementioned enhanced soil attenuation); comparing Fig. 6(a) and (b) with Fig. 4, we note that the higher bandwidth excitation results in an order-of-magnitude reduction in the scattered-field amplitudes.

B. Dielectric BOR Buried Under a Snow Layer

We consider a PVC cylinder of the same shape as above, buried 1 cm under soil, with the soil covered by a 10-cm snow layer. The winter conditions are likely to alter the soil properties (due to freezing), but for simplicity, we consider the soil as in Figs. 4 and 6. Dry snow is often easily penetrated

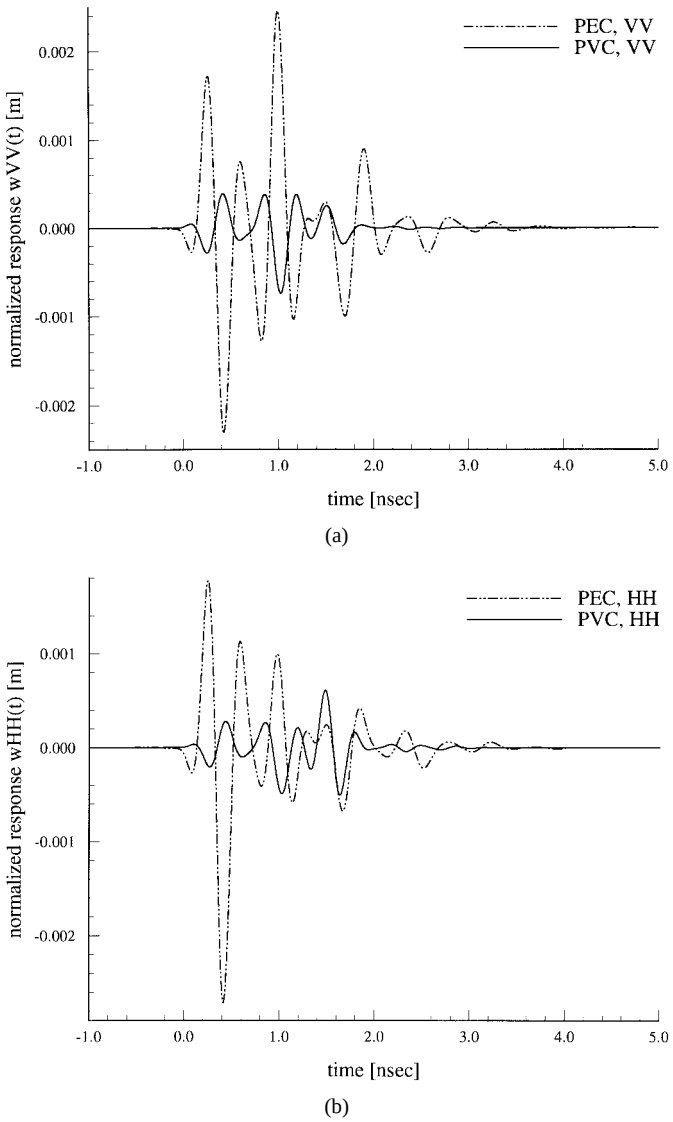


Fig. 6. Same as Fig. 4, but using pulse in Fig. 5. (a) $w_{VV}(t) = w_{\theta\theta}(t)$, (b) $w_{HH}(t) = w_{\psi\psi}(t)$.

(relative to soil) and the shallow target necessitates minimal soil penetration. Therefore, we consider the incident waveform in Fig. 7 (center frequency $f_c = 720$ MHz), characterized by increased bandwidth relative to the pulse in Fig. 3. In Fig. 8 we consider "dry snow," "snow," and "wet snow," characterized, respectively, by $\varepsilon'_{r2} = 1.5$ and $\sigma_2 = 2 \cdot 10^{-4}$ S/m; $\varepsilon'_{r2} = 3$ and $\sigma_2 = 10^{-3}$ S/m; $\varepsilon'_{r2} = 5$ and $\sigma_2 = 10^{-2}$ S/m [40]. For the VV results in Fig. 8(a), the increase in dielectric constant with snow wetness results in an expected temporal shift, but the general shape of the scattered waveform is largely unchanged with snow type. As a comparison, in Fig. 8(b), we consider HH scattering for which the scattered fields are characterized by a more substantial variation with snow type. Interestingly, in Fig. 8(b), the 10-cm layer of "dry snow" results in a stronger scattered field thatn when there is no snow at all; this is attributed to the very low loss of the dry snow and a snow-layer-induced enhanced impedance matching between the soil and air. The difference between the VV and HH cases is demonstrated even more dramatically in Fig. 9 for

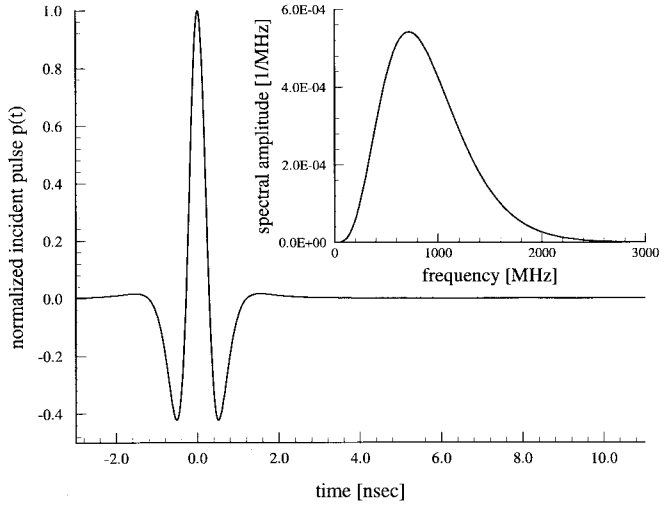


Fig. 7. Incident pulse $p(t)$ (Rayleigh pulse of order $n = 4$, center frequency $f_c = 720$ MHz [39]) and corresponding amplitude spectrum $|P(f)|$ (used in Figs. 8 and 9).

which the backscattered fields are examined as a function of snow depth (0, 5, 10, 15, 20 cm) for the "snow" of Fig. 8. For VV polarization there is verly little reverberation in the snow layer, while for HH polarization, there is a noticeable late-time oscillation manifested with increased snow depth. The minimal VV reverberation is attributed to the fact that there is near-total transmission at the various interfaces (particularly at the air–snow interface), due to near-Brewster-angle excitation [41], while such Brewster effects are absent for the HH case.

IV. Conclusions

A numerical algorithm has been developed for modeling plastic mines buried in a lossy layered medium (e.g., soil with or without a snow layer). To make such an analysis tractable for the wide-band, short-pulse problems of interest, we have assumed that the mine can be modeled as a BOR, which is a good approximation for most mines. Exploitation of this feature results in a significant savings in required computational resources (vis-à-vis finite elements [23] or the finite difference time domain [20]). While the MoM BOR algorithm has been utilized for some time, it is only through application of the method of complex images [27]–[29], for efficient evaluation of the layered-medium Green's function, that the wide-band computations of interest here are tractable. The wide-band scattering results for the PEC and PVC BOR's presented above took typically on the order of one and four hours of CPU time, respectively, on a 200-MHz Pentium Pro personal computer (with the code written in Fortran 77).

As was demonstrated in the results of Section III, the characteristics of the scattered waveform depend strongly on the environment, mine type, and system bandwidth. To deploy a radar system properly for a given application, it is important that the expected scattered signal amplitude and shape be known in advance (there are many plastic mines and soil types for which radar-based detection is *not*

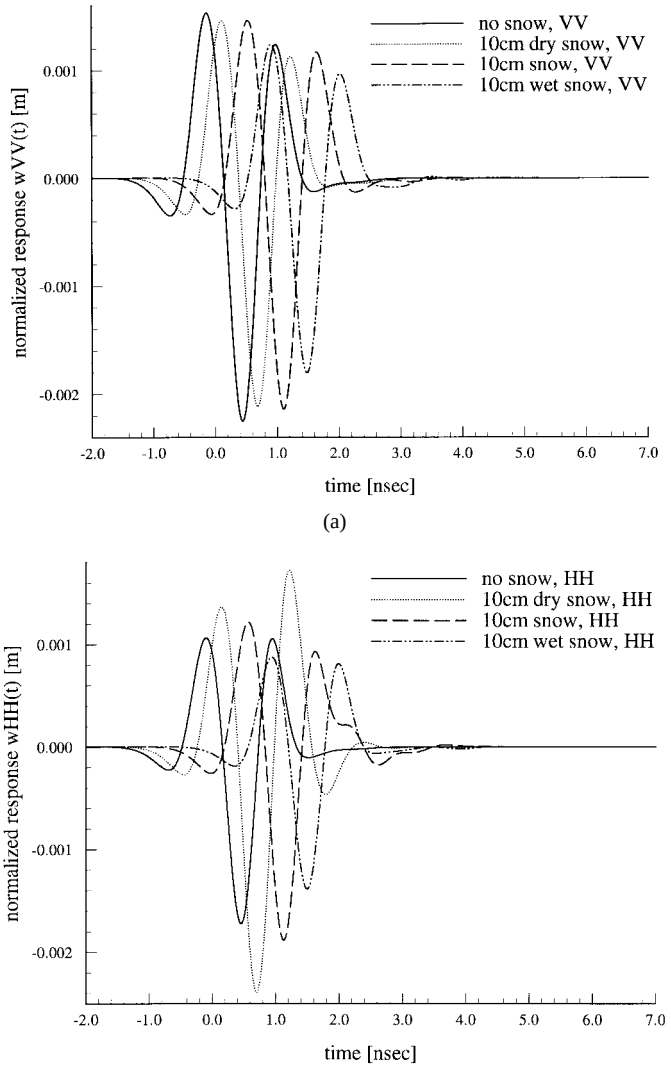
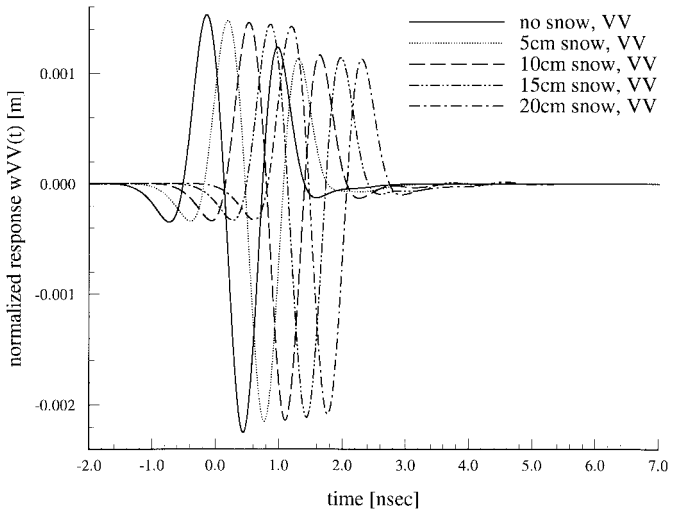


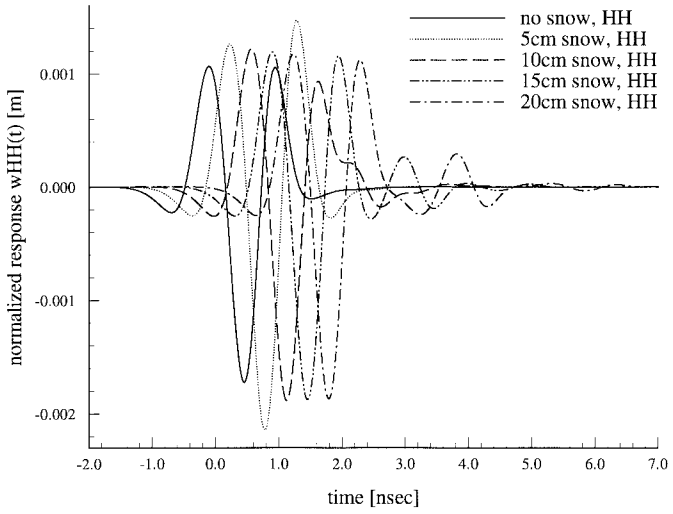
Fig. 8. Normalized time domain response for a buried PVC-cylinder (diameter = 12.6 cm, height = 5.3 cm, $\varepsilon'_{rB} = 2.9, \sigma_B = 10^{-3}$ S/m) including the effect of a 10-cm snow layer (layer 2: "dry snow" with $\varepsilon'_{r2} = 1.5$ and $\sigma_2 = 2 \cdot 10^{-4}$ S/m; "snow" with $\varepsilon'_{r2} = 3$ and $\sigma_2 = 10^{-3}$ S/m; "wet snow" with $\varepsilon'_{r2} = 5$ and $\sigma_2 = 10^{-2}$ S/m) on top of the soil (layer 3: Puerto Rico clay with 10% water content [38]). The top of the cylinder is 1 cm below the ground–snow interface at $z_2 = 0$ cm, the snow–air interface is located at $z_1 = 10$ cm. Results are shown for backscattering and $\theta_{\rm inc} = 60°$. The incident pulse is given in Fig. 7. (a) $w_{VV}(t) = w_{\theta\theta}(t)$. (b) $w_{HH}(t) = w_{\psi\psi}(t)$.

a viable option, especially when the contrast between the dielectric target and soil is small). It is for such purposes that the numerical analysis presented here is of particular importance. Moreover, in the design of optimal detectors [42], one requires *a priori* knowledge of the target signature as a function of soil type and target depth (the latter two are not known exactly in practice and, therefore, must be characterized statistically). For such applications, one must have access to the target signature for a large range of environmental conditions, necessitating the efficient algorithm presented here.

While the layered-medium BOR algorithm is applicable to many problems of interest in mine detection, here we have chosen to present a few representative examples. As expected,



(a)



(b)

Fig. 9. Same as in Fig. 8, but considering variable snow depths ($z_1 = 0, 5, 10, 15, 20$ cm) and "snow" with $\varepsilon'_{r2} = 3$ and $\sigma_2 = 10^{-3}$ S/m. (a) $w_{VV}(t) = w_{\theta\theta}(t)$. (b) $w_{HH}(t) = w_{\psi\psi}(t)$.

plastic targets result in significantly reduced scattered energy (relative to a conducting target of the same shape), with the details of the plastic-target signature dependent on the contrast between the electrical properties of the soil and target. We also examined scattering from targets buried under a snow layer. For dry snow, the attenuation introduced by snow penetration is minimal at the frequencies of interest, and radar provides an effective tool for subsurface sensing. We also witnessed difference in the VV and HH scattered fields in such environments due primarily to Brewster-angle effects.

An advantage of the frequency-domain solution presented here (relative to time-domain techniques, such as the finite difference time domain [20]) is its applicability to the computation of target resonances [18], [21]. The late-time resonances of buried plastic targets have been measured previously, with particularly encouraging results presented in [14]. However, as for the scattering problem addressed here, the details of the buried-plastic-target resonances are strongly dependent on the specifics of the target and background. To quantify the regimes for which resonance-based discrimination is appropriate, it is important to have an accurate model. The MoM analysis of buried-plastic-target (mine) resonances will be the subject of future research.

Member, IEEE

REFERENCES

[1] M. G. Andreasen, "Scattering from bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 13, pp. 303–310, Mar. 1965.
[2] J. R. Mautz and R. F. Harrington, "Radiation and scattering from bodies of revolution," *Appl. Sci. Res.*, vol. 20, pp. 405–435, June 1969.
[3] T. Wu and L. L. Tsai, "Scattering from arbitrarily-shaped lossy dielectric bodies of revolution," *Radio Sci.*, vol. 12, pp. 709–718, Oct. 1977.
[4] J. R. Mautz and R. F. Harrington, "Electromagnetic scattering from a homogeneous material body of revolution," *AEÜ*, vol. 33, pp. 71–80, Feb. 1979.
[5] A. W. Glisson and D. R. Wilton, "Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces," *IEEE Trans. Antenans Propagat.*, vol. 28, pp. 593–603, Sept. 1980.
[6] A. W. Glisson, D. Kajfez, and J. James, "Evaluation of modes in dielectric resonators using a surface integral equation formulation," *IEEE Trans. Microwave Theory Tech.*, vol. 31, pp. 1023–1029, Dec. 1983.
[7] A. W. Glisson, "Integral equation techniques," in *Dielectric Resonators*, D. Kajfez and P. Guillon, Eds. Dedham, MA: Artech House, 1986, pp. 259–325.
[8] D. Kajfez, A. W. Glisson, and J. James, "Computed modal field distributions for isolated dielectric resonators," *IEEE Trans. Micowave Theory Tech.*, vol. 32, pp. 1609–1616, Dec. 1984.
[9] S. D. Gedney and R. Mittra, "The use of the FFT for the efficient solution of the problem of electromagnetic scattering by a body of revolution," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 313–322, Mar. 1990.
[10] R. W. P. King and C. W. Harrison, "The transmission of electromagnetic waves and pulses into the earth," *J. Appl. Phys.*, vol. 39, pp. 4444–4452, Aug. 1968.
[11] D. L. Moffatt and R. J. Puskar, "A subsurface electromagnetic pulse radar," *Geophys.*, vol. 41, pp. 506–518, June 1976.
[12] G. S. Smith and W. R. Scott, "A scale model for studying ground penetrating radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 27, pp. 358–363, July 1989.
[13] C. Liu and L. C. Shen, "Numerical simulation of subsurface radar for detecting buried pipes," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 795–798, Sept. 1991.
[14] L. Peters, Jr., J. J. Daniels, and J. D. Young, "Ground penetrating radar as a subsurface environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.
[15] A. C. Dubey and R. L. Barnard, Eds., "Detection and remediation technologies for mines and minelike targets," *SPIE Proc.*, Orlando, FL, vol. 3079, 1997.
[16] S. Vitebskiy and L. Carin, "Moment-method modeling of short-pulse scattering from and the resonances of a wire buried inside a lossy, dispersive half-space," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 1303–1312, Nov. 1995.
[17] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse plane-wave scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 143–151, Feb. 1996.
[18] S. Vitebskiy and L. Carin, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy dispersive half-space," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.
[19] S. Vitebskiy, L. Carin, M. A. Ressler, and F. H. Le, "Ultra-wide-band, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.
[20] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of a ground-penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–44, Jan. 1996.
[21] N. Geng, D. Jackson, and L. Carin, "On the resonances of a dielectric BOR buried in a dispersive layered medium," submitted to *IEEE Trans. Antennas Propagat.*
[22] L. Carin, R. Kapoor, and C. E. Baum, "Polarimetric SAR imaging of buried landmines," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1985–1988, Nov. 1998..

[23] H. S. Chang and K. K. Mei, "Scattering of electromagnetic waves by buried and partly buried bodies of revolution," *IEEE Trans. Geosci. Remote Sensing*, vol. 23, pp. 596–605, 1985.
[24] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, Part I: Theory," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–344, Mar. 1990.
[25] J. R. Wait, "Image theory of a quasistatic magnetic dipole over a dissipative half-space," *Electron. Lett.*, vol. 5, pp. 281–282, May 1969.
[26] I. V. Lindell and E. Alanen, "Exact image theory for the Sommerfeld half-space problem, Part III: General formulation," *IEEE Trans. Antennas Propagat.*, vol. 32, pp. 1027–1032, Oct. 1984.
[27] Y. L. Chow, J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 588–592, Mar. 1991.
[28] J. J. Yang, Y. L. Chow, and D. G. Fang, "Discrete complex images of a three-dimensional dipole above and within a lossy ground," *Proc. Inst. Elect. Eng.*, vol. 138, pt. H, pp. 319–326, Aug. 1991.
[29] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.
[30] A. J. Poggio, M. L. Van Blaricum, E. K. Miller, and R. Mittra, "Evaluation of a processing technique for transient data," *IEEE Trans. Antennas Propagat.*, vol. 26, pp. 165–173, Jan. 1978.
[31] M. L. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. Antennas Propagat.*, vol. 23, pp. 777–781, Nov. 1975.
[32] Y. Hua and T. K. Sarkar, "Generalized pencil-of-function method for extracting poles of an EM system from its transient response," *IEEE Trans. Antennas. Propagat.*, vol. 37, pp. 229–234, Feb. 1989.
[33] ———, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 814–824, May 1990.
[34] Y. Rahmat-Samii, R. Mittra, and P. Parhami, "Evaluation of Sommerfeld integrals for lossy half-space problems," *Electromagn.*, vol. 1, no. 1, pp. 1–28, 1981.
[35] A. K. Abdelmageed and K. A. Michalski, "Analysis of EM scattering by conducting bodies of revolution in layered media using the discrete complex image method," in *IEEE Antennas Propagat. Symp. Dig.*, Newport Beach, CA, June 1995.
[36] M. I. Aksun, "A robust approach for the derivation of closed-form Green's functions," *Trans. Microwave Theory Tech.*, vol. 44, pp. 651–658, May 1996.
[37] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
[38] J. E. Hipp, "Soil electromagnetic parameters as functions of frequency, soil density, and soil moisture," *Proc. IEEE*, vol. 62, pp. 98–103, Jan. 1974.
[39] P. Hubral and M. Tygel, "Analysis of the Rayleigh pulse," *Geophysics*, vol. 54, pp. 654–658, May 1989.
[40] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, vol. I. Reading, MA: Addison-Wesley, 1982.
[41] C. A. Balanis, *Advanced Engineering Electromagnetics*. New York: Wiley, 1989.
[42] L. Carin, N. Geng, M. McLure, J. Sichina, and L. Nguyen, "Ultra-wide-band synthetic-aperture radar for mine-field detection," *IEEE Antennas Propagat. Mag.*, vol. 41, pp. 18–33, Feb. 1999.

Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Germany, in 1991 and 1996, respectively.

From 1991 to 1996, he was with the Institute for Microwaves and Electronics at the University of Karlsruhe, working on full-wave propagation modeling for radio communication systems. He received the Mannesmann Innovation Award in 1997 for his Ph.D. thesis on this subject. In January 1997 he joined the Department of Electrical and Computer Engineering at Duke University, Durham, NC, as a Visiting Post-Doctoral Fellow for 18 months. Since July 1998 he has been back with the University of Karlsruhe. His current research interests include computational methods in electromagnetics and wave propagation modeling.

Lawrence Carin (SM'96) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department at Polytechnic University, Brooklyn, as an Assistant Professor, and became an Associate professor there in 1994. In September 1995, he joined the Electrical Engineering Department at Duke University, where he is an Associate Professor. His current research interests include short-pulse scattering, propagation and signal processing. He is the principal investigator on an Army Research Office funded Multidisciplinary University Research Initiative (MURI) on demining.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.

Short-Pulse Electromagnetic Scattering from Arbitrarily Oriented Subsurface Ordnance

Norbert Geng and Lawrence Carin

***Abstract*—A rigorous method-of-moments (MoM) analysis is used to model wide-band scattering from general three-dimensional perfectly conducting objects buried in a lossy layered medium. Here, we focus on ordnance buried in a half space (soil). The time-domain fields scattered from a tilted antitank mine are examined in detail as a function of polarization and observation position.**

***Index Terms*—Radar, scattering, short-pulse.**

I. Introduction

The method of moments (MoM) has been used as a modeling tool for nearly three decades [1], and over this time, it has been used to analyze a wide range of scattering problems. Here, we are interested in wide-band (short-pulse) scattering from three-dimensional buried ordinance with application to the detection of buried mines and unexploded ordnance (UXO) [2]–[9]. For such problems, one requires the dyadic layered-medium Green's function [2], [3], [6]–[9], which is often specialized to the half-space case. This Green's function has been discussed in detail in [10], where several alternative formulations have been presented. For wide-band applications, one must efficiently evaluate the Sommerfeld-type integrals required for computation of the dyadic Green's function [10]. Here, we utilize the method of complex images, a technique developed by Chow *et al.* [11], [12] several years ago, and recently extended for general three-dimensional targets [13].

Therefore, the techniques used here to analyze scattering from a general buried perfectly conducting target are not new, the triangular-patch MoM formulation [1], [14], [15], layered-medium Green's function [10], and complex-image technique [11]–[13] having been published elsewhere. However, despite the importance of the buried-ordnance problem, there are few results published in the literature on the phenomenology underlying this application. We seek to fill this void, considering an antitank mine oriented arbitrarily, with short-pulse excitation of general polarization. Previous MoM results for this problem [2], [3], [6]–[9] considered both perfectly conducting and dielectric (plastic) mines, but the problem was simplified by assuming that the mine–(layered medium) *composite* was a body of revolution (BOR). This BOR assumption requires that the axis of the target be perpendicular to the air–ground interface, precluding an examination of the importance of the mine tilt on the scattered fields.

One can also model scattering from general buried targets through use of the finite-difference time-domain (FDTD) method [4]. Previous FDTD results have concentrated on near-field applications, in which interactions between the antenna and target are accounted for rigorously [4]. Here, we are interested in applications for which the sensor is distant from the target (synthetic aperture radar [6], [7]) which, in principle, could also be analyzed via FDTD. We utilize the MoM in lieu of FDTD because the former only requires spatial discretization on the target surface (for a surface MoM formulation [14], [15]), while the latter necessitates discretization of the entire computational domain.

Our principal focus here is on the presentation of example results, which will serve two purposes. First, the detailed phenomenological study will shed light on the physics underlying wide-band scattering from buried ordnance. Second, since, to our knowledge, similar results have not been published previously, the data presented here will be of interest for comparison to future computational algorithms, be they integral-equation (MoM) or differential-equation (FDTD) based.

Manuscript received August 12, 1998; revised October 13, 1998.

N. Geng was with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA. He is now with the Institüt für Hoechstfrequenztechnik und Elektronik, University of Karlsruhe, D-76128 Karlsruhe, Germany.

L. Carin is with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291.

Publisher Item Identifier S 0196-2892(99)03487-7.

II. Phenomenological Study

Consider a cylindrical perfectly conducting target, of 33 cm diameter and 16 cm height, buried in soil with dielectric constant $\epsilon_r = 4$ and conductivity $\sigma = 0.005$ S/m. We have performed detailed analyzes on the effects of soil type on the scattered fields, and the soil properties selected here are representative of a wide class of soils [7], [16]. The MoM analysis can readily handle dispersive soils [2], [3], [6]–[9], although here our focus is on the target itself. The cylindrical target selected closely resembles the M20 antitank mine [6], [7]. Finally, the computer code was validated through comparison with published results [2], [3], [7], in which the special case of *nontilted* bodies of revolution (BOR's) was considered.

It has been shown in a previous correspondence [6] that the cross-polarized fields *backscattered* from a BOR are zero, and this has been validated experimentally through consideration of the *nontilted* M20 mine [6]. Since many forms of anthropic clutter do not satisfy the BOR model, the cross-polarized fields potentially provide a powerful discriminant for distinction of targets and clutter [6]. However, for a buried target, the BOR model is only appropriate if the target axis is perpendicular to the air–ground interface, such that the target–(layered medium) *composite* satisfies the BOR model [6]. In practice, however, one may anticipate that the buried mine may be tilted, such that the BOR-composite model is no longer appropriate. It is of interest to determine the degree to which this phenomenon corrupts the ideal properties of the cross-polarized scattered fields.

A second issue of interest involves the aspect-dependent nature of the fields scattered from a buried tilted mine. If the target–(half space) composite satisfies the BOR model, the scattered fields are only dependent on the spherical coordinates (r, θ), independent of the azimuthal position ϕ, assuming that the center of the coordinate system is along the BOR axis. Now, consider a *tilted* mine, with the center of the coordinate system at the center of the mine, and with the z axis defined as perpendicular to the air–ground interface. If the backscattered fields from the tilted mine are a strong function of the azimuthal angle ϕ, then the difficulty of target detection is enhanced significantly. Moreover, for synthetic aperture radar (SAR) applications [6], [7], the scattered fields from multiple target-sensor orientations are employed to generate a SAR image. Consequently, a strong dependence of the tilted-mine backscattered fields on the target-sensor orientation will have an important effect on the subsequent SAR imagery.

We first consider the polarimetric issue, and in Fig. 1 plot the HV (horizontally polarized receive fields, and vertically polarized excitation fields) for a mine tilted 5°. A pulsed plane wave is assumed incident 30° from grazing, with a pulse shape as shown in Fig. 1; the incident pulse has a center frequency of 500 MHz and

0196–2892/99$10.00 © 1999 IEEE

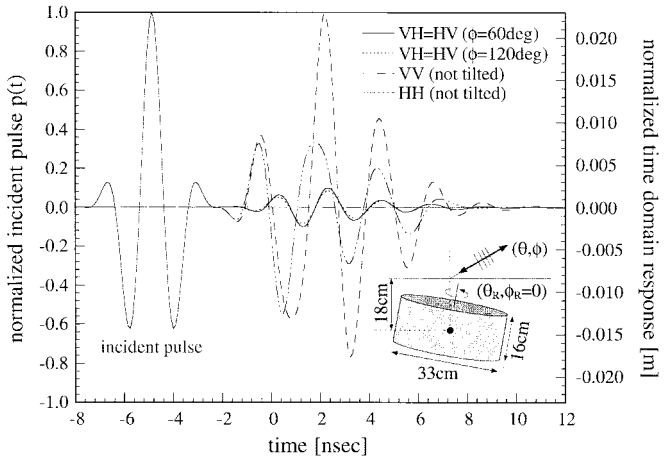


Fig. 1. Time-domain fields backscattered from a tilted perfectly conducting cylinder, which simulates an M20 antitank mine (inset). We consider the target tilted at 5°. Shown are the incidentpulse (scale on left) and four normalized backscattered waveforms (scale on right). The cross-polarized HV fields (which are the same as the VH fields) are shown for azimuthal angles $\phi = 60^\circ$ and $\phi = 120^\circ$. For comparison, results are also shown for the copolarized VV and HH fields for the casewhen the mine is *not* tilted. The soil has $\epsilon_r = 4$ and $\sigma = 0.005$ S/m, and the excitation plane wave is incident at 30° from grazing.

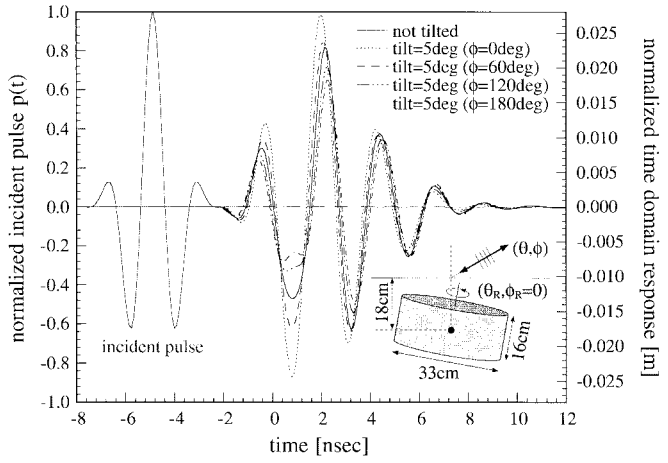


Fig. 2. As in Fig. 1, but for VV polarization. Results are shown for $\phi = 0^\circ, 60^\circ, 120^\circ$, and 180°. Backscattered fields are also shown for the case when the mine is *not* tilted, for which the scattered fields are independent of ϕ.

significant frequencies that extend to approximately 1 GHz. The far-zone, cross-polarized backscattered fields are plotted for azimuthal angles $\phi = 60^\circ$ and $\phi = 120^\circ$, where $\phi = 0^\circ$ defines a bisectional symmetry plane through the tilted target. Note that the cross-polarized backscattered fields vanish if the wave vector is in the $\phi = 0^\circ$ plane [6]. To quantify the relative strength of the cross-polarized backscattered fields, we also plot the VV and HH backscattered fields for the case when the mine is *not* tilted (for which case, as mentioned, there are no cross-polarized fields). Each backscattered far-zone time-domain waveform is normalized through multiplication by the observation distance r (each waveform is also time shifted). We see from Fig. 1 that the tilted mine yields appreciable cross-polarized backscattered fields, even for the relatively slight (5°) tilt considered here. Although not shown due to space constraints, we also considered the same mine tilted 10°, for which we witnessed an approximate doubling of the backscattered HV response. These results demonstrate that the cross-polarized fields for a mine—which

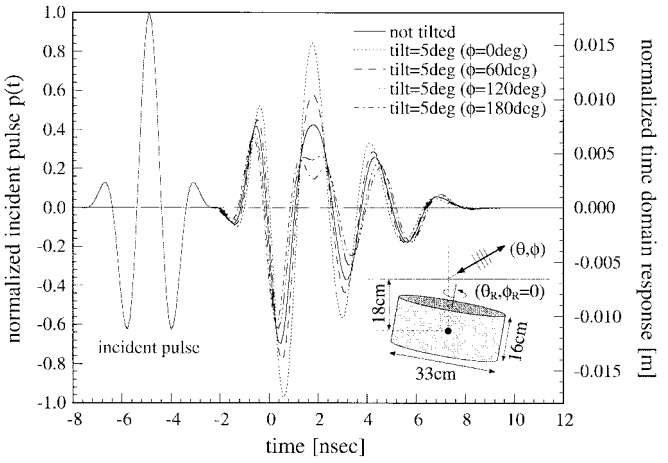


Fig. 3. As in Fig. 2, HH polarization.

are ideally zero if the mine is oriented with its axis perpendicular to the soil—are appreciable even for a modest tilt. This may vitiate the utility of polarimetry for SAR-based target identification. A more conclusive assessment of this issue will require examination of the relative strength of cross-polarized fields from clutter of interest.

We now consider the copolarized fields scattered from a tilted mine as a function of the azimuthal angle ϕ. In Fig. 2, we consider VV backscattering, for a 5° mine tilt, for angles $\phi = 0^\circ, 60^\circ, 120^\circ$, and 180°. For comparison, we also consider the case when the mine is not tilted, for which the backscattered fields are the same for all ϕ. For the $\phi = 0^\circ$ case, the mine is tilted forward toward the sensor, yielding enhanced backscattered signal strength relative to the case of a nontilted mine. Moreover, we see that the backscattered fields are a strong function of the angle ϕ, complicating target identification. Although the detailed shape of the backscattered fields are different for HH polarization (Fig. 3), these are again a strong function of the target-sensor orientation (ϕ). Space limitations preclude showing results for the case of a 10° tilt, but this case again manifests an even stronger difference in the backscattered fields relative to the nontilted mine. In fact, at $\phi = 0^\circ$, the backscattered VV fields were almost twice the strength of the fields for the nontilted case, with similar effects witnessed for HH polarization.

III. Conclusion

A rigorous MoM analysis has been performed on the polarmetric time-domain fields scattered from a tilted antitank mine buried in soil. The mine may initially be placed in the ground in a tilted manner, or such a tilt could be caused by movement of the mine after being in the ground through multiple seasons (many mines have been in the ground for decades). Our results demonstrate that, even for a relatively modest tilt of 5°, the backscattered fields are significantly altered relative to the no-tilt case. In particular, if *not* tilted, BOR's have no cross-polarized fields. However, here we have shown that tilted mines have significant cross-polarized fields, even for a relatively small tilt angle. Moreover, for the tilted mine, the VV and HH scattered fields are a strong function of the target-sensor orientation, complicating target identification and SAR imaging. However, these results also underscore the benefit of rigorous modeling, with which such phenomena can be predicted *a priori* and exploited (or, at least, accounted for) in the context of signal processing. Finally, we note that the dependence of the backscattered fields on the tilt angle is a function of the system bandwidth. For more narrow bandwidth systems than considered here, the dependence on the tilt angle will be less appreciable. However,

in this case, the target appears as a point scatterer, for which little discrimination capability is possible.

REFERENCES

[1] R. F. Harrington, *Field Computation by Moment Methods*. New York: IEEE Press, 1993.

[2] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse plane-wave scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 143–151, Feb. 1996.

[3] S. Vitebskiy and L. Carin, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy dispersive half-space," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.

[4] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of a ground-penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–44, Jan. 1996.

[5] L. Peters, Jr., J. J. Daniels, and J. D. Young, "Ground penetrating radar as a subsurface environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.

[6] L. Carin, R. Kapoor, and C. E. Baum, "Polarimetric SAR imaging of buried landmines," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1985–1988, Nov. 1998.

[7] S. Vitebskiy, L. Carin, M. A. Ressler, and F. H. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.

[8] N. Geng, D. Jackson, and L. Carin, "On the resonances of buried dielectric bodies of revolution," *IEEE Trans. Antennas Propagat.*, submitted for publication.

[9] N. Geng and L. Carin, "Short-pulse scattering from buried dielectric bodies of revolution," *IEEE Trans. Antennas Propagat.*, to be published.

[10] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, Part I: Theory," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–344, Mar. 1990.

[11] Y. L. Chow, J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 588–592, Mar. 1991.

[12] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.

[13] M. I. Aksun, "A robust approach for the derivation of closed-form Green's functions," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 651–658, May 1996.

[14] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering from surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 409–418, May 1982.

[15] D. R. Wilton, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak, and C. M. Butler, "Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains," *IEEE Trans. Antennas Propagat.*, vol. 32, pp. 276–281, Mar. 1984.

[16] J. E. Hipp, "Soil electromagnetic parameters as functions of frequency, soil density, and soil moisture," *Proc. IEEE*, vol. 62, pp. 98–103, Jan. 1974.

Covariance Estimation with Limited Training Samples

Saldju Tadjudin and David A. Landgrebe

***Abstract*— This paper describes a covariance estimator formulated under an empirical Bayesian setting to mitigate the problem of limited training samples in the Gaussian maximum likelihood (ML) classification for remote sensing. The most suitable covariance mixture is selected by maximizing the average leave-one-out log likelihood. Experimental results using AVIRIS data are presented.**

***Index Terms*— Covariance estimation, Gaussian maximum likelihood, leave-one-out log likelihood, regularization.**

I. INTRODUCTION

In the conventional Gaussian maximum likelihood (ML) classifier, the classification rule can be expressed in the form of a discriminant function and a sample is assigned to the class with the largest discriminant function value. A multivariate Gaussian density function is given as

$$f_i(x) = (2\pi)^{-p/2}|\Sigma_i|^{-1/2}\ \exp\Big[-\tfrac{1}{2}(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i)\Big], \qquad 1 \le i \le L$$

where $x \in \Re^p$, μ_i, and Σ_i are the ith class mean vector and covariance matrix, respectively, and L is the number of classes. Assuming a [0, 1] loss function, the ML classification rule then becomes

$$d_i(x) = \min_{i \le i \le L}\ d_i(x)$$

where d_i is the discriminant function given by

$$d_i(x) = (x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i) + \ln|\Sigma_i|.$$

This classification rule is also called a quadratic classifier. A special case occurs when all of the class covariance matrices are identical. It then becomes a linear classifier

$$\Sigma_i = \Sigma, \qquad 1 \le i \le L.$$

In practical situations, the true class distributions are rarely known. Therefore, the sample estimates are computed from the training samples.

The quadratic classifier's performance can be degraded when the number of dimensions is large compared to the training set size due to the instability of sample estimates. In particular, the sample covariance estimate becomes highly variable and may even be singular. One way to deal with the instability of covariance estimate is to employ the linear classifier. By replacing each class covariance estimate with their average, leading to the linear classifier, the number

Manuscript received July 31, 1998; revised October 28, 1998. This work was supported in part by NASA under Grant NAG5-3975 and the Army Research Office under Grant DAAH04-96-1-0444.

The authors are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907-1285 USA (e-mail: landgreb@ecn.purdue.edu).

Publisher Item Identifier S 0196-2892(99)04334-X.

0196–2892/99$10.00 © 1999 IEEE

Fast Multipole Method for Targets Above or Buried in Lossy Soil

N. Geng [*1] and L. Carin [2]

[1] Institut für Höchstfrequenztechnik und Elektronik,
Universität Karlsruhe, Kaiserstraße 12, D-76128 Karlsruhe, Germany

[2] Department of Electrical and Computer Engineering,
Duke University, Box 90291, Durham, NC 27708-0291

1. Introduction

The modeling of electromagnetic scattering has been of longstanding interest, with the method of moments (MoM) [1] representing a prominent example. While the MoM is powerful, it has limitations that restrict its utility. In particular, for N unknowns, the required memory (RAM) is of order $O(N^2)$, while the computational complexity depends on whether a direct (LU-decomposition) or iterative (conjugate gradient CG) solver is applied, the former requiring $O(N^3)$ operations and the latter $O(N^2)$ operations per iteration. To counter these limitations, there has been significant interest in the fast multipole method (FMM) [2-5]. The simplest two-level implementation [2,3] has complexity $O(N^{3/2})$ in RAM and CPU, while a multi-level FMM [4,5] further reduces this to $O(N\log N)$. While the FMM represents a promising tool, it has heretofore been applied primarily to free-space scattering [2-5], 2D analysis in layered media [6], or quasi-planar 3D problems in circuit and antenna design [7]. There are many important applications for which the free-space model or the quasi-planar approximation are not valid. For example, there has been significant interest in radar sensing of buried targets, such as mines or unexploded ordnance (UXO). Moreover, the soil must be included in the analysis of scattering from a target situated above and near the ground (*e.g.*, vehicles). While mines are generally small, and therefore amenable to an MoM analysis, many targets of interest are not. It is therefore desirable to adapt the FMM to the problem of scattering from a target in the vicinity of a half-space interface. Here we concentrate on an extension of the two-level FMM [2, 3], while similar modifications are possible with the multi-level FMM [4,5].

2. Half-space fast multipole method

We utilize the half-space electric field integral equation (EFIE) [1,8]

$$\hat{\mathbf{n}}\times\mathbf{E}^{inc}(\mathbf{x})=\hat{\mathbf{n}}\times j\omega\mu_i\iint_{S'}\left[\bar{\mathbf{I}}+\frac{\nabla\nabla}{k_i^2}\right]\cdot\bar{\mathbf{G}}_{Aii}(\mathbf{x},\mathbf{x}')\cdot\mathbf{J}(\mathbf{x}')dS' \qquad (1)$$

for a general 3D PEC target situated entirely in layer i=1 or i=2 of a half space (Fig. 1). To avoid problems with interior resonances (for closed targets), a combined field integral equation (CFIE) [4] could be used instead. Details on the dyadic half-space Green's function can be found in [8]. Like in the MoM, the surface current $\mathbf{J}(\mathbf{x}')$ is expanded into a set of basis functions $\mathbf{b}_{n'}(\mathbf{x}')$, where we use the RWG basis defined on flat triangles representing the surface [1]. Testing (1) with a set of weighting functions $\mathbf{w}_n(\mathbf{x})$ results in a system of N linear equations (*i.e.*, $[\mathbf{Z}]\cdot\mathbf{I}=\mathbf{V}$) for the current coefficients, with the matrix elements given by

$$Z_{nn'}=j\omega\mu_i\iint_S\iint_{S'}\mathbf{w}_n(\mathbf{x})\cdot\left[\bar{\mathbf{I}}+\frac{\nabla\nabla}{k_i^2}\right]\cdot\bar{\mathbf{G}}_{Aii}(\mathbf{x},\mathbf{x}')\cdot\mathbf{b}_{n'}(\mathbf{x}')dS'dS \quad . \qquad (2)$$

In the FMM [2,3], we divide the computation into "near" and "far" terms. For

0-7803-5639-X/99/$10.00 © 1999 IEEE.

"near" interactions (*i.e.*, MoM part of FMM) we evaluate the half-space Green's function using the method of complex images [9]. Matrix elements representing "near" terms are stored in a sparse matrix [$\mathbf{Z}^{near}$]. For "far" terms we split the dyadic into a "direct" contribution (as in free space, but general *complex* wavenumber k_i) and a remaining dyadic accounting for the interface (Δ is *not* an operator)

$$\bar{\mathbf{G}}_{Aii}(\mathbf{x},\mathbf{x}') = \bar{\mathbf{I}} g_i(\mathbf{x},\mathbf{x}') + \Delta\bar{\mathbf{G}}_{Aii}(\mathbf{x},\mathbf{x}') = \bar{\mathbf{I}} \frac{e^{-jk_i|\mathbf{x}-\mathbf{x}'|}}{4\pi|\mathbf{x}-\mathbf{x}'|} + \Delta\bar{\mathbf{G}}_{Aii}(\mathbf{x},\mathbf{x}') \quad . \tag{3}$$

Consequently, the impedance matrix elements in (2) are split according to

$$Z_{nn'} = Z_{nn'}^{near} + Z_{nn'}^{far} = Z_{nn'}^{near} + Z_{nn'}^{far,hom} + \Delta Z_{nn'}^{far} \quad . \tag{4}$$

Using the addition theorem and a plane wave decomposition of the scalar free-space Green's function, the "far" matrix elements in the 3D *free-space* FMM are written as [2,3] (for a definition of the geometrical parameters see Fig. 1)

$$Z_{nn'}^{far,hom} \approx \frac{\omega\mu_i k_i}{(4\pi)^2} \iint_{4\pi} \mathbf{W}_{m\alpha}(\hat{\mathbf{k}}) \cdot T_L(k_i X_{m'm}, \hat{\mathbf{k}} \cdot \hat{\mathbf{X}}_{m'm})\, \mathbf{B}_{m'\alpha'}(\hat{\mathbf{k}})\, d^2\hat{\mathbf{k}} \tag{5a}$$

$$\mathbf{W}_{m\alpha}(\hat{\mathbf{k}}) = \left[\bar{\mathbf{I}} - \hat{\mathbf{k}}\hat{\mathbf{k}}\right] \cdot \iint_S \mathbf{w}_{n(m,\alpha)}(\mathbf{x}) e^{-jk_i\hat{\mathbf{k}}\cdot(\mathbf{x}-\mathbf{x}_m)}\, dS \tag{5b}$$

$$\mathbf{B}_{m'\alpha'}(\hat{\mathbf{k}}) = \left[\bar{\mathbf{I}} - \hat{\mathbf{k}}\hat{\mathbf{k}}\right] \cdot \iint_{S'} \mathbf{b}_{n'(m',\alpha')}(\mathbf{x}') e^{+jk_i\hat{\mathbf{k}}\cdot(\mathbf{x}'-\mathbf{x}_{m'})}\, dS' \tag{5c}$$

$$T_L(k_i X_{m'm}, \hat{\mathbf{k}} \cdot \hat{\mathbf{X}}_{m'm}) = \sum_{l=0}^{L} (-j)^l (2l+1)\, h_l^{(2)}(k_i X_{m'm})\, P_l(\hat{\mathbf{k}} \cdot \hat{\mathbf{X}}_{m'm}) \quad . \tag{5d}$$

In case of Galerkin testing and a lossless medium, (5b,5c) are complex-conjugate pairs. Consequently, either (5b) *or* (5c) need be stored in memory [2,3]. However, here we are interested in a general lossy half space (*i.e.*, k_i is *complex*), and therefore both must be computed. For an application of (5), the scatterer surface is first partitioned into groups m=1,..,M, each of which has an average number of $A_m \approx N/M$ basis functions. Inside group m the elements are labeled as α=1..A_m. The group information $n(m,\alpha)$ is stored in matrix format. As was shown in [2,3] for the two-level FMM, the optimal number of groups is $M \sim N^{1/2}$. Empirical approximations for the number L of terms needed in (5d) are given in [2,3,5] for a real wavenumber k_i. For complex k_i more terms are required [10]. Therefore, in our implementation L is determined adaptively at the beginning.

Using (5), the complexity with respect to RAM and CPU (for performing matrix-vector multiplies in the CG solver) is reduced from N^2 to $N^{3/2}$ [2,3]. The basic steps described in [2,3] remain valid to account for [$\mathbf{Z}^{near}$] and the "direct" part [$\mathbf{Z}^{far,hom}$] of the "far" interactions (see also final algorithm below).

However, for the half-space it is essential to include the "far" interface interactions represented by [$\Delta\mathbf{Z}^{far}$] in (4). The method of complex images [9] represents the Green's function components as sums of free-space Green's functions (with sources in complex space), which can, in principle, be handled within the FMM. While addition theorem and plane wave expansion remain valid for a complex wavenumber and complex images, the convergence is considerably slower [10]. Therefore, we utilize an approximate but highly accurate method for evaluating

the "far" interface interactions. In particular, if the target is entirely above or below the interface, and if basis and testing functions are distant, the "reflected" term can be evaluated asymptotically [11], leading to an image in real space with the amplitude given by the polarization dependent reflection coefficient.

Generalizing the free-space FMM is now straightforward. In the preprocessing stage [2,3] we include additional calculations of the translation operator (5d) between *image* and observation group centers as well as the spectral Fourier transforms $\mathbf{B}_{m'\alpha'}^{\nu}(\hat{\mathbf{k}})$ of the *image* expansion functions [10]. Extending [2,3], the matrix-vector multiplies in the CG solution are performed according to

$$\mathbf{s}_{m'}^{(\nu)}(\hat{\mathbf{k}}) = \sum_{\alpha'=1}^{A_{m'}} I_{n'(m',\alpha')} \mathbf{B}_{m'\alpha'}^{(\nu)}(\hat{\mathbf{k}}) \tag{6a}$$

$$\mathbf{g}_m(\hat{\mathbf{k}}) = \sum_{m'=1}^{M} T_L(k_i X_{m'm},\ldots)\, \mathbf{s}_{m'}(\hat{\mathbf{k}}) + T_L(k_i X_{m'm}^{\nu},\ldots)\, \bar{\mathbf{r}}_{m'm}^{\nu} \cdot \mathbf{s}_{m'}^{\nu}(\hat{\mathbf{k}}) \tag{6b}$$

$$\sum_{n'=1}^{N} Z_{nn'} I_{n'} = \sum_{n'=1}^{N} Z_{nn'}^{near} I_{n'} + \frac{\omega\mu_i k_i}{(4\pi)^2} \iint_{4\pi} \mathbf{W}_{m\alpha}(\hat{\mathbf{k}}) \cdot \mathbf{g}_m(\hat{\mathbf{k}})\, d^2\hat{\mathbf{k}} \quad , \tag{6c}$$

where the summation in (6b) is over "far" interactions only. The first term in (6b) represents "direct" contributions, the reflection dyadic and translation operator in the second term account for "far" interface interactions [10]. The three steps in (6a)–(6c) are often called aggregation, translation and disaggregation [4], the basic physical interpretation remains unchanged here.

3. Results

We demonstrate the accuracy of the half-space FMM by considering two targets: a model UXO buried under soil (Fig. 2) and a rectangular box situated above the ground (Fig. 3). All relevant parameters can be found in the figure captions. In all examples, the *bistatic* RCS is computed via the FMM developed here, as well as with a rigorous MoM, wherein the half-space Green's function is evaluated rigorously via the method of complex images [9]. Figs. 2 and 3 show an excellent agreement between FMM and MoM results, for co- and cross-polarized RCS. The RAM and CPU requirements of the half-space FMM are about twice to three times compared to the free-space FMM [10] (caused by the extra set of real images and the additional Fourier transforms (5b,5c), needed because (5b) and (5c) are no longer conjugate complex), but the complexity of $N^{3/2}$ is unchanged.

References

[1] S.M. Rao, D.R. Wilton, A.W. Glisson, "Electromagnetic scattering from surfaces of arbitrary shape," *IEEE Trans. AP*, vol. 30, pp. 409-418, May 1982

[2] R. Coifman, V. Rokhlin, S. Wandzura, "The fast multipole method for the wave equation: a pedestrian prescription," *IEEE Ant. Prop. Magazine*, vol. 35, pp. 7-12, June 1993

[3] J.M. Song, W.C. Chew, "Fast multipole method solution using parametric geometry," *Micr. Opt. Techn. Letters*, vol. 7, pp. 760-765, Nov. 1994

[4] J.M. Song, W.C. Chew, "Multilevel fast multipole algorithm for solving combined field integral equations of electromagnetic scattering," *Micr. Opt. Techn. Letters*, vol. 10, pp. 14-19, Sept. 1995

[5] J.M. Song, C.C. Lu, W.C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. AP.*, pp. 1488-1493, Oct. 1997

[6] L. Gürel, M.I. Aksun, "Electromagnetic scattering solution of conducting strips in layered media using the fast multipole method," *IEEE Micr. Guided Wave Let.*, vol. 6, pp. 277-279, Aug. 1996

[7] J.S. Zhao, W.C. Chew, C.C. Lu, E. Michielssen, and J. Song, "Thin-stratified medium fast-multipole algorithm for solving microstrip structures," *IEEE Trans. MTT*, vol. 46, pp. 395-403, Apr. 1998
[8] K.A. Michalski, D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media," *IEEE Trans. AP*, vol. 38, pp. 335-352, March 1990
[9] M.I. Aksun, "A robust approach for the derivation of closed-form Green's functions," *IEEE Trans. MTT*, vol. 44, pp. 651-658, May 1996
[10] N. Geng, A. Sullivan, L. Carin, "Fast multipole method for scattering from an arbitrary PEC target above or buried in a lossy half space," submitted to *IEEE Trans. AP*
[11] I.V. Lindell, *Methods for Electromagnetic Field Analysis*, IEEE Press, 1995

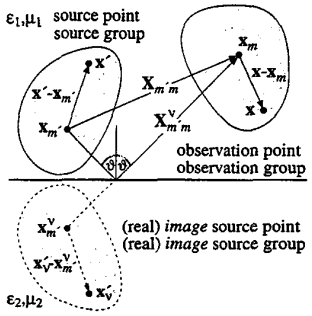


Fig. 1 Source, real image and observation group in 3D half-space FMM.

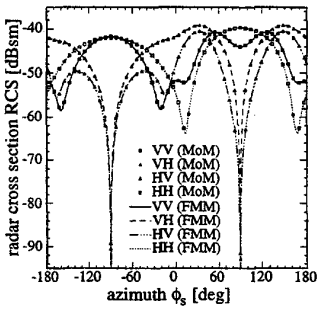


Fig. 2 RCS of a UXO (cylinder with spherical endcap, length=153cm, diameter=40.6cm) buried in Yuma soil with 5% water. The target axis lies in the *yz*-plane, 30° relative to the *z*-axis, with the nose at *z*=-217.5cm. The bistatic RCS is plotted at θ_i=60° for varying azimuth ϕ_s, for a plane wave incident at ϕ_i=-90° and θ_i=60° (frequency *f*=600MHz, *N*=11019 unknowns).

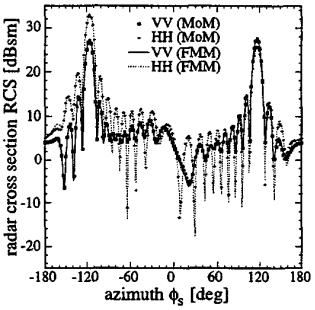


Fig. 3a RCS of a closed rectangular box with size $L_x \times L_y \times L_z$=12×2×2.5m³ situated 50cm above Yuma soil with 10% water. The co-polarized bistatic RCS (VV and HH) is plotted for varying azimuth angle ϕ_s at θ_i=80° (*i.e.*, 10° from grazing), for a plane wave incident at ϕ_i=60° and θ_i=60° (frequency is *f*=150MHz, 3776 triangles, leading to *N*=5664 unknowns).

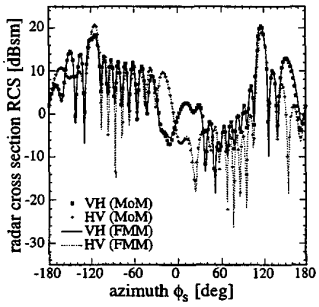


Fig. 3b Same as Fig. 3a, but plotting cross-polarized RCS (VH and HV).

Wide-band VHF Scattering from a Trihedral Reflector Situated Above a Lossy Dispersive Halfspace

Norbert Geng, *Member, IEEE,* Marc A. Ressler, and Lawrence Carin, *Senior Member, IEEE*

***Abstract*—The method of moments (MoM) is used to rigorously analyze wide-band VHF scattering from a perfectly conducting trihedral placed above a lossy, dispersive half space. The method of complex images is employed to evaluate the layered-medium Green's function efficiently and is applied here to the half-space problem. Particular attention is placed on the physics that underlie scattering from such targets (as a function of frequency, incidence angle, polarization, and soil type) for cases in which the target is small or of moderate size relative to wavelength (where high-frequency techniques fail). This problem is of interest when conventional trihedrals are employed to calibrate VHF synthetic aperture radar (SAR) systems.**

***Index Terms*—Calibration, numerical modeling, scattering, trihedrals.**

I. Introduction

IMAGE calibration is an issue of critical importance to synthetic aperture radar (SAR) systems [1]–[3]. Several techniques have been developed for SAR calibration, with nearly all involving canonical scatterers such as a trihedral [4]–[6]. Much of this previous research has been performed over narrow frequency bands, with the operating wavelength relatively small compared to the trihedral (or dihedral) dimensions [1]–[6]. While such operating conditions are appropriate for a wide class of radars, newly emerging systems require modified calibration procedures. For example, foliage-penetrating (FOPEN) and ground-penetrating (GPEN) SAR have been of significant interest recently, with many such systems operating over ultrawide-band (UWB) frequencies [7], [8]. Foliage-penetrating systems are designed to operate at wavelengths at which significant biomass penetration is possible, while retaining sufficient bandwidth to achieve desired resolution. Similar issues hold for soil penetration. For many foliage and soil types, it has been found efficacious to operate in the VHF and low-UHF band [7], [8]. At such frequencies, prohibitively large trihedrals are required if conventional high-frequency scattering models are employed during calibration. Alternatively, if conventional-sized trihedrals are utilized, improved models must be utilized to accurately simulate scattering from trihedrals when the wavelength is no longer small relative to characteristic target dimensions. In this paper, we take the latter approach, in which we rigorously model the VHF and low-UHF fields scattered from conventional-sized trihedrals placed over model ground (a lossy, dispersive half space). The results of this study are of importance for accurate FOPEN and GPEN SAR calibration.

Previous analyses of scattering from trihedral and dihedral scatterers generally have been based on high-frequency techniques [4]–[6] such as the geometrical and uniform theory of diffraction (GTD and UTD, respectively). While such techniques are elegant and often yield simple and therefore useful results [4]–[6], they are only appropriate when the high-frequency approximation is valid (characteristic target dimensions large relative to wavelength). Moreover, most of these previous analyses have considered the trihedral in free space, neglecting the effects of the soil. For conventional-sized canonical targets at the VHF and low-UHF frequencies of interest, such techniques are inappropriate. Therefore, to calibrate a FOPEN or GPEN SAR based on the fields scattered from such targets, we must employ numerical algorithms. Here, we utilize the method of moments (MoM) [9], [10], a well-known technique for modeling electromagnetic scattering at frequencies at which GTD and UTD fail. If the target is in free space, a MoM analysis is relatively straightforward, for one need only employ the scalar free-space Green's function [10]. However, at VHF and low-UHF frequencies, the wavelength is generally large relative to conventional trihedral targets, and the effects of the air–ground interface are not separated easily from those due to the trihedral itself. Consequently, we require rigorous application of the dyadic half-space Green's function, significantly complicating the MoM implementation. This is exacerbated by the fact that the space-domain dyadic Green's function cannot be expressed in closed form, with each component of the dyadic requiring solution of a Sommerfeld-type integral [9]. The kernels of such integrals generally are highly oscillatory and therefore require significant computational resources if conventional numerical-integration techniques are employed.

The components of the layered-medium dyadic Green's function are evaluated here using the recently developed complex-image technique [11]–[13]. This algorithm utilizes the Weyl identity [11]–[13], which equates a particular Som-

Manuscript received September 10, 1998; revised February 10, 1999. This work was supported by the Army Research Laboratory (ARL), Adelphi, MD, and was managed by J. Sichina of ARL. This work was performed in support of the U.S. defense-wide initiative on UWB radar for FOPEN, managed by M. Davis of DARPA.

N. Geng and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA (e-mail: lcarin@ee.duke.edu).

M. A. Ressler is with the Army Research Laboratory, Adelphi, MD 20783 USA.

Publisher Item Identifier S 0196-2892(99)06283-X.

0196–2892/99$10.00 © 1999 IEEE

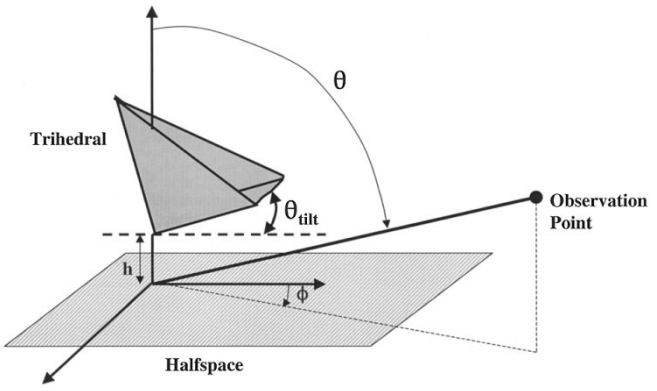


Fig. 1. Schematic of a trihedral fiducial target above a lossy, dispersive half space. The angles Θ and ϕ define the spherical coordinate system used to describe the results. $\Theta_{\rm tilt}$ represents the trihedral tilt, and h is the height of the target above the soil.

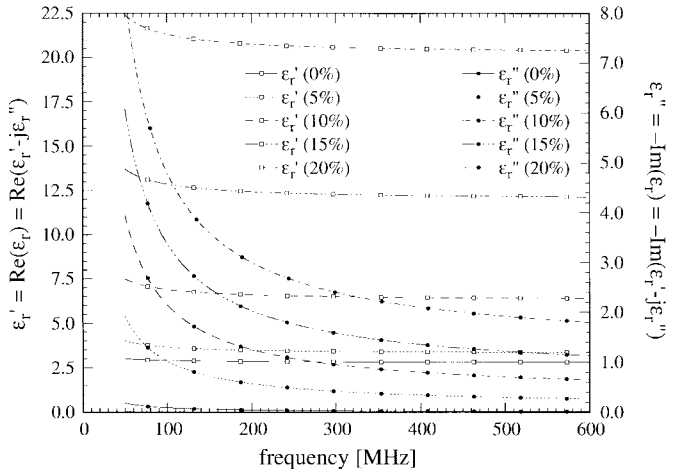


Fig. 2. Real and imaginary parts of the complex dielectric constant as a function of frequency for 0, 5, 10, 15, and 20% water content by percentage weight.

merfeld integral to the free-space Green's function. A parametric fitting algorithm, such as Prony's [14] or the matrix-pencil [15] method is used in the spectral domain to represent each component of the dyadic Green's function in terms of a sum of Sommerfeld integrals, each of which can be evaluated analytically via the Weyl identity [11]–[13]. Consequently, the arduous task of evaluating the oscillatory Sommerfeld integrals numerically is converted to the relatively simple task of performing a parametric fit along an appropriate path in the spectral domain [11]–[13], thereby significantly accelerating the speed at which the components of the dyadic Green's function can be evaluated. This is particularly important for wide-band applications, since the dyadic Green's function must be computed anew for each frequency of interest.

In the work presented here, we provide a summary presentation of the numerical algorithm itself, since its individual components have been presented previously [9]–[15]. Our principal interest is in applying it to the rigorous analysis of wide-band VHF and low-UHF scattering from trihedral fiducial targets, situated above a lossy, dispersive half space (model soil). While this problem is of critical importance for FOPEN and GPEN SAR applications, to our knowledge, such an analysis does not exist in the literature. The results presented here will be of use for accurately calibrating such systems, and potentially for the design of fiducial scatterers that are well tailored to the frequencies of interest in SAR-based detection of concealed targets.

The remainder of the text is organized as follows. In Section II, we summarize the MoM numerical algorithm, followed in Section III by a detailed examination of the fields scattered from trihedrals at VHF and low-UHF frequencies. In these results, we consider scattering from the trihedral when placed in free space and when placed over soil in order to quantify the effects of the latter. We consider trihedral dimensions representative of those generally deployed in practice for SAR applications [1]–[6] and address the variation of scattering from such targets as a function of incidence angle, polarization, frequency, trihedral tilt angle, and soil properties. The conclusions from this work are summarized in Section IV.

II. Summary of Numerical Modeling Algorithm

The electric field $\boldsymbol{E}(\boldsymbol{r})$ at the spatial location $\boldsymbol{r}$ produced by the surface currents $\boldsymbol{J}(\boldsymbol{r'})$ at $\boldsymbol{r'}$ is expressed as [9]

$$\boldsymbol{E}(\boldsymbol{r}) = \left[-j\omega\mu(\boldsymbol{r})\boldsymbol{I} + \frac{1}{j\omega\epsilon(\boldsymbol{r})}\ \nabla\nabla\right] \cdot \int_{S'} \boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r'}) \cdot \boldsymbol{J}(\boldsymbol{r'})\, dS' \tag{1}$$

where $\boldsymbol{I}$ is the unit dyadic, $\mu(\boldsymbol{r})$ is the permeability, and $\varepsilon(\boldsymbol{r})$ is the permittivity of the generally inhomogeneous background. The dyadic Green's function for the background medium is represented by $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r'})$. Here, we consider the special case of a layered medium [9] using "Formulation C" from [9].

We assume that the general 3-D target (e.g., a trihedral) is composed of a perfect electric conductor (PEC) excited by a plane-wave incident from a distant SAR system. The currents induced on the surface of the PEC are solved for using a conventional MoM analysis [9], [10], where here we have employed a triangular-patch [10] discretization of $\boldsymbol{J}(\boldsymbol{r'})$. For the case of free-space scattering, $\mu(\boldsymbol{r}) = \mu_o$ and $\varepsilon(\boldsymbol{r}) = \varepsilon_o$), $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r'}) = (\boldsymbol{xx}+\boldsymbol{yy}+\boldsymbol{zz})\exp(-jk_o|\boldsymbol{r}-\boldsymbol{r'}|)/4\pi|\boldsymbol{r}-\boldsymbol{r'}|$, and therefore the MoM solution of (1) is relatively straightforward [10]. Here, we are interested in a general layered medium, representative of soil. For this problem, each nonzero element of $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r'})$ must be represented as a Sommerfeld integral in the form [9]

$$f(\xi, z, z') = \frac{1}{2\pi}\int_0^\infty \hat{f}(k_\rho, z, z')\, J_o(k_\rho\xi)\, k_\rho\, dk_\rho \tag{2}$$

where $\xi = [(x-x')^2 + (y-y')^2]^{1/2}$. The kernel in (2) is generally highly oscillatory, undermining the efficiency of conventional numerical-integration procedures [9].

In the work presented here, (2) is evaluated by employing the Weyl identity [11]–[13]

$$\frac{e^{-jkR}}{4\pi R} = \frac{1}{2\pi}\int_0^\infty \frac{1}{j2k_z}\exp[(-jk_z h(z, z')]J_o(k_\rho\xi)k_\rho\, dk_\rho \tag{3a}$$

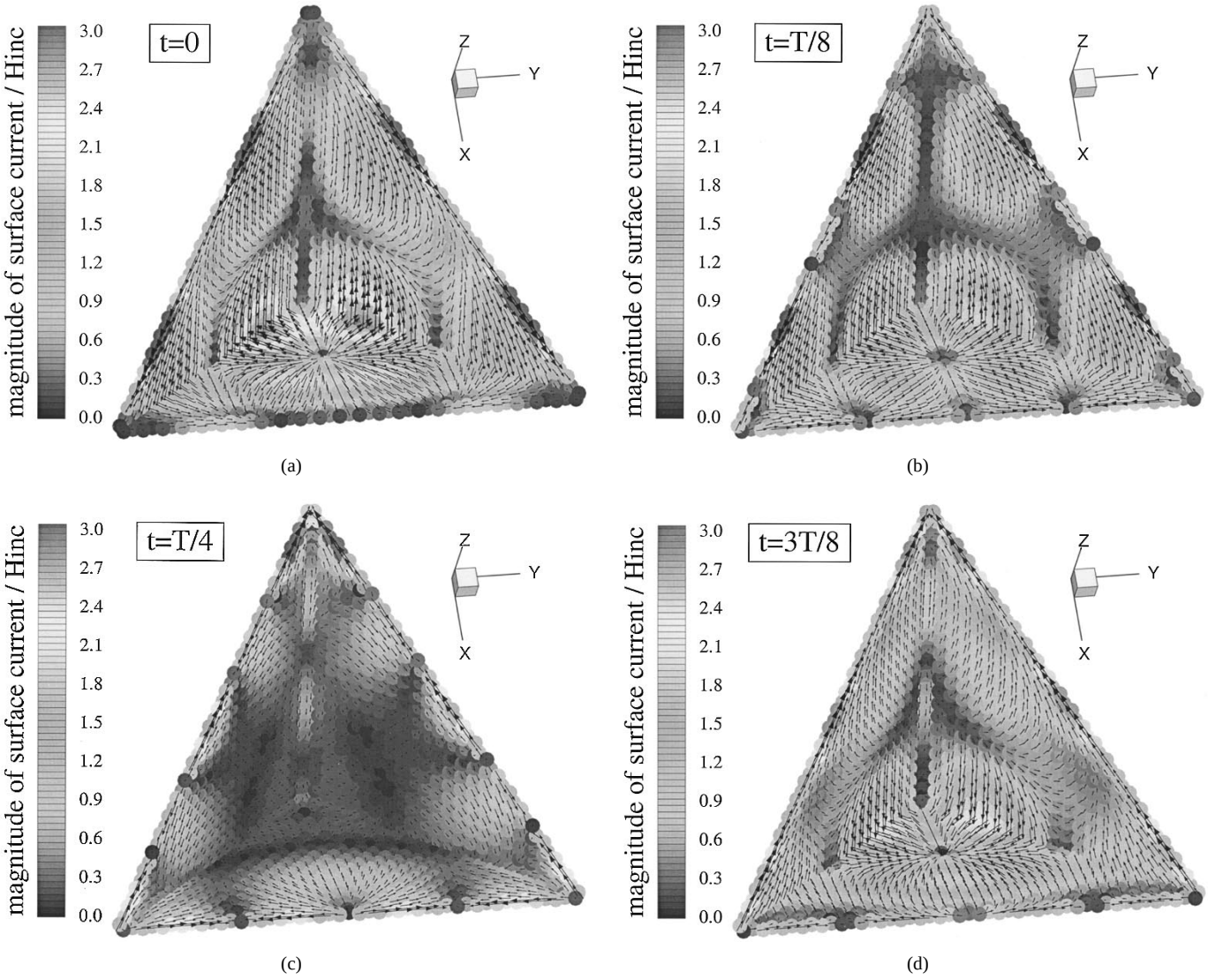


Fig. 3. Normalized currents induced on a perfectly conducting trihedral with a 2.44-m hypotenuse, situated $h = 15$ cm above the soil in Fig. 2, with 10% water content. The trihedral is tilted upward $\Theta_{\text{tilt}} = 35^\circ$ and the vertically polarized plane wave is incident at $\phi = 0^\circ$ and $\Theta = 20^\circ$. The time T is the period of the 300-MHz wave: (a) $t = 0$, (b) $t = T/8$, (c) $t = T/4$, and (d) $t = 3T/8$.

with

$$k^2 = k_\rho^2 + k_z^2, \quad R = \sqrt{(x-x')^2 + (y-y')^2 + h^2(z,\, z')}. \tag{3b}$$

To efficiently evaluate the integral in (2) for a source and observer in layer i, we form the approximation

$$\hat{f}(k_\rho, z, z') = \frac{e^{-jk_{zi}g(z,z')}}{j2k_{zi}}\,\hat{h}(k_\rho) \approx \frac{e^{-jk_{zi}g(z,z')}}{j2k_{zi}}\left[\lim_{k_\rho\to\infty}\hat{h}(k_\rho) + \sum_{m=1}^{M} a_m e^{-k_{zi}b_m}\right] \tag{4}$$

where the coefficients a_m and b_m are estimated along an appropriate line (or lines) in the complex k_{zi} plane [11]–[13], using a technique such as Prony's method [14] or the matrix-pencil [15] method. This procedure is applied separately for each component of the Green's function dyadic. The sum in (4) corresponds to the aggregate response from a collection of sources in complex space and therefore, this procedure is referred to as the method of complex images [11]–[13].

By utilizing (4) in (2) and making use of the identity in (3a), (2) is represented in terms of a sum of components, each of which is integrated in closed form. We have therefore transferred the burden from the numerical evaluation of a highly oscillatory integral to the far simpler task of applying parametric estimation along an appropriate path in the k_{zi} plane [11]–[13]. Through the use of this procedure, it is possible to efficiently characterize wide-band scattering from perfectly conducting targets embedded in a general layered medium. We have augmented this procedure for consideration of dielectric targets as well [16], [17], although such are not addressed here.

III. Wide-band Trihedral Scattering at VHF Frequencies

We consider a trihedral composed of three 90°–45°–45° triangles with hypotenuse equal to 2.44 m. This fiducial target is characteristic of those commonly used for calibration of

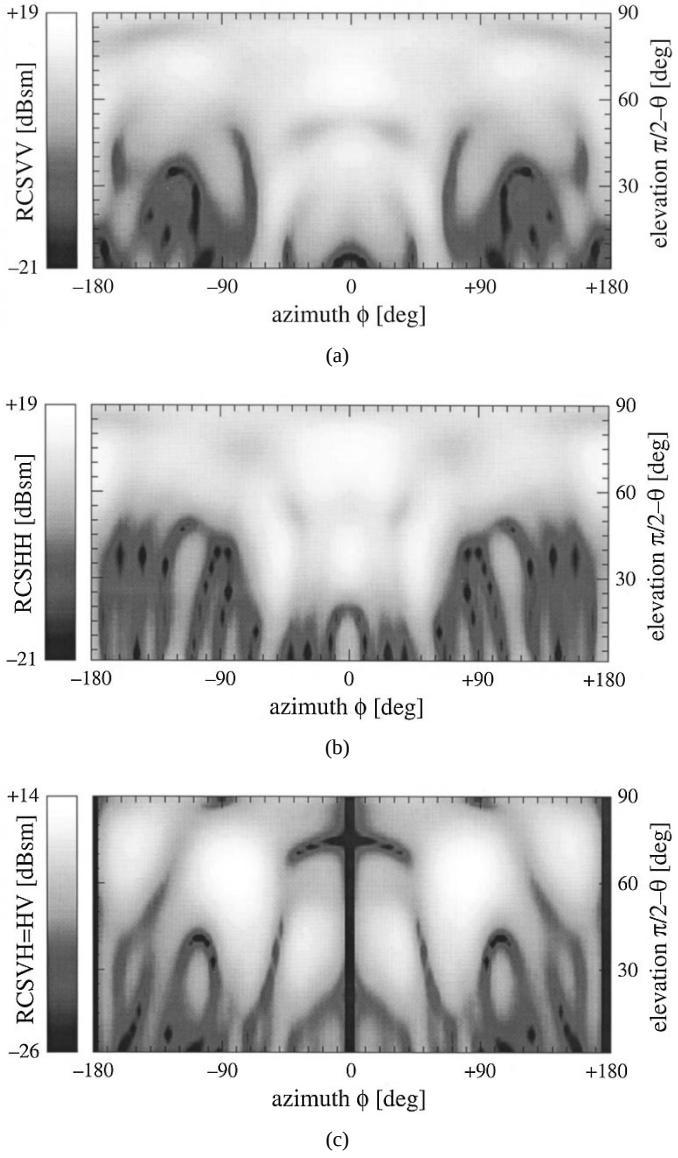


Fig. 4. Backscattered radar cross section (RCS) as a function of ϕ and Θ for a perfectly conducting trihedral with a 2.44-m hypotenuse located in free space. The trihedral is tilted upward $\Theta_{\rm tilt} = 35^\circ$ and results are shown for 300-MHz excitation: (a) VV, (b) HH, and (c) HV.

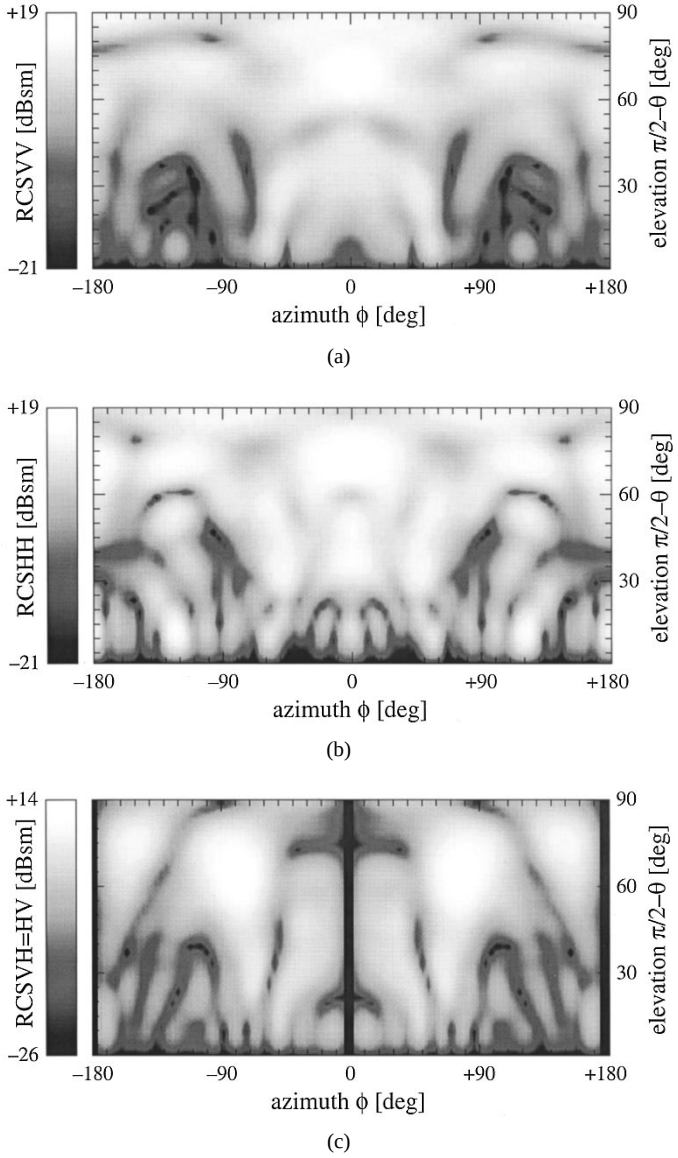


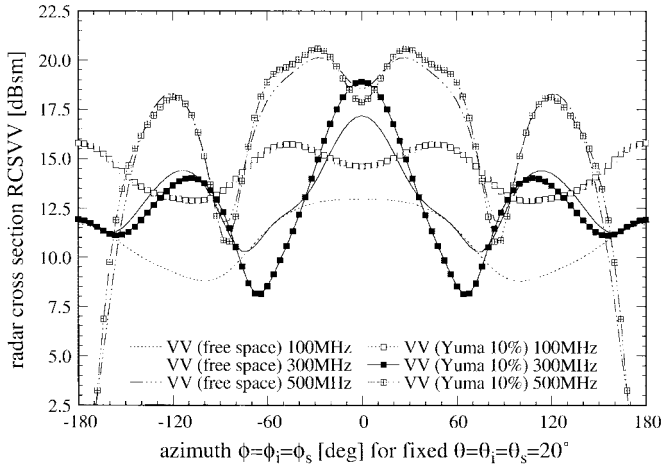
Fig. 5. As in Fig. 4, but the trihedral is now $h = 15$ cm above the soil in Fig. 2 (10% water content).

many SAR systems [1]–[6]. We assume that at the frequencies of interest, the conductor can be simulated as a PEC of infinitesimal thickness. The target is situated over a lossy, dispersive half space (soil), as depicted in Fig. 1. This represents a special case for our MoM algorithm, which is applicable to general layered soils. We have performed SAR measurements [8] at the Yuma Proving Ground, Yuma, AZ (for which we have extensive soil samples) as a function of soil water content. By adjusting the water content, the electrical properties of the soil can be altered significantly [8], [16], [17]. Consequently, we consider the effects of soil moisture content (i.e., soil electrical properties) on the RCS of the trihedral scatterer. The measured real and imaginary parts of the dielectric constant are plotted in Fig. 2 for Yuma soil as a function of water content.

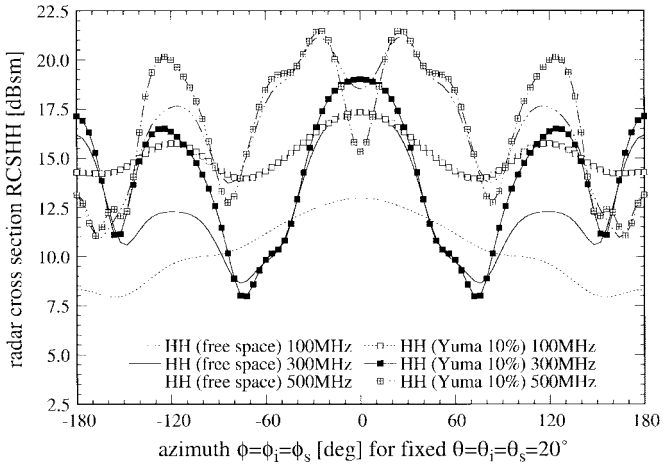
Before proceeding to the results, we note that to our knowledge, no previous rigorous modeling of VHF scattering from trihedrals over soil exists in the literature. While VHF SAR imagery is available [7], [8], that data is uncalibrated due to the lack of accurate modeling of fiducial targets at these frequencies (this of course being what motivates the work presented here). Consequently, to verify the code's accuracy, we have performed tests against results presented in the literature for buried PEC bodies of revolution (BOR's) [18]. Space limitations prohibit showing the comparisons here, but the agreement between our computations and results in [18] is excellent. Our confidence is buttressed by the fact that the BOR results in [18] used a MoM code tailored explicitly to buried BOR's, while here we have employed a code for general PEC targets. While both algorithms employ the half-space Green's function, the details of the algorithm in [18] and that reported here are quite different.

A. Induced Currents

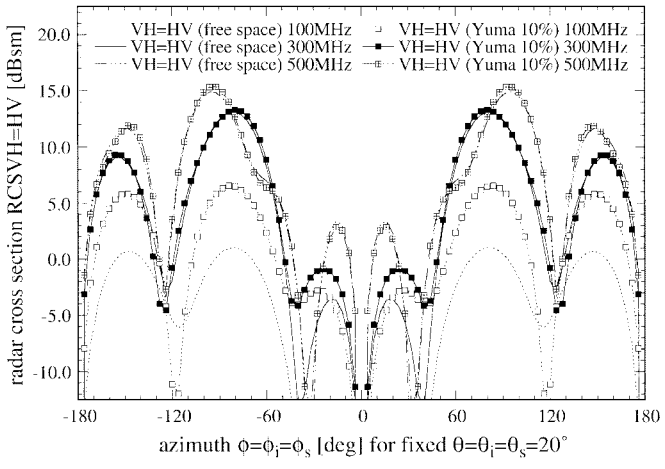
We first consider the currents induced on a trihedral due to linearly polarized plane-wave excitation at 300 MHz. The



(a)



(b)



(c)

Fig. 6. Backscattered RCS as a function of ϕ for $\Theta = 20^\circ$ at frequencies 100, 300, and 500 MHz. Results are shown for the target in free space, and $h = 15$ cm above the soil in Fig. 2 (10% water content). In both cases, the trihedral (2.44-m hypotenuse) is tilted upward $\Theta_{\rm tilt} = 35^\circ$: (a) VV, (b) HH, and (c) HV.

target is placed 15 cm above the 10% Yuma soil, tilted upward 35°, and the vertically (V) polarized radiation is incident at $\phi = 0^\circ$ and $\theta = 20^\circ$ (see Fig. 1). In Fig. 3, the normalized induced currents are plotted at times $t = 0$, $t = T/8$, $t = T/4$, and $t = 3T/8$, where $T = 3.33$ ns is the wave period at 300 MHz. In these plots, the direction of the arrow represents the direction of current flow, with the strength indicated by the arrow length. It is important to emphasize that these are steady-state current distributions at a single frequency. The normalized current magnitude is depicted by the gray-scale. In addition to presenting the intricate current distribution on the target, these results demonstrate that the wavelength (1 m in air) is comparable to the target dimensions, necessitating a rigorous MoM analysis rather than a more typical high-frequency analysis [4]–[6]. Although not shown here in order to conserve space, similar results have been computed for horizontally (H) polarized excitation. As one would expect, the induced currents for this case are markedly different than the V-polarization results in Fig. 3. In particular, for H-polarization excitation, the induced currents flow primarily horizontal to the air–ground interface.

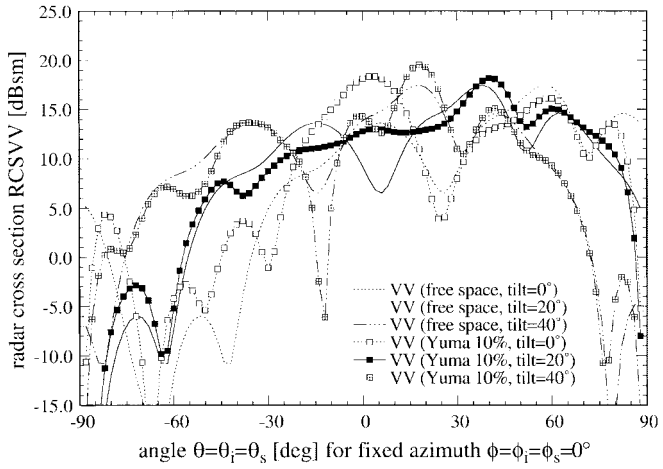
We have considered the induced currents on the same trihedral when situated in free space. The differences between the above-ground and free-space surface currents are very subtle, and therefore, the distinctions in these two cases are addressed quantitatively through consideration of the target radar cross section (RCS).

B. Frequency and Angle Variation

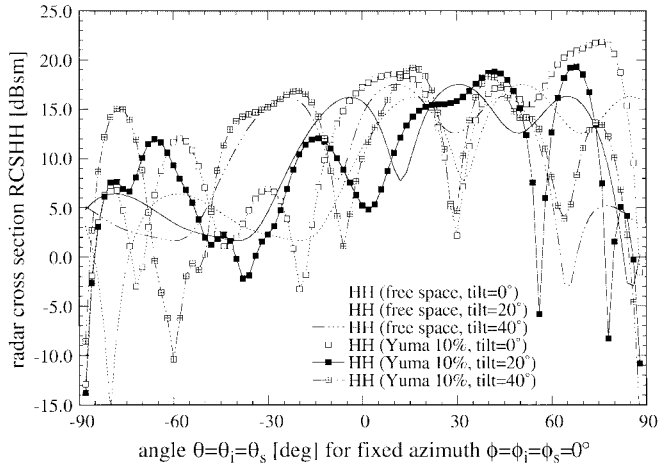
We first plot RCS results for the trihedral in free space tilted 35° upward, considering backscattering at all ϕ and θ for 300-MHz excitation. The VV, HH, and HV backscattered RCS's are presented in Fig. 4(a)–(c) [due to reciprocity, the VH fields are of course equal to the HV results in Fig. 4(c)]. For SAR-calibration purposes, azimuthal angles of $-45^\circ \leq \phi \leq 45^\circ$ are of most relevance. However, in the MoM computations, all ϕ and θ can be considered at modest additional computational cost [9], [10], so these are presented here for completeness. As expected, due to symmetry, the cross-polarized fields vanish at $\phi = 0^\circ$ and $\phi = \pm 180^\circ$.

The influence of the soil on the trihedral RCS is of critical importance for SAR calibration. We address this issue by considering in Fig. 5(a)–(c) the same scenario as discussed above, but with the trihedral placed 15 cm above Yuma soil with 10% water content (see Fig. 2), again tilted upward 35°. For all angles of incidence, the fields penetrate the soil better for V polarization than for H polarization [19]. Consequently, one would anticipate the effects of the soil to be minimum for VV polarization, with the soil affecting the HH and cross-polarized backscattered fields more prominently. This expectation is realized in Figs. 4 and 5, which demonstrate that the RCS is least affected by the soil for VV polarization. However, the gray scale plots are difficult to read for assessment of the subtle differences between the free-space and above-ground cases. We will attempt to make such variations more quantitative.

In particular, in Fig. 6 we plot the RCS as a function of the azimuthal angle ϕ for a fixed incidence angle $\Theta = 20^\circ$ (the Brewster angle for the 10% Yuma soil is approximately 68°). Results are shown for frequencies 100, 300, and 500 MHz for the case of a trihedral placed 15 cm above the soil with a 35°

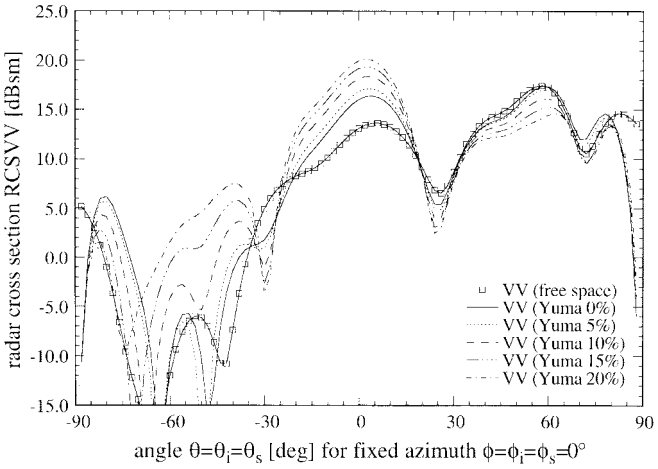


(a)

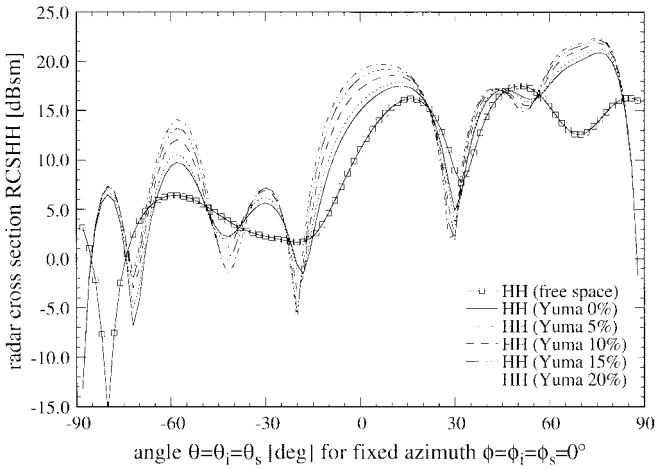


(b)

Fig. 7. Backscattered RCS of a trihedral with a 2.44-m hypotenuse as a function of Θ for $\phi = 0°$. Results are shown for the target in free space, and $h = 15$ cm above the soil in Fig. 2 (10% water content), considering the trihedral tilted upward $\Theta_{\rm tilt} = 0°$, $20°$, and $40°$: (a) VV and (b) HH.



(a)



(b)

Fig. 8. Backscattered RCS of a trihedral with a 2.44-m hypotenuse as a function of Θ for $\phi = 0°$. The trihedral is placed $h = 15$ cm above the soil in Fig. 2 and has a $\Theta_{\rm tilt} = 0°$ tilt. Results are shown for 0%, 5%, 10%, 15%, and 20% water content, as well as for the target in free space: (a) VV and (b) HH.

upward tilt. Considering first the VV results in Fig. 6(a), we see a small difference in the trihedral's above-ground and free-space response, with this difference becoming more prominent at lower frequencies. However, for the HH and HV cases [Fig. 6(b) and (c)], the difference between the free-space and above-ground cases is more significant, consistent with the results in Figs. 4 and 5. We attribute these differences to the aforementioned dichotomy in the soil-penetration properties of vertically and horizontally polarized plane waves. Moreover, the quasioptic picture of trihedral scattering is less appropriate with lowered radar frequency, thus explaining the increased importance of soil interactions as the frequency is lowered.

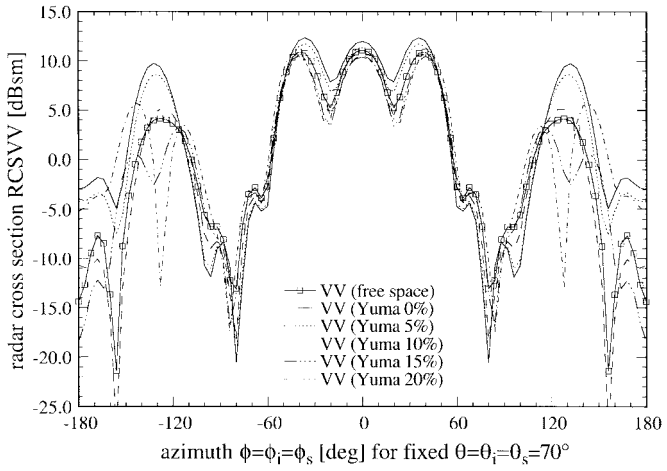
C. Trihedral Tilt Angle

In the above results, we have addressed the soil effects for a fixed trihedral tilt. We now investigate the importance of the trihedral tilt angle (Fig. 1) at 300 MHz for the case of free-space scattering when the target is 15 cm above 10% Yuma soil. In Fig. 7, we consider tilts of 0°, 20°, and 40° for a fixed azimuthal position $\phi = 0°$. Investigating first the VV-polarization case [Fig. 7(a)], we see that for $25° \leq \Theta \leq 75°$, the agreement between the free-space and above-ground results are generally close for all trihedral tilt angles considered. However, the match between the above-ground and free-space cases is best for the largest upward tilt (40°). It is interesting to note that for incidence angles near the Brewster angle (approximately 68°), the agreement between the above-ground and free-space cases is almost exact for all trihedral tilt angles.

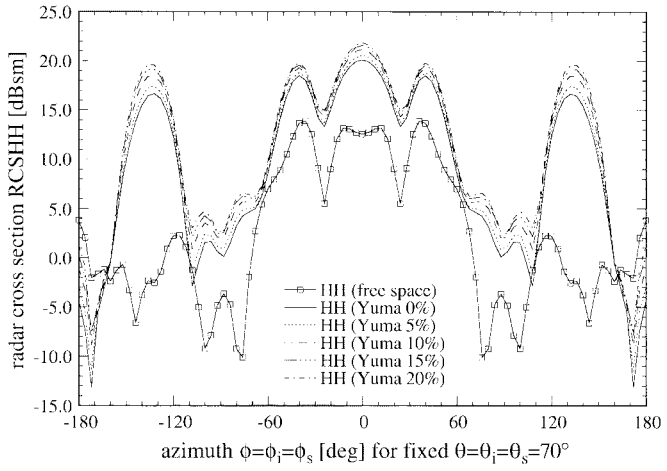
Considering now the HH case [Fig. 7(b)], we notice much more deviation between the above-ground and free-space cases for all tilt angles. The closest agreement appears to hold for the trihedral with the largest upward tilt (40°) in the vicinity of $50° \leq \Theta \leq 62°$. No cross-polarization results are presented for this case, since these vanish at $\phi = 0°$ for all Θ.

D. Soil Variation

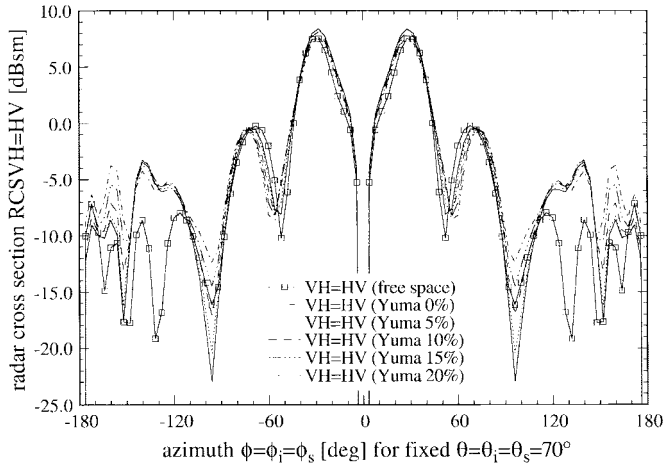
The above results underscore the importance of the soil interface on the RCS of a trihedral over ground. In many



(a)



(b)



(c)

Fig. 9. Backscattered RCS of a trihedral with a 2.44-m hypotenuse as a function of ϕ for $\Theta = 70°$. The trihedral is placed $h = 15$ cm above the soil in Fig. 2, and has $\Theta_{\rm tilt} = 0°$. Results are shown for 0%, 5%, 10%, 15%, and 20% water content, as well as for the target in free space: (a) VV, (b) HH, and (c) HV.

cases, however, the soil electrical properties are not known exactly, due to inhomogeneities in the soil type and moisture content that may change with time. We address this issue by considering the trihedral 15 cm above Yuma soil with no tilt, 0%, 5%, 10%, 15%, and 20% water content (Fig. 2), and 300-MHz radar operation. We consider the backscattered RCS at $\phi = 0°$. Considering the 0, 5, 10, 15, and 20% water content soil, the associated Brewster angles are approximately 59.1°, 62.3°, 68.8°, 74.0°, and 77.7°, respectively. Addressing first the VV response in Fig. 8(a), we see that the RCS of the above-surface and free-space trihedrals are well matched in the regions expected from the corresponding Brewster angles. However, for a given moisture content, at angles away from the corresponding Brewster angle there is often a significant difference between the two results. Considering now the HH results in Fig. 8(b), we see that the distinction between the above-ground and free-space response is significant for most angles Θ. Interestingly, the angles for which the agreement between the free-space and above-ground VV results are closest (approximately $60° \leq \Theta \leq 70°$) have significant differences in the corresponding HH fields .

A trihedral above ground is a relatively complex target, and the scattered fields from such cannot be explained entirely in terms of relatively simple Brewster angle-related phenomena. For example, in considering the VV RCS in Fig. 8(a), note that the above-ground and free-space curves are in close agreement in the vicinity of $\Theta = 30°$, with similar effects manifested at other angles for HH polarization [Fig. 8(b)].

To examine related effects as a function of the azimuthal position, we fix $\Theta = 70°$ and consider variation of the angle ϕ. From Fig. 8, $\Theta = 70°$ is an angle at which the VV fields scattered from the trihedral above ground are relatively close to those of the free-space trihedral for all soil types considered. Moreover, at this angle, the HH fields for the above-ground and free-space trihedral are markedly different. While Fig. 8 only considered the azimuthal position $\phi = 0°$, from Fig. 9(a) we see that the relative sensitivity of the VV fields to the moisture content is relatively independent of ϕ over the principal region $-45° \leq \phi \leq 45°$, with the same holding true for HH polarization [Fig. 9(b)]. Considering the cross-polarized fields, we see the HV fields vanish at $\phi = 0°$, due to symmetry in the trihedral. Moreover, in the region $-45° \leq \phi \leq 45°$, we also note that there is substantial variation between the cross-polarized fields for the free-space and above-ground cases.

E. Time Domain

The above results have focused exclusively on frequency domain operation. However, many SAR systems have wide-band or even time domain operation [7], [8]. We consider time domain scattering from a trihedral placed 15 cm above 10% Yuma soil, tilted upward 35°. The incident pulse is shown in Fig. 10, and it has an associated spectrum that extends to approximately 600 MHz. The time domain scattered fields are viewed in the far zone, scaled by the distance r between the sensor and the target center. Moreover, the waveforms are shifted by the time r/c, where c is the speed of light in a vacuum. It has been demonstrated that if there is sufficient resolution, the fields scattered from a trihedral are characterized by edge diffraction in addition to the specular return from the plate surfaces of the trihedral [4]. Considering the VV results in Fig. 10(a), from four different observation

directions we see that there are smaller wavefronts centered at about $t = 0$ that arrive before and after the principal return. It is clear from these results that the edge-diffracted fields are often characterized by significant amplitudes. Consequently, the time domain trihedral response is not simply a time-shifted replica of the incident pulse. Similar effects are realized for the HH and cross-polarized fields, with the latter vanishing at $\phi = 0°$ due to symmetry.

When performing SAR processing, the subsequent target image is realized by viewing the target from multiple angles [7], [8]. With regard to a wide-band time domain system, the results in Fig. 10 demonstrate that the fields scattered from a trihedral fiducial target are a strong function of the target-sensor orientation, with this angular diversity implicitly incorporated in the subsequent imagery.

IV. Conclusions

A comprehensive study has been performed on wide-band VHF and low-UHF scattering from trihedral fiducial targets placed above soil, with application to SAR calibration. The principal purpose of this study has been to conduct a detailed analysis of the soil effects on the trihedral response as a function of frequency, polarization, trihedral tilt, and soil properties. This has been done in an effort to determine if the free-space trihedral response can be utilized for calibration at the frequencies of interest. While it is recognized that the relatively small data set presented here may be of limited value for calibration of a particular sensor, it is hoped that the insight will have general applicability. Moreover, since the results indicate that the soil effects must be accounted for in many cases of interest, the results presented here can be used to check the accuracy of a numerical model applied for such purposes.

Summarizing the conclusions from this work, the effects of the soil interface become more important as the frequency is lowered, which is of particular interest for VHF applications. Moreover, the soil effects are generally much more prominent for HH and HV scattering, relative to the VV case. With regard to the latter, near the Brewster angle we found the VV scattered fields above soil to be nearly identical to those of the target in free space. Concerning the HH fields, the discrepancy between the above-ground and free-space cases was reduced as the upward tilt of the trihedral was increased. The electrical properties of the soil generally were found to play a prominent role in the trihedral RCS for all polarizations. In particular, variation in the soil properties was found in some cases, which caused several dB variations in the target RCS. Finally, we considered time domain scattering from the trihedral above ground, using frequencies in the VHF and low UHF. For the large (2.44 m) trihedral considered here, the excitation had sufficient resolution to discern edge-diffracted scattering from the principal trihedral specular return. These results demonstrate that the transient response from a trihedral scatterer often is not simply a delayed replica of the incident pulse.

Future research will involve the use of our rigorous trihedral scattering model in the calibration of VHF systems [7], [8] with applications to FOPEN and GPEN SAR. It also will take into consideration algorithms such as those reported in [1]–[3]. After the polarimetric VHF SAR imagery is properly calibrated, high fidelity automatic target detection algorithms can be brought to bear for the detection of concealed targets.

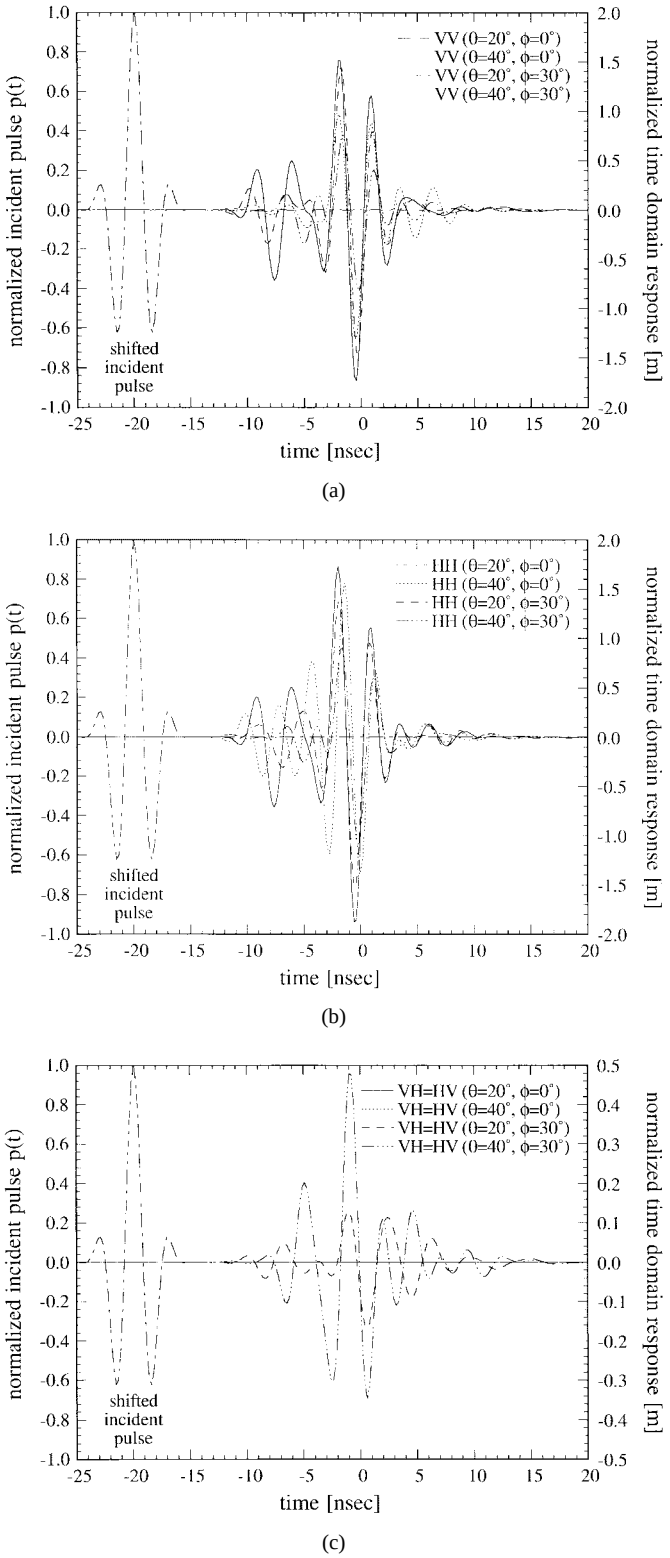


Fig. 10. Backscattered transient fields due to pulsed plane-wave excitation with incident pulse presented as well. The trihedral target has a 2.44-m hypotenuse, is $h = 15$ cm above the soil in Fig. 2 (10% water content), and is tilted upward $\Theta_{\rm tilt} = 35°$. The scattered waveform is scaled and time shifted (see text, Section III-E): (a) VV, (b) HH, and (c) HV.

REFERENCES

[1] J. J. van Zyl, "Calibration of polarimetric radar images using only image parameters and trihedral corner reflector responses," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-28, pp. 337–348, May 1990.

[2] K. Sarabandi, L. E. Pierce, and F. T. Ulaby, "Calibration of a polarimetric imaging SAR," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-30, pp. 540–549, May 1992.

[3] D. R. Sheen, A. Freeman, and E. S. Kasischke, "Phase calibration of polarimetric SAR images," *IEEE Trans. Geosci. Remote Sensing*, vol. 27, pp. 719–731, Nov. 1989.

[4] T. Griesser and C. A. Balanis, "Backscatter analysis of dihedral corner reflectors using physical optics and the physical theory of diffraction," *IEEE Trans. Antennas Propagat.*, vol. AP-35, pp. 1137–1147, Oct. 1987.

[5] G. T. Ruck, D. E. Barrick, W. D. Stuart, and C. K. Krichbaum, *Radar Cross Section Handbook*, vol. 1. New York: Plenum, 1970, pp. 588–597.

[6] K. Sarabandi and T.-C. Chiu, "Optimum corner reflectors for calibration of imaging radars," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1348–1361, Oct. 1996.

[7] M. A. Ressler and J. W. McCorkle, "Evolution of the Army Research Laboratory ultra-wideband test bed," in *Ultra-Wideband Short-Pulse Electromagnetics 2*, L. Carin and L. B. Felsen, Eds. New York: Plenum, 1995, pp. 109–123.

[8] S. Vitebskiy, L. Carin, M. A. Ressler, and F. H. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.

[9] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, Parts I and II," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–352, Mar. 1990.

[10] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering from surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 409–418, May 1982.

[11] Y. L. Chow, J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 588–592, Mar. 1991.

[12] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.

[13] M. I. Aksun, "A robust approach for the derivation of closed-form Green's functions," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 651–658, May 1996.

[14] M. L. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. Antennas Propagat.*, vol. AP-23, pp. 777–781, Nov. 1975.

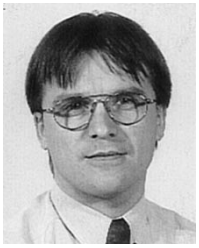
[15] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 814–824, May 1990.

[16] N. Geng and L. Carin, "Ultrawideband, short-pulse scattering from a dielectric body of revolution buried in a lossy, dispersive layered medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Mar. 1999.

[17] N. Geng, D. Jackson, and L. Carin, "On the resonances of dielectric bodies of revolution buried in a lossy, dispersive layered medium," *IEEE Trans. Antennas Propagat.*, vol. 47, Aug. 1999.

[18] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse plane-wave scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas and Propagat.*, vol. 44, pp. 143–151, Feb. 1996.

[19] J. A. Stratton, *Electromagnetic Theory*. New York: McGraw-Hill, 1941, p. 510.



Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Germany, in 1991 and 1996, respectively.

From 1991 to 1996, he was with the Institute for Microwaves and Electronics, University of Karlsruhe, working on full-wave propagation modeling for radio communication systems. In January 1997, he joined the Department of Electrical and Computer Engineering at Duke University, Durham, NC, as a Visiting postdoctoral for 18 months. Since July 1998, he has been back with the University of Karlsruhe. His current research interests include computational methods in electromagnetics and wave propagation modeling.

Dr. Geng received the Mannesmann Innovation Award in 1997 for his Ph.D. thesis on full-wave propagation modeling for radio communication systems.

Marc A. Ressler received the B.S. in electrical engineering from the University of Maryland, College Park, in 1969 and the M.S. degree in electrical and computer engineering from the University of Michigan, Ann Arbor, in 1973.

He has been with the Army Research Laboratory (ARL), Adelphi, MD and one of its predecessors, the Harry Diamond Laboratories, since 1969, where he worked on electronic fuzes for rockets, laser simulators, and instrumentation systems and was a member of the ANSI Y14.26 subcommittee that developed standards for computer-aided preparation of product definition data. He later worked on adaptive antennas for communications, signal processor design and multitarget tracking software for radar, and a microprocessor-based metering system for the electric power industry (for which a patent was awarded). More recently, he has been involved in battlefield data processing and fusion, and a modular microwave radar system that had applications to both airborne and ground-based systems. He is currently assigned to the Sensors and Electron Devices Directorate, where his areas of interest are low-frequency, ultrawideband radar for foliage and ground penetration applications.

Lawrence Carin (SM'96) [redacted] in Washington, DC, and received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department at Polytechnic University, Brooklyn, NY, as an Assistant Professor and became an Associate Professor in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, where he is an Associate Professor. His current research interests include short-pulse scattering, propagation, and signal processing. He is the Principal Investigator on an Army Research Office-funded Multidisciplinary University Research Initiative (MURI) on demining.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.

On the Resonances of a Dielectric BOR Buried in a Dispersive Layered Medium

Norbert Geng, *Member, IEEE*, David R. Jackson, *Fellow, IEEE*, and Lawrence Carin, *Senior Member, IEEE*

***Abstract*— A method-of-moments (MoM) analysis is applied to the problem of determining late-time resonances of dielectric bodies of revolution buried in a lossy layered medium, with application to plastic-land-mine identification. To make such an analysis tractable, we have employed the method of complex images to evaluate the layered-medium Green's function. The application of this method to resonant structures characterized by complex resonant frequencies, introduces numerical issues not manifested at real frequencies (i.e., for driven problems) with such discussed here in detail. Numerical results are presented for several buried targets in which we demonstrate, for example, the spiraling character of the resonant frequencies of particular targets as a function of target depth.**

***Index Terms*—Buried object detection, natural resonances, inhomogeneous media.**

I. Introduction

If a target is excited by a short pulse of electromagnetic energy, fields are diffracted initially at localized scattering centers in and on the target. Subsequently, after the incident wave has departed (at what is termed "late time"), multiple diffractions occur *between* scattering centers and, for a penetrable target, energy reverberates inside as well. Each such multiple diffraction or reverberation is accompanied by energy that radiates away from the scatterer. Consequently, the late-time energy in and on the target decays, as do the associated late-time radiated fields. This late-time phenomenon has been parametrized rigorously in terms of the well-known singularity expansion method (SEM) developed by Baum [1], in which each resonant mode is associated with a pole in the complex frequency plane. Additionally, Felsen *et al.* [2], [3] have rigorously connected the SEM resonances to multiple late-time diffractions and reverberations.

The SEM resonant frequencies, representative of the natural target response, are independent of excitation (i.e., are aspect independent [1]–[5]), although the strength of excitation of a given mode is excitation dependent. Aspect independence of such resonant frequencies has precipitated significant interest in resonant-frequency-based target *identification*. Resonance-based discrimination has been investigated for airborne [6], [7] as well as subsurface [8], [9] targets, using such techniques as Prony's method [10], [11] and the matrix-pencil method [12]. In the work presented here, we are interested in dielectric targets (e.g., plastic land mines) buried in a lossy, dispersive layered medium (simulating soil). There have been previous experimental studies on resonance-based identification of plastic targets, with some results being particularly promising [8], [9]. However, it is well known that the strength of excitation of such modes is strongly dependent on the electrical contrast between the dielectric target and the background soil. Moreover, a recent study [13] on the SEM resonances of buried *conducting* targets has revealed an often strong dependence of the resonant frequencies on the target depth. This is because, for subsurface targets, the resonant frequencies are not dictated by the buried target alone, but by the *total* scattering environment (i.e., the target in the presence of the soil). In particular, if there are strong reverberations between the target and air–ground interface, the target depth plays an important role in the total-target resonant frequencies. This is a very important issue, for the target depth is often not known exactly (e.g., for buried mines), complicating resonance-based identification of such targets. Since it is difficult to comprehensively study this issue experimentally, for a wide range of targets, soil types, and target depths, an accurate theoretical model is required, motivating the work presented here.

Modeling of a target's natural modes requires solution of the source-free target response. Therefore, algorithms which require an excitation such as the finite-difference time-domain (FDTD) [14], are inappropriate. While the target resonant frequencies can in principle be extracted from the late-time fields computed via such a time-domain model (via, for example, the Prony [10], [11] or matrix-pencil [12] algorithms discussed above), modal excitation is dictated by the driving function and, therefore, a range of incident fields would be required to catalogue all the modes of a given structure. Consequently, the natural target response is best analyzed with a frequency-domain algorithm. The finite-element method (FEM) [15] constitutes a frequency domain scheme that can be adapted for such purposes. However, while FEM algorithms are quite general and accurate, they require discretization of the fields throughout the computational domain. For the three-dimensional problems of interest here, in complicated layered environments, the generality of FEM is obviated by the attendant numerical complexity [15]. Therefore, here we utilize the method of moments (MoM), with restriction to targets that can be simulated as a body of revolution (BOR); for the layered-medium problem of interest here, the BOR axis is required to be normal to the layer surfaces (Fig. 1), thereby

Manuscript received February 10, 1998; revised June 16, 1999.

N. Geng and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

D. R. Jackson is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4793 USA.

Publisher Item Identifier S 0018-926X(99)07959-4.

0018–926X/99$10.00 © 1999 IEEE

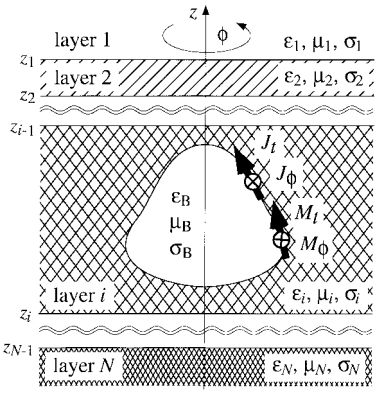


Fig. 1. Schematic of a dielectric body of revolution buried in a layered medium.

preserving rotational symmetry throughout the BOR (layered-medium) composite. While the restriction to BOR targets is clearly a simplification, most land mines are accurately so modeled [16].

The MoM analysis of targets buried in a lossy layered medium requires accurate computation of the layered-medium Green's function with such necessitating the efficient analysis of Sommerfeld spectral integrals [17]–[21]. Numerous techniques have been developed for the analysis of these ubiquitous integrals [17]–[21], with the recently developed method of complex images constituting a particularly attractive option [22]–[25]. This algorithm represents the spectral-domain reflection coefficient inherent to the spectral Green's function in terms of a sum of exponentials, determined via parametric estimation (again, via either of the aforementioned parametric algorithms [10]–[12]). Subsequently, each spectral-domain term in this summation is converted to the space domain analytically, via the Sommerfeld or Weyl identity [26]. The parametric estimation is performed along a proper path in the complex spectral plane with careful consideration of branch cuts and possible surface-wave and leaky-wave poles [27]. While such matters have been addressed in detail for the driven problem, characterized by *real* frequencies, special consideration is required for resonant-frequency computation. In particular, target resonances are characterized by *complex* frequencies, introducing complications with regard to branch-cut and pole locations, with such impacting the path of integration for the Sommerfeld integrals. We address these issues in detail and demonstrate how such are handled within the context of the method of complex images.

The remainder of the text is organized as follows. In Section II we summarize the general features of MoM-based analysis of SEM resonances for dielectric BOR's buried in a lossy, layered medium (e.g., soil). For details on the layered-medium Green's function, recent work on such is cited. However, the application of the method of complex images at resonant (complex) frequencies is new, and is therefore discussed in detail. In Section III, several numerical results are presented as a function of the properties of the target, target depth, and background environment. Finally, conclusions are addressed in Section IV.

II. Numerical Algorithm

A. Method of Moments Formulation

We are interested in the natural (resonant) modes of a dielectric BOR embedded in a lossy, dispersive layered medium (see Fig. 1). By enforcing boundary conditions for the tangential electric and magnetic field components on the BOR surface, one obtains coupled integral equations for the resonant electric and magnetic surface currents $\boldsymbol{J}$ and $\boldsymbol{M}$, respectively. We employ a mixed-potential integral equation (MPIE) formulation from which we have (for *complex* $\omega = \omega' + j\omega''$)

$$\begin{aligned} 0 =& \boldsymbol{n} \times \int_S [j\omega\mu(\boldsymbol{r})\boldsymbol{K_A}(\boldsymbol{r},\boldsymbol{r}') \\ &+ j\omega\mu_B G_B(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{I}] \cdot \boldsymbol{J}(\boldsymbol{r}')\, dS' - \boldsymbol{n} \times \nabla \int_S \\ &\cdot \left[\frac{K_{\phi e}(\boldsymbol{r},\boldsymbol{r}')}{\sigma(\boldsymbol{r}) + j\omega\epsilon(\boldsymbol{r})} + \frac{G_B(\boldsymbol{r},\boldsymbol{r}')}{\sigma_B + j\omega\epsilon_B}\right] \nabla' \cdot \boldsymbol{J}(\boldsymbol{r}')\, dS' \\ &+ \boldsymbol{n} \times \nabla \times \int_S [\boldsymbol{G_F}(\boldsymbol{r},\boldsymbol{r}') + G_B(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{I}] \cdot \boldsymbol{M}(\boldsymbol{r}')\, dS' \end{aligned} \tag{1a}$$

$$\begin{aligned} 0 =& \boldsymbol{n} \times \int_S [(\sigma(\boldsymbol{r}) + j\omega\epsilon(\boldsymbol{r}))\boldsymbol{K_F}(\boldsymbol{r},\boldsymbol{r}') \\ &+ (\sigma_B + j\omega\epsilon_B) G_B(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{I}] \cdot \boldsymbol{M}(\boldsymbol{r}')\, dS' \\ &- \boldsymbol{n} \times \nabla \int_S \left[\frac{K_{\phi m}(\boldsymbol{r},\boldsymbol{r}')}{j\omega\mu(\boldsymbol{r})} + \frac{G_B(\boldsymbol{r},\boldsymbol{r}')}{j\omega\mu_B}\right] \nabla' \cdot \boldsymbol{M}(\boldsymbol{r}')\, dS' \\ &- \boldsymbol{n} \times \nabla \times \int_S [\boldsymbol{G_A}(\boldsymbol{r},\boldsymbol{r}') + G_B(\boldsymbol{r},\boldsymbol{r}')\boldsymbol{I}] \cdot \boldsymbol{J}(\boldsymbol{r}')\, dS' \end{aligned} \tag{1b}$$

where $\boldsymbol{I}$ represents the unit dyad. The layered-medium permittivity, permeability, and conductivity are represented by $\epsilon(\boldsymbol{r}), \mu(\boldsymbol{r})$, and $\sigma(\boldsymbol{r})$, respectively, and ϵ_B, μ_B, and σ_B represent these same parameters for the homogeneous, lossy BOR target. Electric and magnetic charge densities on the surface have been replaced by derivatives of the electric and magnetic surface current densities, respectively, using the continuity relation. Explicit expressions for the spectral-domain layered-medium dyadic kernel $\boldsymbol{K}_A$, the dyadic Green's function $\boldsymbol{G}_A$ (representing the magnetic vector potential produced by an infinitesimal electric dipole), and the electric scalar potential $K_{\phi e}$ of a point charge associated with a horizontal electric dipole have been given by Michalski and Zheng [21], where we use their "Formulation C." The additional terms $\boldsymbol{K}_F, \boldsymbol{G}_F$, and $K_{\phi m}$ are necessitated by the equivalent magnetic currents (not present for perfectly conducting targets [21]) and can be determined via duality. Finally, for calculating the field inside the homogeneous BOR, produced by currents on its surface, we utilize the homogeneous-medium Green's function $G_B = \exp(-jk_B R)/4\pi R$, where R represents the distance between source and observation points and k_B denotes the wavenumber inside the BOR.

The integral equations in (1) are applicable to any dielectric target embedded in a layered medium, but here we restrict ourselves to the case of a BOR to make the numerical analysis tractable. However, as stated in the Introduction, most plastic land mines, the interest of this work, closely approximate BOR's [16]. The MoM analysis of BOR's in free-space is well known [28], [29], with the integral equations in (1) similar

to such, with substitution of the layered-medium Green's function. Therefore, the principal challenge here involves the efficient analysis of the Green's function components $\boldsymbol{K}_A, \boldsymbol{G}_A, K_{\phi e}, \boldsymbol{K}_F, \boldsymbol{G}_F$, and $K_{\phi m}$, necessitating the evaluation of several Sommerfeld-type integrals [17]–[25]. In the subsequent two sections, we address the proper path of integration in the complex spectral plane for the evaluation of such integrals, with application to resonant-frequency calculations. Moreover, we discuss how such integrals are computed efficiently, via a modified form of the method of complex images [22]–[25].

After representation of the surface currents $\boldsymbol{J}$ and $\boldsymbol{M}$ along the BOR generating arc [29], we generate homogeneous matrix equations (one for each azimuthal mode [28]–[30]) of the form

$$\boldsymbol{Zi} = \boldsymbol{0} \tag{2}$$

where $\boldsymbol{Z}$ is an $N \times N$ MoM impedance matrix and $\boldsymbol{i}$ is an $N \times 1$ vector representing the basis-function coefficients for the N expansion functions. The impedance matrix is a function of frequency ω and nontrivial solutions for $\boldsymbol{i}$ occur at frequencies for which the determinant of $\boldsymbol{Z}(\omega)$ vanishes, providing a numerical scheme for computation of the complex resonant frequencies. The relative modal currents are computed subsequently [29], [30]. Details on the efficient computation of $\boldsymbol{Z}(\omega)$ for dielectric BOR's embedded in a layered medium can be found in [31].

B. Path of Spectral Integration: Loss-Free Case

The layered-medium Green's function is of interest in a wide range of problems [21], [31] and, therefore, the evaluation of such has constituted an important area of research. As is well known, layered-medium Green's functions can only be evaluated in closed form in the spectral domain with Sommerfeld-type integrals required for determination of their space-domain counterparts. Numerous numerical and analytic techniques [17]–[25] have been devised for the evaluation of such integrals. Here we exploit the method of complex images [22]–[25]. Previous use of the method of complex images has focused primarily on driven problems at *real* frequencies [22]–[25]. For the evaluation of target resonances, of interest here, we seek the natural (nondriven) modes of the structure, which are characterized by *complex* resonant frequencies [30]. We therefore address the evaluation of Sommerfeld integrals at complex frequencies, with subsequent application to the method of complex images. Moreover, to simplify the presentation, the discussion is restricted initially to the case of *lossless* media with a subsequent generalization for inclusion of loss.

We are interested in evaluating integrals of the form

$$f(\xi, z, z') = \frac{1}{2\pi} \int_0^\infty \hat{f}(k_\rho, z, z') J_o(k_\rho \xi) k_\rho \, dk_\rho \tag{3}$$

where $\xi = [(x-x')^2 + (y-y')^2]^{1/2}$. Because our problem is open above and below, branch points exist at $k_\rho = \pm k_1$ and $k_\rho = \pm k_N$ [32], where k_1 and k_N are the wavenumbers for the top (first) and bottom (Nth) layers, respectively. For an assumed $\exp(j\omega t)$ time dependence, a given complex resonant frequency is of the form $\omega = \omega' + j\omega''$, where $\omega' > 0$ and $\omega'' > 0$ (characteristic of a damped oscillation). Therefore, the branch points reside in the first and third quadrants of the complex k_ρ plane (assuming that the wave velocities in regions 1 and N are real at *complex*, resonant frequencies; the restriction to lossless media implies these velocities are real at *real* frequencies). It is convenient to define branch cuts such that the various surface- and leaky-wave poles are clearly partitioned to particular Riemann sheets [32]. In particular, we utilize the well-known branch cuts [32] defined as (for complex $k_\rho = k'_\rho + k''_\rho$ and $k_n = k'_n + jk''_n$)

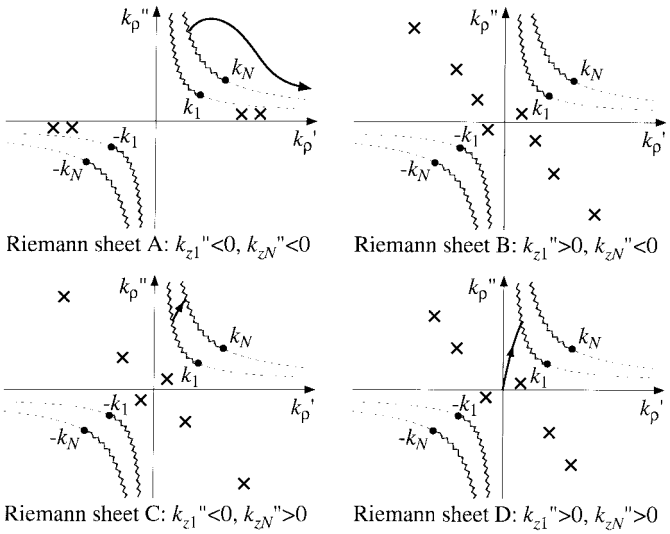


Fig. 2. Four Riemann sheets defined by the choice of branch cuts in (4). The path of integration exists on portions of three sheets, and the crosses denote possible surface- and leaky-wave pole positions. For wavenumbers $k_{z1} = k'_{z1} + jk''_{z1}$ and $k_{zN} = k'_{zN} + jk''_{zN}$ (for the first and Nth layer in Fig. 1, respectively), the Riemann sheets are characterized by $k''_{z1} < 0$ and $k''_{zN} < 0$ on sheet A, $k''_{z1} > 0$ and $k''_{zN} < 0$ on sheet B, $k''_{z1} < 0$ and $k''_{zN} > 0$ on sheet C, and $k''_{z1} > 0$ and $k''_{zN} > 0$ on sheet D.

$$k'_1 k''_1 = k'_\rho k''_\rho, \quad k'^2_\rho - k''^2_\rho < k'^2_1 - k''^2_1 \tag{4a}$$

$$k'_N k''_N = k'_\rho k''_\rho, \quad k'^2_\rho - k''^2_\rho < k'^2_N - k''^2_N. \tag{4b}$$

The branch cuts in (4a) and (4b) are necessitated by $k_{z1} = [k_1^2 - k_\rho^2]^{1/2}$ and $k_{zN} = [k_N^2 - k_\rho^2]^{1/2}$, respectively, both of which appear in the spectral Green's function (we define $k_{z1} = k'_{z1} + jk''_{z1}$ and $k_{zN} = k'_{zN} + jk''_{zN}$). These branch cuts result in the four Riemann sheets shown in Fig. 2, where sheets A through D are characterized, respectively, by $k''_{z1} < 0$ and $k''_{z1} < 0, k''_{z1} > 0$ and $k''_{zN} < 0, k''_{z1} < 0$ and $k''_{zN} > 0$, and $k''_{z1} > 0$ and $k''_{zN} > 0$. Recognizing that the fields radiated away in regions 1 and N are characterized by $\exp(-jk_{z1}z)$ and $\exp(jk_{zN}z)$, respectively, we can attribute physical meaning to the modes (poles) that may exist on each of the four Riemann sheets. (It should be noted that such modes, characterized by poles in the k_ρ plane, are the waveguide modes supported by the layered medium [32], at a given complex resonant frequency, not the modes of the resonant structure itself; the poles of the resonant structure occur in the complex *frequency* plane [1]). On Riemann sheet

A, the fields decay exponentially with increasing z in region 1 and with decreasing z in region N and, therefore, this surface is characterized by a finite number of poles, representative of surface waves. At the complex frequency characteristic of a resonant mode, the fields *grow* exponentially with transverse distance ξ away from the source [30] and, therefore, (for an $\exp(-jk_\rho\xi)$ dependence for large ξ) the surface-wave poles reside in the first and third quadrants of Riemann sheet A.

On Riemann sheet $B, k''_{z1} > 0$ and $k''_{zN} < 0$, and, therefore, the layered-medium modes represented by poles on this sheet are characterized by exponential growth with z in region 1, while decaying exponentially in region N. This phenomenon is characteristic of a leaky (improper) mode [27], [32], with leakage into region 1. One must be careful to assess the position of these leaky-wave poles in the complex k_ρ plane. We begin by considering resonant frequencies $\omega' + j\omega''$ for which $\omega'' \to 0$, characteristic of a very high-Q resonator. This represents the limiting case of nondamped time-harmonic excitation, for which the leaky-wave pole positions are well known [27], [32]. In particular, for the time-harmonic case such modes are characterized by exponential growth in region 1 $(k''_{z1} > 0)$, and exponential decay with distance ξ away from the source $(k''_\rho < 0)$ [27], [32], the decay in ξ manifested by energy lost to leakage. Therefore, under such circumstances, the leaky-wave poles reside in the forth quadrant of this Riemann sheet (the negatives of these leaky-wave poles also exist, in the second quadrant) [27], [32]. There are an infinite number of such leaky-wave poles [27], [32], with the rate of leakage increasing with increasing $|k''_\rho|$. We now consider what happens as ω'' increases, reflecting a lowering of the resonator Q. The energy emitted from the resonant source decays with time, with a corresponding growth in energy with distance ξ from the source [30]. If the rate of leakage is sufficiently small, the spatial growth associated with the radiated resonant energy will overcome the exponential decay in ξ attributed to leakage and, therefore, such weakly leaky modes are characterized by $k''_\rho > 0$ (i.e., they grow with ξ). Therefore, it is possible that on sheet B, a *finite* number of leaky-wave poles may exist in the first (and third) quadrant. However, as mentioned above, for the harmonic case the rate of leakage increases with increasing $|k''_\rho|$ and, therefore, there will be an infinite number of leaky-wave poles for which the rate of leakage overcomes the spatial growth associated with the resonator, leading to leaky-wave fields that decay with ξ (i.e., $k''_\rho < 0$); the associated poles of such modes are situated in the forth (and second) quadrant of the complex k_ρ plane (as for the case of harmonic excitation).

This phenomenon can be understood by making an analogy to a source in the presence of an active medium that supports leaky waves (e.g., a leaky transmission line [33]–[35] loaded with active elements [36]). In the absence of leakage, the fields on the transmission line grow with distance from the source due to the active elements (corresponding to the resonant fields in the layered media, which grow with distance from the resonator). However, due to leakage, the transmission line mode loses energy with distance. If the rate of leakage exceeds the rate of active-element-induced gain, the fields will decay with distance from the source, with the opposite true in the case of weak leakage. If, as for the layered medium, there are an infinite number of leaky waves, with increasing rates of leakage, there will only be a finite number of such modes with small enough leakage for the fields to grow with distance from the source, while there are an infinite number of leaky waves for which the rate of leakage exceeds the rate of gain, causing energy decay with distance.

Returning to the layered medium of interest here, we address the other types of leaky waves that can be supported in such a medium [27], [32]. In particular, the poles on Riemann sheet C are characterized by leakage (exponential growth) into layer N and exponential decay in region 1, while the poles of Riemann sheet D are characterized by leakage (exponential growth) into both regions 1 and N. The leaky-wave pole positions for sheets C and D are as discussed above for Riemann sheet B.

Having detailed the four Riemann sheets of the complex k_ρ plane, we now address the path of integration for the integral in (3). Recall that on sheet D the fields are radiated (and grow exponentially) out of regions 1 and N, on sheet C the fields decay in region 1 but are radiated into region N, and finally on sheet A the fields decay in both regions 1 and N. Assuming that medium N is denser than medium 1 (e.g., if region 1 is air and region N soil), the path of integration is as shown in Fig. 2 on sheets D, C, and A. The portion of the integral on sheet D represents the radiation spectrum of the source, at wavenumbers k_ρ for which radiation emanates into layers 1 and N; after crossing the branch cut from sheet D to C the source is still in its radiation regime, but total internal reflection occurs at the interface of layers 1 and 2, such that the fields emanating into region 1 decay exponentially; finally, after crossing the branch cut to sheet A, the fields decay exponentially in both layers 1 and N, characteristic of total internal reflection at the top and bottom layers. The selected branch cuts and path of integration are similar to those in [37] for which there were only two Riemann sheets (two branch cuts) because the microwave structure of interest there was bound below by a conducting plane. However, in [37] the possibility of leaky-wave poles in quadrant one (and three) of the complex k_ρ was not mentioned. We have chosen a path of integration that resides above the finite number of surface- and leaky-wave poles in quadrant one (on the appropriate Riemann sheets), motivated by the pole positions in the limit $\omega'' \to 0$.

C. Method of Complex Images at Complex Frequencies

To effect the method of complex images, we utilize the identity

$$\frac{e^{-jkR}}{4\pi R} = \frac{1}{2\pi}\int_0^\infty \frac{1}{j2k_z} e^{-jk\cdot/z-z'/} J_o(k_\rho\xi) k_\rho \, dk_\rho \quad \text{(5a)}$$

with

$$k^2 = k_\rho^2 + k_2^2, \quad R = \sqrt{x-x')^2 + (y-y')^2 + |z-z'|^2} \quad \text{(5b)}$$

Therefore, to efficiently evaluate the integral in (3), for a

source in layer i, we form the approximation

$$\hat{f}(k_\rho, z, z') = \frac{e^{-jk_{zi}g(z,z')}}{j2k_{zi}}(\hat{h}k_\rho) \approx \frac{e^{-jk_{zi}(z,z')}}{j2k_{zi}} \cdot \left[\lim_{k_\rho \to \infty} \hat{h}(k_\rho) + \sum_{m=1}^{M} a_m e^{-k_{zi}b_m}\right] \quad (6)$$

where the coefficients a_m and b_m are estimated along an appropriate line in the complex k_{zi} plane [22]–[25]. To this end, we change from integration in the variable k_ρ to integration in k_{zi}. The path of integration in the k_ρ plane of Fig. 2 is converted to a linear path in the complex k_{zi} plane (Fig. 3), along which parametric estimation of a_m and b_m can be readily effected. For simplicity, in Fig. 3 we only show the branch cuts in quadrants 1 and 4, in which the integration path resides; similar branch cuts exist for the negatives of the branch points shown. We have branch points at $k_{zi} = \pm k_i$ due to $k_\rho = [k_i^2 - k_{zi}^2]^{1/2}$, at $k_{zi} = \pm[k_i^2 - k_1^2]^{1/2}$ due to $k_{z1} = [k_1^2 - k_i^2 + k_{zi}^2]^{1/2}$, and at $k_{zi} = \pm[k_i^2 - k_N^2]^{1/2}$ due to $k_{zN} = [k_N^2 - k_i^2 + k_{zi}^2]^{1/2}$. The branch cut emanating from $k_{zi} = k_i$ has been selected such that $\mathrm{Im}(k_\rho) > 0$ over the entire linear path of integration (as it was over the entire path of integration in Fig. 2). The branch cuts emanating from $[k_i^2 - k_1^2]^{1/2}$ and $[k_i^2 - k_N^2]^{1/2}$ are selected as follows. In Fig. 3, along the solid path of integration starting at $k_{zi} = k_i$, we have $k_\rho'' > 0, k_{z1}'' > 0$, and $k_{zN}'' > 0$, as for the initial path of integration on sheet D in Fig. 2. After this solid path of integration crosses the branch cut emanating from $[k_i^2 - k_1^2]^{1/2}$, $k_\rho'' > 0, k_{z1}'' < 0$, and $k_{zN}'' > 0$, and the dashed portion of integration represents the integration on sheet C of Fig. 2. Finally, after crossing the branch cut emanating from $[k_i^2 - k_N^2]^{1/2}$, $k_\rho'' > 0, k_{z1}'' < 0$, and $k_{zN}'' < 0$, and the dashed-dot path of integration in Fig. 3 represents integration on sheet A of Fig. 2. We perform the parametric fit in (6), using Prony's method [10], along the composite path in Fig. 3, which actually represents integration along three different k_{zi}-plane Riemann sheets. After so effecting the approximate representation on the right side of (6), the identity in (5a) is used to evaluate the space-domain Green's function analytically. We note that depending on the relative densities of layers 1, N, and i, the branch points may change their relative location and/or move into the forth quadrant of the k_{zi}-plane, and similar branch cuts are realized (in Fig. 3 we assume layer i is denser than layer 1 (air) and layer N, with layer N also denser than layer 1).

D. Modifications Due to Lossy Media

The previous discussion was restricted to the case of lossless media (real-wave velocities), for the purpose of simplifying the analysis. However, for the case of targets buried in soil, losses must be accounted for. If the loss in layer N is sufficiently large, the spatial growth with $|z|$ of the resonant fields in that layer will be overcome by the loss due to material dissipation and k_N will move into the forth quadrant of the k_ρ plane (cf. Fig. 2). To achieve the same four Riemann surfaces as in Fig. 2 (dictated by the properties of k_{z1}'' and k_{zN}''), the branch cut associated with k_N is also in quadrant four (with the branch cut associated with its negative residing in quadrant two). Under this circumstance, we take the same general integration path as in Fig. 2, but only cross one branch cut (associated with k_1). In this manner, we account for exponential growth of resonant fields into layer 1 (air) while realizing exponential decay in layer N (we only integrate on two Riemann sheets, as in [37], rather than integration on the three sheets considered in Fig. 2). Such modifications have straightforward implications with regard to integration in the k_{zi} plane (Fig. 3).

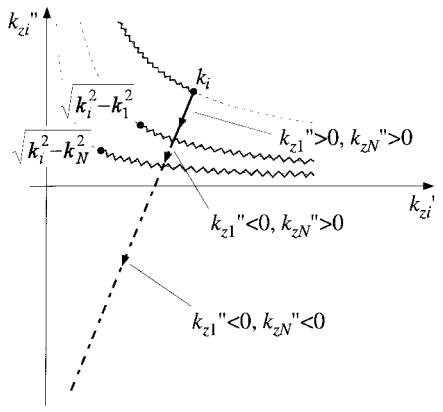


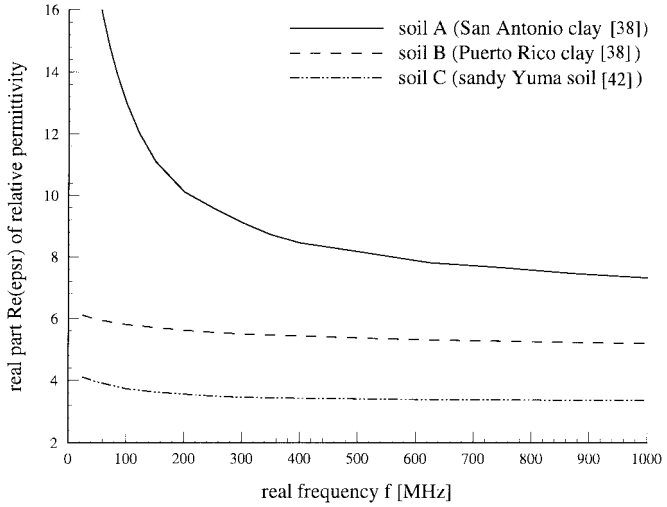
Fig. 3. Path of integration in the k_{zi} plane for source and observer in layer i (see Fig. 1). Along the solid path of integration $k_{z1}'' > 0$ and $k_{zN}'' > 0$ (characteristic of sheet D in Fig. 2), along the dashed path of integration $k_{z1}'' < 0$ and $k_{zN}'' > 0$ (characteristic of sheet C in Fig. 2), and along the dashed-dot path of integration $k_{z1}'' < 0$ and $k_{zN}'' < 0$ (characteristic of sheet A in Fig. 2).

Therefore, when accounting for loss associated with real soils, care must be taken to track the location of the branch cuts (and poles) in the complex k_ρ and k_{zi} planes, but the general framework introduced in Sections II-B and C remains principally unchanged.

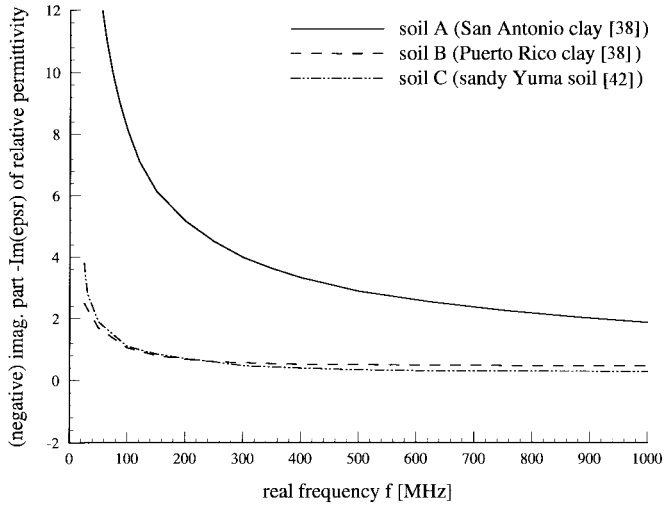
III. Example Resonant-Frequency Computations

Below we address the resonant behavior of several buried plastic targets, as computed by the algorithm outlined in Section II. Unfortunately, to our knowledge, there are no previous such computations existing in the literature (*measured* data have been presented [9], although these examples did not present sufficient details, e.g., soil properties, for numerical comparison). However, we have performed exhaustive tests on the accuracy of the results presented here, by carefully verifying, for example, the accuracy of the complex-image technique for Green's function evaluation at the complex (resonant) frequencies of interest.

In the first two examples, we consider a lossless cylindrical target of 8-cm diameter and 5-cm height, with dielectric constant $\epsilon_{rB} = 20$. This target is situated in a halfspace of clay, with electrical parameters of soil B described in Fig. 4 (representative of Puerto Rico clay with 10% water content, as reported in [38]). From Fig. 4, we see that the dielectric contrast between the target and soil background is significant, leading one to anticipate support of relatively high-Q resonances. We consider this example first because the high-Q resonance case is beset by less ambiguity in the soil electrical parameters. In particular, the frequency-dependent soil parameters in Fig. 4 are only valid on the *real*



(a)



(b)

Fig. 4. Complex dielectric constant of three types of soil [38], [42] used in resonant-frequency computations. (a) ϵ_r'. (b) ϵ_r''.

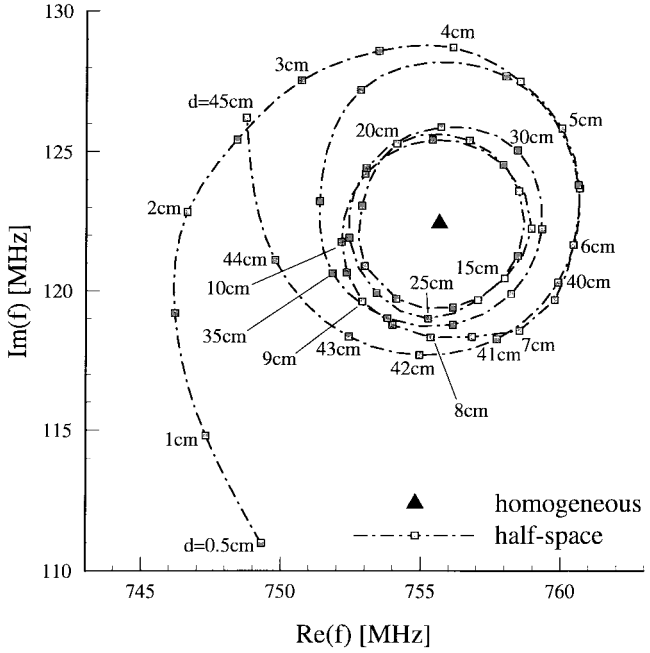
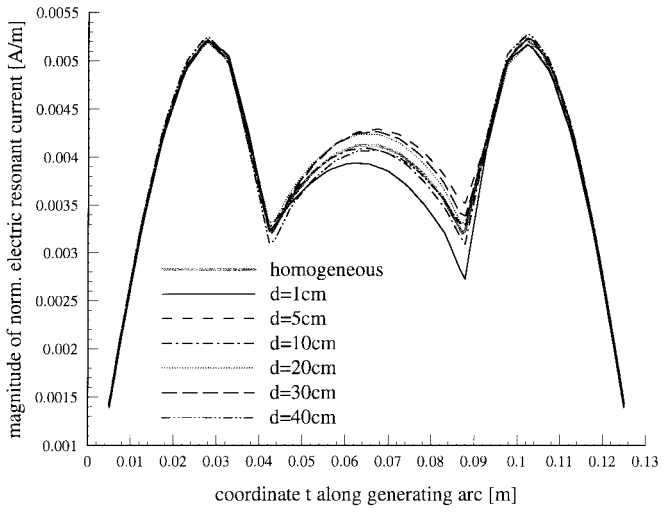


Fig. 5. Resonant frequency of a dielectric cylinder ($\epsilon_{rB} = 20$) of 8-cm diameter and 5-cm height buried in a half-space with electrical parameters described by soil B in Fig. 4. The resonant frequency is plotted as a function of depth d from the top of the target. Also plotted is the resonant frequency of the target situated in a homogeneous medium with the same electrical parameters.

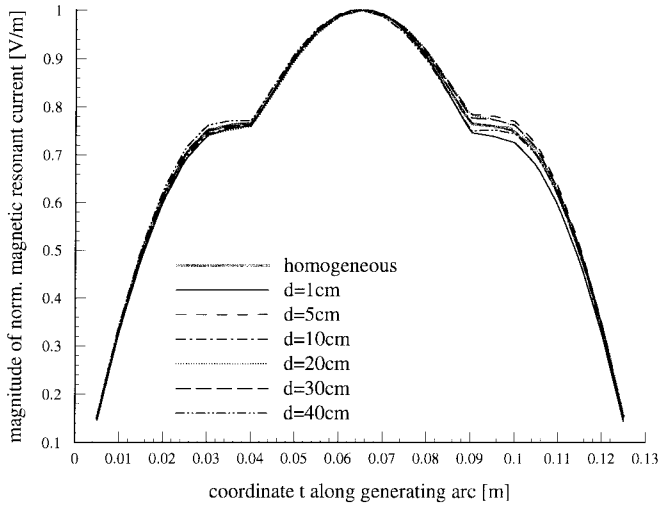
frequency axis. One requires an analytic expression for the complex dielectric constant, valid at *all* frequencies to perform analytic continuation into the complex frequency plane [39]. Unfortunately, soil measurements are only performed over a very limited frequency band, undermining derivation of such an analytic function. Therefore, in the work presented here, the complex dielectric constant of the soil [needed in (1)] is set to the dielectric constant at ω' (i.e., the real part of the complex resonant frequency). This approximation should be most appropriate for resonant frequencies near the real ω axis (ω'' small relative to ω'), representative of the high-Q resonant target we consider first.

In Fig. 5 we plot the resonant frequency of the fundamental mode as a function of depth from the top of the target, from 0.5 to 45 cm. In this plot, we also identify the resonant frequency of the same target situated in a homogeneous medium characterized by the same properties as the soil. As the target depth increases, one would anticipate that the resonant frequency would approach that of the same target in the homogeneous environment (i.e., that the air–soil interface will be of less importance with increasing target depth). We see in Fig. 5 that the buried-target resonant frequency does indeed approach that of the target in a homogeneous environment. However, an interesting spiraling effect is manifested with increasing target depth. A similar phenomenon has been witnessed for a wire above a lossy plane [40], a ring above a lossy halfspace [41], and for a wire buried in a lossy half-space [42]. All these examples, as well as the results in Fig. 5, can be explained in the same manner. In particular, assume that the target has (complex) resonant frequency ω_1 at depth d_1. If the target is lowered to a depth $d_1 + n\lambda_1/2$, where λ_1 is the approximate resonant wavelength at ω_1 and n is an integer, the impedance seen by the target looking toward the air–ground interface is approximately unchanged (at frequency ω_1). Therefore, if the target resonates at frequency ω_1 for depth d_1, then it should also resonant at ω_1 for depths $d_1 + n\lambda_1/2$. Note that the resonant frequencies initially spiral inward (toward the homogeneous-medium resonance) with increasing depth, followed by a subsequent outward spiral. Similar effects were seen in [40]–[42] for perfectly conducting targets, with such attributed to mode coupling between the self modes of the target itself and "image" modes produced by reverberations between the target and the air–ground interface. We believe the results presented here are the first to show this effect also occurs for dielectric targets.

The algorithm in Section II allows computation of the resonant fields, in addition to the resonant frequencies of Fig. 5. As an example, in Fig. 6 we plot the resonant surface currents $\boldsymbol{J}$ and $\boldsymbol{M}$ of (1), representative of the tangential magnetic and electric fields, respectively, on the surface of the BOR target. For a set of depths considered in Fig. 5, we see



(a)



(b)

Fig. 6. Resonant currents $\boldsymbol{J}$ and $\boldsymbol{M}$ for the resonances in Fig. 5. (a) magnitude of J_ϕ. (b) magnitude of M_t.

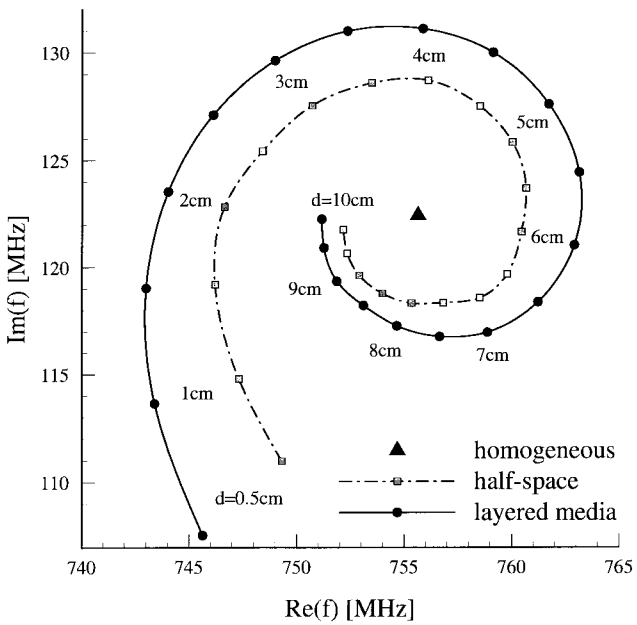


Fig. 7. Resonant frequency of the target considered in Fig. 5, centered in a layer of soil of thickness $2d + 5$ cm, with the electrical properties of the layer described by soil B in Fig. 4. Beneath this layer of soil is a half-space with electrical properties described by soil C in Fig. 4. The target resonant frequency is plotted as a function of d. Also plotted is the resonant frequency of the target situated in a half-space environment (from Fig. 5) and in a homogeneous medium of soil B in Fig. 4.

that there is minimal change in the resonant surface currents (J_ϕ and M_t for the mode considered here) with increasing target depth. In Figs. 5 and 6 we have considered the properties of a single (lowest-order) resonant mode; an infinite set of higher-order modes exist [6].

To demonstrate an example of greater complexity than the half-space problem of Figs. 5 and 6, we consider the same target *centered* in a layer of thickness $2d+5$ cm (i.e., there is a distance d from the 5-cm-thick target to the top and bottom of the layer). Moreover, this layer of soil is characterized by the same electrical parameters as considered in Figs. 5 and 6 (soil type B in Fig. 4). Beneath this layer is a half-space characterized by soil type C in Fig. 4 (taken from [42], for 5% water content). This example may simulate a buried target, for which the disturbed soil has electrical parameters different than those of the background (undisturbed) soil. The resonant frequencies of this target are demonstrated in Fig. 7 for d ranging from 0.5 to 10 cm. For direct comparison, the corresponding results of the target in a half-space (Fig. 5) are also plotted. A spiraling behavior similar to that in Fig. 5 is manifested with increasing d with an analogous explanation. We see that the resonant frequencies are quantitatively very similar for these examples (for the same distance from the target to the air–soil interface). Moreover, note that the differences between the resonant frequencies diminishes as the distance from the target to both interfaces increases.

The examples in Figs. 5–7 considered a target of dielectric constant $\epsilon_{rB} = 20$ such that relatively high-Q resonances could be supported. As discussed, the approximations with regard to soil properties at complex frequencies are more appropriate under such circumstances. Nevertheless, most buried plastic targets of interest are composed of dielectric constants much smaller than considered above. We therefore consider an example using parameters that may be expected of practical radar problems. In particular, we consider a "PMN2" plastic antipersonnel mine, with dimensions shown in the inset of Fig. 8. This mine is principally plastic, with dielectric constant $\epsilon_{rB} = 2.9$. We consider this target buried in a half-space, using soil type A in Fig. 4. Because the electrical contrast between the target and background is now smaller than that considered in Figs. 5–7, one would anticipate this to be a lower Q target. The results in Fig. 8 are characterized by a spiraling behavior similar to that in Figs. 5 and 7. It is interesting to note that, for the depths considered here, the resonant frequencies of this target continually spiral outward with increasing depth (cf. Fig. 5).

IV. Conclusions

A rigorous method-of-moments (MoM) algorithm has been devised for modeling the resonances of targets buried in a lossy, layered medium, representative of soil. This research

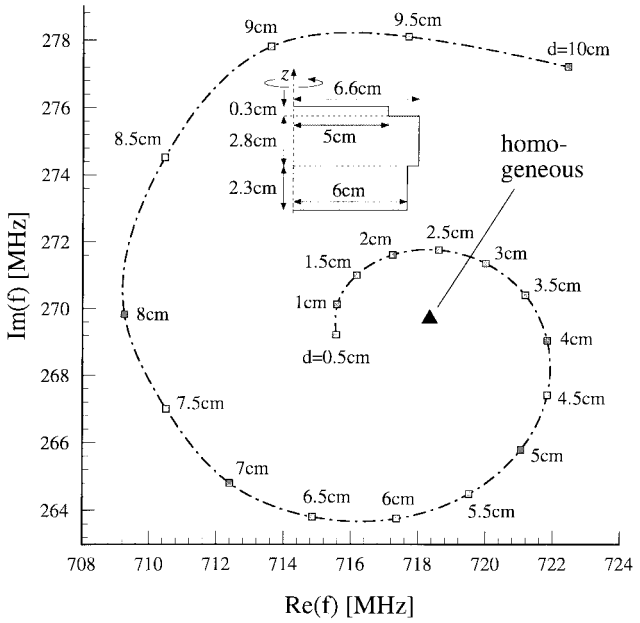


Fig. 8. Resonant frequency of a "PMN2" plastic land mine ($\epsilon_{rB} = 2.9$, with geometry shown inset) buried in a halfspace with electrical parameters described by soil type A in Fig. 4. Results are plotted as a function of depth d from the top of the target. For comparison, the resonant frequency of the target situated in a homogeneous medium of the same parameters is also plotted.

has been motivated by the use of resonances as discriminants for identifying buried plastic land mines. We have therefore restricted our analysis to the case of targets that can be simulated as a dielectric body of revolution since most plastic mines can be so approximated. Even with this simplification, such a MoM analysis is computationally challenging. In particular, one must evaluate the layered-medium Green's function at complex (resonant) frequencies. The layered-medium Green's function can only be expressed analytically in the spectral domain, with Sommerfeld-type integrals required for conversion to the space domain. Such integrals are computationally expensive if traditional integration techniques are employed and, therefore, here we have utilized the recently developed method of complex images [22]–[25].

The method of complex images is based on performing a parametric fit to the spectral Green's function along an appropriate path in the spectral domain, after which each term in the expansion can be converted to the space domain analytically via the Sommerfeld or Weyl identity [26]. While the method of complex images has been in use for several years, it is believed that this work is its first application to dielectric targets. We therefore required use of additional Green's function components than those used previously for purely perfectly conducting targets [21]–[25]. Moreover, the complex frequencies of interest required careful attention to the spectral-domain path of integration.

Several numerical examples have been presented. We initially considered relatively high-dielectric targets, for which high-Q resonances are supported. Additionally, we also examined the lowest-order resonant frequency of an actual plastic land mine. It was demonstrated that the low electrical contrast between the soil and mine results in low-Q resonances which undermine the utility of resonance-based discrimination.

REFERENCES

[1] C. E. Baum, "On the singularity expansion method for the solution of electromagnetic interaction problems," Air Force Weapons Lab. Interaction Notes, Note 88, 1971.

[2] E. Heyman and L.B. Felsen, "A wavefront interpretation of the singularity expansion method," *IEEE Trans. Antennas Propagat.*, vol. AP-33, pp. 706–718, July 1985.

[3] H. Shirai and L. B. Felsen, "Modified GTD for generating complex resonances for flat strips and disks," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 779–790, June 1986.

[4] L. Marin, "Natural mode representation of transient scattered fields," *IEEE Trans. Antennas Propagat.*, vol. AP-21, pp. 809–818, Nov. 1973.

[5] F. M. Tesche, "On the analysis of scattering and antenna problems using the singularity expansion technique," *IEEE Trans. Antennas Propagat.*, vol. AP-21, pp. 53–62, Jan. 1973.

[6] B. L. Merchant, P. L. Moser, A. Nagl, and Überall, "Complex pole patterns of the scattering amplitude for conducting spheroids and finite-length cylinders," *IEEE Trans. Antennas Propagat.*, vol. 36, pp. 1769–1777, Dec. 1988.

[7] E. J. Rothwell, J. Baker, K.-M. Chen, and D. P. Nyquist, "Approximate natural response of an arbitrary shaped thin wire scatterer," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 1457–1462, Oct. 1991.

[8] L. Peters, Jr. and J. D. Young, "Applications of subsurface transient radars," in *Time-Domain Measurements in Electromagnetics*, E. K. Miller, Ed. New York: Van Nostrand Reinhold, 1986.

[9] L. Peters, Jr., J. J. Daniels, and J. D. Young, "Ground penetrating radar as an environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.

[10] M. L. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. Antennas Propagat.*, vol. AP-23, pp. 777–781, Nov. 1975.

[11] D. G. Dudley, "Parametric modeling of transient electromagnetic systems," *Radio Sci.*, vol. 14, pp. 387–396, 1979.

[12] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 814–824, May 1990.

[13] S. Vitebskiy and L. Carin, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy, dispersive half space," *IEEE Trans. Antennas Propagat.*, vol. 28, pp. 1575–1583, Dec. 1996.

[14] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of ground penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–28, Jan. 1996.

[15] H. S. Chang and K. K. Mei, "Scattering of electromagnetic waves by buried and partly buried bodies of revolution," *IEEE Trans. Geosci. Remote Sensing*, vol. GRS-23, pp. 596–592, July 1985.

[16] A. C. Dubey and R. L. Barnard, Eds., "Detection and remediation technologies for mines and minelike targets," *Proc. SPIE*, vol. 3079, 1997.

[17] J. R. Wait, "Image theory of a quasistatic magnetic dipole over a dissipative half-space," *Electron. Lett.*, vol. 5, no. 13, pp. 281–282, June 1969.

[18] Y. Rahmat-Samii, R. Mittra, and P. Parhami, "Evaluation of Sommerfeld integrals for lossy half-space problems," *Electromagn.*, vol. 1, no. 1, pp. 1–28, 1981.

[19] S. F. Mahmoud and A. D. Metwally, "New image representations for dipoles near a dissipative earth," *Radio Sci.*, vol. 21, pp. 605–616, Nov. 1981.

[20] I. V. Lindell and E. Alanen, "Exact image theory for the Sommerfeld half-space problem, Part III: General formulation," *IEEE Trans. Antennas Propagat.*, vol. AP-32, pp. 1027–1032, Oct. 1984.

[21] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in Layered media, Parts I and II," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–352, Mar. 1990.

[22] J. J. Yang, Y. L. Chow, D. G. Fang, "Discrete complex images of a three-dimensional dipole above and within a lossy ground," *Proc. Inst. Elect. Eng.*, vol. 138, pt. H, no. 4, pp. 319–326, Aug. 1991.

[23] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.

[24] Y. L. Chow, J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 588–562, Mar. 1991.

[25] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 28, pp. 143–151, Feb. 1996.
[26] W.C. Chew, *Waves and Fields in Inhomogeneous Media*. Piscataway, NJ: IEEE Press, 1995.
[27] T. Tamir and A. A. Oliner, "Guided complex waves—Fields at an interface: Part I; Relation to radiation patterns: Part II," *Proc. Inst. Elect. Eng.*, vol. 110, pp. 310–334, 1963.
[28] S. R. Vechinski and T. H. Shumpert, "Natural resonances of conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 1133–1136, July 1990.
[29] A. W. Glisson, D. Kajfez, and J. James, "Evaluation of the modes in dielectric resonators using a surface integral equation formulation," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-31, pp. 1023–1029, Dec. 1983.
[30] S. Vitebskiy and L. Carin, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy dispersive halfspace," *IEEE Trans. Antennas Propagat.*, vol. 28, pp. 1575–1583, Dec. 1996.
[31] N. Geng and L. Carin, "Wideband electromagnetic scattering from a dielectric BOR buried in a layered lossy, dispersive medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.
[32] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
[33] N. K. Das and D. M. Pozar, "Full-wave spectral-domain computation of material, radiation and guided-wave losses in infinite multilayered printed transmission lines," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 54–63, Jan. 1991.
[34] L. Carin and N. K. Das, "Leaky waves on broadside-coupled microstrip," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 58–66, Jan. 1992.
[35] M. Tsuji and H. Shigesawa, "Packaging of printed circuit lines: A dangerous cause of narrow pulse distortion," *IEEE Trans. Microwave Theory Tech.*, vol. 42, pp. 1784–1790, Sept. 1994.
[36] C. J. Madden, M. J. W. Rodwell, R. A. Marsland, Y. C. Pao, and D. M. Bloom, "Generation of 3.5 ps fall-time shock waves on a monolithic GaAs nonlinear transmission line," *IEEE Electron. Device Lett.*, vol. 9, p. 303, 1988.
[37] K. A. Michalski and D. Zheng, "Analysis of microstrip resonators of arbitrary shape," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 112–119. Jan. 1992.
[38] J. E. Hipp, "Soil electromagnetic parameters as functions of frequency, soil density, and soil moisture," *Proc. IEEE*, vol. 62, pp. 98–103, Jan. 1974.
[39] K. E. Oughstun and G. C. Sherman, *Electromagnetic Pulse Propagation in Causal Dielectrics*. New York: Springer-Verlag, 1994, vol. 16.
[40] L. S. Riggs and T. H. Shumpert, "Trajectories of singularities of a thin wire scatterer parallel to lossy ground," *IEEE Trans. Antennas Propagat.*, vol. AP-27, pp. 864–868, Nov. 1979.
[41] E. J. Rothwell and M. J. Cloud, "On the natural frequencies of an annular ring above a conducting half space," *J. Electron. Waves Appl.*, vol. 10, pp. 155–179, Feb. 1996.
[42] S. Vitebskiy, L. Carin, M. A. Ressler, and F. H. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.

Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl.Ing. and Dr.Ing. degrees in electrical engineering from the University of Karlsruhe, Germany, in 1991 and 1996, respectively.

From 1991 to 1996, he was with the Institute for Microwaves and Electronics at the University of Karlsruhe, working on full-wave propagation modeling for radio communication systems. In January 1997 he joined the Department of Electrical and Computer Engineering at Duke University, Durham, NC, for 18 months, in a Visiting Postdoctoral capacity. Since July 1998 he has been back with the University of Karlsruhe. His current research interests include computational methods in electromagnetics and wave propagation modeling.

Dr. Geng received the Mannesmann Innovation Award in 1997 for his Ph.D. dissertation on full-wave propagation modeling for radio communication systems.

David R. Jackson (S'83–M'85–SM'95–F'99) [redacted] He received the B.S.E.E. and M.S.E.E. degrees from the University of Missouri, Columbia, in 1979 and 1981, respectively, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles, in 1985.

From 1985 to 1991, he was an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Houston, TX. From 1991 to 1998 he was an Associate Professor in the same department and since 1998 has been a Professor there. He is on the editorial board of the *International Journal of RF and Microwave Computer-Aided Engineering*. His current research interests include computer-aided design of microstrip antennas and circuits, microstrip antenna analysis and design, periodic structures, leaky-wave antennas, leakage effects in microwave integrated circuits, and bioelectromagnetics.

Dr. Jackson currently serves as the secretary for URSI United States Commission B and is a member of ADCOM for the IEEE Antenna and Propagation Society. He is on the editorial board of IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES, is an Associate Editor for the IEEE Press Series on Electromagnetic Waves, and is a past Associate Editor for the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.

Lawrence Carin (SM'96) [redacted] He received the B.S., M.S., and Ph.D. degrees, all in electrical engineering, from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department at Polytechnic University, Brooklyn, NY, as an Assistant Professor, and became an Associate Professor there in 1994. In September 1995 he joined the Electrical Engineering Department at Duke University, Durham, NC, where he is an Associate Professor. His current research interests include short-pulse scattering, propagation, and signal processing. He is the principal investigator on an Army Research Office funded Multidisciplinary University Research Initiative (MURI) on demining.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.

On the Extended-Born Technique for Scattering from Buried Dielectric Targets

Nilanjan Dasgupta, Norbert Geng,
Traian Dogaru, and Lawrence Carin

***Abstract*—The extended Born technique is an approximate nonlinear method for analyzing scattering from a weak discontinuity. Moreover, when applied to the low-frequency (electromagnetic induction) applications for which it was developed originally, extended Born has accurately modeled scattering from inhomogeneities considerably larger than those appropriate for the standard linear Born technique. In this letter, we examine the extended Born technique at radar frequencies, considering three-dimensional (3-D) scattering from a dielectric target buried in a lossy half space.**

***Index Terms*—Born approximation, buried object detection.**

There are many applications for which one would like to consider electromagnetic scattering from a weak inhomogeneity in a background medium. For example, plastic anti-personnel land mines are often composed of materials with dielectric constants close to those of the background soil [1]. Such a scattering problem can be solved rigorously via the method of moments (MoM) [2], taking proper account of the half-space or layered-medium Green's function [2], [3] used to model the soil. However, while the MoM is highly accurate, it is often computationally expensive. In the context of the weak-scattering problem of interest here, we can avail ourselves of approximate modeling algorithms, which can yield accurate results while simultaneously being relatively inexpensive computationally. The Born approximation [4] is a well-known example of such a technique, it assuming that the fields inside the scatterer are the same as the incident fields, in the absence of the inhomogeneity. While this is a popular technique, it is only appropriate for relatively small inhomogeneities [4], [5]. Recently, Habashy *et al.* [5] have developed a new algorithm, which they demonstrated can handle inhomogeneities considerably larger than appropriate for the Born approximation. However, their previous research concentrated on very-low-frequency (kilohertz) applications, of interest in geophysics [5]. Moreover, previous results assumed a homogeneous background medium. Here we consider radar-frequency operation, for a target buried in a half-space.

Consider a nonmagnetic background medium characterized by the generally inhomogeneous permittivity $\epsilon_b(\boldsymbol{r})$. Further, assume that the permittivity $\epsilon(\boldsymbol{r})$ characterizes the medium when a dielectric inhomogeneity is present (i.e., $\epsilon(\boldsymbol{r}) - \epsilon_b(\boldsymbol{r}) \neq 0$ for $\boldsymbol{r} \in V$, where V is the volume of the discontinuity, while $\epsilon(\boldsymbol{r}) - \epsilon_b(\boldsymbol{r}) = 0$ for $\boldsymbol{r} \notin V$). If $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$ characterizes the dyadic Green's function for the medium $\epsilon_b(\boldsymbol{r})$, then the total electric field at $\boldsymbol{r}$ is [4], [5]

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}^{inc}(\boldsymbol{r}) + \int_V \boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}') \cdot \boldsymbol{E}(\boldsymbol{r}')[\epsilon(\boldsymbol{r}') - \epsilon_b(\boldsymbol{r}')]\, d^3\boldsymbol{r}'. \quad (1)$$

The well-known Born approximation is characterized by assuming $\boldsymbol{E}(\boldsymbol{r}') \approx \boldsymbol{E}^{\rm inc}(\boldsymbol{r}')$ inside the integral, which yields a direct but approximate solution for $\boldsymbol{E}(\boldsymbol{r})$. This solution is linear in the inhomogeneity $\epsilon(\boldsymbol{r}) - \epsilon_b(\boldsymbol{r})$, which implies that it neglects multibounce interactions inside the scatterer [4], [5].

Manuscript received October 27, 1998; revised August 23, 1999.

N. Dasgupta, T. Dogaru, and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

N. Geng was with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA. He is now with the University of Karlsruhe, Institut fuer Hoechstfrequenztechnik und Elektronik D-76128 Karlsruhe, Germany.

Publisher Item Identifier S 0018-926X(99)09951-2.

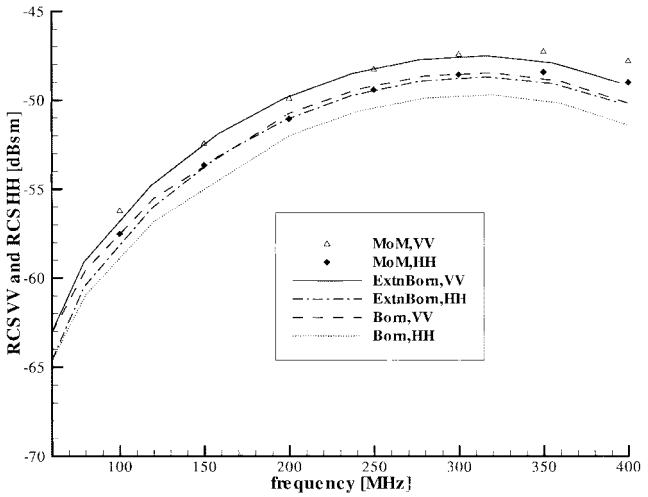


Fig. 1. Backscattered radar cross section (RCS) of a 10-cm radius sphere with its center placed 15 cm beneath the air–ground interface due to vertically (V) or horizontally (H) polarized plane-wave excitation, 50° from grazing. The lossless sphere is characterized by dielectric constant $\epsilon_r = 3$ and the lossy background soil is characterized by $\epsilon_r = 4$ and $\sigma = 0.01$ S/m.

The extended Born approximation is based on recognizing that $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$ is large for $\boldsymbol{r}$ in the vicinity of $\boldsymbol{r}'$, while being relatively small for $\boldsymbol{r}$ distant from $\boldsymbol{r}'$. Consequently,

$$\boldsymbol{E}(\boldsymbol{r}) \approx \boldsymbol{E}^{inc}(\boldsymbol{r}) + \int_V \boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}') \cdot \boldsymbol{E}(\boldsymbol{r})[\epsilon(\boldsymbol{r}') - \epsilon_b(\boldsymbol{r}')]\, d^3\boldsymbol{r}', \qquad \text{for } \boldsymbol{r} \in V. \quad (2)$$

Rearranging (2), we find

$$\boldsymbol{E}(\boldsymbol{r}) \approx \boldsymbol{M}(\boldsymbol{r})^{-1} \cdot \boldsymbol{E}^{inc}(\boldsymbol{r}), \qquad \text{for } \boldsymbol{r} \in V$$
$$\boldsymbol{M}(\boldsymbol{r}) \equiv \boldsymbol{I} - \int_V \boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')[\epsilon(\boldsymbol{r}') - \epsilon_b(\boldsymbol{r}')]\, d^3\boldsymbol{r}'. \quad (3)$$

The expression in (3) yields an approximation for the electric fields inside the target, which can be used in (1) to find the fields everywhere (i.e., for $\boldsymbol{r} \notin V$); this approximation has been termed "extended Born" [5]. Note that the inverse of $\boldsymbol{M}$ approximately accounts for multiple interactions, yielding a nonlinear algorithm in $\epsilon(\boldsymbol{r}) - \epsilon_b(\boldsymbol{r})$.

If the integral term in (3) is ignored, extended Born reduces to the classical Born approximation. Moreover, the extended Born approximation involves the inversion of a 3×3 matrix $\boldsymbol{M}(\boldsymbol{r})$, for each $\boldsymbol{r} \in V$. These matrices are computed by representing $\epsilon(\boldsymbol{r}) - \epsilon_b(\boldsymbol{r})$ as an aggregate of N three-dimensional cubes (or other appropriate basis function) [5], with cube dimensions small relative to wavelength; here the cubes have dimensions of at most $\lambda/10$, where λ is the wavelength in the medium for which $\epsilon(\boldsymbol{r}) - \epsilon_b(\boldsymbol{r}) \neq 0$. We compute $\boldsymbol{M}(\boldsymbol{r})$ at the $N\boldsymbol{r}_n$ that constitute the centers of the cubes. Consequently, we need invert N 3×3 matrices. This can be contrasted with a volumetric MoM solution of (1), in which the *fields* $\boldsymbol{E}(\boldsymbol{r}'), \boldsymbol{r}' \in V$ are expanded in a basis. The basis functions must again be small relative to wavelength, but in the case of MoM the constituent to be expanded is a vector, in comparison to the scalar $\epsilon(\boldsymbol{r}) - \epsilon_b(\boldsymbol{r})$. Moreover, in a volumetric MoM solution of this same

0018–926X/99$10.00 © 1999 IEEE

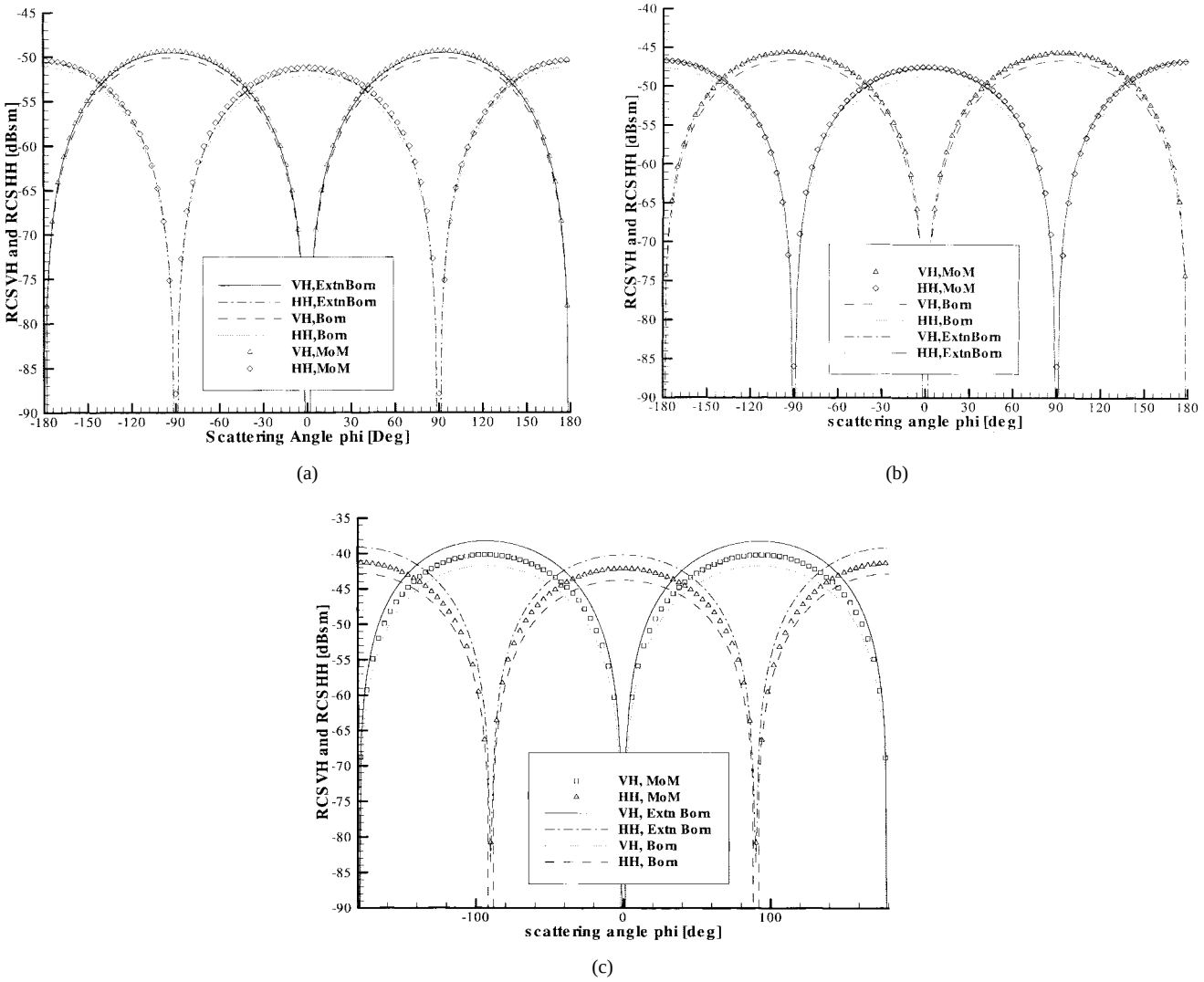


Fig. 2. Bistatic RCS of a 10-cm radius sphere with its center placed 15 cm beneath the air–ground interface. The incident plane wave is incident 50° from grazing at $\phi = 0°$ and bistatic-scattering results are plotted as a function of ϕ, 50° from grazing. The lossy background soil is characterized by $\epsilon_r = 4$ and $\sigma = 0.01$ S/m and the operating frequency is 300 MHz corresponding to $d/\lambda = 0.4$, where d is the sphere diameter and λ is the wavelength in the soil. Results are plotted for variable dielectric constants for the lossless sphere. Sphere dielectric constant (a) $\epsilon_r = 3.5$, (b) $\epsilon_r = 3.0$, and (c) $\epsilon_r = 2.0$.

problem, we invert a single $3N \times 3N$ (three vector field components, each represented by a separate cube), rather than N 3×3 matrices.

As with the MoM, the extended Born technique requires accurate computation of the dyadic Green's function $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$. While this is relatively straightforward for the free-space case [4], [5], for the layered-medium problem of interest here, accurate and efficient computation of $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$ is challenging. We have utilized the complex-image technique, which represents the dyadic Green's function as a sum of homogeneous-medium Green's functions with generally complex source points [6].

Below we present a series of results for a lossless dielectric target buried in a lossy half-space. In addition to the results computed via Born and extended Born, reference MoM results are presented. The MoM results were computed via a surface-integral-equation formulation (rather than the volumetric MoM procedure alluded to above). A detailed discussion of this MoM formulation can be found in [2] and [7].

Before proceeding, we note that the loss in the half-space is used to represent the characteristics of soil, but it is also critical to the accuracy of the extended Born algorithm. In particular, recall that the approximation in (2) (which drives the extended-Born algorithm) is based on the assumption that $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$ is peaked near $\boldsymbol{r} = \boldsymbol{r}'$ and decays quickly with increasing $|\boldsymbol{r} - \boldsymbol{r}'|$. Essentially, this implies that the fields due to a point source at $\boldsymbol{r}'$ diminish quickly with increasing $|\boldsymbol{r} - \boldsymbol{r}'|$; this phenomenon is aided by the presence of loss in the background medium in which the discontinuity is immersed. In particular, we tested the extended Born algorithm for a lossless background medium (e.g., vacuum), at the radar frequencies of interest—for this case, we did not see any advantage of the extended Born algorithm *vis-à-vis* conventional linear Born.

The first set of results consider a sphere of 10-cm radius and dielectric constant $\epsilon_r = 3.0$, buried in a half-space with dielectric constant $\epsilon_r = 4.0$ and conductivity $\sigma = 0.01$ S/m. The center of the sphere is buried 15 cm from the air–ground interface, and the incidence angle is 50° from grazing. Far-zone backscatter results are shown in Fig. 1, for VV and HH polarization (vertical polarization incidence, vertical-polarization receive; horizontal-polarization incidence, and horizontal-polarization receive). For this symmetric target, there are no cross-polarized *backscattered* fields [8]. Results are plotted as a function of operating frequency and we see that the Born and extended Born solutions are in close agreement at frequencies less than approximately 100 MHz ($d/\lambda < 0.137$, where

λ is the approximate wavelength in the *lossy* soil and d is the sphere diameter), while the extended Born solution stays in close agreement with the MoM results up to approximately 300 MHz (i.e., for $d/\lambda \leq 0.4$). These results demonstrate the superiority of extended Born relative to conventional Born. In particular, for a fixed inhomogeneity, these results indicate that extended Born allows consideration of larger targets (electrically), implying accurate performance at higher operating frequencies.

The results in Fig. 1 considered a single inhomogeneity profile at variable frequencies. To further address the accuracy of the extended Born and Born algorithms relative to MoM, we fix the operating frequency at 300 MHz and vary the target-induced inhomogeneity for the same half-space background. The sphere radius is again 10 cm placed 15 cm from the interface; the target is situated close to the surface to enhance target-surface interaction (this being handled rigorously via MoM) approximately via extended Born, and not at all via classical Born. In Fig. 2(a)–(c), we consider the target with dielectric constant 3.5, 3.0, and 2.0 (recall the soil has a dielectric constant $\epsilon_r = 4$ and a conductivity of $\sigma = 0.01$ S/m). The far-zone *bistatic* results are plotted as a function of the azimuthal angle ϕ, 50° from grazing, for an incident field propagating 50° from grazing, at $\phi = 0°$. For a target of dielectric constant $\epsilon_r = 3.5$ and $\epsilon_r = 3.0$, the agreement between extended Born and the MoM is excellent for all bistatic angles. By comparison, for these same cases a noticeable discrepancy (1 dB) is seen between MoM and classical Born, attributed to the degree of inhomogeneity as well as the proximity of the target to the interface. For this example, we see that the extended Born breaks down appreciably when the target dielectric constant is reduced further to $\epsilon_r = 2$ [Fig. 2(c)]. Finally, close inspection of Fig. 2(a) and (b) implies that the extended Born is more accurate for a target with $\epsilon_r = 3.0$, relative to $\epsilon_r = 3.5$. We attribute this anomaly to inaccuracies in the MoM rather than in extended Born. In particular, it has been mentioned that the surface-integral-equation MoM formulation used here can often be slightly inaccurate for very weak inhomogeneities [9]. Finally, while space limitations prohibit our showing the results here, the extended Born-predicted VV fields were found to be as accurate as the HH and VH results in Fig. 2.

It is of interest to address the memory and central processing unit (CPU) requirements of the extended Born solution, relative to the Born and surface MoM solutions shown here for comparison. In Fig. 3, we plot the CPU and RAM requirements of the extended Born, Born, and MoM solutions for the problem in Fig. 1. With regard to the MoM solution, results are shown for a general triangular-patch model for the surface currents [7] and, in all results, the subsectional basis functions were designed to discretizate at a level of 20 basis functions per wavelength (for all methods). As expected, the simple Born solution requires the least CPU and memory. However, it is also important to note that the extended Born solution, which provides very accurate results relative to the benchmark MoM solution (see Fig. 1), requires significantly less CPU and RAM than the MoM. The electrical size of the sphere surface area scales as f^2 (order $(r/\lambda)^2$, where r is the constant sphere radius and λ the variable wavelength), as does the number of patch basis functions N required in the surface MoM solution, where f is the frequency. Consequently, the MoM RAM is of order f^4 (if N is the number of unknowns, the RAM required for the MoM impedance matrix is order N^2). Further, at low frequencies the MoM CPU is order f^4 required to fill the $N \times N$ MoM impedance matrix, with this becoming order f^6 at higher frequencies, where the LU decomposition matrix solver (order N^3) becomes the primary computational burden. The memory requirements of the extended Born solution are almost frequency independent, since for all frequencies we need only store consecutive 3×3 matrices. At

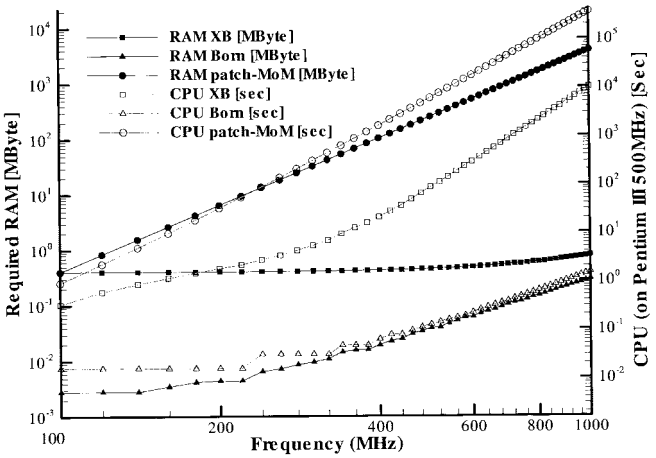


Fig. 3. Comparison of the RAM and CPU requirements of the extended Born, Born, and general triangular-patch method-of-moments (patch MOM) solutions [7], for the example in Fig. 1. All results were run on a 500-MHz Pentium III personal computer.

low frequencies the extended Born CPU is approximately linear in the number of unknowns and, therefore, is of order f^3 (extended Born is a volumetric method *vis-à-vis* the surface MoM solution). Note that as the frequency increases, the slope of the extended Born CPU changes. We attribute this to the computational burden of computing the half-space Green's function at higher frequencies [2], [7].

In conclusion, the extended Born algorithm has been utilized to model electromagnetic scattering at radar frequencies for dielectric targets buried in a lossy half-space. In general, for the class of problems considered here, the extended Born algorithm yields results that are at least as accurate as conventional Born, and often the results are considerably better (relative to a rigorous MoM solution). However, the computational complexity of extended Born is higher than that of classical Born, since the extended-Born algorithm requires accurate computation of the background medium (half-space) Green's function, as well as multiple volumetric near-field integrations. We also note that we have focused on frequency-domain integral-equation formulations, although for time-domain applications, the finite difference time domain may be more appropriate [10], it also applicable to a more general problem class (not just layered media). Finally, all results presented here have been for far-zone scattering. Future research will involve consideration of extended Born for near-zone radar scattering in a half space. Such an extended-Born solution has been successfully implemented for electromagnetic induction (EMI) problems [11].

REFERENCES

[1] L. Carin, N. Geng, M. McClure, J. Sichina, and L. Nguyen, "Ultrawide-band synthetic aperture radar for mine-field detection," *IEEE Antennas Propagat. Mag.*, vol. 41, pp. 18–33, Feb. 1999.

[2] N. Geng and L. Carin, "Wideband electromagnetic scattering from a dielectric body of revolution buried in a lossy, dispersive layered medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.

[3] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, parts I and II," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–352, Mar. 1990.

[4] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. Oxford, U.K.: Oxford Press, 1996.

[5] T. M. Habashy, R. W. Groom, and B. Spies, "Beyond the Born and Rytov approximations: A nonlinear approach to electromagnetic scattering," *J. Geophys. Res.*, vol. 98, pp. 1759–1775, 1993.

[6] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-

spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.
[7] J. He, N. Geng, T. Yu, and L. Carin, "Electromagnetic scattering from a general dielectric target embedded in a multilayered medium," *Radio Sci.* to be published.
[8] L. Carin, R. Kapoor, and C. E. Baum, "A Polarimetric SAR imaging of buried landmines," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1985–1988, Nov. 1998.
[9] J. R. Mautz and R. F. Harrington, "Electromagnetic scattering from a homogeneous material body of revolution," *AEU*, vol. 33, pp. 71–80, 1979.
[10] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of a ground-penetrating radar: FDTD theory compared with measurements," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–44, Jan. 1996.
[11] T. Yu and L. Carin, "Extended-Born analysis of the electromagnetic-induction response of a dielectric target embedded in a lossy layered medium," *Radio Sci.*, to be published.

Buried Object Detection and Location Estimation from Electromagnetic Field Measurements

George A. Tsihrintzis, Peter Meincke Johansen, and Anthony J. Devaney

***Abstract*—A translation property is derived describing the field scattered from a known buried object placed at distinct locations. The result is used to derive the optimum algorithm for detecting the known buried object and estimating its location from noisy scattered electromagnetic field measurements.**

***Index Terms*—Ground penetrating radar, Lippmann–Schwinger equation, object detection, wave scattering.**

I. Introduction

The problem of detecting known buried objects and estimating their location from electromagnetic field measurements is relevant in many technological areas such as demining, buried waste clean up, excavation planning, and archaeological investigations. In all of the above applications, serious challenges arise, mainly due to physical limitations such as: 1) significant losses due to moist soil that limit the signal-to-noise ratio; 2) presence of a large number of randomly distributed unwanted objects, returns from which obscure the return from the object of interest (volume clutter); and 3) random roughness of the air/soil interface that results in incoherent (random) returns (surface clutter). As a result of these limitations, standard threshold detection algorithms [1] may not successfully address the buried object detection problem, especially when the object is small and the return signal weak.

Manuscript received August 27, 1998; revised August 16, 1999. The work of G. A. Tsihrintzis and A. J. Devaney were supported in part by ARO Grant DAAG55-97-1-0013 and by AFOSR Grant F49620-96-1-0028. The work of P. M. Johansen was supported by the Danish Technical Research Council.

G. A. Tsihrintzis is with the Department of Informatics, University of Piraeus, Piraeus, 185 34 Greece.

P. M. Johansen is with the Department of Electromagnetic Systems, Technical University of Denmark, Lyngby, DK-2800 Denmark.

A. J. Devaney is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA.

Publisher Item Identifier S 0018-926X(99)09952-4.

The signal processing strategy with the highest potential to address the buried object detection problem is the one that utilizes field models for the air/soil/object environment. Due to the complicated nature of this environment, however, certain simplifications need to be made, namely the assumption of a planar air/soil interface and of no interactions between the object of interest and the distribution of unwanted objects. These assumptions allow, as illustrated in this letter, the solution to the buried object detection problem within the framework of exact electromagnetic field scattering theory and the derivation of a computationally efficient optimum detection algorithm. More specifically, the letter is organized as follows. Section II is devoted to statement of the basic field equations governing the interaction of probing fields with the air/soil/object environment and derivation of a field translation property. Section III addresses the buried object detection problem and presents the maximum likelihood algorithm for the solution to it. Finally, Section IV is a discussion of the results of the letter with suggestions for possible future research avenues.

II. Configuration and Scattering Equations

Consider the configuration in Fig. 1 in which a planar interface separates air (medium 1) from soil (medium 2). The orthogonal coordinate system $\mathbf{r} = \mathbf{x} + \hat{z}z$ is defined, with $\mathbf{x}$ indicating a two-dimensional (2-D) coordinate on the interface and the z axis directed so that $z > 0$ is air and $z < 0$ is soil. Buried in the soil is a known target object, but its coordinates $\mathbf{r}_c = \mathbf{x}_c + \hat{z}z_c$ with $z_c \leq 0$ are unknown. A monochromatic plane wave with electric field $\mathbf{E}^i(\mathbf{r}) = \mathbf{E}_0 e^{i\mathbf{k}_1 \cdot \mathbf{r}}$ and wavevector $\mathbf{k}_1$ is incident upon the planar interface. As the incident field reaches the interface, it partially reflects back into air and partially refracts into the soil where it interacts with the buried object. The interaction produces a scattered field, part of which refracts back into air where it is measured.

The total field $\mathbf{E}(\mathbf{r};\mathbf{r}_c)$ at position $\mathbf{r}$ is given by

$$\mathbf{E}(\mathbf{r};\mathbf{r}_c) = \mathbf{E}^b(\mathbf{r}) + \mathbf{E}^s(\mathbf{r};\mathbf{r}_c) \tag{1}$$

where $\mathbf{E}^b(\mathbf{r})$ is the electric field in the absence of the buried object (background field) consisting of the incident and the reflected field in air and the refracted field in soil and $\mathbf{E}^s(\mathbf{r};\mathbf{r}_c)$ is the scattered field due to the presence of the buried object at the unknown location$\mathbf{r}_c$

$$\mathbf{E}^s(\mathbf{r};\mathbf{r}_c) = \begin{cases} \mathbf{E}_2^s(\mathbf{r};\mathbf{r}_c) \equiv \displaystyle\int_{z'<0} \overline{\mathbf{G}}_{22}(\mathbf{r},\mathbf{r}') \cdot [\mathbf{E}^b(\mathbf{r}') \\ \qquad +\mathbf{E}_2^s(\mathbf{r}';\mathbf{r}_c)] O(\mathbf{r}';\mathbf{r}_c)\, d^3r', & z < 0 \text{ (soil)} \\ \mathbf{E}_1^s(\mathbf{r};\mathbf{r}_c) \equiv \displaystyle\int_{z'<0} \overline{\mathbf{G}}_{12}(\mathbf{r},\mathbf{r}') \cdot [\mathbf{E}^b(\mathbf{r}') \\ \qquad +\mathbf{E}_2^s(\mathbf{r}';\mathbf{r}_c)] O(\mathbf{r}';\mathbf{r}_c)\, d^3r', & z > 0 \text{ (air)}. \end{cases} \tag{2}$$

In (2), the object function $O(\mathbf{r};\mathbf{r}_c)$ is defined as

$$O(\mathbf{r};\mathbf{r}_c) = k^2(\mathbf{r};\mathbf{r}_c) - k_2^2 \tag{3}$$

where $k(\mathbf{r};\mathbf{r}_c)$ is the complex wavenumber at point $\mathbf{r}$ and k_2 is the complex wavenumber of the soil. Equation (2) is the Lippmann–Schwinger equation for the scattering problem and maps the object function to the corresponding scattered field nonlinearly.

The dyadic Green functions $\overline{\mathbf{G}}_{12}$ and $\overline{\mathbf{G}}_{22}$ are defined for the source point in soil and the observation point in air and soil, respectively. Explicit expressions for them can be found in [2]; however, the critical information for this context is that both Green

0018–926X/99$10.00 © 1999 IEEE

Analysis of the Electromagnetic Inductive Response of a Void in a Conducting-Soil Background

Tiejun Yu, *Member, IEEE,* and Lawrence Carin, *Senior Member, IEEE*

***Abstract*—A lossless dielectric object situated in a lossy dielectric medium (soil) constitutes a void in a conducting background, which can be detected via an electromagnetic-induction (EMI) sensor operating at appropriate frequencies. The electromagnetic character of this void is dependent on the target and soil properties, as well as on the frequency of operation. We utilize the rigorous method of moments (MoM) and the approximate extended-Born technique to model this three-dimensional (3-D) problem. The modeling algorithms are discussed in detail, with a focus on efficient computation of the dyadic Green's function at the frequencies of interest. The MoM results are used to calibrate the accuracy of the approximate extended-Born solution, over a wide range of operating conditions. Furthermore, the computer simulations are used to perform a detailed phenomenological study.**

***Index Terms*—Electromagnetic induction, extended-Born, scattering.**

I. Introduction

THERE are many applications for which one may be interested in sensing a low-loss dielectric target embedded in a lossy host medium. For example, buried plastic land mines [1]–[4], plastic pipes [5], and underground tunnels generally constitute voids in the presence of a lossy soil background. While such targets can in principle be detected by a radar sensor [1]–[5], radar suffers well-known difficulties due to soil-induced attenuation [1]. An electromagnetic-induction (EMI) sensor [6], [7] on the other hand, affords the potential for significant soil penetration. Such sensors are typically used as metal detectors, with example applications including detection of buried *conducting* land mines and unexploded ordnance [8]–[10]. As elucidated here, such sensors can also be used to detect the *absence* of conductivity (a void) in the presence of a conducting host medium. Our focus here is on generally small, shallow targets (e.g., plastic mines), although the computer model is quite general. For such a problem, while the target's electrical properties cannot be changed, those of the host soil can. We therefore address the character of the EMI conductivity-void signature as a function of soil water content. The objective is to increase the soil conductivity by increasing the soil water content [1], thereby enhancing the conductivity contrast between the plastic target and soil, yielding an increased target signature. Clearly, many types of clutter may also manifest a void in a conducting background (e.g., tree roots, rocks, etc.). Therefore, the objective is to use the EMI-based void detector as a queuing device for other devices such as chemical sensors [11], these also being applicable to land mines with low or no metal content. Furthermore, although not pursued here, our forward model could be applied in the context of inverse scattering [6] to yield improved information on the details of the subsurface target (void).

The rigorous modeling of EMI interaction with a buried void in a conducting host medium presents several challenges. Focusing on a frequency-domain, integral-equation analysis [2], [12]–[15], one must carefully evaluate the dyadic layered-medium Green's function [2], [12]–[15] for a layered-medium soil model. For radar-based problems, the method of complex images [2], [14], [15] has proven a very efficient and accurate tool for the analysis of the requisite Sommerfeld integrals [2], [12]–[15]. However, we have found this technique less useful for the EMI frequencies of interest here. Consequently, we have employed efficient interpolation-based Green's-function computation, similar in spirit to that presented in [12].

Having realized an efficient and accurate technique for Green's-function computation, we are still left with the problem of solving the integral equation. We have here applied a volumetric electric field integral equation (EFIE) [12], in which the unknowns are the total electric fields inside the dielectric target for a known heterogeneity between the target and the background (soil) medium [12]. Such an EFIE is often solved via a volumetric method of moments (MoM) solution [12]. While a MoM analysis is accurate, it generally requires the inversion of a large matrix. This task is circumvented by employing the extended-Born method [16], [17]. This is an approximate method that only requires inversion of simple 3×3 matrices. The attendant speed enhancement accrued by extended-Born, *vis-a-vis* the MoM, is critical if the model is to be utilized in the context of signal processing and/or inverse scattering [18]. An example of employing sophisticated scattering models in the context of target detection and classification can be found in [1].

While the reduced computational complexity of extended-Born is attractive, it is important that the results not suffer markedly in accuracy. Therefore, in the work reported here, we perform a detailed frequency-dependent comparison between the rigorous MoM and approximate extended-Born solutions. This test is performed using parameters from an actual EMI sensor [7] for several different target and soil scenarios. In the context of these comparisons, we also investigate the phenomenology underlying EMI-based buried-void detection.

Manuscript received June 4, 1999; revised November 30, 1999. This work was supported in part by the U.S. Joint Unexploded Ordnance Coordination Office, Adelphi, MD, and in part by the DoD Multidisciplinary University Research Initiative (MURI) in demining.

The authors are with Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0196-2892(00)03926-7.

0196–2892/00$10.00 © 2000 IEEE

The remainder of the paper is organized as follows. In Section II, we review the electric-field integral equation characteristic of the buried-void problem. We discuss the integral-equation solution in the context of both the rigorous MoM and the approximate extended-Born methods. In this section, we also elucidate the technique applied for Green's-function evaluation. A detailed comparison is performed in Section III between the MoM and extended-Born solutions, wherein we also address the associated wave phenomenology. Finally, conclusions are addressed in Section IV.

II. Theory

A. Electric Field Integral Equation and MoM Solution

The volumetric EFIE for the layered-earth model has been derived in [12], and therefore here is only provided a summary. Assume that $\boldsymbol{E}(\boldsymbol{r})$ represents the electric field at the position $\mathbf{r}$ and that $\boldsymbol{E}^{inc}(\boldsymbol{r})$ represents the incident fields produced by the excitation in the absence of the target. Using $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$ to denote the dyadic, layered-medium Green's function, we have [12]

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}^{inc}(\boldsymbol{r}) - j\omega \int_V [\varepsilon_t(\boldsymbol{r}') - \varepsilon_b(\boldsymbol{r}')]\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}') \cdot \boldsymbol{E}(\boldsymbol{r}')d^3\boldsymbol{r}' \tag{1}$$

where an $\exp(j\omega t)$ time dependence is assumed and suppressed. In (1), ε_b represents the inhomogeneous dielectric constant of the background layered medium, and ε_t represents the dielectric constant of the generally inhomogeneous target, with the latter assumed (for numerical simplicity) to reside entirely within a single layer of ε_b. In general, the layered background medium and the target are lossy and therefore, both ε_b and ε_t are complex [19].

For a known excitation $\boldsymbol{E}^{inc}(\boldsymbol{r})$, (1) implies that knowledge of the total electric field inside the target can be used to calculate $\boldsymbol{E}(\boldsymbol{r})$ everywhere. Therefore, the problem reduces to solving for the total electric field $\boldsymbol{E}(\boldsymbol{r})$ for $\boldsymbol{r}$ inside the target volume V. In the MoM solution, we expand the unknown $\boldsymbol{E}(\boldsymbol{r})$ for $\boldsymbol{r} \in V$ in terms of known basis functions with unknown coefficients

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{x}\sum_{n=1}^{N} a_n e_{xn}(\boldsymbol{r}) + \boldsymbol{y}\sum_{n=1}^{N} b_n e_{yn}(\boldsymbol{r}) + \boldsymbol{z}\sum_{n=1}^{N} c_n e_{zn}(\boldsymbol{r}) \tag{2}$$

where

$\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$	unit vectors in Cartesian coordinates;
a_n, b_n, and c_n	unknown basis-function coefficients;
e_{xn}, e_{yn}, and e_{zn}	known basis functions for the electric field in the x, y and z directions, respectively.

As in [12], we utilize simple cube expansion functions for all e_{xn}, e_{yn} and e_{zn}.

Employing (2) inside the integral in (1) and applying point matching [20] inside each of the $3N$ basis-functions (cubes), we attain the MoM matrix equation

$$\boldsymbol{Z}\boldsymbol{i} = \boldsymbol{v} \tag{3}$$

where

$\boldsymbol{Z}$	$3N \times 3N$ MoM "impedance" matrix;
$\boldsymbol{i}$	$3N \times 1$ vector representing the unknown coefficients a_n, b_n, and c_n;
$\boldsymbol{v}$	$3N \times 1$ vector representing the point-matched incident electric field (at the center of each cubic basis function).

One can use (3) to solve for the unknown i, this constituting the procedure applied in [12]. The MoM, representing a rigorous analysis, constitutes a reference solution. However, we are also interested in developing a more efficient, albeit approximate, solution. This is motivated by the fact that an LU-decomposition solution of (3) is of order N^3 complexity, while an iterative, conjugate-gradient (CG) solution [21] has complexity of order PN^2, where P is the number of CG iterations (P may approach N if a good preconditioner is not used). For large N, such computations can become prohibitive. Independent of the matrix solver, order N^2 complexity is required to fill the matrix $\boldsymbol{Z}$. Some of the complexity of the MoM solution can be reduced by exploiting convolutional properties of the integral equation [22] through use of the fast Fourier transform (FFT).

B. Extended-Born Solution

The extended-Born method is an approximate technique first developed in [16], [17]. Here we extend this algorithm to three-dimensional layered-media, for a realistic excitation. The algorithm is summarized as follows. The "extended-Born" approximation [16], [17] is based on recognizing that $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$ is large for $\boldsymbol{r}$ in the vicinity of $\boldsymbol{r}'$, while being relatively small for $\boldsymbol{r}$ distant from $\boldsymbol{r}'$. Consequently, from (1) we have

$$\boldsymbol{E}(\boldsymbol{r}) \approx \boldsymbol{E}^{inc}(\boldsymbol{r}) - j\omega \int_V [\varepsilon_t(\boldsymbol{r}') - \varepsilon_b(\boldsymbol{r}')]\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}') \cdot \boldsymbol{E}(\boldsymbol{r})\, dv' \quad \text{for } \boldsymbol{r} \in V. \tag{4}$$

Rearranging (4), we find

$$\boldsymbol{E}(\boldsymbol{r}) \approx \boldsymbol{M}(\boldsymbol{r})^{-1} \cdot \boldsymbol{E}^{inc}(\boldsymbol{r}) \qquad \text{for } \boldsymbol{r} \in V$$
$$\boldsymbol{M}(\boldsymbol{r}) \equiv \boldsymbol{I} + j\omega \int_V \boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')[\varepsilon(\boldsymbol{r}') - \varepsilon_t(\boldsymbol{r}')]d^3\boldsymbol{r}' \tag{5}$$

where $\boldsymbol{I}$ is a 3×3 diagonal matrix with $\boldsymbol{I}_{kk} = 1$. The expression in (5) yields an approximation for the electric fields inside the target, which can be used in (1) to find the fields everywhere (i.e., for $\boldsymbol{r} \notin V$). Note that the inverse of $\boldsymbol{M}$ approximately accounts for multiple interactions, yielding a nonlinear algorithm in $\varepsilon_t(\boldsymbol{r}) - \varepsilon_b(\boldsymbol{r})$. Note that the aforementioned property of $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$ driving the extended-Born solution (namely, that $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$ is large for $\boldsymbol{r}$ in the vicinity of $\boldsymbol{r}'$) is true for general Green's functions, including that for the layered-media considered here. Moreover, the approximation is enhanced if the medium in which $\boldsymbol{r}'$ resides is lossy, such as for a target buried in lossy soil.

If the integral term in (5) is ignored, extended-Born reduces to the classical Born approximation. Moreover, the extended-Born approximation involves the inversion of a 3×3 matrix $\boldsymbol{M}(\boldsymbol{r})$ for each $\boldsymbol{r} \in V$. These matrices are computed by representing $\varepsilon_t(\boldsymbol{r}) - \varepsilon_b(\boldsymbol{r})$ as an aggregate of N three-dimensional (3-D) cubes [16], [17], similar to (2), except now the dielectric inhomogeneity is a scalar, rather than vector electric field. We compute $\boldsymbol{M}(\boldsymbol{r})$ at the $N\boldsymbol{r}_n$ that constitute the centers of the cubes. Consequently, we need to invert $N3 \times 3$ matrices, constituting order N complexity. Since we need only store and then discard a distinct 3×3 matrix for each of the N points inside the target,

memory requirements of extended Born are also substantially less than those of the MoM.

C. Green's Function Evaluation

As described in [12], [13], each term in the dyadic Green's function can be expressed in terms of functions of the form

$$S_n[f(k_\rho;\, z,\, z')] = \int_0^\infty f(k_\rho;\, z,\, z') J_n(k_\rho \rho) k_\rho^{n+1}\, dk_\rho \tag{6}$$

where the path of integration is the traditional Sommerfeld contour [12], [13]. Each dyadic component generally consists of a "direct" term, corresponding to source radiation in a homogeneous medium, plus a "reflection" term, corresponding to interaction with the layer boundaries. Each direct term is integrated in closed form using the Sommerfeld or Weyl identity [23]

$$\frac{e^{-jkr}}{r} = -j \int_0^\infty dk_\rho k_\rho J_o(k_\rho \rho) \frac{\exp\left(-j\sqrt{k^2 - k_\rho^2}\,|z|\right)}{\sqrt{k^2 - k_\rho^2}}. \tag{7}$$

A useful tool for integration of the remaining "reflection" term is the complex image technique [2], [14], [15]. While this has proven a very powerful tool for radar applications [2], [15], we have found it much less reliable for the EMI frequencies (kHz) of interest here.

To circumvent these difficulties, we have employed a technique related to that discussed in [12]. In particular, note from (6) that the (generally complicated) functions $f(k_\rho;\, z,\, z')$ are independent of the radial distance ρ, the latter accounted for by the Bessel function. We therefore precompute the required $f(k_\rho;\, z,\, z')$ for the "reflection" term of each dyadic component for the range of z and z' of interest. This tabulation is performed along a path in the complex k_ρ plane [24], along which the $f(k_\rho;\, z,\, z')$ are generally slowly varying (see Fig. 1), thereby reducing the number of discrete complex k_ρ that need be considered to accurately represent $f(k_\rho;\, z,\, z')$. This deformed path of integration therefore reduces the required spectral-domain tabulation. The spectral-domain table thus established, we now evaluate the requisite integrals of the form in (6) for the range of ρ required (dictated by the target size). We thereby generate a table of the dyadic Green's function components in the $(z,\, z',\, \rho)$ domain, discarding the $f(k_\rho;\, z,\, z')$ one-by-one, as they are no longer needed. Dividing Green's function tabulation into a separate, initial spectral-domain computation, followed by conversion to the space domain, results in significant savings, for we need only calculate the requisite $f(k_\rho;\, z,\, z')$ once for a given z and z' (for all ρ). Interpolation in the $(z,\, z',\, \rho)$ space is performed as in [12].

III. Example Results

A. Preliminaries

All computations are performed using parameters characteristic of an actual wideband EMI sensor, termed the GEM-3 [7]. The parameters of the GEM-3 are detailed in Fig. 2. The two excitation loops are designed with radii and currents such that the normal component of the magnetic field vanishes along the coil axis (if the sensor is in a vacuum). A third, smaller loop is placed along the sensor axis. This is used to measure the induced voltage. Consequently, the induced voltage is ideally zero unless the sensor is placed in the vicinity of a field-disturbing medium (e.g., a soil half space and possibly a buried target). In the computations presented here, the incident fields due to the two current loops, in the presence of a lossy layered medium, are computed rigorously by integrating the loop currents with the dyadic layered-medium Green's function [12], [13]. Moreover, the induced voltage is computed via Faraday's law as

$$V = -j\omega \int_S \boldsymbol{B} \cdot \boldsymbol{n}\, dS \tag{8}$$

where integration is performed over the planar surface S of the inner loop (with unit normal $\boldsymbol{n}$). The induced magnetic fields $\boldsymbol{B}$ are computed by integrating the electric fields induced inside the target (computed via MoM or extended-Born) with the dyadic Green's function. In general, the currents on the excitation loops change in the presence of the induced EMI fields (due to the presence of a target), although this relatively small effect is ignored here.

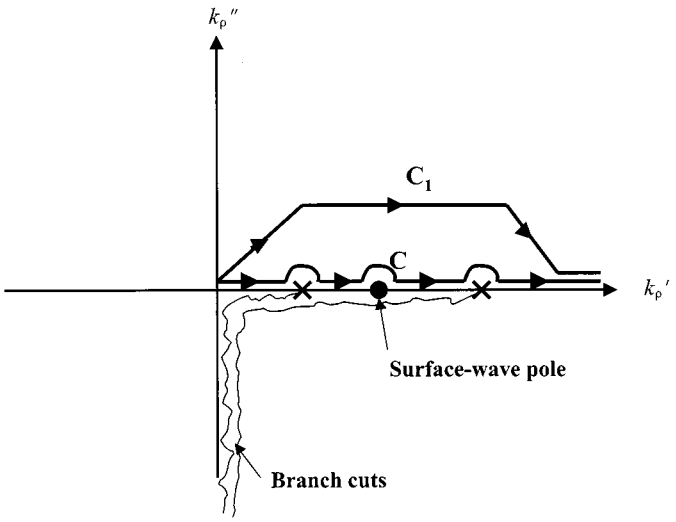


Fig. 1. Schematization of the path of integration in the complex k_ρ plane ($k_\rho = k_\rho' + jk_\rho''$). Contour C is the original path of integration, with consideration of the branch points and surface-wave poles (for layered medium). Contour C_1 is chosen such that the poles and branch points are avoided, and the variation of the spectral domain Green's function is relatively smooth. The left branch cut emanates from branch point k_o and the right from $k_o(\varepsilon_{rN})^{1/2}$, where k_o is the free-space (top layer) wavenumber, and ε_{rN} is the relative dielectric constant of the bottom-most layer in a multilayered environment.

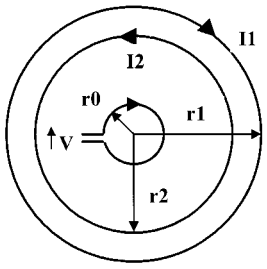


Fig. 2. Schematization of the current loops employed by the GEM-3 EMI sensor [7]. The radii are $R0 = 4$ cm, $R1 = 26$ cm, and $R2 = 15.84$ cm, and the currents are $I1 = 10$ A and $I2 = 6$ A. The loop of radius $R0$ is wound 600 times.

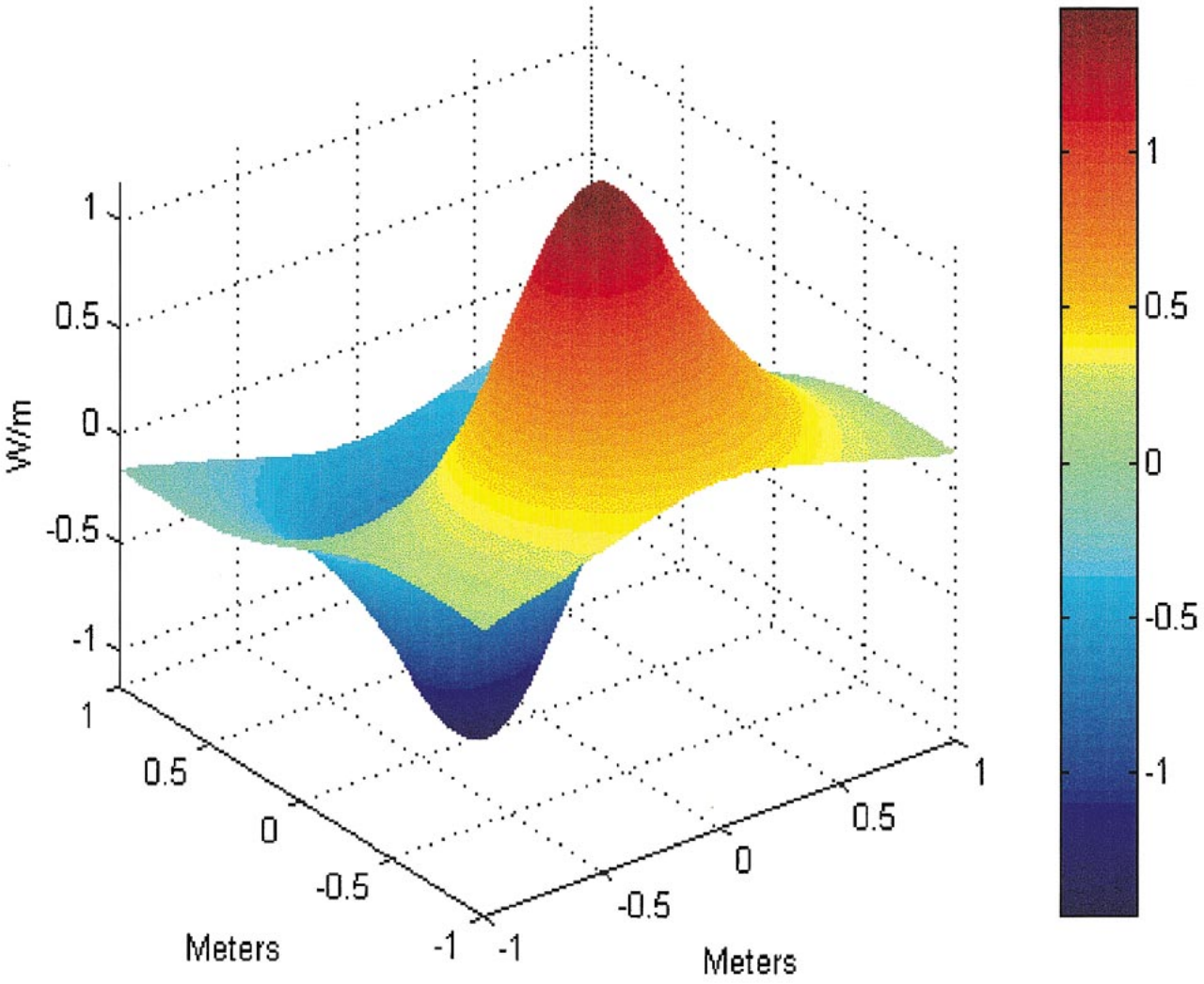


Fig. 3. Space-dependent electric fields $\boldsymbol{E}_y$ (V/m) 10 cm beneath and parallel to the surface of the GEM-3 sensor. The spatial dimensions are meters, and the imaginary part of $\boldsymbol{E}_y$ is plotted. The real part is similar. Note that, as designed, the GEM-3 fields are very small along the sensor axis (origin of this plot).

In the subsequent set of results, we consider several target/soil scenarios, with the induced electric fields inside the target computed by both the extended-Born (approximate) and MoM (reference) methods. The results of the MoM allow quantitative assessment of the extended-Born accuracy for the problem of interest here. Further, the set of examples cover a range of wave phenomenology.

Before proceeding to the results, we note a numerical issue that arose while addressing the computations. It is well known that, for subsectional basis functions, we typically require ten basis functions per wavelength to accurately solve integral equations of the type considered here. For EMI problems, the wavelength is generally much larger than the target dimensions, so this condition on the number of basis functions does not present an issue. From Section II, however, it is also clear that the spatial discritization must be sufficiently fine to accurately represent the fields excited by the EMI sensor $\boldsymbol{E}^{inc}$ in the absence of a target. As shown in Fig. 2, the GEM-3 has two excitation coils, which generate fast field gradients at positions away from the loop axis. This issue must be considered carefully when discretizing the target. The gradients in $\boldsymbol{E}^{inc}$ are most significant for shallow targets (near the excitation loops), with such becoming less severe for deeper targets. An example of $\boldsymbol{E}^{inc}$ at a depth of 10 cm is shown in Fig. 3 (at 500 KHz). In the results presented here, we systematically increased the number of subsectional basis functions until numerical convergence was achieved.

B. Single Target

In our first set of results, we consider the GEM-3 response due to a fixed buried target as a function of the soil properties and the EMI frequency. With regard to the soil properties, measurements have been performed for soil samples taken from the Yuma Proving Ground, Yuma, AZ [1]. These measurements are extrapolated to the EMI frequencies of interest here. We have found that the frequency-dependent relative permittivity is well modeled by $\varepsilon_r = \varepsilon' - j\varepsilon'' - j\sigma/\omega\varepsilon_o$, where ε' and ε'' are constants, and σ is the conductivity (S/m). The conductivity term accounts for the frequency dependence of the complex ε_r. The

TABLE I
SOIL PROPERTIES EXTRAPOLATED FROM DATA IN [1] FOR SOIL FROM YUMA, AZ. THE RELATIVE DIELECTRIC CONSTANT IS EXPRESSED AS $\varepsilon_r = \varepsilon_r' - j\varepsilon_r'' - j\sigma/\omega\varepsilon_o$, WHERE ε_r' AND ε_r'' ARE CONSTANTS AND σ IS THE CONDUCTIVITY (S/m). THE ANGULAR FREQUENCY IS REPRESENTED BY ω. THE DATA IS TABULATED AS A FUNCTION OF THE SOIL MOISTURE CONTENT, BY PERCENTAGE WEIGHT

Water Content	ϵ_r''	ϵ_r''	σ (S/m)
5 %	5.0	0.2	0.01
10 %	8.5	0.5	0.03
15 %	14.5	0.9	0.06
20 %	24	1.5	0.1

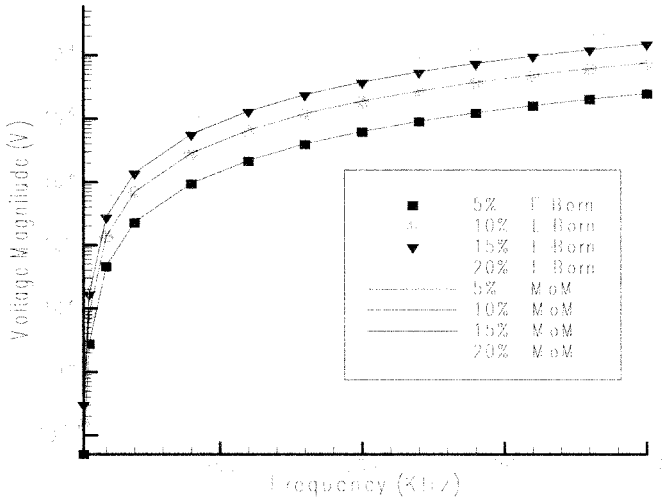


Fig. 4. Magnitude of the voltage induced by a lossless cube target with 16 cm sides and dielectric constant $\varepsilon_r = 2.0$, buried to a depth of 12 cm between the air-ground interface and the target's top surface. The GEM-3 coils are parallel to the soil interface at a height of 4 cm above the soil, with the sensor axis oriented through the target center. Results are plotted for various soil types, distinguished by the percentage of water content by weight (see Table I). The curves represent the MoM solution, and the points are the results of the extended-Born method.

ε', ε'' and σ used here are tabulated in Table I as a function of the soil moisture content (percentage water by weight).

In Fig. 4, we plot the frequency-dependent GEM-3 response due to a lossless cube target of 16 cm sides and dielectric constant $\varepsilon_r = 2.0$, buried to a depth of 12 cm between the air-ground interface and the target's top surface. The GEM-3 coils are parallel to the soil interface, at a height of 4 cm above the soil, with the sensor axis oriented through the target center. In Fig. 4, the curves represent the results of the MoM, where the points denote computations from the extended-Born method. The results in Fig. 4 are for the magnitude of the voltage. For these results, the imaginary component of the voltage is two orders of magnitude below its real counterpart. The three principal observations from Fig. 4 are the following.

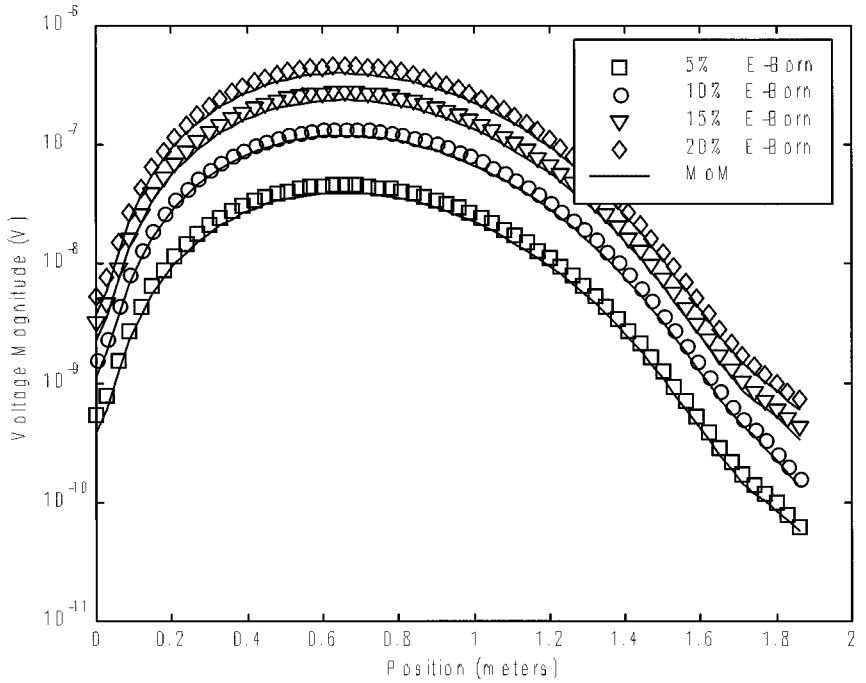
1) The induced voltage increases with EMI frequency.
2) The extended-Born method yields results that are in close accord with the reference MoM results.
3) The increased target-soil contrast with increasing water content yields the expected increase in the target signature.

Note that at 20% water content (see Table I), there is a substantial contrast between the target and soil, but the extended-Born method still yields quite accurate results. The extended-Born solutions are also significantly more efficient than the MoM computations. For these computations, we utilized $N = 512$ cubes (1536 unknowns). For computation of the electric fields induced inside the target, less than 1 minute was required for the extended-Born solution, while greater than 12 min were required for the MoM computations. All computations reported here were performed on a typical Pentium II personal computer. For both the MoM and extended-Born solutions, we precomputed the dyadic Green's function table (see Section II-C), and this required approximately 5 min of CPU time on a personal computer.

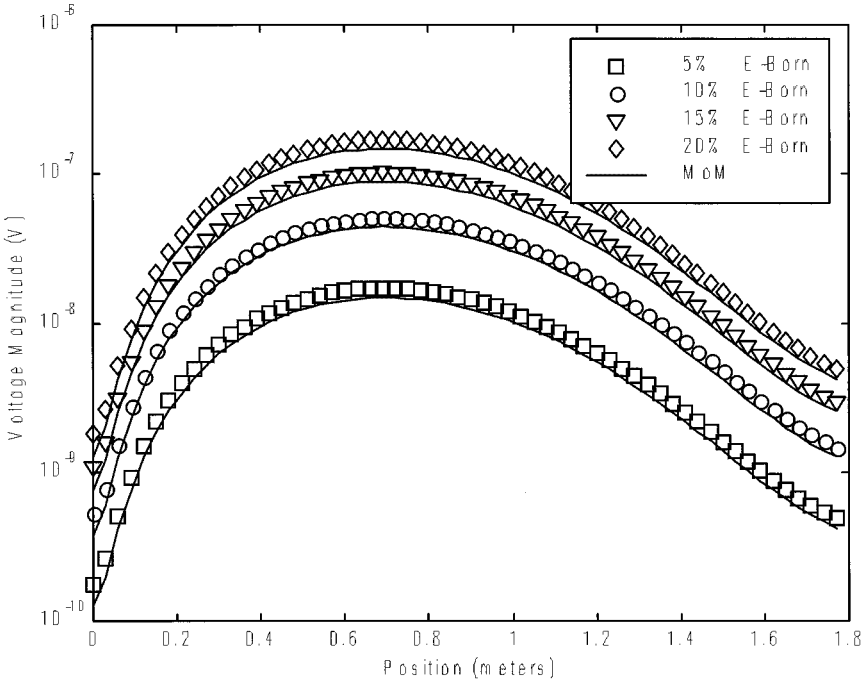
For classification of the subsurface target, the spatial variation of the EMI signature is also of importance. In Fig. 5(a) and (b), we plot the space-dependent EMI signature of the GEM-3 sensor at 10 kHz for target depths of 25 and 45 cm. The sensor coils are parallel to the air-ground interface, at a height of 2 cm above the soil. For this example, we consider a lossless spherical target of 10 cm diameter and $\varepsilon_r = 3.5$, and the target depth is defined relative to the target center. As before, the solid curves represent the MoM results, while the points denote the extended-Born solution. These results were computed using $N = 257$ cubes, and the relative computational requirements of the MoM and extended-Born results were similar to those required for Fig. 4. The voltage magnitude is plotted, and again, the imaginary part of the voltage is two orders of magnitude smaller than that of the real part. For the target considered here, we note that the maximum signal response does not occur when the sensor axis is aligned with the target center. This is attributed to the details of the fields emitted by the sensor (see Fig. 2). In particular, the maximum excitation fields $\boldsymbol{E}^{inc}$ occur slightly away from the sensor center, as dictated by the sensor design (dictated by the goal of achieving a null voltage when the sensor is in a vacuum). Fig. 5 indicates the expected reduction in the sensor response with increased target depth. Moreover, the discrepancies between the MoM and extended-Born solutions are more apparent in the space-dependent results of Fig. 5. In particular, note that the extended-Born error increases as the soil moisture content increases. This is representative of an increased contrast between the target and background. Nevertheless, the extended-Born results faithfully reproduce the principal characteristics of the EMI signature.

C. Multiple Targets

The MoM and the extended-Born software are easily adapted to targets of arbitrary shape, including composite targets. In the next set of results, we consider two distinct objects in the sensor's field of view. As indicated in Fig. 6, we consider the GEM-3 response due to proximate square and cylindrical targets, with the sensor operating at 5 MHz (this frequency is chosen arbitrarily). For this example, the complex dielectric constant of the soil $\varepsilon_r = \varepsilon' - j\varepsilon'' - j\sigma/\omega\varepsilon_o$ is characterized by $\varepsilon' = 5$, $\varepsilon'' = 0$, and $\sigma = 0.02$ S/m, and both lossless targets have a real dielectric constant $\varepsilon_r = 2.25$. The space-dependent EMI responses, as computed via MoM and extended-Born, are shown in Fig. 7. We note that the MoM and extended-Born so-



(a)



(b)

Fig. 5. Space-dependent EMI signature of the GEM-3 sensor at 10 KHz, for a lossless spherical target of 10 cm diameter and $\varepsilon_r = 3.5$. The sensor is 2 cm above the air-ground interface, and the target depth is defined relative to the target center. Results are plotted for various soil types, distinguished by the percentage of water content by weight (see Table I). The curves represent the MoM solution, and the points are the results of the extended-Born method: (a) 25-cm deep target and (b) 45-cm deep target.

lutions are in close agreement. Furthermore, the geometrically larger cylindrical target yields an EMI signature of considerably larger amplitude than the smaller square target. In Fig. 7, we again plot the magnitude of the EMI voltage, the imaginary component of which is again two orders of magnitude smaller than the real component. Similar space-dependent signatures and agreement between MoM and extended-Born solutions was found for all multiple-target tests we considered, although space limitations preclude demonstration of all such examples.

IV. Conclusions

A rigorous MoM and approximate extended-Born analysis have been developed for the EMI response of a buried low-loss

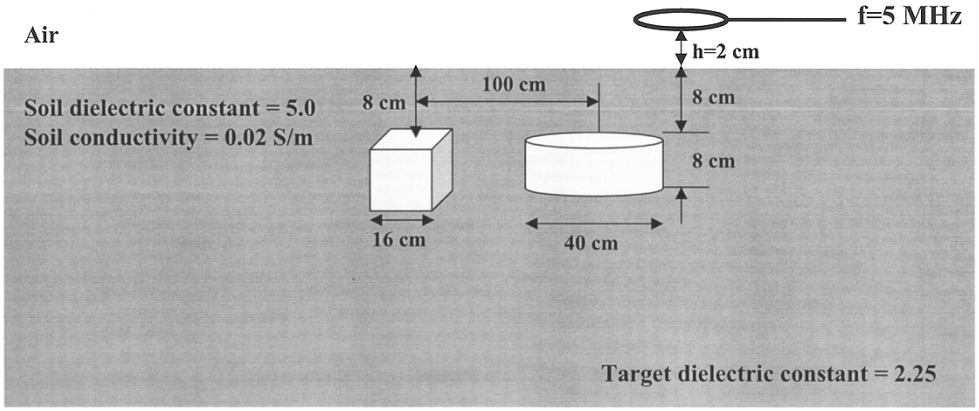


Fig. 6. Schematization of two buried lossless dielectric targets. The sensor is deployed at a height of 2 cm above the interface at a frequency of 5 MHz.

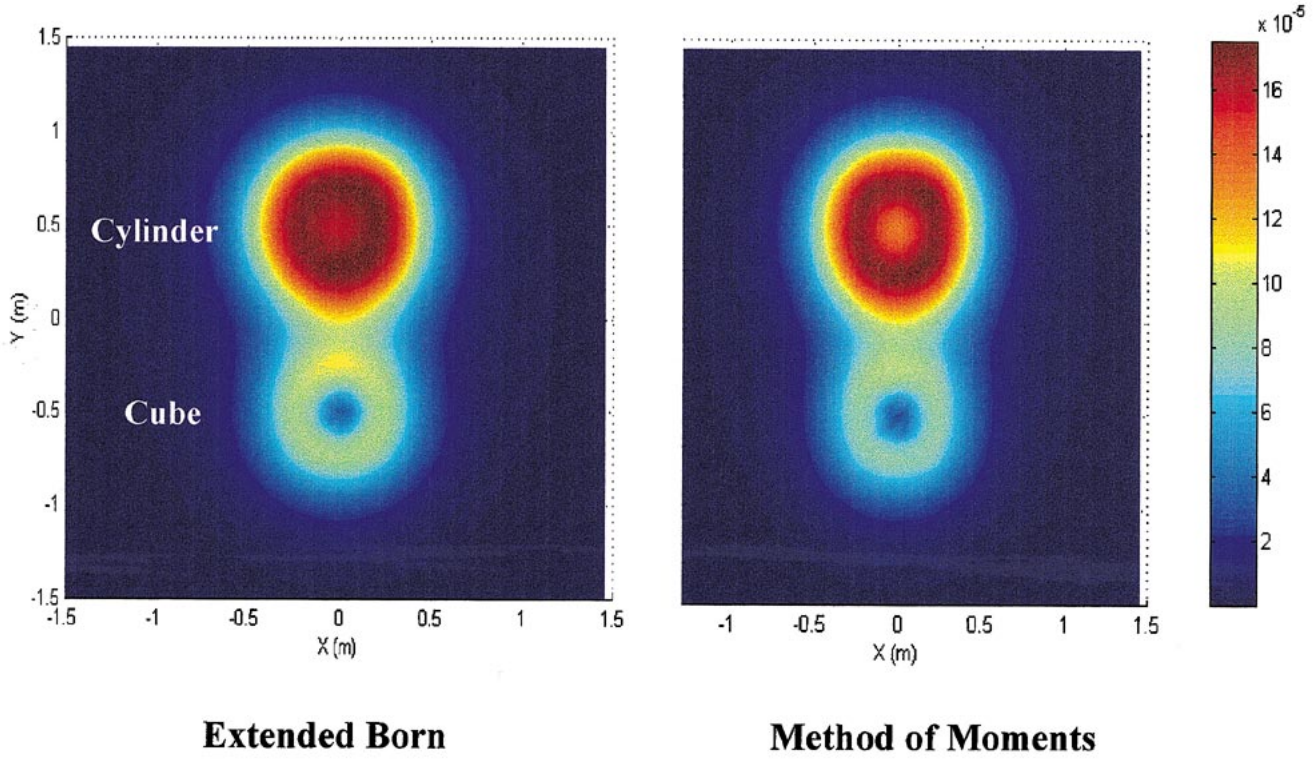


Fig. 7. Space-dependent induced voltage for the two targets and sensor parameters considered in Fig. 6. The left result is produced by the extended-Born method and the right by MoM.

target embedded in a lossy half space (soil). This research is applicable to many scenarios, in which an EMI sensor may be deployed to sense low-metal-content buried targets (e.g., buried plastic land mines). It has been demonstrated that the highly efficient extended-Born method consistently yields space- and frequency-dependent results that are in close agreement with the reference MoM solution. While the properties of the buried target cannot be changed, those of the soil can. In particular, we have demonstrated the expected strong enhancement of the EMI signature with an increase in the soil water content. It is important to emphasize that this phenomenon (increased water content) is also expected to enhance the electrical contrast for radar applications. However, unlike EMI, the radar excitation is generally severely attenuated as the soil water content increases [1].

We note from Table I that relatively strong electrical contrasts have been considered here and that for such cases, the extended-Born method has been in good agreement with the MoM reference solution. We attribute this accuracy to the loss in the background soil (from Table I for the cases examined, the conductivity increases with increasing soil dielectric constant), which enhances the accuracy of the approximation to $\boldsymbol{G}(\boldsymbol{r}, \boldsymbol{r}')$, discussed in Section II-B. Nevertheless, extended-Born is an approximation, and it will breakdown as the contrast and target electrical size increases. In this paper, we have sought to consider soil contrasts and target sizes consistent with what may be found in practice, for which the extended-Born method appears to be a particularly attractive solution method.

Having demonstrated the accuracy of the extended-Born algorithm for this problem class, there are several directions for

future research. For example, we are investigating the use of the extended-Born method in inverse-scattering problems [18]. Furthermore, the extended-Born computed EMI signatures are also appropriate for an optimal Bayesian detection algorithm [1].

REFERENCES

[1] L. Carin, N. Geng, M. McClure, J. Sichina, and L. Nguyen, "Ultra-wide-band synthetic-aperture radar for mine-field detection," *IEEE Antennas Propagat. Mag.*, vol. 41, pp. 18–33, Feb. 1999.

[2] N. Geng and L. Carin, "Wideband electromagnetic scattering from a dielectric body of revolution buried in a layered-medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.

[3] L. Peters Jr., J. J. Daniels, and J. D. Young, "Ground penetrating radar as an environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.

[4] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of ground penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 28–36, Jan. 1996.

[5] X. X. Zeng and McMechan, "AGPR characterization of buried tanks and pipes," *Geophysics*, vol. 62, pp. 797–806, May/June 1997.

[6] A. Witten, I. J. Won, and S. J. Norton, "Subsurface imaging with broadband electromagnetic induction," *Inv. Prob.*, vol. 13, pp. 1621–1639, Dec. 1997.

[7] I. J. Won, D. A. Keiswetter, and D. R. Hanson, "GEM-3: A monostatic broadband electromagnetic induction sensor," *J. Environ. Eng. Geophys.*, vol. 2, pp. 53–64, Mar. 1997.

[8] Y. Das, J. E. McFee, J. Toew, and G. C. Stuart, "Analysis of an electromagnetic induction detector for real-time location of buried objects," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 278–288, May 1990.

[9] Y. Das and J.E. McFee, "A simple analysis of the electromagnetic response of buried conducting objects," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 342–344, Mar. 1991.

[10] N. Geng, C. E. Baum, and L. Carin, "On the low-frequency natural response of conducting and permeable targets," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 347–359, Jan. 1999.

[11] M. S. Freund and N. S. Lewis, "A chemically diverse conducting polymer-based electronic nose," in *Proc. Nat. Acad. Sci. U.S.A.*, vol. 92, 1995, p. 2652.

[12] P. E. Wannamker, G. W. Hohmann, and W. A. SanFilipo, "Electromagnetic modeling of three-dimensional bodies in layered earths using integral equations," *Geophys.*, vol. 49, pp. 60–74, 1984.

[13] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media—Parts I and II," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–352, March 1990.

[14] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. AP-28, pp. 143–151, Feb. 1996.

[15] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.

[16] T. M. Habashy, R. W. Groom, and B. R. Spies, "Beyond the born and Rytov approximations—A nonlinear approach to electromagnetic scattering," *J. Geophys. Res. Solid Earth*, vol. 98, pp. 1759–1775, Feb. 1993.

[17] C. Torres-Verdin and T. M. Habashy, "Rapid 2.5-dimensional forward modeling and inversion via a new nonlinear scattering approximation," *Radio Sci.*, vol. 29, pp. 1051–1079, Jul./Aug. 1994.

[18] E. L. Miller and A. S. Willsky, "Wavelet-based methods for the nonlinear inverse scattering problem using: The extended born approximation," *Radio Sci.*, vol. 31, pp. 51–65, Jan./Feb. 1996.

[19] C. A. Balanis, *Advanced Engineering Electromagnetics*. New York: Wiley, 1989.

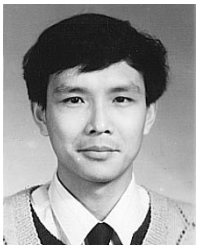
[20] R. F. Harrington, *Field Computation by Moment Methods*. New York: IEEE Press, 1993.

[21] T. K. Sarkar, X. P. Yang, and E. Arvas, "A limited survey of various conjugate-gradient methods for solving complex matrix equations arising in electromagnetic-wave interactions," *Wave Motion*, vol. 10, pp. 527–546, Dec. 1988.

[22] T. J. Cui and W. C. Chew, "Fast algorithm for electromagnetic scattering by buried 3-D dielectric objects of large size," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2597–2608, Sept. 1999.

[23] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. Oxford, U.K.: Oxford Univ. Press, 1996.

[24] K. A. Michalski, "On the efficient evaluation of integrals arising in the Sommerfeld halfspace problem," *Proc. Inst. Elect. Eng. F*, vol. 132, pp. 312–318, Aug. 1985.



Tiejun Yu (M'99) was [redacted] He received the B.S. degree in electrical engineering from the National Defense University of Technology, Changsha, China, in 1988, and the M.S.E. and Ph.D. degrees in electronic engineering,, both from Tsinghua University, Beijing, China, in 1991 and 1996, respectively.

From 1991 to 1996, he was a Lecturer and Assistant Professor at Tsinghua University. From 1997 to 1998, he worked in the Department of Mathematics, University of North Carolina, Charlotte, as a Postdoctoral Researcher. Since then, he has been with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, as a Research Associate. His current research interests include parameter extraction and simulation for VLSI and packaging, electromagnetic scattering, and high-Tc superconducting applications.

Dr. Yu was the winner of the "Best Paper Award" at the 36th Design Automation Conference (DAC), San Francisco, CA, in June 1998.

Lawrence Carin (SM'96) [redacted] in Washington, DC, and received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor, and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, Durham, NC, where he is an Associate Professor. His current research interests include short pulse scattering, subsurface sensing, and wave-based signal processing. He is the Principal Investigator on the U.S. Army Multidisciplinary University Research Initiative (MURI) on demining.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies. He is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.

Classification of Landmine-Like Metal Targets Using Wideband Electromagnetic Induction

Ping Gao, *Student Member, IEEE,* Leslie Collins, *Member, IEEE,* Philip M. Garber, *Student Member, IEEE,* Norbert Geng, *Member, IEEE,* and Lawrence Carin, *Senior Member, IEEE*

***Abstract*—In our previous work, we have shown that the detectability of landmines can be improved dramatically by the careful application of signal detection theory to time-domain electromagnetic induction (EMI) data using a purely statistical approach. In this paper, classification of various metallic landmine-like targets via signal detection theory is investigated using a prototype wideband frequency-domain EMI sensor. An algorithm that incorporates both a theoretical model of the response of such a sensor and the uncertainties regarding the target/sensor orientation is developed. This allows the algorithms to be trained without an extensive data collection. The performance of this approach is evaluated using both simulated and experimental data. The results show that this approach affords substantial classification performance gains over a standard approach, which utilizes the signature obtained when the sensor is centered over the target and located at the mean expected target/sensor distance, and thus ignores the uncertainties inherent in the problem. On the average, a 60% improvement is obtained.**

***Index Terms*—Bayes procedures, electromagnetic induction, object detection, pattern classification.**

I. Introduction

A PERSISTENT problem with traditional narrowband EMI sensors involves not just detection of metal objects, but discrimination of targets from clutter. In most fielded sensors, the energy in the output of such sensors is calculated, and a decision regarding the presence or absence of a target is made using this statistic [1]. This approach leads to excessively large false alarm rates. When each piece of buried metal must be excavated in order to determine whether it is a target of interest, significant costs are incurred both due to lost time and costs associated with digging. The false alarm issue is particularly problematic in real world landmine-detection scenarios. In order to facilitate the discrimination of targets of interest from other pieces of metal, several modifications to traditional EMI sensors have been considered [1]–[8]. For instance, the late time EMI fields are characterized by an exponential decay in the time-domain [2], [3], [7], [8]. The decay rate has been used for target identification, because it strongly depends on the target conductivity, permeability, shape, and orientation. Alternatively, a promising approach is to operate the sensor in the frequency-domain by utilizing wideband excitation. The frequency dependence of the induced fields excited by buried conducting targets can then be exploited by a detector.

A second problem that besets statistical algorithms is the need for adequate training data. A wide range of targets and clutter signatures must be obtained at all possible object/sensor orientations. Classically, this problem has been addressed by using data measured in the field [1], [9], [10]. However, a lack of sufficient training data can severely degrade performance [1], [11]. In general, it is difficult if not impossible to obtain such data for all possible object/sensor orientations. An alternative is to train the algorithms with data produced by computational models, which have only recently become general enough to consider such problems [8].

In this paper, we consider the problem in which we assume an object has been detected, and a decision as to "target" or "clutter" is required. In this approach, we use the complex frequency-dependent EMI response as a signature. A full-wave model is developed for the wideband EMI response of targets, specialized to the case of a body of revolution, thus obviating the need for training on field-collected data. Subsequently, a Bayesian classification algorithm is developed, which incorporates the wave model and the target/sensor position uncertainty. Substantial improvements are achieved via this approach over a processor, which ignores the orientation uncertainties.

This paper is organized as follows. In Section II, we describe a new prototype wideband frequency-domain EMI sensor, the GEM-3 [9]. In Section III, we discuss a model that calculates the wideband EMI responses. A model-based Bayesian approach for discriminating targets is discussed in Section IV. The process used to generate the simulated data and the experiment performed to collect the measured data are described in Section V. Next, the results from both simulated and measured data are shown. Finally, we summarize our major findings based on these results.

II. Sensor Overview

When operating an EMI sensor in the frequency-domain, it has been shown that the frequency-dependent induced fields can differ significantly depending on the target shape and conductivity [12]. This variability may be exploited to enhance discrimination performance. Therefore, data from a prototype wideband EMI sensor, the GEM-3, developed by Geophex Ltd., was selected for this analysis. The validity of a numerical model that predicts the wideband EMI responses (discussed in the next section) can be tested using data collected with the GEM-3. Furthermore, a decision-theoretic discrimination algorithm can be applied to both simulated data generated based on the model

Manuscript received November 23, 1998; revised June 15, 1999. This work was supported by the Army Research Office, Research Triangle Park, NC, under Grants DAAH04-96-1-0448 (Demining MURI) and DAAG55-98-1-0340.

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA (e-mail: lcollins@ee.duke.edu).

Publisher Item Identifier S 0196-2892(00)02849-7.

0196–2892/00$10.00 © 2000 IEEE

predictions and real data measured with the sensor. In this section, the sensor operating principles are briefly described.

The design of the transmitting coils of the GEM-3, two concentric circular coils, results in a magnetic cavity in the center zone of the two coils, i.e., an area where the primary magnetic flux vanishes. A small receiving coil is located in this magnetic cavity [13]. Therefore, the receiver can sense a weak, secondary field returned from the earth and any buried objects. By using two transmitting coils connected in series, the coils carry the same amount of current. However, current flows in opposite directions. The magnetic cavity is created by carefully choosing the radius of the two coils and the number of turns of the coils. The sensor records the real and imaginary parts (in-phase and quadrature) of the induced complex voltage at the receiving coil, relative to that on the transmitting coils.

Instead of using a pulse excitation, such as that used by time-domain EMI systems, the transmitting coils of the wideband frequency-domain EMI sensor transmit a continuous, complex waveform consisting of multiple frequencies predefined by the operator [9], [13]. Thus, the sensor is only subject to the noise at the frequencies of interest, not within the whole frequency band, as is the case for time-domain EMI sensors. Therefore, frequency-domain EMI sensors can operate at much higher SNR's than time-domain systems. In addition to the improved SNR, theoretical calculations and experimental data have shown that the frequency-domain EMI signatures differ significantly across targets [12], which provides the underlying physical mechanisms important for discriminating, identifying, or classifying targets.

III. Model for Wideband Frequency-Domain EMI Responses

In this paper, a model-based Bayesian decision-theoretic approach is investigated to discriminate four manmade metal targets of different shapes, sizes, and metal types under conditions where the target/sensor relative position is unknown. In order to model the signature of these targets, a method of moment (MoM) analysis is used to predict the theoretical response from the target [12], [14]. The calculation provides the theoretical induced voltage (magnitude and phase, or in-phase and quadrature components) for each target and frequency considered. Later in this paper, it is shown that by incorporating the model into the detector formulation, the classification performance is improved dramatically when the relative target/sensor position is uncertain, as it is in field operations.

We consider the fields induced by a highly (but not perfectly) conducting and/or permeable target in free space, due to EMI excitation at kHz frequencies. The problem is solved via a frequency-domain boundary-integral equation formulation. Moreover, to make such an analysis tractable, we specialize to a particular class of targets: those that can be modeled as a body of revolution (BOR) [15] (i.e., targets possessing rotational symmetry). Although here we consider near-field effects for metallic and ferrous targets, the general formulation is very similar to those used previously for far-zone scattering from low-loss dielectric targets [16], [17]. In particular, the problem is formulated in terms of the tangential electric $\mathbf{E}$ and magnetic $\mathbf{H}$ fields on the target surface or equivalently, in terms of electric and magnetic surface currents $\mathbf{J} = \mathbf{n} \times \mathbf{H}$ and $\mathbf{K} = \mathbf{E} \times \mathbf{n}$, respectively, where $\mathbf{n}$ is the outward unit normal. If $\mathbf{E}_1$ and $\mathbf{H}_1$ represent, respectively, the electric and magnetic fields inside the target, and $\mathbf{E}_2$ and $\mathbf{H}_2$ represent the "scattered" fields outside the target, boundary conditions at the interface yield the relationships (enforced at the boundary)

$$\begin{aligned}
&\mathbf{n} \times \left[\mathfrak{L}_1^{EJ}(-\mathbf{J}) + \mathfrak{L}_1^{EK}(-\mathbf{K})\right] \\
&\quad = \mathbf{n} \times \left[\mathfrak{L}_2^{HJ}(-\mathbf{J}) + \mathfrak{L}_2^{EK}(-\mathbf{K})\right] + \mathbf{n} \times \mathbf{E}^i \\
&\mathbf{n} \times \left[\mathfrak{L}_1^{HJ}(-\mathbf{J}) + \mathfrak{L}_1^{HK}(-\mathbf{K})\right] \\
&\quad = \mathbf{n} \times \left[\mathfrak{L}_2^{HJ}(-\mathbf{J}) + \mathfrak{L}_2^{HK}(-\mathbf{K})\right] + \mathbf{n} \times \mathbf{H}^i
\end{aligned} \tag{1}$$

where $\mathbf{E}^i$ and $\mathbf{H}^i$ represent the incident fields. The operators $\mathfrak{L}_n$ involve well-known manipulations of the homogeneous-media Green's function [17]–[19] for medium parameters inside ($n = 1$) and outside ($n = 2$) the body. The problem therefore reduces to solving for $\mathbf{J}$ and $\mathbf{K}$ for particular incident fields $\mathbf{E}^i$ and $\mathbf{H}^i$. In the MoM solution for BOR [15]–[20], $\mathbf{J}$, $\mathbf{K}$, $\mathfrak{L}_n$, $\mathbf{E}^i$, and $\mathbf{H}^i$ are expanded in a Fourier series in the azimuthal variable ϕ, and for each Fourier component, $\mathbf{J}$ and $\mathbf{K}$ are expanded in terms of one-dimensional (1-D) basis functions along the BOR generating arc (see Fig. 1). In this paper, we use subsectional basis functions and testing functions, as in [18]–[20]. The interested reader is referred to [18]–[20] for details concerning implementation of the general algorithm, while here we focus on issues of particular relevance to the EMI problem.

We first consider requirements concerning the subsectional-basis-function discretization of $\mathbf{J}$ and $\mathbf{K}$. For scattering from low-loss targets, it is well known that approximately ten basis functions are required per wavelength [17]. In such problems, this rule is applied to the smallest wavelength of interest in the problem, generally corresponding to the medium inside the target. For the highly conducting targets of interest here, the wavenumber inside the target approximately satisfies $k_1 = (1 - j)/\delta$, where δ is the skin depth. To sample the Green's function phase $\exp(-jk_1R)$ sufficiently, we require $\Delta t/\delta \ll 2\pi$ and $\Delta t/\delta \ll 1$ (for the real and imaginary parts of k_1, respectively), where Δt is the basis-function width. These constraints are usually sufficient to satisfy the outer region sampling requirements (k_2 generally representing the *free-space* wavenumber). Extensive numerical experiments indicate that accurate results are obtained if $\Delta t < \delta/3$.

As discussed above, for EMI applications, we are generally interested in current loop excitation, as distinguished from the plane-wave fields considered for radar problems. While the fields due to a current loop are well known [21], [22], we discuss how such are placed into the BOR framework, as well as appropriate approximations for the EMI problem. In particular, the incident fields are derived from the vector potential [21], [22]

$$A_\phi(\rho, z) \approx \frac{\mu_0 Ia}{4\pi} \int_0^{2\pi} \frac{d\phi' \cos\phi'}{\sqrt{\rho^2 + a^2 + z^2 - 2a\rho\cos\phi'}} \tag{2}$$

where the origin of the local cylindrical coordinate system (ρ, ϕ, z) is situated at the loop center, with axis parallel to z, and I and a are the loop current and radius, respectively. The

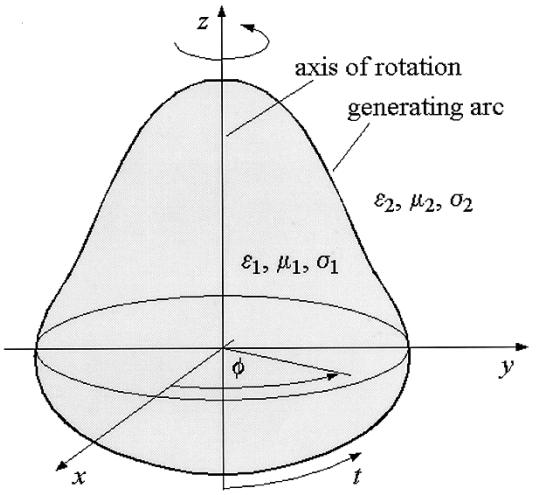


Fig. 1. Equivalent electric and magnetic surface currents for modeling electromagnetic interaction with a highly conducting and/or permeable body of revolution (BOR).

expression in (2) invokes a quasi-static approximation, since, at the wavelengths of interest (in air and soil), the electrical distance between the sensor and target is infinitesimal. Similar approximations can be used (but have not been here) with regard to the Green's function components in the air region ($\mathfrak{L}_2$ in (1)), while the very high conductivity and/or permeability inside the target necessitates a rigorous analysis (i.e., a rigorous formulation of $\mathfrak{L}_1$). The incident electric and magnetic fields are readily computed as

$$H_\rho^i(\rho,z) = -\frac{1}{\mu}\frac{\partial A_\phi}{\partial z},$$
$$H_z^i(\rho,z) = -\frac{1}{\mu\rho}\frac{\partial(\rho A_\phi)}{\partial \rho},$$
$$E_\rho^i(\rho,z) \approx -j\omega A_\phi \tag{3}$$

and these fields are finally expressed in terms of complete elliptical integrals. If the loop axis and the BOR axis are aligned, the fields in (3) can be applied directly to the BOR MoM solution, and only the lowest-order Fourier series mode is excited (reflecting azimuthal symmetry). If the axes are not aligned, a Fourier series representation of the incident tangential fields on the BOR surface is generally required (with the BOR problem solved separately using the incident fields from each such mode [15]–[18]). While the Fourier components for plane-wave incidence can be expressed in closed form [15], we have not found such a simple representation for the loop-induced fields. Therefore, in the general case, we must numerically determine the Fourier coefficients. For example

$$\hat{H}_{mt}(t) = \frac{1}{2\pi}\int_0^{2\pi} H_t^i(t,\phi)\exp(-jm\phi)\,d\phi \tag{4}$$

where H_t^i represents the incident magnetic field along the generating arc (Fig. 1), and (t,ϕ) represents the local BOR coordinate system. Thus, while the space domain fields for the loop can be expressed in closed form, the requisite Fourier components are evaluated numerically. However, the incident fields are generally slowly varying in ϕ and integrals of the type in (4) do not present a significant numerical challenge.

Before proceeding to a comparison of theoretical and measured results, we note that the EMI fields induced by a conducting and/or ferrous target are generally measured in the near zone. Therefore, when calculating the induced fields, we cannot invoke the simplifying far zone approximation generally used for radar-scattering problems [15], [17], [23]. We therefore calculate the EMI "scattered" fields via a rigorous convolution of the calculated currents $\mathbf{J}$ and $\mathbf{K}$ with the free-space Green's function (e.g., with $\mathfrak{L}_2$), performing integrals similar to those used in calculating the components of the MoM impedance matrix. Additionally, we note that an actual EMI sensor does not measure the induced fields, but rather the electromotive force induced on a sensing current loop. To calculate such, we have used appropriate magnetic field components, integrated over the aperture of the sensing loop, to generate a theoretical induced voltage for each target and frequency considered.

Using data collected from a prototype wideband frequency-domain EMI sensor, the GEM-3, the effectiveness of the numeric model is tested. A comparison of the theoretical model and measurements is shown in Section V-B.

The simulation outputs from the model can be used to calibrate the frequency-domain EMI sensor. Let $c(\omega)$ represent the calibration constant for frequency ω, the $K \times 1$ vector $\mathbf{M}$ represent a set of measurements obtained at several (K) positions, and the $K \times 1$ vector $\mathbf{B}$ represent model outputs for the same target and positions. We have the relation that $\mathbf{B}c(\omega) = \mathbf{M}$, and a least-squares method is used to obtain the calibration constants as a function of frequency.

IV. Forward Model-Based Bayesian Classifier Formulation

In a real-world classification scenario, the uncertainty inherent in the sensor output is not only due to additive noise, but also to the fact that the relative position between the sensor and the target is unknown at the point when the measurements are obtained. In this work, we investigate the classification performance of a Bayesian classifier that incorporates modeled wideband EMI signatures as well as position uncertainties and compare its performance to an approach that ignores these uncertainties and assumes the target is at a fixed position corresponding to the mean assumed position.

In this paper, we consider the task of classifying data from one of four known metal objects. It is always true that one of the objects is present, and our goal is to decide which object is present. In real world situations such as landmine detection, it is often the case that a metal object can be located. The task is then to determine whether it is a target or a clutter object. In this case, a library of targets of interest can be established and typical clutter can also be modeled. Alternatively, a statistical model could be imposed for clutter based on localized measurements and the target models can be used as is described here. Thus, this approach can also be applied to an extended set of objects in practice.

In this work, four metal objects are considered (a more detailed description can be found in Section V). Signals used

to measure classification performance are either the modeled wideband frequency-domain EMI responses or measured EMI responses from the GEM-3. Since any sensor is subject to noise, which is usually assumed to follow a Gaussian distribution, the distribution of the sensor outputs (obtained data set of discrete frequencies) while the target/sensor is at a known height and horizontal position, is a Gaussian random vector. The mean of this response is the theoretical response, and the variance is equal to that of the additive noise. Let H_i represent the hypothesis that the ith target is present, where $i = 1, 2, 3, 4$. The received data from the ith target at a known position can be modeled as

$$r_{ij} = A_{ij} + n_j \tag{5}$$

where j corresponds to the discrete frequencies of interest, $j = 1, 2, \ldots, N$, r_{ij} is the received data from the sensor, A_{ij} is the predicted response obtained from the model (as described in detail in Section III, the model can calculate the theoretical frequency-domain EMI responses for a well specified BOR object at a known position) for the ith target at the jth frequency at a known depth and horizontal position relative to the center of the sensor, and n_j is Gaussian noise with zero mean and variance of $\sigma_{n_j}^2$. We assume that n_j's are independent. However, their variance is a function of frequency. Let q_i represent the *a priori* probability that hypothesis H_i is true. We further assume that the cost of a correct decision is zero, and the cost of any wrong decision equals 1. Bayes' solution for this classification problem [24], [25] is to decide that H_i is true if

$$\frac{p(H_i|r)}{p(H_k|r)} = \frac{q_i p(r|H_i)}{q_k p(r|H_k)} > 1 \tag{6}$$

is satisfied for any $k \neq i$. Here $p(H_i|r)$ is the *a posteriori* distribution or discriminant function [26], $p(r|H_i)$ is the probability density or likelihood function of data r given H_i, and r is the received data from the sensor. Assuming the magnitude and the phase of the frequency response are independent, $\mathbf{r}$ is a vector containing both the magnitude and phase information. Therefore, when the sampled data r is received, we decide in favor of hypothesis H_i, where

$$q_i p(\mathbf{r}|H_i) = \max_k \{q_k p(\mathbf{r}|H_k)\} \quad k = 1, 2, 3, 4 \tag{7}$$

Thus, we decide in favor of a hypothesis that has the largest *a posteriori* probability or largest discriminant function at r among all four possible pdf's. Since we usually have no *a priori* knowledge of q_i (in other words, we do not know the probability that a particular target is going to be present), an equal probability assumption for each target is made (i.e., $q_i = 1/4$). Based on the uniform *priori* on q_i, (7) can be further understood as seeking a hypothesis that provides the maximum likelihood among the four possible values. Thus, it can also be referred to as a maximum likelihood (ML) classifier. Since any monotonically increasing function of $p(H_i|\mathbf{r})$ is also a valid discriminant function [26], an alternative discriminant function based on the above assumptions is

$$p(\mathbf{r}|H_i) = (2\pi)^{-N}|\Sigma|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{r}-\mathbf{A}_i)^T\Sigma^{-1}(\mathbf{r}-\mathbf{A}_i)\right] \tag{8}$$

TABLE I
PROBABILITY OF CORRECT CLASSIFICATION OF THE OPTIMAL CLASSIFIER WHEN TARGETS ARE AT A FIXED KNOWN POSITION AS THE NOISE VARIANCE IS INCREASED FROM σ_n^2 TO $2^9\sigma_n^2$

NOISE VARIANCE	PROBABILITY OF CORRECT CLASSIFICATION			
	TARGET 1	TARGET 2	TARGET 3	TARGET 4
σ_n^2	1.0000	1.0000	1.0000	1.0000
$2\sigma_n^2$	1.0000	1.0000	1.0000	1.0000
$2^2\sigma_n^2$	0.9993	1.0000	0.9990	1.0000
$2^3\sigma_n^2$	0.9862	1.0000	0.9877	1.0000
$2^4\sigma_n^2$	0.9365	1.0000	0.9394	1.0000
$2^5\sigma_n^2$	0.8604	1.0000	0.8640	1.0000
$2^6\sigma_n^2$	0.7809	0.9977	0.7747	1.0000
$2^7\sigma_n^2$	0.7134	0.9791	0.6854	1.0000
$2^8\sigma_n^2$	0.6545	0.9254	0.5756	0.9981
$2^9\sigma_n^2$	0.6101	0.8552	0.4554	0.9827

where N is the total number of frequencies used, $\mathbf{r}$ and $\mathbf{A}_i$ are $2N$ by 1 vectors, and Σ is the covariance matrix of $\mathbf{r}$. Given the assumptions on the noise process, Σ is a diagonal matrix with $\sigma_{n_j}^2$ on the jth diagonal, where j corresponds to frequency. Since the coefficient of the exponential term of (8) is the same for all the hypotheses, it can be neglected. After taking the logarithm, the alternative discriminant function simplifies to

$$\log p'(\mathbf{r}|H_i) = -(\mathbf{r}-\mathbf{A}_i)^T\Sigma^{-1}(\mathbf{r}-\mathbf{A}_i) \tag{9}$$

where $(\mathbf{r}-\mathbf{A}_i)^T\Sigma^{-1}(\mathbf{r}-\mathbf{A}_i)$ is often referred to as the Mahalanobis distance from r to $\mathbf{A}_i$ [26]. If Σ is a diagonal matrix with each diagonal element $\sigma_{n_j}^2$, (9) can be expressed as

$$\log p''(\mathbf{r}|H_i) = -\sum_{j=1}^{2N}(r_j - A_{ij})^2 \Big/ \sigma_{n_j}^2 \tag{10}$$

The discriminant function obtained above [(10)] is valid if the height and horizontal position of the object are both known, and the noise is assumed to be independent at each frequency. This solution is optimal only under the assumptions that all the parameters are known, and the sensor is subject only to Gaussian noise. This formulation differs from a bank of matched filters since the noise is not identically distributed and the variance of the noise is a function of frequency, and the signals are not of equal energy. These two facts result in a formulation, which is similar to, but not identical to, the traditional matched filter $r^T A_i$, which is the result of $i.i.d$ additive white Gaussian noise [25], [27].

The performance of the classifier given by (10) is a function of the noise variance and the modeled response. Table I lists the theoretical performance of the classifier as the noise variance is increased from σ_n^2 to $2^9\sigma_n^2$, where σ_n^2 is a vector that contains the noise variance of the magnitude and phase as a function of frequency obtained from experimental data (see Section. V-B). As expected, an increase in the noise variance results in a decrease in the classification performance. This analysis provides

insight into how the classification performance is affected by the sensor noise. Once the sensor is manufactured and well calibrated, the sensor noise cannot be changed artificially. Thus, only simulated data was used to illustrate this effect.

A more realistic assumption for the classification problem is that the height and horizontal position are uncertain, since the exact sensor position where measurements are obtained relative to the underground objects is unknown in practice. In this case, the previously derived processor [as in (10)], which assumed a "fixed" target/sensor orientation, is not the optimal solution. Hence, in order to obtain the optimal discriminant function for the received data, the effect of these random factors must be integrated out, i.e.,

$$p(\mathbf{r}|H_i) = \iiint p(\mathbf{r}|H_i, h, x, y)p(h)p(x, y)\, dh\, dx\, dy \quad (11)$$

where h represents the height of the sensor from the target, x and y represent the horizontal position of the sensor relative to the center of the target, $p(h)$ and $p(x, y)$ are the *a priori* distributions of the position factors, and

$$p(\mathbf{r}|H_i, h, x, y) = \frac{1}{(2\pi)^N |\Sigma|^{1/2}} \exp\Big(-\frac{1}{2}(\mathbf{r} - \mathbf{A}_i(h, x, y))^T \times \Sigma^{-1}(\mathbf{r}\mathbf{A}_i(h, x, y))\Big)$$

where $\mathbf{A}_i(h, x, y)$ is the model prediction (described in detail in Section III) of the ith target response when it is located at the position (h, x, y) relative to the sensor. The model predicts the theoretical frequency-domain EMI response as a function of constitutive parameters, exact dimensions of the object and the horizontal and vertical distance from the center of sensor to that of the object. Monte Carlo integration was implemented in order to calculate the integral in (11).

V. Simulated and Experimental Data

The performance of both the fixed-position processor (10) and the optimal classifier (11) is investigated by using both simulations and measurements for the GEM-3 sensor. In this section, the method used to generate simulated data, the experimental design, and the methods used to take the measurements are described.

Four metal targets are considered for both the simulations and experimental measurements: an aluminum barbell, an aluminum disk, a thick brass disk, and a thin brass disk. The dimensions of these targets are as follows. The diameter of each of the targets is 5.08 cm. The heights of the targets are 2.897 cm, 2.667 cm, 2.34 cm, and 0.3175 cm for the aluminum bar-bell, the aluminum disk, the thick and the thin brass disk, respectively. The response from a target depends on the constitutive parameters, geometry of the target, as well as the horizontal and vertical distance from the center of the sensor to that of the target. In the calculations, 21 linearly spaced frequencies were chosen, ranging from 3990 Hz to 23 970 Hz. These frequencies are within the range that the GEM-3 operates.

A. Simulations

In order to test whether the classification performance is improved by incorporating the model into the classification formulation, several cases were considered. These cases are:

1) fixed position;
2) random height but fixed horizontal position;
3) random horizontal position but fixed height;
4) both height and horizontal position random.

It is assumed that the distribution of the height, h, follows a Gaussian distribution with a mean of 20 cm and variance of 1.53^2 cm^2 and the horizontal position is uniformly distributed in a 20 cm by 20 cm square. To generate the simulated data we specify the constitutive parameters and the dimensions of the target, then generate 10 000 random sets of height (h) and horizontal position (x, y), which follow the distributions described above. Based on these parameters, the wideband EMI response is calculated by the model for each set of h, x, and y. After obtaining the theoretical responses of each target at all the specified positions (which are used in the formulation of the detector), Gaussian noise is added to the theoretical responses to create the simulated data set. Both the processor, which assumes a fixed target/sensor orientation (10), and the optimal classifier (11) are then applied to these data. Results of these classifiers are discussed in Section VI.

B. Measured Data

Using synthetic data to evaluate the performance of the classifier provides useful insight regarding performance bounds, but limiting the analysis to simulated data is not sufficient. Therefore, measurements of the wideband frequency-domain EMI response from the four metal targets were taken using the GEM-3 in order to evaluate the improvement of the classification algorithm in a more realistic scenario. First we consider whether data taken from objects in air is comparable to data obtained when the objects are buried in soil. Figs. 2 and 3 provide the wideband frequency-domain EMI responses measured using the GEM-3 from two metal landmines, a Valmara (an antipersonnel metal landmine) and a VS50 (an antipersonnel metal mine), in air and in North Carolina clay soil (buried 1 in below the surface). These figures indicate that the soil effects can be neglected at least for large metal objects. Therefore, the measured data used to evaluate the performance of various classification techniques was taken in free space.

The experimental set up is as follows. The GEM-3 was mounted on a wooden rack with the sensor head, approximately 1.8 m above the wooden base of the platform. Both rack and platform contained no metal parts. The rack assembly allows placement of a target on a wooden shelf at various distances beneath the sensor head.

First, in order to obtain an estimate of the noise variance associated with the sensor, $\sigma_{n_j}^2$, 100 measurements were taken with the sensor at a fixed position and no target present. We refer to an individual measurement taken without a target present as a background response. This response is subtracted from the responses measured with the target present to estimate the response due to the target alone. Fig. 4 shows a typical plot of the background

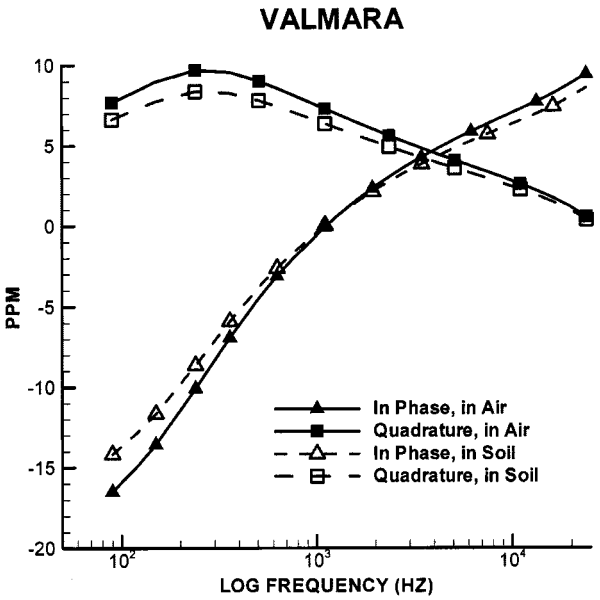


Fig. 2. Wideband frequency-domain response of a Valmara in free space and buried 1 in below the surface of the ground in North Carolina clay soil. The units "ppm" reflect the sensor output multiplied by 10^6.

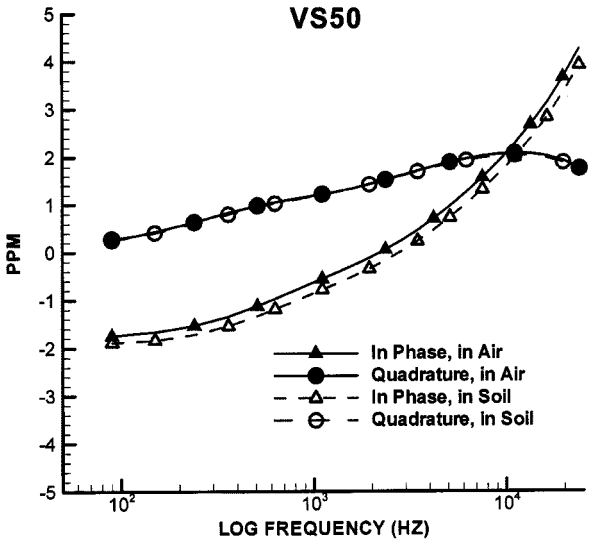


Fig. 3. Wideband frequency-domain response of a VS50 in free space and buried 1 in below the surface of the ground in North Carolina clay soil. The units "ppm" reflect the sensor output multiplied by 10^6.

response. The background response is not the same at all frequencies, and the noise variance is also a function of frequency. Table II lists the variance of the noise for the magnitude and phase of the complex response and the ratio between the mean value of the response and the standard deviation of the noise as a function of frequency, respectively. These estimates were used in the classifier given by (10) and (11).

As described in Section III, to calibrate the sensor, measurements for the four targets were taken so the calibration coefficients could be calculated. Each target was placed beneath the center of the sensor head at distances of 17 cm, 19 cm, 20 cm, 21 cm, and 23 cm. Using these 20 measurements, calibration coefficients were calculated by the least-squares method. Fig. 5

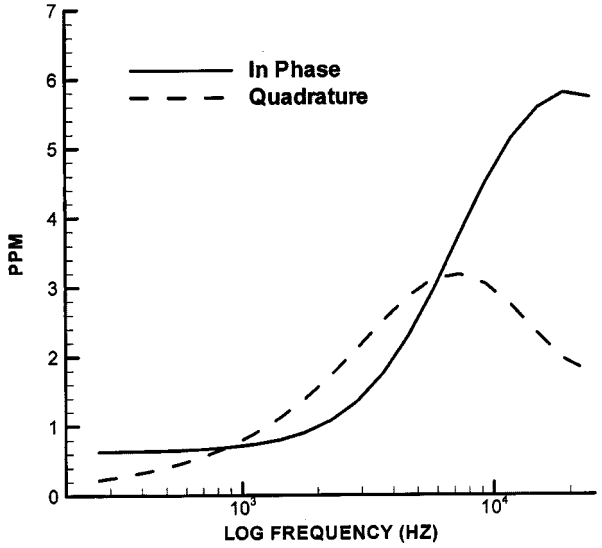


Fig. 4. Wideband frequency-domain response from Earth without the presence of any targets or background response. The units "ppm" reflect the sensor output multiplied by 10^6.

shows the comparison of the theoretical model predictions and the measurements. As has been noted previously, the model predicts the GEM-3 response well [12].

To obtain the data used to evaluate algorithm performance, measurements were taken from each target at seven heights from 17 cm to 23 cm in 1 cm increments. The distribution of height is assumed to be Gaussian with a mean of 20 cm and a variance of 1.53^2 for the simulations. At each height, between 11 and 36 measurements were taken. The exact count was calculated based on the assumed distribution. At each height, the position of each measurement is uniformly distributed within a 20 cm $\times$ 20 cm square. For each target, there were a total of 328 measurements taken. These data were not used to train the algorithm, only to evaluate performance.

VI. Results

We exploit Bayesian decision theory to formulate an optimal classifier to discriminate these targets. In order to show the improvement of the optimal classifier, the performance of a processor that assumes a fixed target/sensor orientation was also evaluated. It was assumed that the sensor is subject to a small amount of additive Gaussian noise. This assumption is verified by the experimental data (see Section V-B). The performance of these classifiers, shown in this section, was evaluated using both synthetic data and experimental measurements.

A. Simulation Results

1) Fixed Height and Horizontal Position: First, the case where all the position parameters are known exactly is considered. The model of each target at the same position and all desired frequencies is calculated. Then, by adding Gaussian random noise with zero mean and variance obtained based on the experimental data (see Table II), 10 000 realizations of simulated data for each target are generated. The decision of which target is present is made based on (7) by using the

TABLE II
VARIANCE OF THE BACKGROUND NOISE AND THE RATIO OF THE MEAN OF THE RESPONSE AND ITS STANDARD DERIVATION AS A FUNCTION OF FREQUENCY FOR THE MAGNITUDE AND PHASE COMPONENTS, RESPECTIVELY

Frequency (Hz)	$\sigma^2_{magnitude}$	$\frac{mean(magnitude)}{\sigma_{magnitude}}$	σ^2_{phase}	$\frac{mean(phase)}{\sigma_{phase}}$
3,990	8.25E-05	3.62E+02	8.86E-02	1.74E+02
5,010	8.99E-05	4.12E+02	5.55E-02	2.00E+02
5,970	9.41E-05	4.52E+02	3.91E-02	2.16E+02
6,990	9.47E-05	4.95E+02	2.83E-02	2.27E+02
8,010	8.96E-05	5.46E+02	2.19E-02	2.31E+02
8,970	8.35E-05	5.96E+02	1.77E-02	2.31E+02
9,990	7.47E-05	6.58E+02	1.44E-02	2.28E+02
11,010	6.57E-05	7.27E+02	1.21E-02	2.22E+02
11,970	5.73E-05	8.00E+02	1.05E-02	2.14E+02
12,990	5.44E-05	8.42E+02	8.99E-03	2.05E+02
14,010	5.70E-05	8.40E+02	8.05E-03	1.92E+02
14,970	6.23E-05	8.18E+02	7.20E-03	1.80E+02
15,990	9.25E-05	6.82E+02	6.23E-03	1.70E+02
16,950	1.11E-04	6.31E+02	5.14E-03	1.64E+02
17,970	1.55E-04	5.42E+02	4.87E-03	1.46E+02
18,990	1.49E-04	5.62E+02	9.35E-03	9.00E+01
19,950	1.63E-04	5.40E+02	1.39E-02	6.25E+01
20,970	7.26E-04	2.60E+02	3.64E-03	1.01E+02
21,950	3.90E-04	3.58E+02	9.92E-03	4.99E+01
22,950	7.81E-04	2.58E+02	9.23E-03	3.99E+01
23,970	1.21E-03	2.10E+02	8.32E-03	3.10E+01

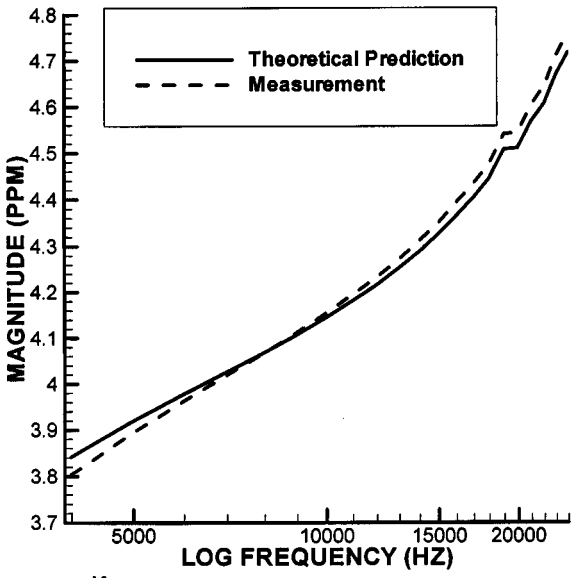


Fig. 5. Comparison of measurements and theoretical predictions for the thin brass disk when the distance from the target to the sensor is 20 cm.

processor expressed in (10), which is optimal for this case. Because of the fact that the wideband EMI signature of these targets is significantly different [12] and the experimentally derived σ_n^2's are low, the performance is perfect.

2) Height Uncertain, Fixed Horizontal Position: Next, the case where only the height of the sensor from the target is unknown and the target is located under the center of the sensor is considered. This situation occurs in a real detection scenario when the sensor operator can accurately center the sensor, but the burial depth of the mine is unknown. The height of the sensor was modeled as a Gaussian distributed random variable with a mean of 20 cm and a variance of 1.53^2 cm^2. Fig. 6 shows the performance of a processor that assumed a fixed target/sensor orientation along with that of the optimal classifier. For the former, it is assumed the target is at the mean height of 20 cm. Clearly, substantial improvements in classification performance are achieved by the optimal classifier over a processor that assumes a fixed target/sensor orientation. This performance is achieved for a relatively small level of uncertainty in the height. The average performance improvement is over 70%.

3) Horizontal Position Uncertain, Fixed Height: Thirdly, we simulate the case where horizontal position is uncertain. It is assumed that the sensor is located at a known, fixed height. Because the exact positions of mines are unknown to the sensor operator during detection, we assumed a uniform distribution in the horizontal plane. Fig. 7 shows the simulation results of the processor that assumes a fixed target/sensor orientation and the optimal classifier when the horizontal positions of targets are uniformly distributed. For the former, it was assumed that the target was at the mean horizontal position and was directly under the sensor. Again, the performance of the optimal classifier is substantially better than that of the processor, which ignores the target/sensor orientation uncertainty. It improves on average by 60%.

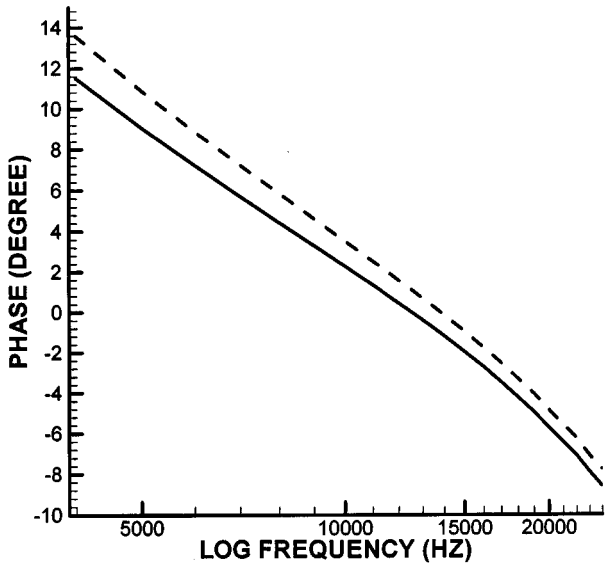


Fig. 6. Comparison of the processor, which ignores target/sensor orientation uncertainty ("fixed position" processor) and the optimal processor under uncertain height, fixed-horizontal position conditions for simulated data.

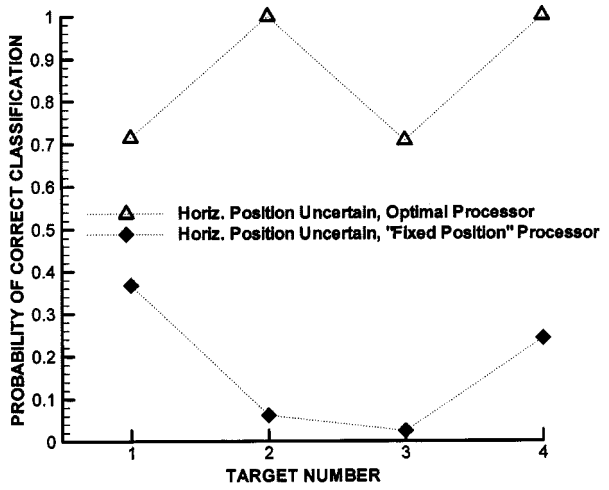


Fig. 7. Comparison of the "fixed position" processor and optimal processor performance under the uncertain horizontal position, but fixed height condition for simulated data.

4) Both Height and Horizontal Position Uncertain: In the final simulation, both height and horizontal position are uncertain. The height is assumed to follow a Gaussian distribution with mean of 20 cm and variance of 1.53^2. The horizontal position follows a uniform distribution (within a 20 cm $\times$ 20 cm square). Fig. 8 illustrates the performance of the two processors. The "fixed orientation" processor assumes that the target is located at the mean height and horizontal position. Performance improves under these conditions by an average of 70% over that of the "fixed position" processor. The results in Figs. 6–8 indicate that for the fixed position processor the performance becomes progressively worse as the position uncertainty increases. Clearly, incorporating the uncertainty of these environmental parameters into the processor affords a significant performance gain over a processor, which ignores this uncertainty.

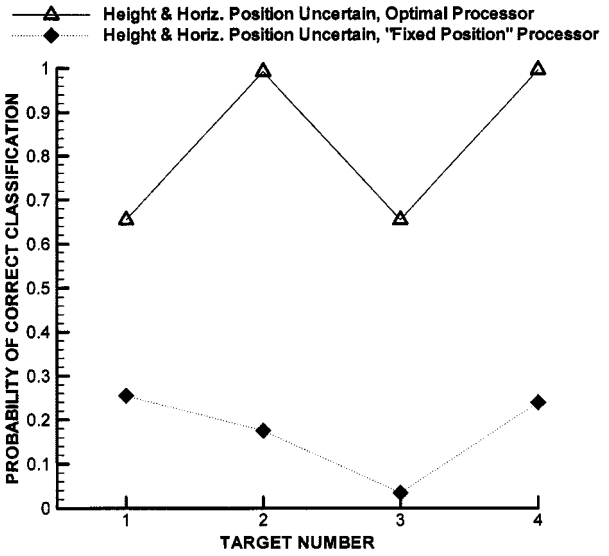


Fig. 8. Comparison of the "fixed position" processor and optimal processor performance when both height and horizontal position are uncertain for simulated data.

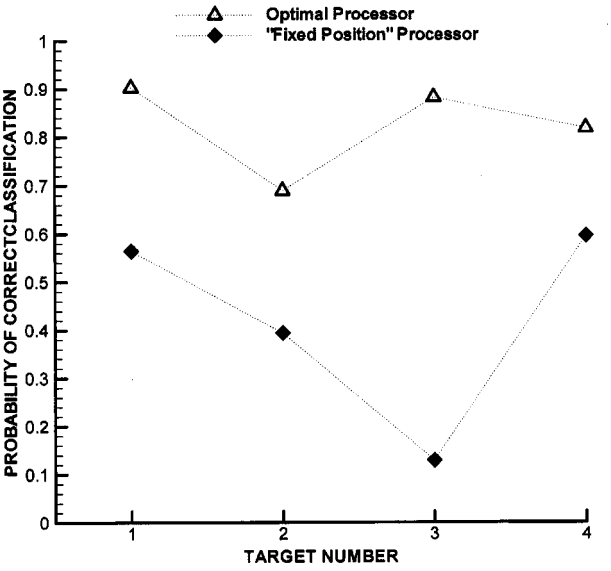


Fig. 9. Comparison of the "fixed position" processor and optimal processor performance under the condition of both height and horizontal position unknown for measured data.

B. Experimental Data

Simulations have shown that significant performance improvements can be achieved when the position uncertainty is incorporated into the classifier. To verify this result, measured data were collected using the GEM-3, as described in Section V-B. In this section, the results of implementing these processors using the measured data are shown.

The same two signal processing algorithms that were applied to simulated data: a fixed position processor, which assumes each target at the mean position, and the optimal classifier, which incorporates the position uncertainty into the processor, were applied to the experimental data. Fig. 9 illustrates the

performance achieved by each of these two algorithms. Clearly, better performance is achieved by the optimal processor. Performance improves on the average by 60%. This improvement, obtained on the measured data, is consistent with that observed in the simulated data set.

VII. Conclusion

In this paper, we utilize a Bayesian decision-theoretic approach to classify metal targets using wideband EMI data. Four manmade metal targets were used. Results from both simulation and measured data, shown in Section V, indicate that incorporating the uncertainty associated with the target/sensor relative position into the processor affords a significant performance gain over a processor that matches to the predicted response at the mean expected target position. It is also noted that, as expected, under conditions of uncertainty the performance of both the fixed orientation processor and the optimal processors drops compared to that of the signal-known-exactly case. Though the optimal classifier can improve performance under uncertain conditions over processors that ignore the uncertainties, it will never achieve the performance obtained when no uncertainty is present.

As expected, simulations have shown that the SNR's of the signal affect the performance of a classifier. Since frequency-domain systems can achieve high SNR's compared to time-domain EMI systems, essentially, it will improve the classification performance.

Our work indicates that we can effectively discriminate different metal targets using wideband EMI signals by incorporating an accurate physical model and models of the uncertainty regarding environmental parameters into the classifier. Performance can be dramatically improved over the standard approach, which ignores environmental uncertainty. In addition, extensive libraries of target signatures do not have to be measured experimentally in order to train the classifier.

This technique can be extended to apply to other applications such as landmine detection and unexploded ordnance (UXO) detection, since in these applications, targets of interests need to be discriminated from metallic clutter in order to reduce false alarm rates. The standard algorithms for these applications, such as matched filters, do not take the uncertainties associated with the target/sensor orientation into account, and only partially exploited the underlying physical nature of the outputs from the sensor. The work shown in this paper provides a promising technique, which integrates both the uncertainties associated with target/sensor orientation and a forward model exploiting the physical signature of wideband frequency-domain EMI response. By developing a model for other signals and sensor modalities, this algorithm can be further applied to other applications that require classification of different targets.

Acknowledgment

The authors would like to thank Dr. D. Keiswetter and Dr. I. J. Won for their assistance with the GEM-3. They would also like to thank R. Weaver for supplying the landmines and for his support of this work through the Joint UXO Coordinate Office (JUXOCO), Fort Belvoir, VA. They would also like to thank the reviewers, Y. Tan for her assistance with the data collection, and Dr. L. Nolte and Dr. S. Tantum for useful discussions regarding this work.

References

[1] L. Collins, P. Gao, and L. Carin, "An improved bayesian decision theoretic approach for land mine detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 811–819, Mar. 1999.

[2] C. E. Baum, "Low Frequency Near-Field Magnetic Scattering from Highly, but Not Perfectly Conducting Bodies," Phillips Laboratory, Interaction Note 499, Nov. 1993.

[3] G. D. Sower and S. P. Cave, "Detection and identification of mines from natural magnetic and electromagnetic resonances," in *Proc. SPIE*, Orlando, FL, 1995.

[4] A. H. Trang, P. V. Czipott, and D. A. Waldron, "Characterization of small metallic objects and non-metallic anti-personnel mines," in *Proc. SPIE*, Orlando, FL, 1997.

[5] S. Vitebskiy and L. Carin, "Short-pulse plane wave scattering from a buried perfectly conducting body of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 112–120, Jan. 1996.

[6] S. Vitebskiy and L. Carin, "Late-time resonant frequencies of buried bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.

[7] C. E. Baum, N. Geng, and L. Carin, "Integral Equations and Polarizability for Magnetic Singularity Identification," Phillips Lab., Interaction Note 524, Mar. 1997.

[8] N. Geng, C. E. Baum, and L. Carin, "On the low-frequency natural response of conducting and permeable targets," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 347–359, Jan. 1999.

[9] I. J. Won, D. A. Keiswetter, and D. R. Hansen, "GEM-3: A monostatic broadband electromagnetic induction sensor," *J. Env. Eng. Geophys.*, vol. 2, pp. 53–64, Aug. 1997.

[10] D. Keiswetter, I. J. Won, B. Barrow, and T. Bell, Object Identification Using Multifrequency EMI Data, UXO Forum, Atlanta, Georgia, May 1999.

[11] P. Gader, H. Frigui, G. Vaillette, B. Nelson, and J. Keller, "New results in fuzzy set based detection of landmines with GPR," in *Detection and Remediation Technologies for Mines and Minelike Targets IV Conf., Int. Symp. Aerospace/Defense Sensing and Controls*, Orlando, FL, Apr. 1999.

[12] N. Geng *et al.*, "Wideband Electromagnetic Induction for Metal-Target Identification: Theory, Measurement and Signal Processing," Duke Univ., Technical Report, Sept. 1997.

[13] I. J. Won, D. Keiswetter, G. R. A. Fields, and L. C. Sutton, "GEM-2: A new multifrequency electromagnetic sensor," *J. Env. Eng. Geophys.*, vol. 1, pp. 129–137, Mar. 1996.

[14] R. F. Harrington, *Field Computation by Moment Methods*. New York: Macmillan, 1968.

[15] J. R. Mautz and R. F. Harrington, "Radiation and scattering from bodies of revolution," *Appl. Sci. Res.*, vol. 20, pp. 405–435, June 1969.

[16] T. Wu and L. L. Tsai, "Scattering from arbitrarily-shaped lossy dielectric bodies of revolution," *Radio Sci.*, vol. 12, pp. 709–718, Sept. 1977.

[17] J. R. Mautz and R. F. Harrington, "Electromagnetic scattering from a homogeneous material body of revolution," *AEÜ*, vol. 33, pp. 71–80, Feb. 1979.

[18] A. W. Glisson and D. R. Wilton, "Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces," *IEEE Trans. Antennas Propagat.*, vol. 28, pp. 593–603, Sept. 1980.

[19] A. W. Glisson, D. Kajfez, and J. James, "Evaluation of modes in dielectric resonators using a surface integral equation formulation," *IEEE Trans. Microw. Theory Technol.*, vol. 31, pp. 1023–1029, Dec. 1983.

[20] A. W. Glisson, "Integral equation techniques," in *Dielectric Resonators*, D. Kajfez and P. Guillon, Eds. Dedham, MA: Artech House, 1986, pp. 259–325.

[21] C. A. Balanis, *Antenna Theory—Analysis and Design*. New York: Wiley, 1982.

[22] K. Simonyi, *Theoretische Elektrotechnik*. Berlin, Germany: Springer-Verlag, 1980.

[23] C. A. Balanis, *Advanced Engineering Electromagnetics*: Wiley, 1989.

[24] J. C. Hancock and P. A. Wintz, *Signal Detection Theory*. New York: McGraw Hill, 1966, pp. 80–84.

[25] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.

[26] R. J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*. New York: Wiley, 1992, pp. 34–43.
[27] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. San Diego, CA: Academic, 1990, p. 127.



Ping Gao (S'97) received the B.S. and M.E. degrees (honors) in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1993 and 1996, respectively, and the M.S.E.E. from Duke University, Durham, NC, in 1997, where she is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering.

Since 1997, she has worked as a Research Assistant at Duke University. Her main research interests are in signal processing, detection and estimation theory, statistical signal modeling and analysis, and applied electromagnetics.

Leslie Collins (M'96) [redacted] She received the B.S.E.E. degree from the University of Kentucky, Lexington, in 1985, and the M.S.E.E. and Ph.D. degrees in electrical engineering, both from the University of Michigan, Ann Arbor, in 1986 and 1995, respectively.

She was a Senior Engineer with the Westinghouse Research and Development Center, Pittsburgh, PA, from 1986 to 1990. In 1995, she became an Assistant Professor in the Electrical and Computer Engineering Department, Duke University, Durham, NC. Her current research interests include incorporating physics-based models into statistical signal processing algorithms, and she is pursuing applications in subsurface sensing as well as enhancing speech understanding by hearing impaired individuals.

Dr. Collins is a member of the Tau Beta Pi, Eta Kappa Nu, and Sigma Xi honor societies.

Philip M. Garber (S'99) [redacted] He received the B.S. degree in electrical engineering from Duke University, Durham, NC, in 1999.

He is currently a Graduate Student with the Radiation Laboratory, University of Michigan, Ann Arbor. His interests are in microwave circuits, communications, and wireless systems.



Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Germany in 1991 and 1996, respectively.

From 1991 to 1996, he was with the Institute for Microwaves and Electronics, University of Karlsruhe, working on full-wave propagation modeling for radio communication systems. In January 1997, he joined the Department of Electrical and Computer Engineering, Duke University, Durham, NC, as a Visiting Postdoctoral Fellow for 18 months. Since July 1998, he has been with the University of Karlsruhe. His current research interests include computational methods in electromagnetics and wave propagation modeling.

Dr. Geng received the Mannesmann Innovation Award in 1997 for his Ph.D. thesis on full-wave propagation modeling for radio communication systems.

Lawrence Carin (SM'96) [redacted] and earned the B.S., M.S., and Ph.D. degrees in electrical engineering at the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, New York, as an Assistant Professor, and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, Durham, NC, where he is an Associate Professor. His current research interests include short pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies. He is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.

Multilevel Fast-Multipole Algorithm for Scattering from Conducting Targets Above or Embedded in a Lossy Half Space

Norbert Geng, *Member, IEEE*, Anders Sullivan, *Member, IEEE*, and Lawrence Carin, *Senior Member, IEEE*

***Abstract*—An extension of the multilevel fast multipole algorithm (MLFMA), originally developed for targets in free space, is presented for the electromagnetic scattering from arbitrarily shaped three-dimensional (3-D), electrically large, perfectly conducting targets above or embedded within a lossy half space. We have developed and implemented electric-field, magnetic-field, and combined-field integral equations for this purpose. The nearby terms in the MLFMA framework are evaluated by using the rigorous half-space dyadic Green's function, computed via the method of complex images. Non-nearby (far) MLFMA interactions, handled efficiently within the multilevel clustering construct, employ an approximate dyadic Green's function. This is expressed in terms of a direct-radiation term plus a single real image (representing the asymptotic far-field Green's function), with the image amplitude characterized by the polarization-dependent Fresnel reflection coefficient. Examples are presented to validate the code through comparison with a rigorous method-of-moments (MoM) solution. Finally, results are presented for scattering from a model unexploded ordnance (UXO) embedded in soil and for a realistic 3-D vehicle over soil.**

***Index Terms*—Fast algorithms, method of moments (MoM), numerical methods, scattering.**

I. Introduction

ELECTROMAGNETIC scattering from surface and subsurface targets has constituted a problem of long-term interest. Considering numerical modeling of such problems, there has been interest in integral-equation [1]–[3] and differential-equation [4]–[6] based methods. Investigators have considered method-of-moments (MoM) [1]–[3], [7]–[9] and finite-element method (FEM) [6], [9], [10] solutions. The principal focus in the context of time-domain differential-equation solvers has been the finite difference time domain method (FDTD) [4], [5], [11] and more recently, the multiresolution time domain method (MRTD) [12], [13]. Each of these numerical techniques has its relative strengths and weaknesses, and each is applicable to relatively small targets (with characteristic target dimensions small relative to a few wavelengths). There are many applications for which the target may comprise a large number of wavelengths, for example, scattering from vehicles, trees, and large unexploded ordnance. The aforementioned MoM, FEM, FDTD and MRTD algorithms can handle such targets in principle, but memory requirements and computation time become excessive.

There has consequently been significant interest recently in the development of a new generation of fast algorithms to cover the range of frequencies for which the above computational tools are inappropriate and for which high-frequency asymptotic techniques (e.g., PO, GTD, and UTD) are not applicable or are difficult to implement. For example, the fast multipole method (FMM) was extended to the case of electromagnetic fields by Rhoklin and his colleagues [14]. More recently, Chew and colleagues have implemented the FMM [15] and extended it to a multilevel framework [16], resulting in the multilevel fast multipole algorithm (MLFMA) [17]–[20]. The MLFMA has been developed for electric-field, magnetic-field, and combined-field integral equations for both perfectly conducting [17]–[19] and dielectric targets [20].

The FMM and MLFMA exploit a particular expansion of the free-space Green's function [14], [15], [21] and therefore, virtually all implementations to date have been relegated to the case of targets in free space. There have been layered-medium results presented for the case of two-dimensional (2-D) scattering [22], as well as quasi-planar three-dimensional (3-D) problems in circuit and antenna design [23]. The difficulty of directly applying the FMM/MLFMA formalism to general 3-D layered-medium problems resides in the fact that the layered-medium Green's function is a dyadic, each term of which is represented in terms of a generally complicated Sommerfeld integral [8], [24]. One can use the complex-image technique [25]–[27] to represent each term of the dyadic in terms of a sum of images, generally located in complex space. Therefore, in principle, one can apply the fundamental FMM/MLFMA expansion to each term of this expansion. However, we have demonstrated [28], [29] that for the general case, the convergence (for complex image points) is often considerably slower than when source and observer are in real space. This mitigates the attractiveness of the underlying formalism.

In [28], [29], we therefore developed an approximate means of handling the dyadic half-space Green's function, with application to the FMM. In particular, the near FMM terms are evaluated via the use of the exact dyadic Green's function, the latter evaluated efficiently via the complex-image technique [25]–[27]. This part of the model is exactly as in a rigorous MoM analysis of scattering from targets in the vicinity of a half space [3], [8]. The far terms, evaluated efficiently via the clustering algorithm, employ the asymptotic form of the dyadic Green's function [30]. As elucidated further in

Manuscript received August 31, 1999; revised March 27, 2000.

N. Geng is with the Institut für Höchstfrequenztechnik und Elektronik, University of Karlsruhe, Karlsruhe, Germany.

A. Sullivan and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0196-2892(00)05895-2.

0196–2892/00$10.00 © 2000 IEEE

the following, each component of the approximate Green's function is expressed in terms of the direct-radiation term plus radiation from an image source in real space [28]–[30]. The former accounts for the radiation of currents into the medium in which it resides, while the latter accounts for interactions with the half-space interface. The above approximation exploits the fact that the near interactions are most sensitive to accurate Green's-function evaluation, while the far interactions should be less so (the far interactions within the MLFMA formalism are always at least about half a wavelength apart [17]–[20]). Therefore, the half-space MLFMA presented here differs from the free-space MLFMA [17]–[20] in several ways. First, the nearby interactions are handled rigorously through use of the complete dyadic half-space Green's function, since these are likely to be most susceptible to approximations in the Green's function. For far interactions, we utilize the standard MLFMA clustering, with the interactions between such constituting the aforementioned direct radiation contribution. Finally, a separate set of image clusters, at real spatial positions, constitute the reflected component of the radiated fields. As is well known, the free-space FMM has O($N^{3/2}$) memory and computational (CPU) requirements (per iteration) [14], [15], where N is the number of unknowns for representation of the target, while the free-space MLFMA reduces these further to O($N\lg N$) [17]–[20]. With the above construct, the real images introduce a new set of source clusters, and therefore, the algorithm requires slightly (typically about 20–60%) more memory and about twice the computation time compared to the free-space version. However, the computational complexity of O($N\lg N$), both in RAM and CPU (per iteration), remains unchanged.

The half-space FMM presented previously in [28], [29], which uses the aforementioned scheme for the efficient evaluation of the dyadic half-space Green's function, employed an electric-field integral equation, which limited its applicability somewhat, due to possible interior resonances for closed targets [9], [17]. In the work presented here, we therefore extend our previous work to an MLFMA formalism, employing electric-field (EFIE), magnetic-field (MFIE), and combined-field integral equations (CFIE), the latter eliminating interior resonances [9], [17].

The remainder of the paper is organized as follows. In Section II, we present our MLFMA formalism for surface or buried perfectly conducting targets. The model is validated in Section III-A through comparison with results from a rigorous body-of-revolution MoM algorithm [1], [2]. In Section III-B and Section III-C, we present several results for realistic targets of interest to the remote sensing community. In particular, we consider scattering from large buried unexploded ordnance (UXO) and scattering from a realistic 3-D vehicle above ground. Memory and computational requirements for the half-space MoM, FMM, and MLFMA are compared. Finally, conclusions and suggestions for future work are discussed in Section IV.

II. Theory

A. Integral Equation and Half-Space MoM Formulation

For solving the problem of scattering from an arbitrarily shaped 3-D, perfectly electric conducting (PEC) target situated above (i.e., target in layer $i = 1$) or buried (i.e., target in layer $i = 2$) in a lossy half space (Fig. 1), we utilize the CFIE [9], [17]

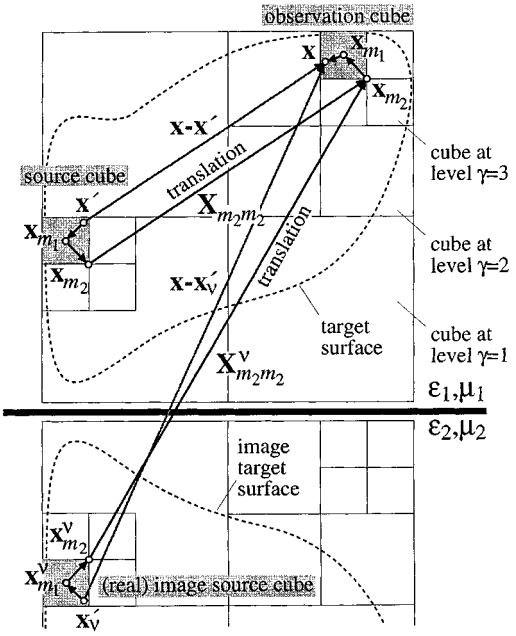


Fig. 1. Multilevel clustering in 3-D MLFMA and generalization to a half-space environment using real images to account for far interface interactions.

$$\begin{aligned}\hat{\boldsymbol{t}}\cdot[\alpha\boldsymbol{E}^{inc}(\boldsymbol{x})+\eta_i(1-\alpha)\hat{\boldsymbol{n}}\times\boldsymbol{H}^{inc}(\boldsymbol{x})]\\ =\hat{\boldsymbol{t}}\cdot[-\alpha\boldsymbol{E}^{scat}(\boldsymbol{x})+\eta_i(1-\alpha)\boldsymbol{J}(\boldsymbol{x})-\eta_i(1-\alpha)\hat{\boldsymbol{n}}\\ \times\boldsymbol{H}^{scat}(\boldsymbol{x}\in S^+)]\end{aligned}\tag{1}$$

with the scattered fields (inside half space i, in which the target is located) given by [8]

$$\begin{aligned}\boldsymbol{E}^{scat}(\boldsymbol{x})&=-j\omega\mu_i\left[\overleftrightarrow{\boldsymbol{I}}+\frac{\nabla\nabla}{k_i^2}\right]\cdot\iint_{S'}\overleftrightarrow{\boldsymbol{G}}_{Aii}(\boldsymbol{x},\boldsymbol{x}')\cdot\boldsymbol{J}(\boldsymbol{x}')\,dS'\\ &=-j\omega\mu_i\iint_{S'}\overleftrightarrow{\boldsymbol{K}}_{Aii}(\boldsymbol{x},\boldsymbol{x}')\cdot\boldsymbol{J}(\boldsymbol{x}')\,dS'\\ &\quad+\frac{\nabla}{j\omega\varepsilon_i}\iint_{S'}K^{ii}_{\phi e}(\boldsymbol{x},\boldsymbol{x}')\,\nabla'\cdot\boldsymbol{J}(\boldsymbol{x}')\,dS'\end{aligned}\tag{2a}$$

$$\boldsymbol{H}^{scat}(\boldsymbol{x})=\nabla\times\iint_{S'}\overleftrightarrow{\boldsymbol{G}}_{Aii}(\boldsymbol{x},\boldsymbol{x}')\cdot\boldsymbol{J}(\boldsymbol{x}')\,dS'.\tag{2b}$$

The unit vectors $\hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{t}}$ are perpendicular and tangential to the scatterer surface, respectively, $\boldsymbol{x}$ is on and $\boldsymbol{x}\in S^+$ is an infinitesimal distance outside the (closed) target surface. $\varepsilon_i = \varepsilon_i' - j\sigma_i/\omega$, μ_i, and k_i represent (in general complex) the permittivity, permeability, and wavenumber of the medium in which the target resides, and ω is the angular frequency (with a time dependence exp $j\omega t$ assumed and suppressed). Details on the Green's function dyadics $\overleftrightarrow{\boldsymbol{G}}_{Aii}$ and $\overleftrightarrow{\boldsymbol{K}}_{Aii}$, as well as the scalar Green's function $K^{ii}_{\phi e}$ have been given by Michalski and Zheng in [8], where we use their formulation C. The CFIE in (1) is

valid for a more general layered medium [8], but here we only consider the half-space problem for simplicity.

The CFIE (1) includes the electric-field integral equation (EFIE, $\alpha = 1$) [9], which must be solved for an open target, and the magnetic-field integral equation (MFIE, $\alpha = 0$) [9] as special cases. To avoid the possible spurious interior resonances of the EFIE and MFIE solutions [9], we typically choose α in the range of $\alpha = 0.2$–0.5 if analyzing a closed target. It has been proven [9], [17], that the CFIE with $0 < \alpha < 1$ eliminates the effects of interior resonances.

In the conventional MoM solution, the unknown surface current $\boldsymbol{J}$ is expanded into a set of N basis (expansion) functions $\boldsymbol{b}_{n'}(\boldsymbol{x}')$

$$\boldsymbol{J}(\boldsymbol{x}') = \sum_{n'=1}^{N} I_{n'} \boldsymbol{b}_{n'}(\boldsymbol{x}') \tag{3}$$

where we use the triangular patch Rao–Wilton–Glisson (RWG) basis introduced in [7]. Testing the CFIE (1) with a set of weighting (testing) functions $\boldsymbol{w}_n(\boldsymbol{x})$ tangential to the target surface results in N linear equations

$$[Z] \cdot \boldsymbol{I} = \boldsymbol{V} \tag{4}$$

for the unknown coefficients $I_{n'}$. The elements of the driving vector $\boldsymbol{V}$ (representing the incident field "tested" on the target surface) and the impedance matrix $[\boldsymbol{Z}]$ are given by [14], [15], [17]

$$V_n = \iint_S \boldsymbol{w}_n(\boldsymbol{x}) \cdot [\alpha \boldsymbol{E}^{inc}(\boldsymbol{x}) + \eta_i (1-\alpha) \hat{\boldsymbol{n}} \times \boldsymbol{H}^{inc}(\boldsymbol{x})] \, dS \tag{5}$$

and

$$\begin{aligned} Z_{nn'} = {} & \alpha \, j\omega\mu_i \iint_S \iint_{S'} \boldsymbol{w}_n(\boldsymbol{x}) \cdot \left[\overleftrightarrow{\boldsymbol{I}} + \frac{\nabla\nabla}{k_i^2} \right] \\ & \cdot \overleftrightarrow{\boldsymbol{G}}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') \cdot \boldsymbol{b}_{n'}(\boldsymbol{x}') \, dS' \, dS + \eta_i (1-\alpha) \iint_S \boldsymbol{w}_n(\boldsymbol{x}) \\ & \cdot \left[\boldsymbol{b}_{n'}(\boldsymbol{x}) - \hat{\boldsymbol{n}} \times \nabla \times \iint_{S'} \overleftrightarrow{\boldsymbol{G}}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') \right. \\ & \left. \cdot \boldsymbol{b}_{n'}(\boldsymbol{x}') \, dS' \Big|_{\boldsymbol{x} \in S^+} \right] dS \end{aligned} \tag{6}$$

respectively.

In an MLFMA analysis (Sections II-B and II-C), we divide the computation of interactions into near and far terms, where it must be emphasized that for far interactions, basis and weighting functions do not have to be in each other's far zone, but rather, these terms are handled by exploiting a plane wave spectral representation of the free-space Green's function (which requires that basis and weighting functions are "far enough" apart [14], [15], [19], but not necessarily in the far zone, see Section II-B). For near interactions (MoM part of the MLFMA), the evaluation of the impedance matrix elements (6), including the proper handling of self-term singularities and near singularity extraction, is done similar to the MoM literature [7], [31], [32]. The dyadic half-space Green's function is evaluated rigorously using the method of discrete complex images [25]–[27], thereby avoiding direct numerical evaluation of Sommerfeld integrals [8], [24]. Impedance matrix elements representing these near interactions are stored in a sparse matrix called $[\boldsymbol{Z}^{near}]$. The method of complex images and explicit equations for the MoM impedance matrix elements can be found in the literature, and therefore, we do not repeat the details.

B. Free-Space and Half-Space MLFMA

The half-space dyadic Green's function can be split into a term $\overleftrightarrow{\boldsymbol{I}} g_i$ representing the "direct" radiation between source and observation point (as in free space, but using in general a complex wave number k_i) and a remaining dyadic $\Delta\overleftrightarrow{\boldsymbol{G}}_{Aii}$ accounting for interactions with the interface (i.e., here Δ is not an operator) [8], [28], [29].

$$\begin{aligned} \overleftrightarrow{\boldsymbol{G}}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') &= \overleftrightarrow{\boldsymbol{I}} g_i(\boldsymbol{x}, \boldsymbol{x}') + \Delta\overleftrightarrow{\boldsymbol{G}}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') \\ &= \overleftrightarrow{\boldsymbol{I}} \frac{e^{-jk_i|\boldsymbol{x}-\boldsymbol{x}'|}}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|} + \Delta\overleftrightarrow{\boldsymbol{G}}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') \end{aligned} \tag{7}$$

While the dyadic $\Delta\overleftrightarrow{\boldsymbol{G}}_{Aii}$ needs some further investigation (see below), the *free-space* FMM [14], [15] or MLFMA[17]–[20] can be applied to the "direct" term, with only minor changes due to the (in general) lossy background. The *free-space* FMM and MLFMA are based on the addition theorem [14], [21], leading to the (propagating) plane wave representation [14], [15], [21]

$$\begin{aligned} & g_i(\boldsymbol{x}, \boldsymbol{x}') \\ & \quad = \frac{e^{-jk_i|\boldsymbol{x}-\boldsymbol{x}'|}}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|} \\ & \quad \approx \frac{-jk_i}{(4\pi)^2} \iint_{4\pi} e^{-jk_i\hat{\boldsymbol{k}}\cdot(\boldsymbol{x}-\boldsymbol{x}_m)} T_L(k_i X_{m'm}, \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{X}}_{m'm}) \\ & \qquad \cdot e^{+jk_i\hat{\boldsymbol{k}}\cdot(\boldsymbol{x}'-\boldsymbol{x}_{m'})} \, d^2\hat{\boldsymbol{k}} \end{aligned} \tag{8a}$$

$$\begin{aligned} & T_L(k_i X_{m'm}, \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{X}}_{m'm}) \\ & \quad = \sum_{l=0}^{L} (-j)^l (2l+1) \, h_l^{(2)}(k_i X_{m'm}) \, P_l(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{X}}_{m'm}) \end{aligned} \tag{8b}$$

$$\hat{\boldsymbol{X}}_{m'm} = \boldsymbol{X}_{m'm} / X_{m'm} = \boldsymbol{X}_{m'm} \Big/ \sqrt{\boldsymbol{X}_{m'm} \cdot \boldsymbol{X}_{m'm}} \tag{8c}$$

of the scalar Green's function. The distance vector $\boldsymbol{R} = \boldsymbol{x} - \boldsymbol{x}'$ has been subdivided into a vector from the source point $\boldsymbol{x}'$ to the center $\boldsymbol{x}_{m'}$ of a "source group," a vector from an "observation group" center $\boldsymbol{x}_m$ to the observation point $\boldsymbol{x}$, and a vector $\boldsymbol{X}_{m'm}$ connecting the group centers (Fig. 1, where in this Section II-B, we write $m' = m'_\gamma$, $m = m_\gamma$, and $\boldsymbol{X}_{m'm} = \boldsymbol{X}_{m'_\gamma m_\gamma}$ for simplification)

$$\boldsymbol{R} = \boldsymbol{x} - \boldsymbol{x}' = (\boldsymbol{x} - \boldsymbol{x}_m) + \boldsymbol{X}_{m'm} - (\boldsymbol{x}' - \boldsymbol{x}_{m'}) = \boldsymbol{X}_{m'm} + \boldsymbol{d}. \tag{9}$$

Details on the convergence of the addition theorem [14], [21], resulting in approximate formulas for the minimum number L

of terms needed for a desired accuracy, can be found in the literature [14], [15], [19] for a real wavenumber $k_i = k_0$. In the case of a lossy background (e.g., soil) more terms are required [29]. Therefore, in our implementation, L is determined adaptively for a predefined minimum separation of the group centers, the latter typically chosen in the range $1.5\,d_{\max}$–$2\,d_{\max}$, where $d_{\max}$ is the maximum group diameter. For groups situated closer together, the interactions are included in the sparse matrix [$\boldsymbol{Z}^{near}$] (Section II-A). A more detailed discussion was given in [29]. We should also mention that for all examples in Section III, the upper limit L in the series expansion has been determined to provide an error of less than 1% in the underlying addition theorem [14], [21]. For the efficient numerical evaluation of (8a) over the solid angle 4π, a Gaussian quadrature integration with $K = 2L^2$ plane wave directions $\hat{\boldsymbol{k}}$ is applied [14], [15], although more efficient quadrature rules are available [33].

To incorporate the expansion (8) into a fast algorithm for the iterative solution of the matrix equation (4), the scatterer surface has to be partitioned first into groups. While in the single-stage FMM, one level of $M \sim \sqrt{N}$ groups is applied [14], [15], the MLFMA first encloses the object in a large cube of edge length d, and then the cube (or subcube) is recursively divided into eight smaller cubes until the edge length $d/2^g$ at the finest level $\gamma = g$ is approximately half a wavelength [17]–[20] (Fig. 1). Only information for the nonempty cubes is stored using tree-structured data. Therefore, the memory and computational costs depend on nonempty cubes only [17]–[20]. In the following, basis/weighting functions numbered globally as $n(m, \alpha) = 1 \cdots N$ are also labeled according to a group (cube) index $m = 1 \cdots M$ and an additional index $\alpha = 1 \cdots A_m$ within each group (cube). For a specific MLFMA level γ, these variables will be denoted as m_γ, M_γ, α_γ, and A_{m_γ}, respectively.

Using the expansion (8), the elements of the *far* interaction impedance matrix (i.e., for $X_{m'm}$ sufficiently large) in the context of a free-space scattering problem can be written as [17]

$$Z_{nn'} \approx \frac{\omega\mu_i k_i}{(4\pi)^2} \iint_{4\pi} \boldsymbol{W}_{m\alpha}(\hat{\boldsymbol{k}}) \cdot T_L(k_i X_{m'm}, \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{X}}_{m'm}) \cdot \boldsymbol{B}_{m'\alpha'}(\hat{\boldsymbol{k}})\, d^2\hat{\boldsymbol{k}} \tag{10}$$

$$\boldsymbol{B}_{m'\alpha'}(\hat{\boldsymbol{k}}) = \left[\overleftrightarrow{\boldsymbol{I}} - \hat{\boldsymbol{k}}\hat{\boldsymbol{k}}\right] \cdot \iint_{S'} \boldsymbol{b}_{n'(m',\alpha')}(\boldsymbol{x}') \cdot e^{+jk_i\hat{\boldsymbol{k}}\cdot(\boldsymbol{x}'-\boldsymbol{x}_{m'})}\, dS' \tag{11}$$

$$\boldsymbol{W}_{m\alpha}(\hat{\boldsymbol{k}}) = \alpha \left[\overleftrightarrow{\boldsymbol{I}} - \hat{\boldsymbol{k}}\hat{\boldsymbol{k}}\right] \cdot \iint_{S} \boldsymbol{w}_{n(m,\alpha)}(\boldsymbol{x}) \cdot e^{-jk_i\hat{\boldsymbol{k}}\cdot(\boldsymbol{x}-\boldsymbol{x}_m)}\, dS + (1-\alpha)\hat{\boldsymbol{k}} \times \iint_{S} \hat{\boldsymbol{n}} \times \boldsymbol{w}_{n(m,\alpha)}(\boldsymbol{x}) \cdot e^{-jk_i\hat{\boldsymbol{k}}\cdot(\boldsymbol{x}-\boldsymbol{x}_m)},\, dS \tag{12}$$

where (11) and (12) represent the far-field radiation pattern of the basis functions and weighting pattern of the testing functions, respectively. Thus, the expansion (8) is first used to translate each plane wave component of the field radiated by the individual basis functions within a source group into a common center $\boldsymbol{x}_{m'}$ (aggregation). The spectral components $\boldsymbol{B}_{m'\alpha'}(\hat{\boldsymbol{k}})$ are then shifted individually to the common center $\boldsymbol{x}_m$ of a group of weighting functions using the operators T_L (translation). Finally, the plane waves are weighted according to the receiving patterns $\boldsymbol{W}_{m\alpha}(\hat{\boldsymbol{k}})$, including the redistribution of the fields from the common center to the location of the weighting functions (disaggregation).

For a half-space FMM/MLFMA, it is essential to include the effects of the far interface interactions. In the complex-image technique, each component of the dyadic $\Delta\overleftrightarrow{\boldsymbol{G}}_{Aii}$ is expressed in terms of a sum of free-space Green's functions with image sources located in complex space [25]–[27]. Therefore, in principle, the expansion (8) (which remains valid for general complex source points [21]) could also be applied for far interface interactions. However, we have shown recently [28], [29], that the number of terms L required for convergence can be prohibitively large for general complex source points, undermining the efficiency of using (8) for far interface interactions in the context of the discrete complex-image technique.

In [28] and [29] we therefore described an alternative (though approximate) formulation. While the half-space dyadic Green's function is rigorously accounted for in the near interaction matrix, far interactions are often less sensitive to approximations in the Green's function [28]–[30]. The FMM has been successfully extended to the scattering from a PEC object above or buried in a half space by employing the asymptotic form of the Green's function [30] for far interactions. The asymptotic form of the Green's function is represented utilizing a single *real* image at $[\overleftrightarrow{\boldsymbol{I}} - 2\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}]\cdot\boldsymbol{x}'$ (assuming the interface at $z = 0$) with its polarization dependent magnitude given by the reflection dyadic [30]

$$\overleftrightarrow{\boldsymbol{r}}(\hat{\boldsymbol{k}}) = \hat{\boldsymbol{h}}\hat{\boldsymbol{h}}\, R_{\mathrm{TE}}(\hat{\boldsymbol{k}}) + \left[\overleftrightarrow{\boldsymbol{I}} - \hat{\boldsymbol{h}}\hat{\boldsymbol{h}}\right] R_{\mathrm{TM}}(\hat{\boldsymbol{k}}) \quad \text{with } \hat{\boldsymbol{h}} = \frac{\hat{\boldsymbol{z}} \times \hat{\boldsymbol{k}}}{|\hat{\boldsymbol{z}} \times \hat{\boldsymbol{k}}|}. \tag{13}$$

Therefore, in addition to the accurate calculation of the half-space dyadic Green's function in the near interaction matrix [$\boldsymbol{Z}^{near}$] (Section II-A), the half-space MLFMA only requires the definition of a single set of real image sources (Fig. 1), which can be handled similarly to real sources. Generalizing the free-space MLFMA [17]–[19] to a half-space MLFMA is now straightforward. Besides some additional operations in the matrix-vector product (Section II-C), the preprocessing stage has to include calculations of the translation operators

$$T_L(k_i X^{\nu}_{m'm}, \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{X}}^{\nu}_{m'm}) = \sum_{l=0}^{L} (-j)^l (2l+1)\, h_l^{(2)}(k_i X^{\nu}_{m'm})\, P_l(\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{X}}^{\nu}_{m'm}) \tag{14}$$

between image cube and observation cube centers for all nonnearby cubes at all levels, as well as the Fourier transforms

$$\boldsymbol{B}^{\nu}_{m'\alpha'}(\hat{\boldsymbol{k}}) = \left[\overleftrightarrow{\boldsymbol{I}} - \hat{\boldsymbol{k}}\hat{\boldsymbol{k}}\right] \cdot \left[\overleftrightarrow{\boldsymbol{I}} - 2\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}\right] \cdot \iint_{S'} \boldsymbol{b}_{n'(m',\alpha')}(\hat{\boldsymbol{x}}') \cdot e^{+jk_i\hat{\boldsymbol{k}}\cdot\left[\overleftrightarrow{\boldsymbol{I}}-2\hat{\boldsymbol{z}}\hat{\boldsymbol{z}}\right]\cdot(\boldsymbol{x}'-\hat{\boldsymbol{x}}_{m'})}\, dS' \tag{15}$$

of the image expansion functions $[\overleftrightarrow{I} - 2\hat{z}\hat{z}] \cdot \boldsymbol{b}_{n'}(\hat{\boldsymbol{x}}')$, where $\boldsymbol{X}^{\nu}_{m'm}$ is a vector from the image source group center $\boldsymbol{x}^{\nu}_{m'}$ to the observation group center (Fig. 1), and the dyadic $\overleftrightarrow{I} - 2\hat{z}\hat{z}$ accounts for the relative orientation and location of the image expansion functions, respectively.

Before proceeding, we should add a few comments on the image sources and their respective clustering. One might argue that we could include *all* sources (i.e., sources and real image sources) in a single big box, to which we could then apply the conventional free-space MLFMA (with general complex wavenumber). However, each image magnitude depends on the plane-wave direction (or, equivalently, on the relative location of source and observer). This is in contrast to the behavior of the sources in a free-space MLFMA. Another interesting issue arises from a comparison of the outgoing plane wave expansions in (11) and (15). Some reductions in computational complexity and memory are possible when computing and storing only one set of vectors, although this approach has not been included in our numerical implementation up to now.

C. Matrix-Vector Product in Half-Space MLFMA

The MLFMA accelerates the solution of the integral equation (1) by reducing the complexity of the required RAM and the CPU needed for a single matrix-vector multiplication in an iterative conjugate gradient solution of (4) [34], [35] from O(N^2) for the MoM to O($N \lg N$) for the MLFMA [17]–[20]. We investigated several iterative solvers, including the standard conjugate gradient method operating on the normal equation (CGN), the biconjugate gradient (BiCG), the biconjugate gradient stabilized (BiCGStab), the conjugate gradient squared (CGS), and the transpose free quasi-minimum residual (TFQMR) methods [34], [35]. The latter three have the advantage that they only perform matrix-vector products $[\boldsymbol{Z}] \cdot \boldsymbol{I}$, while the CGN and BiCG methods also require the product $[\boldsymbol{Z}]^{\dagger} \cdot \boldsymbol{I}$ with the adjoint (i.e., conjugate transpose) matrix. Interestingly, for the problems we investigated up to now, the CGN method was often the best choice (see [36]), at least without a preconditioner. However, while the reduction in the number of iterations in the CGN when using a block-diagonal preconditioning scheme [19], [35] was moderate, we observed a significant speedup in the BiCGStab, CGS, and TFQMR, typically reducing the number of iterations by a factor of two to four.

Following the discussion in [17] for the free-space MLFMA,we briefly summarize the basic steps for performing the matrix-vector product $\boldsymbol{y} = [\boldsymbol{Z}] \cdot \boldsymbol{I}$ in the half-space MLFMA (the adjoint-vector product $[\boldsymbol{Z}]^{\dagger} \cdot \boldsymbol{I}$ can be derived in a similar way, if needed in the iterative solver). It consists of two sweeps and a final weighting by the plane wave expansion of the testing functions.

For a given vector $\boldsymbol{I}$, the first sweep calculates the outgoing plane wave expansions (i.e., the radiation patterns) of all nonempty cubes at all levels $\gamma = 1 \cdots g$, where the number $K_\gamma = 2L^2_\gamma$ of plane waves necessarily increases with increasing cube dimensions [14], [15], [17], [19] (i.e., it increases when going from the finest level $\gamma = g$ to the coarsest level $\gamma = 1$). While the outgoing plane wave expansions at the finest level g are calculated directly using the radiation patterns $\boldsymbol{B}_{m'\alpha'}(\hat{\boldsymbol{k}})$ and $\boldsymbol{B}^{\nu}_{m'\alpha'}(\hat{\boldsymbol{k}})$ of the basis and image basis function, respectively

$$\boldsymbol{s}^{(\nu)}_{m'_g}(\hat{\boldsymbol{k}}_{\kappa_g}) = \sum_{\alpha'_g=1}^{A_{m'_g}} I_{n'(m'_g,\alpha'_g)} \boldsymbol{B}^{(\nu)}_{m'_g\alpha'_g}(\hat{\boldsymbol{k}}_{\kappa_g}) \quad \text{for} \begin{cases} m'_g = 1, 2, \cdots, M_g \\ \kappa_g = 1, 2, \cdots, K_g \end{cases} \tag{16a}$$

the plane wave expansions at all remaining levels are obtained using shifting and interpolation [16], [17]. The interpolation step is necessary because the required number of plane waves K_γ at the coarser level γ is larger than the number of plane waves $K_{\gamma+1}$ at level $\gamma + 1$. Therefore, the expansions $\boldsymbol{s}^{(\nu)}_{m'_{\gamma+1}}$ are first interpolated (using a sparse interpolation matrix $[\boldsymbol{W}]$ according to a third-order Lagrange interpolation) and are then shifted to the cube center at the coarser level γ, and finally, the contribution from all child cubes is added

$$\boldsymbol{s}^{(\nu)}_{m'_\gamma}(\hat{\boldsymbol{k}}_{\kappa_\gamma}) = \sum_{\substack{\text{child} \\ \text{cubes}}} e^{-jk_i \hat{\boldsymbol{k}}_{\kappa_\gamma} \cdot (\boldsymbol{x}_{m'_\gamma} - \hat{\boldsymbol{x}}_{m'_{\gamma+1}})} \cdot \sum_{\kappa_{\gamma+1}=1}^{K_{\gamma+1}} W_{\kappa_\gamma \kappa_{\gamma+1}} \boldsymbol{s}^{(\nu)}_{m'_{\gamma+1}}(\hat{\boldsymbol{k}}_{\kappa_{\gamma+1}}) \quad \text{for} \begin{cases} \gamma = g-1, \cdots, 2, 1 \\ m'_\gamma = 1, 2, \cdots, M_\gamma \\ \kappa_\gamma = 1, 2, \cdots, K_\gamma \end{cases} . \tag{16b}$$

The second sweep calculates the incoming plane wave expansions at the finest level $\gamma = g$, starting with a direct calculation at the coarsest level $\gamma = 1$ and then recursively going to finer levels using shifting and anterpolation [16], [17]. At the coarsest level, the incoming plane wave expansions with phase reference at the cube centers are given by

$$\boldsymbol{g}_{m_1}(\hat{\boldsymbol{k}}_{\kappa_1}) = \sum_{m'_1 \in D_{m_1}} T_L(k_i X_{m'_1 m_1}, \hat{\boldsymbol{k}}_{\kappa_1} \cdot \hat{\boldsymbol{X}}_{m'_1 m_1}) \boldsymbol{s}_{m'_1}(\hat{\boldsymbol{k}}_{\kappa_1}) + T_L(k_i X^{\nu}_{m'_1 m_1}, \hat{\boldsymbol{k}}_{\kappa_1} \cdot \hat{\boldsymbol{X}}^{\nu}_{m'_1 m_1}) \overleftrightarrow{\boldsymbol{r}}(\hat{\boldsymbol{k}}_{\kappa_1}) \cdot \boldsymbol{s}^{\nu}_{m'_1}(\hat{\boldsymbol{k}}_{\kappa_1}) \quad \text{for} \begin{cases} m_1 = 1, 2, \cdots, M_1 \\ \kappa_1 = 1, 2, \cdots, K_1 \end{cases} \tag{17a}$$

$$\boldsymbol{g}_{m_\gamma}(\hat{\boldsymbol{k}}_{\kappa_\gamma}) = \sum_{m'_\gamma \in D_{m_\gamma}} T_L(k_i X_{m'_\gamma m_\gamma}, \hat{\boldsymbol{k}}_{\kappa_\gamma} \cdot \hat{\boldsymbol{X}}_{m'_\gamma m_\gamma}) \boldsymbol{s}_{m'_\gamma}(\hat{\boldsymbol{k}}_{\kappa_\gamma}) + T_L(k_i X^{\nu}_{m'_\gamma m_\gamma}, \hat{\boldsymbol{k}}_{\kappa_\gamma} \cdot \hat{\boldsymbol{X}}^{\nu}_{m'_\gamma m_\gamma}) \overleftrightarrow{\boldsymbol{r}}(\hat{\boldsymbol{k}}_{\kappa_\gamma}) \cdot \boldsymbol{s}^{\nu}_{m'_\gamma}(\hat{\boldsymbol{k}}_{\kappa_\gamma}) + \sum_{\kappa_{\gamma-1}=1}^{K_{\gamma-1}} \frac{w_{\kappa_{\gamma-1}}}{w_{\kappa_\gamma}} W_{\kappa_{\gamma-1}\kappa_\gamma} \boldsymbol{g}_{m_{\gamma-1}}(\hat{\boldsymbol{k}}_{\kappa_{\gamma-1}}) \cdot e^{+jk_i \hat{\boldsymbol{k}}_{\kappa_{\gamma-1}} \cdot (\boldsymbol{x}_{m_{\gamma-1}} - \boldsymbol{x}_{m_\gamma})} \quad \text{for} \begin{cases} m_\gamma = 1, 2, \cdots, M_\gamma \\ \kappa_\gamma = 1, 2, \cdots, K_\gamma \end{cases} \tag{17b}$$

where D_{m_γ} denotes all cubes well-separated (far) from cube m_γ at level γ (but not well-separated at the parent level $\gamma - 1$). The incoming plane wave expansions at successively finer levels γ are given by a contribution from all well-separated cubes at this level (but not well-separated at the parent level) and the

contribution from well-separated cubes at the parent level, the latter calculated using shifting and anterpolation as described in [16], [17], where the factors w_{κ_γ} are the weighting coefficients of the $K_\gamma = 2L_\gamma^2$-point Gaussian quadrature integration applied at level γ.

Finally, the resulting vector $\boldsymbol{y} = [\boldsymbol{Z}] \cdot \boldsymbol{I}$ of the matrix-vector multiplication is given by

$$\begin{aligned} y_n &= y_{n(m_g, \alpha_g)} \\ &= \sum_{n'=1}^{N} Z_{nn'}^{\text{near}} I_{n'} + \frac{\omega \mu_i k_i}{(4\pi)^2} \iint_{4\pi} \boldsymbol{W}_{m_g \alpha_g}(\hat{\boldsymbol{k}}) \cdot \boldsymbol{g}_{m_g}(\hat{\boldsymbol{k}})\, d^2\hat{\boldsymbol{k}} \\ &\approx \sum_{n'=1}^{N} Z_{nn'}^{\text{near}} I_{n'} + \frac{\omega \mu_i k_i}{(4\pi)^2} \sum_{\kappa_g=1}^{K_g} w_{\kappa_g} \boldsymbol{W}_{m_g \alpha_g}(\hat{\boldsymbol{k}}_{\kappa_g}) \\ &\quad \cdot \boldsymbol{g}_{m_g}(\hat{\boldsymbol{k}}_{\kappa_g}) \quad \text{for } n = 1, 2, \cdots, N \end{aligned} \tag{18}$$

where the first term accounts for near interactions (utilizing the sparse matrix $[\boldsymbol{Z}^{\text{near}}]$ introduced in Section II-A). Far interactions, i.e., the second term in (18), are included by an appropriate weighting (applying the receiving pattern of the weighting function and a Gaussian quadrature integration) of the incoming plane wave expansion with phase reference at the center $\boldsymbol{x}_{m_g}$ of the cube in which the weighting function $\boldsymbol{w}_n$ is located.

As can be easily seen, the first and second sweep require approximately twice the number of operations compared to a free-space MLFMA, and the final (and computationally less intensive) step (18) is unchanged. Therefore, the number of operations needed for one matrix-vector product only increases by a factor of about two compared to the free-space MLFMA. While the same is true for the calculation of the translation operators in the preprocessing stage (Section II-B), the increase in CPU for calculating the Fourier transforms of basis and weighting functions is even smaller (only about 50%). There is only one additional set of Fourier transforms (15) for the image basis functions but none for the weighting functions. Only the number of operations required for the evaluation of the sparse matrix $[\boldsymbol{Z}^{\text{near}}]$ does (in general) not scale with a factor of two (or less) compared to the free-space MLFMA, depending on the number of complex images needed for an accurate representation of the half-space dyadic Green's function. Nevertheless, for large scattering problems the CPU time for the iterative solution often dominates, and therefore an overall increase in the CPU time on the order of two was observed in most of our calculations relative to the free-space case. Before leaving the theoretical part of this paper, it should be mentioned also that the increase in RAM is always less than 100% (typically on the order of 20–60%) compared to the free-space case. An additional set of translation operators (14) and plane wave expansions of the image basis functions (15) must be stored in memory, but the memory required for the sparse matrix $[\boldsymbol{Z}^{\text{near}}]$ is unchanged. Therefore, the overall complexity in RAM as well as CPU (for one step in the iterative solver) remains O(Nlg$N$) like in the free-space MLFMA [17]–[20].

III. Results

All results are presented for perfectly conducting (PEC) targets in the presence of a half space, with the electrical properties

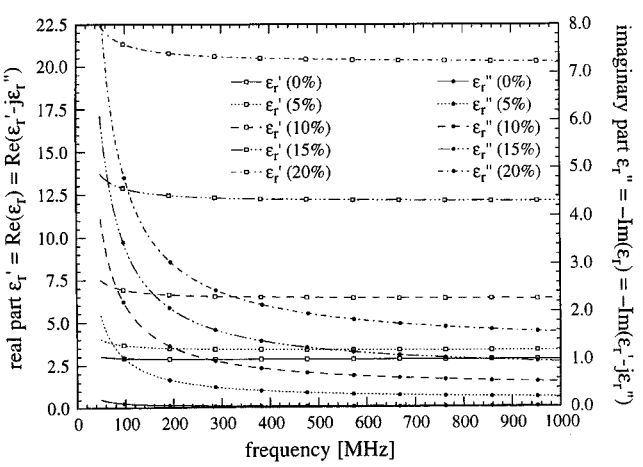


Fig. 2. Frequency dependent permittivity (real and imaginary part) of Yuma soil for different water content (percentage by weight) [37].

of the soil representative of soil samples taken from Yuma, AZ [37]. The soil electrical properties over the frequency range of interest are shown in Fig. 2 as a function of water content by percentage weight. We consider plane-wave excitation and all scattered fields are computed in the far zone. Results are shown for three target types, each of interest to remote sensing. First, we consider scattering from a perfectly conducting cylinder over a half space, with the cylinder axis normal to the half-space interface. Such rotationally symmetric targets, which give rise to no cross-polarized backscattered fields [38], are convenient calibration targets. The accuracy of the MLFMA for this target is verified through comparison with results computed via a MoM algorithm specialized to the case of a body of revolution (BOR). We next consider scattering from a buried missile-like target, of interest for the detection of buried UXO. In that example, we compare the MLFMA results to MoM and FMM solutions. Finally, we consider scattering from a vehicle-like target, of interest for military remote sensing. Due to the large size of that target, we do not have separate MoM results for comparison, although the previous two examples are meant to validate the accuracy of the MLFMA. We have found that (for closed targets) conjugate-gradient convergence is significantly improved by use of a CFIE formulation, *vis-à-vis* an EFIE or MFIE-based solution. This is in addition to the other salutary properties of a CFIE (no interior resonances [9], [17]). Therefore, all MLFMA computations employ the CFIE formulation.

A. Cylinder Over a Half Space

Consider a PEC cylinder situated over a half space, as shown inset in Fig. 3. The cylinder is 3 m long, has a diameter of 1 m, and is situated 20 cm over the soil interface. Electromagnetic scattering from this target can be analyzed via a specialized MoM solution tailored to a BOR [1], [2]. The BOR-MoM provides a good model for comparison to the MLFMA, because the BOR basis functions are entirely different than the RWG triangular patch basis functions [7] employed in the MLFMA. In particular, the BOR solution exploits the rotational symmetry of the target-(half space) composite, expanding both the Green's function and surface currents in a Fourier-series representation.

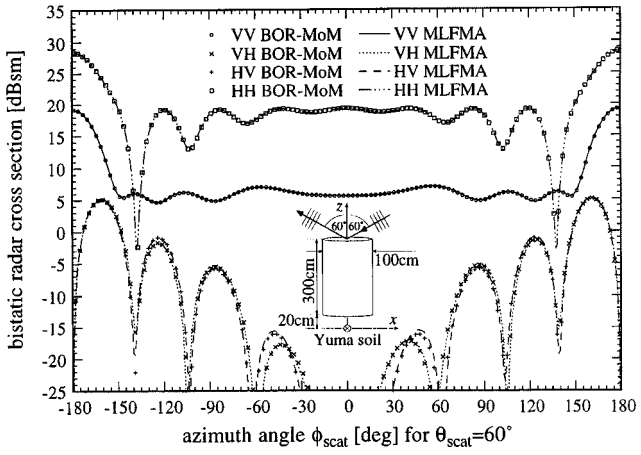


Fig. 3. RCS of a conducting cylinder of height 300 cm and diameter 100 cm (see inset and mesh in Fig. 4) situated 20 cm above Yuma soil of 10% water content (see Fig. 2). A plane wave is incident from $\theta_{inc} = 60^\circ$ (i.e., 30° from grazing) and $\phi_{inc} = 0^\circ$. The figure shows the bistatic RCS for $\theta_{scat} = 60^\circ$ and a frequency of $f = 600$ MHz. The MLFMA results (CFIE with four levels, $N = 15\,519$ unknowns, 548 MB of RAM) are compared to the RCS values computed via a rigorous BOR MoM (EFIE) [1], [2].

Consequently, the BOR solution is characterized by a current expansion along the generating arc only, *vis-à-vis* the MLFMA, which employs 3-D triangular patches to cover the whole target surface. The MLFMA basis functions are expanded at a rate of about 10 per wavelength, while for the purposes of providing a benchmark reference solution, the BOR basis functions are expanded at a rate of 20 subsectional basis functions per wavelength along the generating arc. In these examples the BOR-MoM results employ an EFIE [1], [8], while the MLFMA employs a CFIE.

In Fig. 3, we consider operation at a frequency of 600 MHz, which implies that the target is six wavelengths long, constituting a relatively large target, electrically. In particular, the MLFMA solution is characterized by $N = 15\,519$ unknowns (548 MB of RAM), for which we employ a four-level solution. The incident plane wave direction is fixed at $\phi_{inc} = 0^\circ$ and $\theta_{inc} = 60^\circ$, and the bistatic RCS is observed at $\theta_{scat} = 60^\circ$ for all ϕ_{scat}. Note that at $\phi_{scat} = 0^\circ$, the results correspond to backscatter, for which the cross-polarized fields vanish for this rotationally symmetric target [38]. In addition to noting the very close agreement between the BOR-MoM and MLFMA solutions, several other interesting observations can be gleaned from these results. In particular, the results presented here are for soil properties characteristic of Yuma soil with 10% water content (see Fig. 2). For such, the Brewster angle corresponds to approximately 68.8° (for lossy soil, there is no exact Brewster angle). The plane-wave incident angle considered here is close enough to 68.8° that very good soil penetration (and therefore little reflection) is achieved for vertical polarization. This explains why the HH response is considerably larger than its VV counterpart, the dihedral response of the latter being reduced significantly by soil penetration.

Before proceeding, we consider the cross-polarization results in Fig. 3. As expected, the cross-polarized fields are considerably smaller than their copolarized counterparts. An interesting issue involves reciprocity, from which we expect the VH and HV scattered fields to be identical. This is due to the specific incident and (bistatic) scattering directions used in this example, even though this is not true for the general bistatic case. Above, approximately -15 dBsm (about 45 dB below the maximum copolarized RCS), reciprocity is satisfied well. However, below -15 dBsm there is a noticeable difference between the VH and HV scattered fields. Interestingly, this does not appear to be a problem associated specifically with the MLFMA, since it is in very good agreement with the BOR-MoM results for both cross-polarized fields considered. We therefore attribute the errors in satisfying reciprocity to inaccuracies in computing the Green's function. The BOR-MoM model employs the method of complex images [25]–[27] for all (expansion function)–(testing function) interactions, while the MLFMA uses this same technique for the near interactions (Section II-A). In the complex-image technique, we fit each individual dyadic component of the spectral-domain Green's function to an exponential expansion with complex arguments, with the fit achieved via a least-square Prony's method [39]. We therefore set a predefined acceptable relative error between the exact and Prony representation of the spectral Green's function, with that error set to 1% in all our computations. Small errors in the expansion give rise to the errors in Fig. 3, manifested in the cross-polarized RCS. Nevertheless, the results demonstrate a dynamic range of approximately 45 dB, with the significant cross-polarization errors occurring primarily at angles for which the VH and HV RCS are very small.

We note that for a sampling rate of 20 basis functions per wavelength along the generating arc, the BOR-MoM results reported in Fig. 3 required 101 MB of RAM and about 70 min of CPU time (these numbers reduce to 26 MB and approximately 18 min for ten expansion functions per wavelength). By comparison, the MLFMA required 548 MB and nearly 90 min of CPU, at a sampling rate of ten basis functions per wavelength, with both the BOR-MoM and MLFMA computations performed on a SGI R10000/195 MHz processor. This raises an interesting question: For specialized targets such as the BOR in Fig. 3, when is it better to treat scattering with a specialized BOR-MoM code versus a more general but also highly efficient MLFMA scheme? In the BOR-MoM computations of Fig. 3, azimuthal modes $e^{\pm jm\phi}$ with m ranging from $m = 0$ to $m = 12$ were employed, based on the empirical relation in [2]. As the frequency increases, the number of azimuthal modes required increases, and the BOR-MoM formalism becomes less efficient. Numerical experiments have revealed that when the frequency of operation is increased above 1 GHz (see Fig. 4), the MLFMA becomes faster than the BOR-MoM solution. Such issues are addressed in greater detail in the future.

The results in Fig. 3 considered bistatic scattering at a single frequency. To address the performance of the MLFMA as a function of frequency, in Fig. 4 we present backscattered RCS for the incident angles considered in Fig. 3. The MLFMA and BOR-MoM employed a sampling of approximately ten basis functions per wavelength (recall that the MLFMA and BOR-MoM solutions are characterized by distinct basis-function types). As above, the BOR-MoM employs an EFIE, while the MLFMA is based on a CFIE formulation. We see in Fig. 4

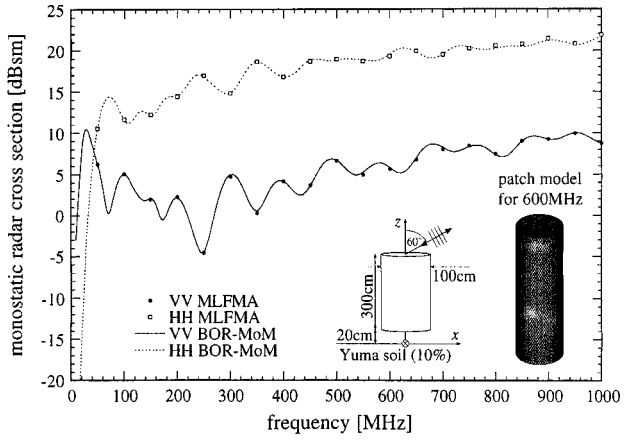


Fig. 4. As in Fig. 3, but plotting the *monostatic* RCS as a function of frequency. The discretization was fixed to about $\lambda 10$, for both the MLFMA and the body-of-revolution MoM. At $f = 1000$ MHz the MLFMA used $N = 42\,873$ unknowns, five levels, and a total of 1.57 GB of RAM.

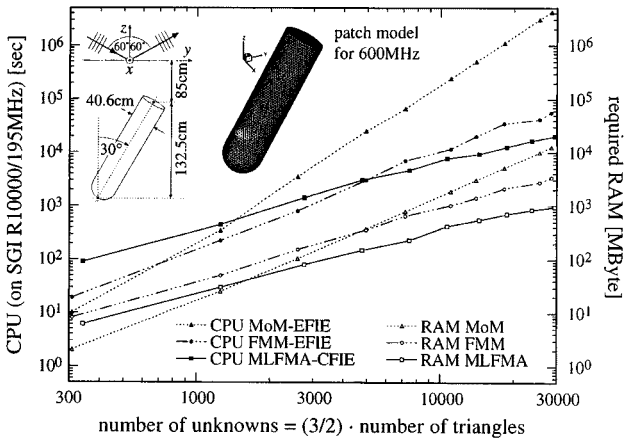


Fig. 6. Required RAM and overall CPU time (on SGI R10000/195 MHz) for bistatic RCS calculations as in Fig. 5 (polarimetric RCS matrix for 1800 scattering angles), for the frequency range 100–1000 MHz (corresponding to $N = 309$–$28\,539$ unknowns) using MoM (with LUD), FMM, and MLFMA. The CPU time for the MoM solution for $f > 600$ MHz ($N > 10\,653$) has been extrapolated.

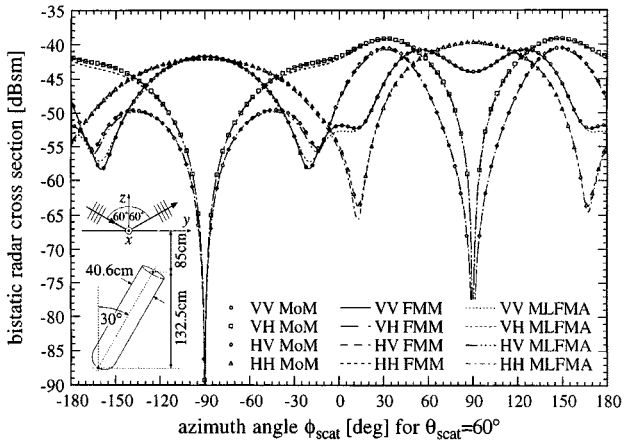


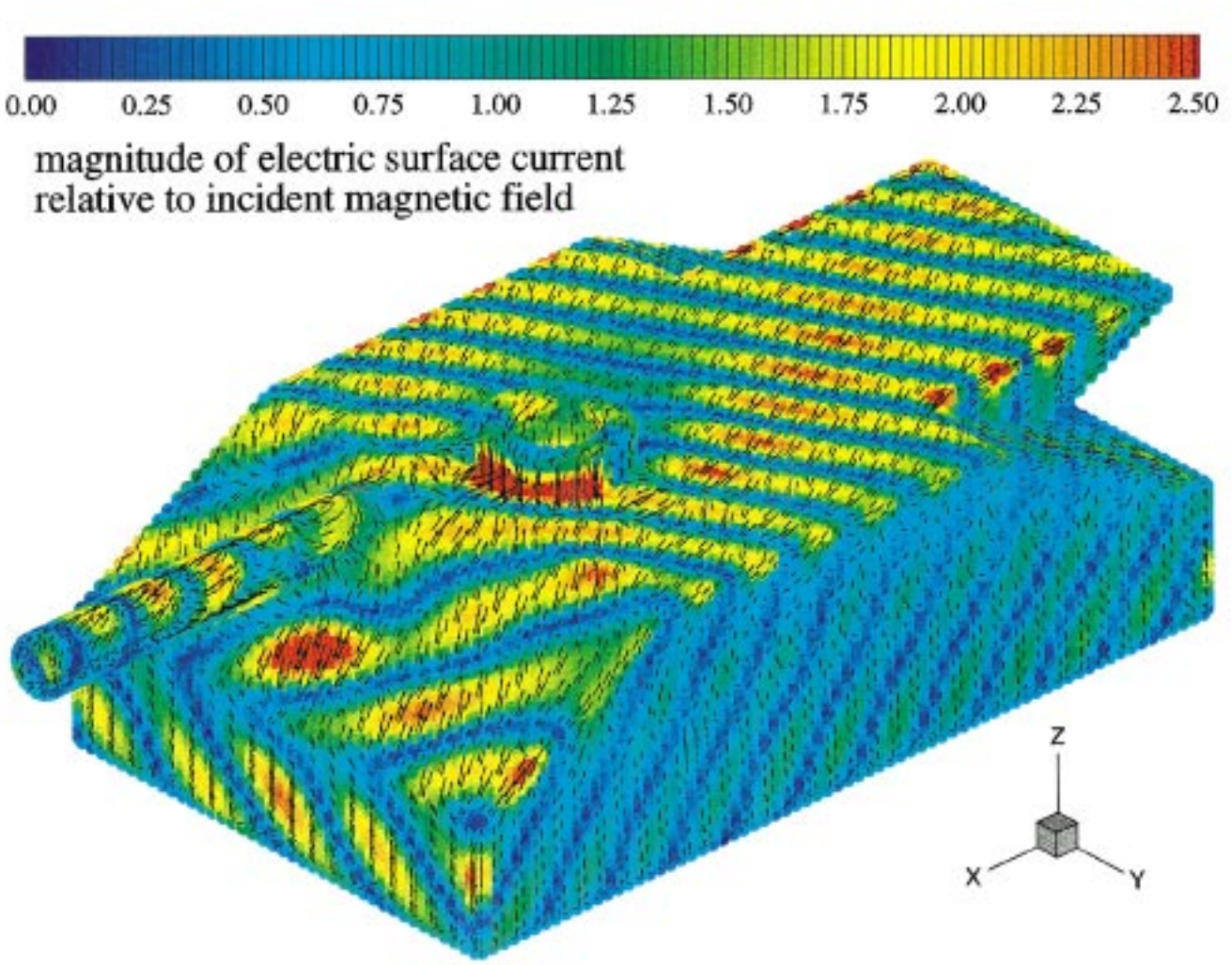
Fig. 5. Bistatic RCS of a model UXO (cylinder with hemispherical endcap) of length 153 cm and diameter 40.6 cm (see inset). The model UXO is buried in Yuma soil of 5% water content (see Fig. 2) with the target symmetry axis in the yz-plane, at an angle of 30° relative to the z-axis and the nose down at a depth $z = -217.5$ cm. The RCS is plotted for varying ϕ_{scat} at a fixed angle $\theta_{scat} = 60°$ and a frequency of 600 MHz ($N = 10\,653$ unknowns, see mesh in Fig. 6), assuming a plane wave incident at $\phi_{inc} = -90°$ (in the yz-plane) and $\theta_{inc} = 60°$. Results from an MoM (using EFIE), FMM (using EFIE), and MLFMA (using CFIE and four levels) code are compared.

that the agreement between the MLFMA and BOR-MoM is excellent, for all frequencies considered. Results are only shown for copolarization, the cross-polarized fields vanishing in backscatter, as discussed above. To get a feel for the complexity, at 300 MHz the MLFMA required 161 MB of RAM and 20 min of CPU (3936 unknowns, three levels), while at 1 GHz, these numbers are 1.57 GB and 227 min (42 873 unknowns, five levels), where all computations were run on a single SGI R10000/195 MHz processor.

B. Buried Unexploded Ordnance

In our next example we consider the ordnance-like target inset in Fig. 5, with the soil characterized by 5% water content (see Fig. 2). We compare results from the MLFMA, computed with a four-level CFIE, to results computed by the MoM and FMM, the latter two calculated using an EFIE. Note that, although the UXO target is rotationally symmetric, the target-(half space) composite in Fig. 5 is not, and therefore cannot be modeled with the BOR-MoM software. Consequently, both the MoM and FMM computations employ the same triangular-patch basis functions as the MLFMA model. Using the coordinate system shown inset in Fig. 5, the plane wave is incident at $\phi_{inc} = -90°$ and $\theta_{inc} = 60°$, and the bistatic fields are observed at $\theta_{scat} = 60°$ and all ϕ_{scat}. Due to bisectional symmetry, the cross-polarized fields should vanish at $\phi_{scat} = \pm 90°$. From Fig. 5, we see that the three solutions are in good agreement, for all bistatic angles considered. We also note that, with Figs. 3 and 4 for comparison, the scattering of such a buried ordnance is quite small at the frequency (600 MHz) considered, underscoring the difficulty of radar-based sensing of such targets. However, in a synthetic aperture radar (SAR) mode [37], the addition of signatures from multiple positions along the SAR aperture may provide a sufficient signature in the image for target detection. The MLFMA will be used in the future to perform such phenomenological studies as a function of target type, soil type, system bandwidth, aperture length, target orientation, and depression angle. Such a comprehensive set of data is difficult to acquire experimentally. The computations in Fig. 5 are characterized by 10 653 unknowns (see Fig. 6 for RAM and CPU requirements).

Such phenomenological studies have motivated development of the MLFMA model, and it is therefore of interest to assess the CPU and RAM requirements of this technique versus the other algorithms presented in Fig. 5, namely, the MoM and FMM. In Fig. 6, we present the CPU and RAM requirements for bistatic results as in Fig. 5, covering the frequency range 100–1000 MHz (corresponding to $N = 309$–$28\,539$ unknowns). As expected [17]–[20], the MLFMA requires order O($N \lg N$) RAM and CPU, while for the FMM these parameters are approximately order O($N^{3/2}$). The slight undulations in the RAM/CPU complexity of the single-stage FMM are caused by a suboptimal clustering employed in our implementation of the FMM [28], [29]. We have found in general that the half-space



(a)

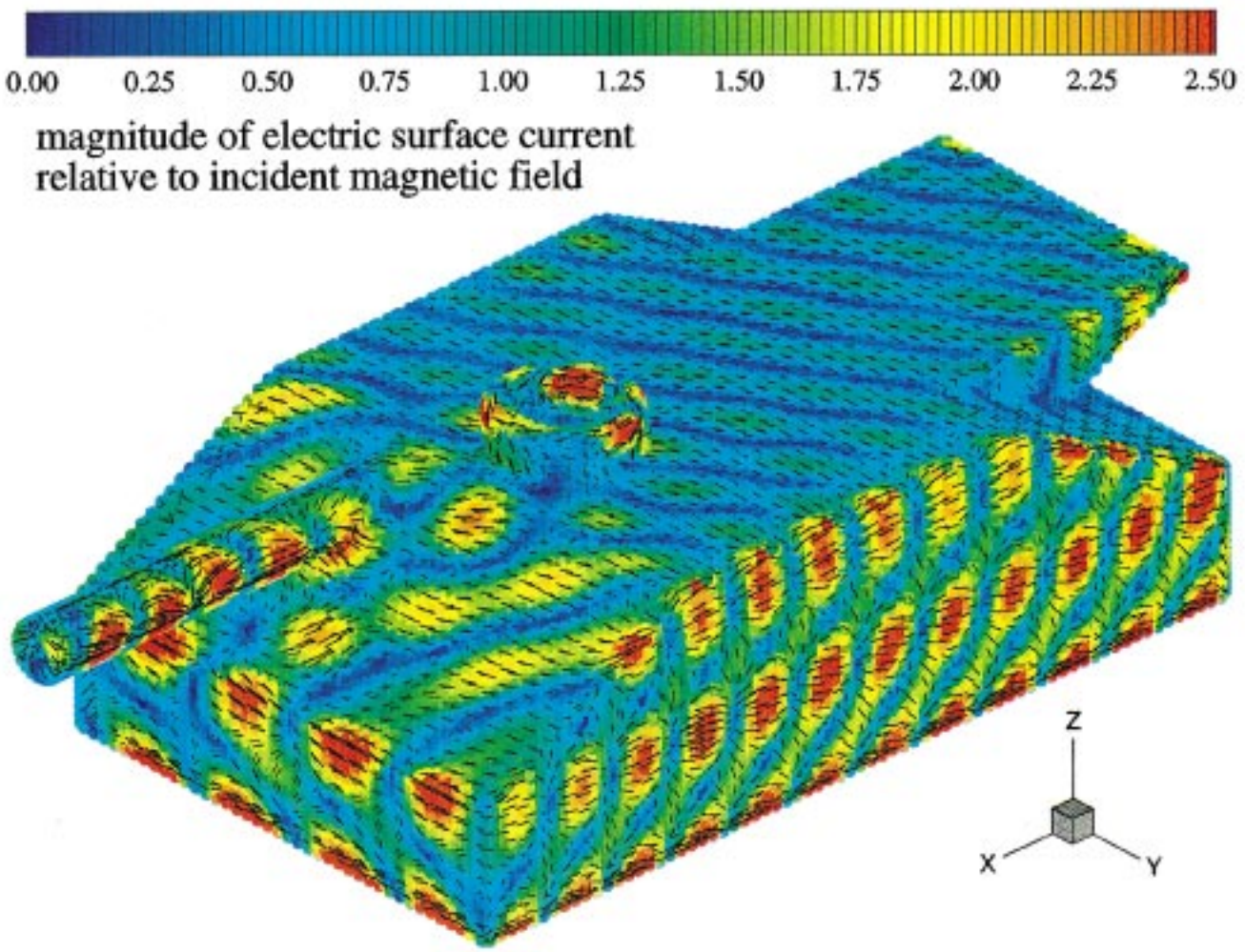
Fig. 7. Snapshot of the induced electric surface current on a model tank at a frequency of 700 MHz, for a plane wave incident from $\theta_{inc} = 60^\circ$ and $\phi_{inc} = 30^\circ$. The magnitude of the surface current is coded in color, the arrows indicate the direction of the current flow. The five-level MLFMA solved a CFIE with $N = 59\,253$ unknowns in about 7 h (both polarizations together) on a single SGI R10000/195 MHz processor using 2.68 GB (compared to 56.2 GB for a conventional MoM). (a) V-polarization incident.

MLFMA, as formulated here, requires typically 20–60% more RAM than its free-space counterpart and approximately twice the CPU time (see also discussion Section III-C). Fig. 6 clearly demonstrates the efficacy of the MLFMA, *vis-à-vis* the MoM and FMM, this CPU and RAM enhancement being particularly important when numerous phenomenological studies are required of electrically large problems.

C. Tank-Like Target

In our final example, we present results of interest to radar-based sensing of vehicles over soil, by considering the tank-like target in Fig. 7. The target mesh is shown inset in Fig. 8 (for a frequency of 300 MHz), along with its associated dimensions. The target is 20 cm above the half space, here characterized as Yuma soil with 10% water content (see Fig. 2). The currents induced on the target are shown in Fig. 7, for V-polarization (a) and H-polarization (b) plane-wave excitation at $\theta_{inc} = 60^\circ$ and $\phi_{inc} = 30^\circ$ (see coordinate system inset in the figures). The arrows indicate the direction of the current flow, and the colors represent the normalized current magnitude. These results were computed via a five-level CFIE MLFMA with 59 253 unknowns, requiring 7 h of total CPU run time (for both polarizations) on a SGI R10000/195 MHz processor, and 2.68 GB of RAM (a MoM solution requires 56.2 GB). The results presented here are for 700 MHz, corresponding to more than nine wavelengths along the length of the target.

The most striking aspect of the currents in Fig. 7 is the large electrical size of this target, reflected in the many undulations in current amplitude. An important distinction between the V-polarization and H-polarization results is reflected in the character of the currents induced on the flat portions of the target, which constitute one half of a dihedral reflector (the soil interface constituting the other half of the dihedral). Considering Fig. 7(b), we see the front and side faces of the target are characterized by a complicated interference pattern. This pattern reflects the superposition of currents induced directly by the incident plane wave (as if the target were in free space) plus the currents induced by the ground bounce of the incident fields. A similar pattern is manifested on the "cannon" portion of the target. By



(b)

Fig. 7. (*Continued.*) Snapshot of the induced electric surface current on a model tank at a frequency of 700 MHz, for a plane wave incident from $\theta_{inc} = 60^{\circ}$ and $\phi_{inc} = 30^{\circ}$. The magnitude of the surface current is coded in color, the arrows indicate the direction of the current flow. The five-level MLFMA solved a CFIE with $N = 59\,253$ unknowns in about 7 h (both polarizations together) on a single SGI R10000/195 MHz processor using 2.68 GB (compared to 56.2 GB for a conventional MoM). (b) H-polarization incident.

contrast, this pattern is substantially reduced for the V-polarization in Fig. 7(a). As discussed when presenting the results in Fig. 3, the near-Brewster excitation results in significant penetration of the incident wave into the soil, markedly reducing the strength of the ground bounce.

One of the fundamental challenges of detecting large man-made targets similar to that considered in Fig. 7 is the strong signature variation as a function of target-sensor orientation. Such targets are generally concealed, and therefore, the target aspect is generally unknown *a priori*, significantly complicating radar-based target detection and identification. The modeling of such phenomena is particularly important to the development of signal processing algorithms, since a comprehensive set of measurements, for all target-sensor orientations, frequencies, incident angles, polarizations, and soil types are generally difficult to attain.

In Fig. 8, we plot the monostatic RCS of the tank-like target, for the copolarized fields, at a frequency of 300MHz and an angle $\theta = 60^{\circ}$. These results involved $N = 13\,734$ unknowns, and the four-level MLFMA required 646 MB of RAM (an MoM solution requires 3.1 GB) and 129 h of CPU on one SGI R10000/195 MHz processor (for all 360 monostatic angles). The results were performed consecutively in azimuth, sampling ϕ at an increment of $\Delta\phi = 1^{\circ}$. To accelerate convergence of the conjugate-gradient solver, the current solution from the previous azimuthal angle was used as the initial guess for the subsequent solution, reducing the number of iterations per angle and polarization from approximately 90 (for zero initial guess) to an average of approximately 35. As indicated when discussing Fig. 3, the model's ability to satisfy reciprocity is one indicator of the dynamic range over which the results can be trusted. The cross-polarized fields for this target (not shown here) indicate that the results are highly accurate down to approximately -10 dBsm. Further accuracy can be achieved, if desired, by evaluating the Green's function more accurately for the near MLFMA terms (see the discussion associated with Fig. 3), by reducing the relative residual error at which the iterative solver is deemed to converge (for Figs. 8 and 9 we used 1%) and by using more terms in the various plane wave representations in the MLFMA (i.e., increasing the number of

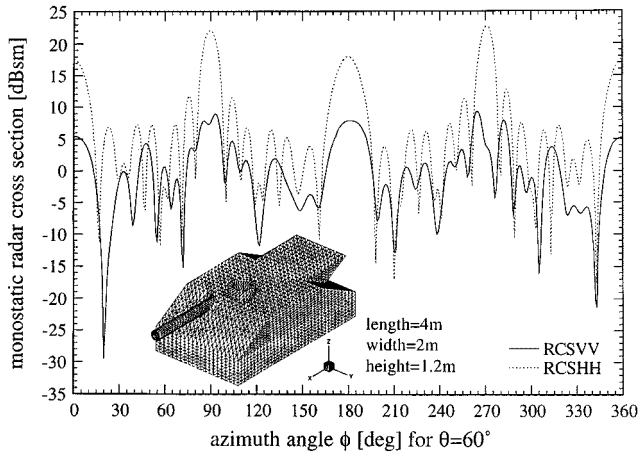


Fig. 8. Radar cross section (here we only plot VV and HH) of a model tank with $N = 13\,734$ unknowns (see inset) situated above Yuma soil of 10% water content (see Fig. 2) at a frequency of 300 MHz. *Monostatic* RCS is plotted for $\theta = 60^\circ$ (30° from grazing) and varying azimuth angles ϕ. The four-level MLFMA required a total of 646 MB of RAM (compared to 3.1 GB for a conventional MoM) and approximately 129 h of CPU time (on a single SGI R10000/195 MHz processor) for 360 monostatic angles.

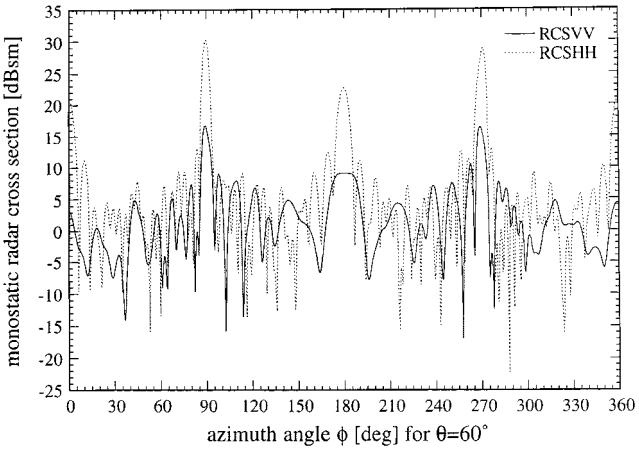


Fig. 9. As in Fig. 8, but for a frequency of 700 MHz and $N = 59\,253$ unknowns (the mesh in Fig. 8 was refined). The five-level MLFMA required 2.68 GB of RAM (compared to 56.2 GB for a conventional MoM solution) and 1022 h of total CPU time (on a single SGI R10000/195 MHz processor) for 720 monostatic angles.

terms L in the addition theorem expansion beyond the required L for an error less than 1%, which has been applied for all examples presented here) [17]–[20].

In Fig. 9, we again consider the monostatic RCS of the tank-like target, now computed at a frequency of 700 MHz. For these results, the azimuthal angle was sampled at a rate of $\Delta\phi = 0.5^\circ$, to accurately capture the faster variation of the RCS at higher frequencies. For these computations the target was characterized by $N = 59\,253$ unknowns, and the five-level MLFMA required 2.68 GB of RAM. As in Fig. 8, we consider $\theta = 60^\circ$. The computation of the full-polarimetric RCS (i.e., VV, VH, HV, and HH), at 720 different angles, required 1022 hours of total CPU time on a single SGI R10000/195 MHz processor. The RAM and CPU requirements (compared to the 300 MHz case) are consistent with the theoretical complexity of O($N\lg N$), when noting that the number of angles in the 700 MHz calculation was doubled. As expected, relative to the 300 MHz example of Fig. 8, the RCS results are characterized by substantially faster variation with azimuthal angle.

IV. Conclusions

An extension of the MLFMA has been presented for the analysis of electromagnetic scattering from electrically large perfect conductors situated in the presence of a half space. The structure of the half-space MLFMA is similar in many ways to its free-space counterpart [17]–[20]. In particular, we exploit properties of the free space Green's function to compute far (expansion function)–(testing function) interactions cumulatively through use of a multi-level clustering procedure [17]–[20]. Within the context of nonnearby cluster computations, interaction with the half space is accounted for approximately via polarization-dependent images located at real spatial locations. Hence, the first escalation in complexity of the half-space MLFMA relative to the free-space version is the introduction of image clusters, located in real space. In addition, the near interactions in the half-space MLFMA are treated via a rigorous analysis of the dyadic Green's function, here computed via the method of complex images [25]–[27]. We have found such a rigorous analysis of the near interactions essential for generating accurate results.

The MLFMA was calibrated by considering a special problem applicable to a MoM solution tailored to a BOR. This provides a good test, because the BOR-MoM solution is numerically rigorous, and because the BOR-MoM and MLFMA solutions employ distinct basis functions. The agreement between the MLFMA and BOR-MoM results was excellent, for all examples considered. Results were also presented for a UXO embedded in a lossy half space. For that problem the accuracy of the MLFMA was favorably compared with MoM and FMM solutions. Moreover, CPU and RAM comparisons clearly demonstrated the efficacy of the MLFMA technique, *vis-à-vis* the MoM and FMM. The final example considered a tank-like target, for which several results were presented.

The results for scattering from the tank-like target were insightful, while also demonstrating that, for such a large target, the memory requirements of the MLFMA are substantially smaller than those of the MoM. However, these results also clearly demonstrate that significant work is still required on the development of modeling algorithms for wideband, multi-aspect scattering from such targets. In particular, we noted that the multi-aspect computations in Figs. 8 and 9 required substantial CPU (although much less than would be required using a MoM model), and in practice, we would likely also require data over a wide frequency range. These observations point to directions for future research. The half-space MLFMA represents a numerically rigorous analysis of Maxwell's equations, within the context of the Green's-function approximation (for far interactions) discussed previously. Consequently, we can use the MLFMA results at high frequencies (such as in Fig. 9) to examine the accuracy of approximate, asymptotic solutions. The MLFMA can then be used to fill the frequency gap for which the MoM is intractable and for which asymptotic solutions are inappropriate. In addition, all results presented here were run on a single processor. It is of interest to develop a

scaleable, parallel version of the half-space MLFMA. Finally, in all MLFMA results, we have employed the RWG triangular-patch basis functions [7]. These basis functions are quite general, but for electrically large problems, they result in a very significant number of unknowns. Researchers have, within the context of MoM, developed hybrid techniques that combine subsectional basis functions with asymptotic traveling- wave and edge-diffraction basis functions [40]. The use of such basis functions substantially reduces the number of unknowns, reducing the problem size. It is of interest to incorporate such techniques into the MLFMA framework.

REFERENCES

[1] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse plane-wave scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 143–151, Feb. 1996.

[2] N. Geng and L. Carin, "Wideband electromagnetic scattering from a dielectric BOR buried in a layered lossy, dispersive medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.

[3] N. Geng, M. Ressler, and L. Carin, "Wideband VHF scattering from a trihedral reflector situated above a lossy dispersive half space," , to be published.

[4] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of a ground-penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–44, Jan. 1996.

[5] ——, "A complete electromagnetic simulation of the separated-aperture sensor for detecting buried land mines," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 1419–1426, Oct. 1998.

[6] H. S. Chang and K. K. Mei, "Scattering of electromagnetic waves by buried and partly buried bodies of revolution," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-23, pp. 596–605, 1985.

[7] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering from surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 30, pp. 409–418, May 1982.

[8] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, Parts I and II," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–352, Mar. 1990.

[9] A. F. Peterson, S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*. Piscataway, NJ: IEEE Press, 1998.

[10] J. M. Jin, *The Finite Element Method in Electromagnetics*. New York, NY: Wiley, 1993.

[11] A. Taflove, *Computational Electromagnetics—The Finite-Difference Time-Domain Method*. Boston, MA: Artech House, 1995.

[12] T. Dogaru and L. Carin, "Application of multi-resolution time-domain to two-dimensional electromagnetic scattering problems," , submitted for publication.

[13] A. Taflove, Ed., *Advances in Computational Electromagnetics—The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 1998.

[14] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas Propagat. Mag.*, vol. 35, pp. 7–12, June 1993.

[15] J. M. Song and W. C. Chew, "Fast multipole method solution using parametric geometry," *Microw. Opt. Technol. Lett.*, vol. 7, pp. 760–765, Nov. 1994.

[16] A. Brandt, "Multilevel computations of integral transforms and particle interactions with oscillatory kernels," *Comput. Phys. Commun.*, vol. 65, pp. 24–38, 1991.

[17] J. M. Song and W. C. Chew, "Multilevel fast multipole algorithm for solving combined field integral equations of electromagnetic scattering," *Microw. Opt. Technol. Lett.*, vol. 10, pp. 14–19, Sept. 1995.

[18] J. M. Song, C. C. Lu, W. C. Chew, and S. W. Lee, "Fast Illinois solver code (FISC)," *IEEE Trans. Antennas Propagat. Mag.*, vol. 40, pp. 27–33, June 1998.

[19] J. M. Song, C. C. Lu, and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 1488–1493, Oct. 1997.

[20] X. Q. Sheng, J. M. Jin, J. Song, W. C. Chew, and C. C. Lu, "Solution of combined-field integral equation using multilevel fast multipole algorithm for scattering by homogeneous bodies," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 1718–1726, Nov. 1998.

[21] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1970.

[22] L. Gürel and M. I. Aksun, "Electromagnetic scattering solution of conducting strips in layered media using the fast multipole method," *IEEE Microwave Guided Wave Lett.*, vol. 6, pp. 277–279, Aug. 1996.

[23] J. S. Zhao, W. C. Chew, C. C. Lu, E. Michielssen, and J. Song, "Thin-stratified medium fast-multipole algorithm for solving microstrip structures," *IEEE Trans. Microw. Theory Tech.*, vol. 46, pp. 395–403, Apr. 1998.

[24] Y. Rahmat-Samii, R. Mittra, and P. Parhami, "Evaluation of Sommerfeld integralsor lossy half-space problems," *Electromagnetics*, vol. 1, no. 1, pp. 1–28, 1981.

[25] J. J. Yang, Y. L. Chow, and D. G. Fang, "Discrete complex images of a three-dimensional dipole above and within a lossy ground," *Proc. Inst. Elect. Eng. H*, vol. 138, pp. 319–326, Aug. 1991.

[26] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.

[27] M. I. Aksun, "A robust approach for the derivation of closed-form Green's functions," *Trans. Microw. Theory Tech.*, vol. 44, pp. 651–658, May 1996.

[28] N. Geng, A. Sullivan, and L. Carin, "Fast multipole method for scattering from 3D PEC targets situated in a half-space environment," *Microw. Opt. Technol. Lett.*, vol. 21, pp. 399–405, June 1999.

[29] ——, "Fast multipole method for scattering from an arbitrary PEC target above or buried in a lossy half space," , to be published.

[30] I. V. Lindell, *Methods for Electromagnetic Field Analysis*. Piscataway, NJ: IEEE Press, 1995.

[31] D. R. Wilton, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak, and C. M. Butler, "Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains," *IEEE Trans. Antennas Propagat.*, vol. 32, pp. 276–281, Mar. 1984.

[32] R. E. Hodges and Y. Rahmat-Samii, "The evaluation of MFIE integrals with the use of vector triangle basis functions," *Microw. Opt. Technol. Lett.*, vol. 14, pp. 9–14, Jan. 1997.

[33] A. D. McLaren, "Optimal numerical integration on a sphere," *Math. Comput.*, vol. 17, pp. 361–383, 1963.

[34] T. K. Sarkar and E. Arvas, "On a class of finite step iterative methods (conjugate directions) for the solution of an operator equation arising in electromagnetics," *IEEE Trans. Antennas Propagat.*, vol. GE-33, pp. 1058–1066, Oct. 1985.

[35] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Boston, MA: PWS, 1996.

[36] N. M. Nachtigal, S. C. Reddy, and L. N. Trefethen, "How fast are nonsymmetric matrix iterations?," *SIAM J. Matrix Anal. Appl.*, vol. 13, pp. 778–795, July 1992.

[37] S. Vitebskiy, L. Carin, M. A. Ressler, and F. H. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.

[38] L. Carin, R. Kapoor, and C. E. Baum, "Polarimetric SAR imaging of buried landmines," , to be published.

[39] M. L. Van Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. Antennas Propagat.*, vol. 23, pp. 777–781, Nov. 1975.

[40] A. Sullivan and L. Carin, "A hybrid technique combining the moment method with physical optics and uniform asymptotics for scattering from 2D cylinders," *Microw. Opt. Technol. Lett.*, vol. 21, pp. 117–121, Apr. 1999.

Norbert Geng (S'91–M'96) was on born May 14, 1965 in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Germany, in 1991 and 1996, respectively.

From 1991 to 1996, he was with the Institute for Microwaves and Electronics, University of Karlsruhe, working on full-wave propagation modeling for radio communications systems. From 1997 to 1998, he was with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, as a Visiting Postdoctoral Researcher. From 1998 to 1999, he was with the University of Karlsruhe. His research at both of these universities focused on numerical techniques in computational electromagnetics for planar-stratified media. Since January 2000, he has been with SIEMENS Corporate Technology, Munich, Germany, as a Research Scientist on the design of broadband wireless communications systems.

Dr. Geng received the Mannesmann Innovation Award in 1997 for his Ph.D. dissertation on this subject.

Anders Sullivan (M'93) [redacted] in Staten Island, NY. He received the B.S. and M.S. degrees in aerospace engineering from the Georgia Institute of Technology, Atlanta, GA, in 1985 and 1987, respectively, and the Ph.D. degree in electromagnetics from Polytechnic University, Brooklyn, NY, in 1997.

From 1988 through May 1998, he was with the Air Force Research Lab, Eglin Air Force Base, FL. From June 1998 to September 1999, he was with the Electrical Engineering Department, Duke University, Durham, NC, as a Research Associate. Since September 1999, he has been with the Army Research Lab, Adelphi, MD. His current research interests include modeling complex targets and short-pulse scattering.

Dr. Sullivan is a member of the Tau Beta Pi and Sigma Gamma Tau honor societies.

Lawrence Carin (SM'96) [redacted] He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor, and became an Associate Professor in 1994. Since September 1995, he has been with the Electrical Engineering Department, Duke University, Durham, NC, where he is an Associate Professor. He is the Principal Investigator on a Multidisciplinary University Research Initiative (MURI) on demining. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION. He is a member of Tau Beta Pi and Eta Kappa Nu.

Ultrawide-Band Synthetic Aperture Radar for Detection of Unexploded Ordnance: Modeling and Measurements

Anders Sullivan, *Member, IEEE*, Raju Damarla, *Member, IEEE*, Norbert Geng, *Member, IEEE*, Yanting Dong, *Student Member, IEEE*, and Lawrence Carin, *Senior Member, IEEE*

***Abstract*—Electromagnetic (EM) scattering from subsurface unexploded ordnance (UXO) is investigated both theoretically and experimentally. Three EM models are considered: the multilevel fast multipole algorithm (MLFMA), the method of moments (MoM), and physical optics (PO). The relative accuracy of these models is compared for several scattering scenarios. Moreover, the model results are compared to data measured by an experimental synthetic-aperture radar (SAR) system. SAR images have been generated for subsurface UXO targets, in particular 155-mm shells. We compare SAR images from the measured data with theoretical images produced by the MoM and PO simulations, using a standard back-projection imaging technique. In addition to such comparisons with measurement, we consider additional buried-UXO scattering scenarios to better understand the underlying wave phenomenology.**

***Index Terms*—Buried object detection, ground-penetrating radar, synthetic aperature radar, ultrawide-band (UWB) radar.**

I. Introduction

THE electromagnetic (EM) characterization of surface and subsurface targets has been of interest to researchers for many years [1]–[11]. Recent development of wide-band high-resolution synthetic aperture radar (SAR) technology has shown that it may be possible to detect targets buried close to the ground surface over very large open areas [12]–[15], in a high standoff mode. In general, there are two limiting factors influencing the practicality of using wide-band SAR for wide-area target detection. First, the presence of strong ground clutter due to roughness, soil inhomogeneities, and foliage may limit the radar's ability to resolve the target from clutter. Likewise, target depth can also play a major role. For deeply buried targets, the incident wave may experience significant attenuation as it penetrates the lossy soil. Consequently, while wide-band SAR detection of targets close to the ground surface is generally considered to be feasible, detection of targets buried more than 1 m may be possible only in low-loss soils and/or for large targets [16]. In this paper, our model assumes a flat air–ground interface so we concentrate exclusively on the effects of target depth and orientation, as a function of sensor parameters. The dispersion and loss associated with the soil are modeled rigorously.

In the work presented here, we consider three models for simulating EM scattering from conducting unexploded ordnance (UXO): 1) a method-of-moments (MoM) analysis for arbitrary perfectly conducting targets in a layered medium, with the lossy dispersive layers representing the typical layered character of many soils; 2) a multilevel fast multipole algorithm (MLFMA) [17] model for electrically large conducting targets above or embedded within a lossy half-space; and 3) a physical optics (PO) model [18] for perfectly conducting UXO above or below a half space. In addition to addressing this scattering problem numerically, SAR images from UXO are presented for data collected with the U.S. Army Research Laboratory BoomSAR, Adelphi, MD [15], [19]. Comparisons are made between measured and computed SAR images, the latter simulated via the MoM and PO forward solvers. In addition to addressing the model accuracy through comparisons to measurements, we subsequently use the models to examine variation of the physical parameters to determine how such impact the SAR image.

The remainder of the text is organized as follows. In Section II, we summarize the numerical models developed for simulating wide-band scattering from surface and subsurface UXO and briefly describe the experimental SAR system employed in the measurements. A comparison between results from the three models is presented in Section III. We also perform comparisons between computed and measured SAR imagery, followed by numerical experiments to address the variation of the target signature as a function of variable physical parameters. The work is summarized and conclusions drawn in Section IV.

II. Theoretical Models and Measurement System

A. MoM and MLFMA Models

MoM [20], [21] and MLFMA [17], [22], [26] applied here have been discussed elsewhere. We, therefore, provide only a summary of the principal issues involved in such models. Our MoM and MLFMA models both solve a combined-field integral equation (CFIE) [27], by employing triangular-patch basis functions [28]. The MoM analysis is applicable to arbitrarily shaped perfectly conducting targets in an arbitrary layered medium, while the MLFMA is restricted to the half-space case. For both problems a principal challenge involves computation

Manuscript received September 2, 1999; revised March 14, 2000. This work was supported by the U.S. Department of Defense Strategic Environmental Research and Development Program (SERDP).

A. Sullivan and R. Damarla are with the Army Research Laboratory, Adelphi, MD 20783 USA.

N. Geng is with the Institute for Microwaves and Electronics, University of Karlsruhe, 76128 Karlsruhe, Germany.

Y. Dong and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0018-926X(00)09375-3.

0018–926X/00$10.00 © 2000 IEEE



Fig. 1. Photograph of a 155-mm shell.

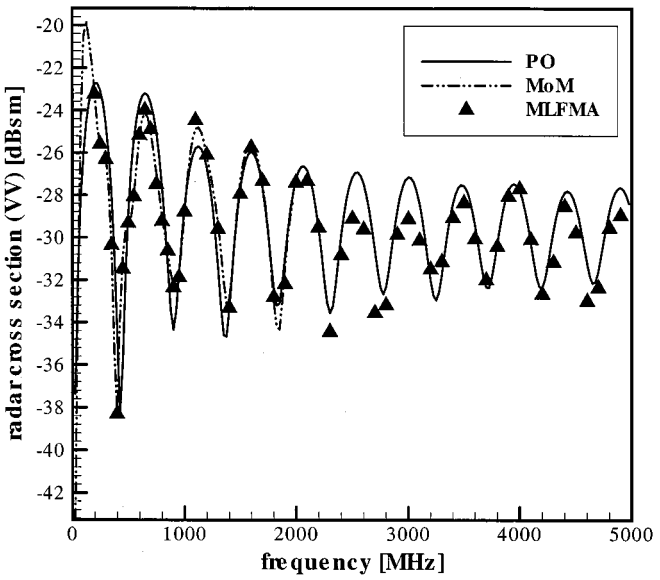


Fig. 2. Backscattered radar cross section for a 155-mm shell buried 2.54 cm beneath the soil surface, with the lossy dispersive soil characterized by Yuma 5% [15]. The excitation plane wave is incident 30° from grazing, propagating 45 from the target axis [see Fig. 6(b)]. Results are computed from MoM, MLFMA, and PO. The incident field is vertically polarized and we consider the VV-polarized backscattered fields.

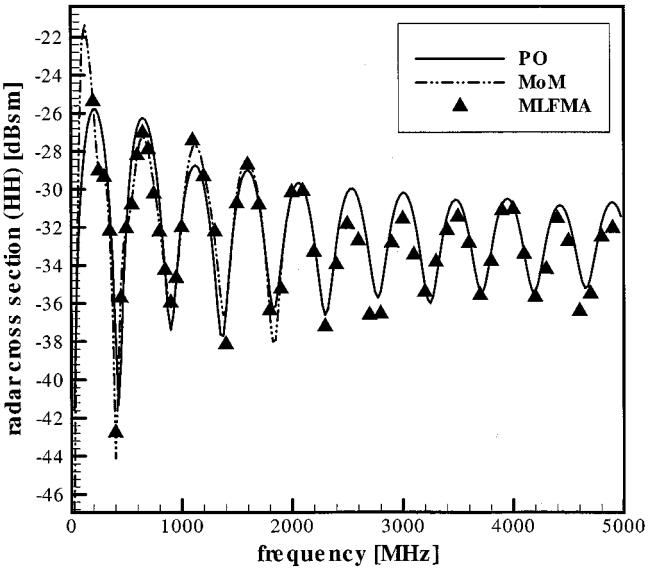


Fig. 3. As in Fig. 2, but for HH-polarization excitation and reception.

of the dyadic Green's function, each component of which involves Sommerfeld integrals. These integrals are solved here via the method of complex images [29]. In the MoM one computes an $N \times N$ matrix, for N expansion functions, with each matrix component involving a rigorous evaluation of the dyadic Green's function. In the MLFMA analysis, expansion and testing functions are partitioned into a set of multilevel clusters, with the "far" (expansion function)–(testing function) interactions treated efficiently within the MLFMA construct [17], [26]. For this portion of the model the dyadic half-space Green's function is evaluated approximately in terms of a set of images in *real* space [17], [22]. However, for basis and testing functions that are sufficiently close (typically half a wavelength or less) the interactions are treated rigorously, through application of the complex-image technique.

B. PO Solution

While the MLFMA is more efficient than the MoM for large N, both of these models are computationally expensive for electrically large targets. We have, therefore, also developed an approximate PO solution [18]—this significantly more efficient numerically than the MoM and MLFMA models. As applied here, the UXO is partitioned into a set of triangular sections, analogous to the triangular-patch basis functions applied in the MoM and MLFMA. The surface currents on all triangles in the lit region are approximated via the PO approximation, with the currents on all unlit triangles set to zero. The far-zone scattered fields from these approximate currents are computed through a rigorous use of the asymptotic half-space Green's function. Concerning the incident fields used to compute the PO currents, all reflection and transmission at the half-space interface is accounted for rigorously (although multiple interactions between the target and soil interface are neglected).

C. Experimental System

The U.S. Army Research Laboratory has developed an experimental time-domain SAR system, with instantaneous bandwidth covering 50–1200 MHz. In this system, four horn antennas are placed atop a telescoping boom lift capable of moving at approximately 1 km/h while the basket is elevated to 45 m [14], [15]. For typical collection geometries, down-look angles to the target vary from 45° to approximately 10° (from grazing), depending on the range to the target and the height of the boom. The details of this fully polarimetric system are found in [14] and [15], as is a discussion on SAR-image formation.

III. Theoretical and Measured Data

A. Comparison of Theoretical Models

We first consider a direct comparison between the three theoretical models presented in Section II. In particular, we consider the 155-mm shell in Fig. 1 (155 mm maximum diameter). The

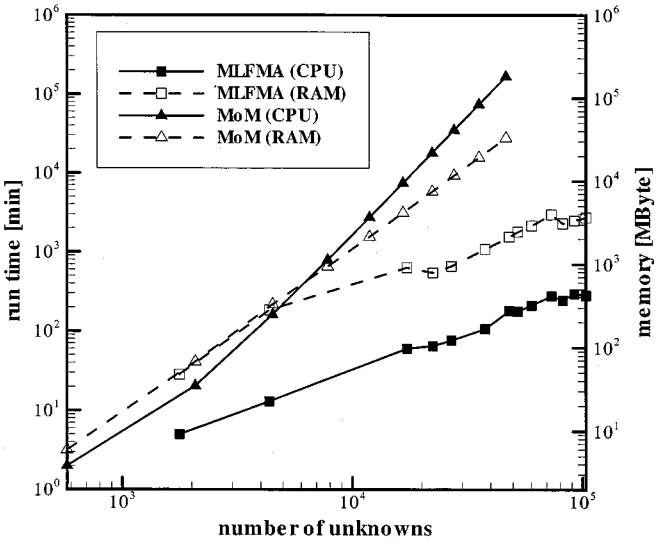


Fig. 4. For the results in Fig. 2, a comparison between the central processing unit (CPU) and memory (RAM) of the MoM and MLFMA results. The MoM results above 2 GHz are extrapolated.

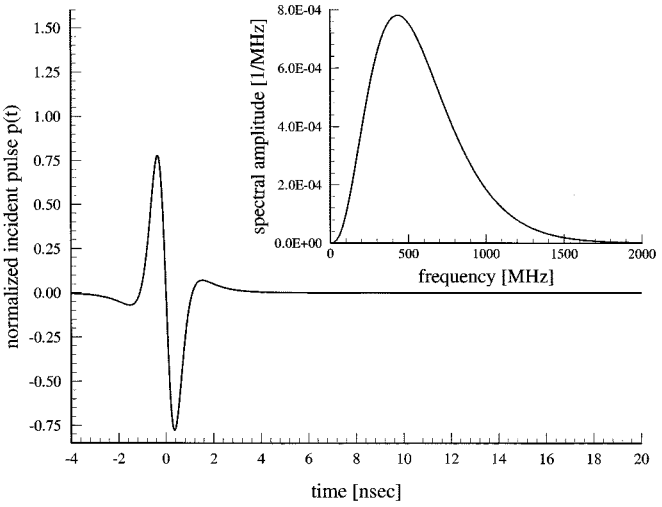


Fig. 5. Pulse shape and pulse spectrum (inset) for the time-domain plane-wave excitation used in the computed SAR images (Figs. 6 and 7).

length of the shell is 0.58 m, representing nearly four free-space wavelengths at 2 GHz (the *buried* target is of course even larger electrically). The length of the UXO is parallel to the air–ground interface, and the top of the UXO is 2.54 cm beneath the interface, the soil representative of Yuma soil with 5% water content [15]. In Figs. 2 and 3 we plot the backscattered RCS for vertical (VV) and horizontal (HH) polarization, respectively, for incidence 30° from grazing, directed 45° azimuthally to the target broadside [see Fig. 6(b)]. In these results we plot data up to 5 GHz, with the MLFMA and PO results plotted for all frequencies and the MoM results plotted only up to 2 GHz. The CPU and RAM requirements of the MoM are prohibitive when the target becomes electrically large, as it does above 2 GHz (see Fig. 4). Hence, the MLFMA is used to examine the accuracy of PO when the target electrical size becomes large, corresponding to where the PO is expected to perform well.

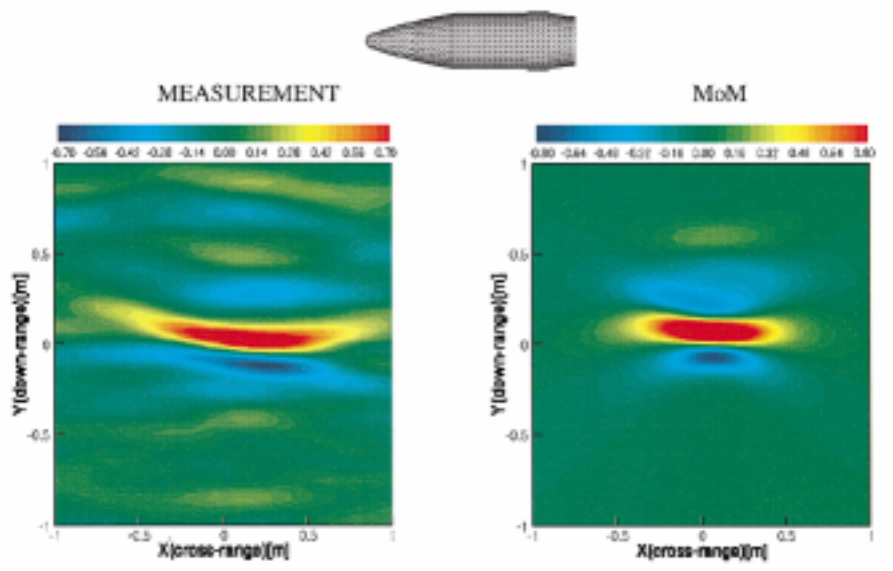
The PO solution is in good agreement with the much more computationally expensive MoM and MLFMA solutions, for almost all frequencies. The computation time of the PO results are infinitesimal (almost instantaneous) relative to MoM and MLFMA. With regard to the MoM and MLFMA solutions, the agreement between these models is almost exact, for all frequencies for which both are considered. However, as indicated in Fig. 4, the CPU and RAM requirements of the MoM are substantially higher than those of the MLFMA. In Fig. 4, for frequencies greater than 2 GHz, the MoM CPU and RAM are extrapolated from results at and below 2 GHz, assuming the expected order N^3 and N^2 variation, respectively (the MoM matrix equation was solved via a direct LU-decomposition, this of order N^3 complexity).

Close inspection of Figs. 2 and 3 reveals slight differences between the PO and MLFMA solutions, at frequencies above 2 GHz. This may appear somewhat anomalous, for the PO approximation is expected to become nearly exact at high frequencies. This latter expectation, however, is based on experience from free-space scattering. For the half-space problem, the PO solution does not account for multiple interactions between the target and the soil, regardless of the frequency. To this we attribute the small differences between the PO and MLFMA solutions at high frequencies. This shortcoming of the PO solution is examined further below, in the context of SAR images for the buried UXO. In Figs. 2 and 3 we only consider a single target-sensor orientation. In the SAR image we implicitly view the target from multiple orientations, and the aspect-dependent character of the signatures is also addressed more fully in Section IV.

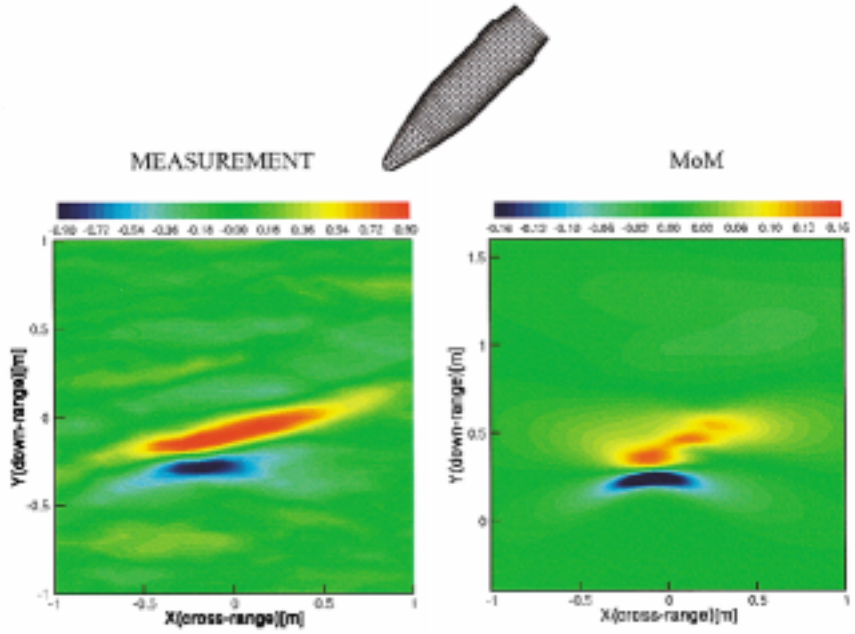
B. Measured and Computed SAR Imagery

We compare the models to measurements by forming theoretical SAR images. All SAR-image examples below, both theoretical and experimental, are for VV polarization, approximately 30° incidence from grazing. In computing the time-domain scattered waveform $s_k(t)$ from the kth aperture position, we account for the k-dependent angle of incidence to a given image pixel, assuming plane wave excitation, pure polarization properties and a fixed incident-pulse shape (see Fig. 5). The latter two conditions are dictated by our incomplete knowledge of the antenna properties over ultrawide-band (UWB) frequencies for, in reality, the radiated pulse shape changes with variable target-antenna orientation, as does the polarization. On the other hand, the plane wave incidence approximation is typically appropriate for the SAR applications of interest since the distance between the sensor and targets is generally substantial. Finally, the waveforms $s_k(t)$ used to form the theoretical images are weighted as a function of aperture position, to approximately reflect the properties of the antenna pattern (see [15]).

The need to compute time-domain scattered waveforms at a large number of target-sensor orientations, for formation of a theoretical SAR image, places a significant burden on the computational requirements of the scattering model. In particular, the time-domain computations for each $s_k(t)$ used in the SAR image required 80 frequency calculations (25-MHz increments) over the 50–1200 MHz bandwidth of the experimental SAR system (from Fig. 5 we note that, while the usable SAR bandwidth is 50–1200 MHz, calculations are required up to 2 GHz to cover the full bandwidth of the incident pulse).



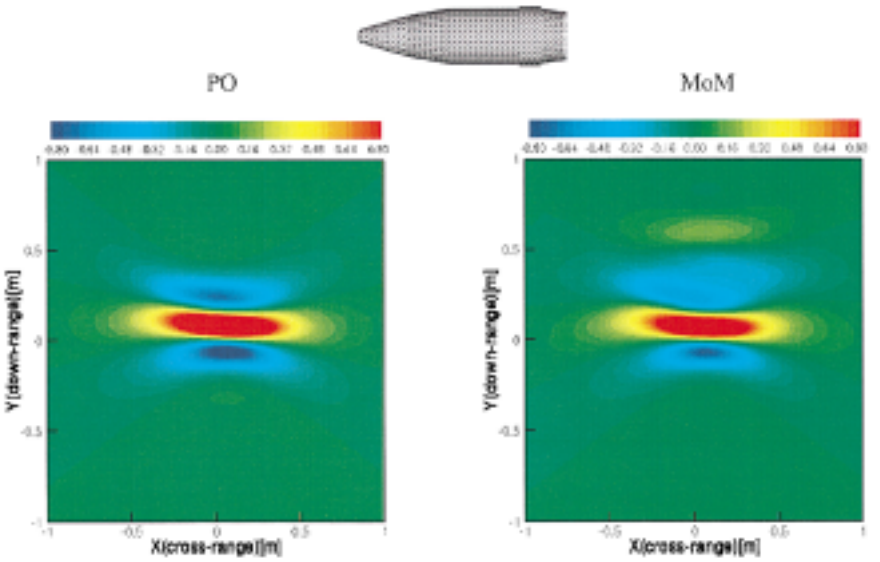
(a)



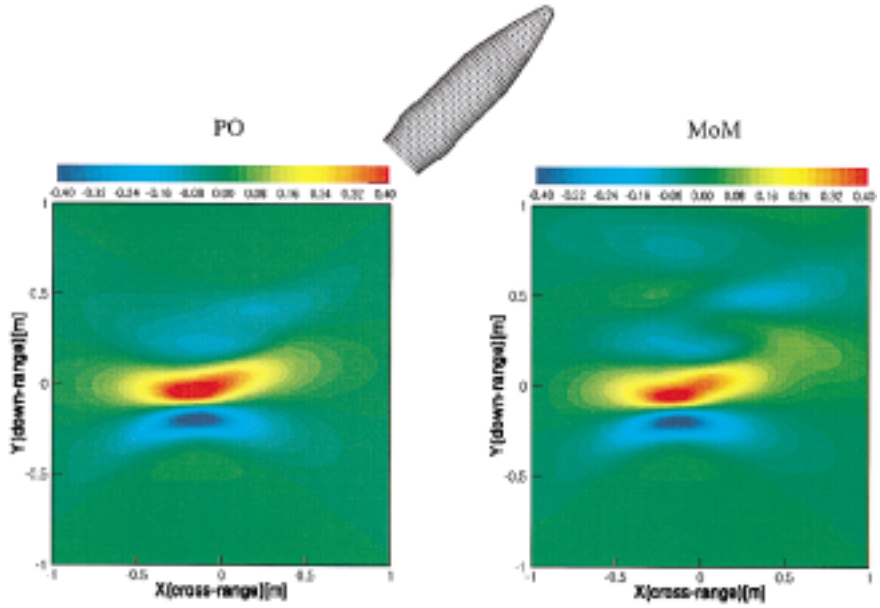
(b)

Fig. 6. Comparison of computed and measured SAR images for the 155-mm shell (Fig. 1) flush-buried in 5% Yuma soil [15]. The computations use pulsed plane-wave excitation at 30° from grazing, characterized by the pulse in Fig. 5. The measured and computed images use an aperture length that yields a 60° angle between the target center and aperture. (a) Shell axis parallel to linear SAR aperture. (b) Shell axis 45° from linear SAR aperture.

Moreover, for computations of the SAR image, we employed an azimuthal sampling rate of 1°. We view the target from different azimuthal positions, as we traverse the linear SAR aperture and here we have precomputed time-domain scattered waveforms at a 1° azimuthal sampling rate. Over the linear SAR aperture we must, in principle, also consider variation of the incident angle θ, although for the aperture sizes considered this variation is negligible. Note that for rotationally symmetric land mines [12]–[14], [21], the target signature is independent of the azimuthal orientation, substantially reducing the number of scattering computations required to form the image [15]. The UXO half-space composite generally does not possess such symmetry. Consequently, in forming all the theoretical SAR images presented below, we employ MoM and PO as the numerical forward solvers. If the MoM matrix equation is inverted via a direct LU-decomposition then, at a given frequency, multiple incident fields can be considered with little additional computational burden. By contrast, with the MLFMA the matrix equation is solved iteratively via the conjugate-gradient method [30]. In such iterative solvers, multiple incident fields must generally be computed separately, vitiating the use of MLFMA for theoretical SAR-image formation. We should note, however, that the MoM matrix multiplication is an order N^2 operation and for the image formation this must be done for all backscatter angles of interest (usually a large number). Therefore, overall, for the imaging problem



(a)



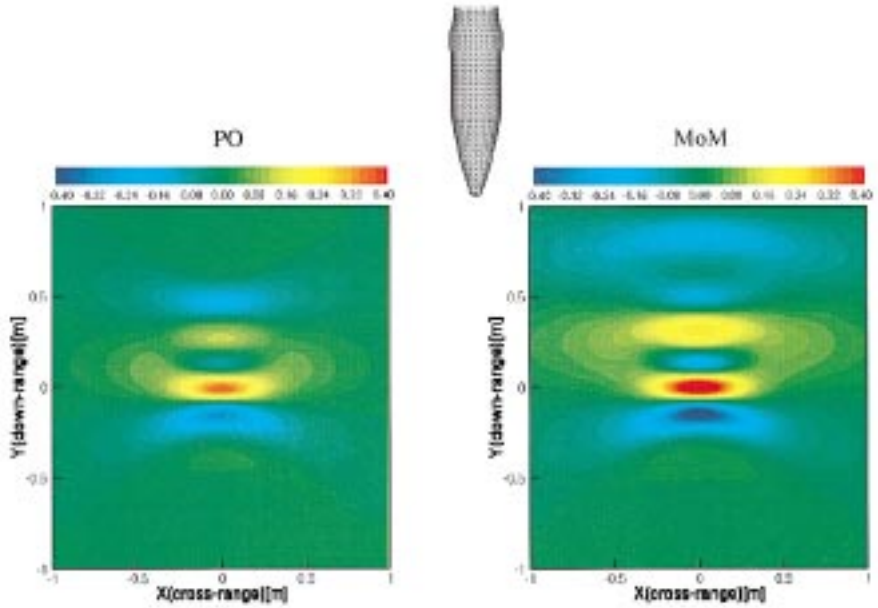
(b)

Fig. 7. Comparison of computed SAR images for the 155-mm shell (Fig. 1) buried 2.54 cm in soil characterized by $\epsilon_r = 5$ and $\sigma = 0.003$ S/m, with the target axis parallel to the air–soil interface. The computations use pulsed plane-wave excitation at 30° from grazing, characterized by the pulse in Fig. 5.The images use an aperture length that yields a 60° angle between the target center and aperture. The images are produced by scattered fields from MoM and PO forward solvers. (a) Incidence normal to the target axis. (b) Incidence 45° from target axis.

of interest, MoM and MLFMA have similar computational burden when N is large. In this context, the principal benefit of MLFMA *vis-á-vis* the MoM is in the area of RAM; the finite memory of a given computer limits the maximum target electrical size that can be considered, the MLFMA allowing consideration of significantly larger targets than the MoM (cf. Fig. 4).

The measured data is presented for the 155-mm shell shown in Fig. 1, for a test site at Yuma Proving Ground in Yuma, AZ. The Yuma soil was characterized by approximately 5% water content, with the associated electrical properties described in [15]. All measured results are for the UXO buried just under the air–ground interface, with the target axis parallel to the interface (flush buried). The gridded model of the target, used in the numerical computations, is shown in Fig. 6.

In Fig. 6(a) we consider the shell-oriented parallel to the linear SAR aperture, while in Fig. 6(b) the ordnance is oriented at 45°. For both images the total angle spanned by the linear aperture relative to the target center is 60°. The agreement between the model and measured SAR image is reasonable, especially considering the complexity of the experimental system and the uncertainty in the angle-dependent incident-wave polarization and pulse shape. As might be expected, when the target broadside is parallel to the linear aperture, the image is characterized primarily by a single strong return. Moreover, when tilted at 45° the scattered return is characterized by two



(c)

Fig. 7. (*Continued.*) Comparison of computed SAR images for the 155-mm shell (Fig. 1) buried 2.54 cm in soil characterized by $\epsilon_r = 5$ and $\sigma = 0.003$ S/m, with the target axis parallel to the air–soil interface. The computations use pulsed plane-wave excitation at 30° from grazing, characterized by the pulse in Fig. 5.The images use an aperture length that yields a 60° angle between the target center and aperture. The images are produced by scattered fields from MoM and PO forward solvers. (c) Incidence parallel to target axis.

diagonally offset returns, characteristic of diffraction from the target front and back. Note that, in the model results in Fig. 6(b), one can almost distinguish the two (diagonally offset) scattering mechanisms, while in the measured response these appear to merge into a single diagonal response. This may be due to an over estimation in the model of the SAR's cross-range resolution. Moreover, random motion of the sensor (for example, due to wind) also undermines the resolution of the measured image.

While the principal features of the measured and computed imagery (Fig. 6) are similar, there is a slight difference in the color scale of the images. This scaling difference is due to the fact that the experimental system is not calibrated, this constituting a very challenging task for the UWB system under consideration. Addressing this point further, to do system calibration, one requires knowledge of the theoretical response from fiducial targets to which the radar will be calibrated. Accurate models for the large fiducial targets of interest [20] is a nontrivial issue, it constituting a separate application of the MoM and MLFMA models.

While the theoretical SAR images in Fig. 6 capture most of the features in their measured counterparts, we notice several features in the measured images that are not seen in the computed data, this particularly true in Fig. 6(b). Fig. 6(b) is more sensitive to clutter because the target signature is driven by relatively weak edge diffraction, as compared to the specular reflection characteristic of Fig. 6(a). The image features not modeled by the theory are attributed to clutter, this, as discussed in Section I, constituting one of the principal challenges to SAR-based UXO detection (and why, as discussed further below, SAR is best for UXO near the soil interface). The measurements were taken in a relatively benign environment, with minimal foliage cover on the flat terrain. Consequently, the clutter in Fig. 6 is attributed primarily to subsurface soil inhomogeneities. It is very difficult to include such effects in a model for comparison to measured scattering data since the subsurface soil characteristics are often altered upon excavation (if due to localized soil moisture, for example). Clutter induced by subsurface inhomogeneities must be handled statistically, employing an appropriate statistical model for the subsurface soil properties.

Based on the accuracy of the PO results in Figs. 2 and 3, we next perform a comparison between a SAR image computed via PO with one computed via MoM, with both corresponding to the 155-mm shell buried 2.54 cm under the Yuma-soil surface (5% water content [15]), as in Figs. 2 and 3. We see in Fig. 7 that the agreement between the PO and MoM images is consistent with the comparisons in Figs. 2 and 3. We note, however, that there are noticeable small differences between the MoM and PO solutions. In particular, consider in Fig. 7(a) a linear cut in the image, running parallel to the vertical axis, through the image center (through the zero cross-range position). Increasing down-range corresponds to further distance between the target and sensor, and therefore increasing downrange in the image can also be viewed as increasing time (reflections viewed at the sensor at later times are represented in the image as being further from the sensor). Note that with increasing time (increasing downrange position), the MoM is characterized in Fig. 7(a) by a sequence of three bumps, followed by a faint fourth return. By contrast, the PO solution in Fig. 7(a), along the same cut, does not have this fourth faint return. This is explained as follows. The PO solution only includes the direct reflected waveform, with no multiple interaction between the buried target and the air–soil interface. By contrast, the rigorous MoM solutions should capture all phenomena, assuming sufficient numerical accuracy (e.g., enough basis functions). Consequently, the small differences between the PO and MoM solution in Fig. 7(a) are attributed to mul-

tiple interaction between the target and the interface. Similar small differences are witnessed in Fig. 7(b) and (c). Nevertheless, these multiple interactions appear to constitute a small effect for this target. Such multiple interactions between the target and interface may also explain the small differences between the PO and rigorous solutions (MoM and MLFMA) seen in Figs. 2 and 3. Before leaving Fig. 7, it is important to emphasize that the PO solution captures most of the important shapes in the SAR image, relative to MoM. However, there is a noticeable difference in the amplitude of the PO solution, compared to MoM. We attribute this to the fact that the PO images use the PO forward solver at all frequencies, this being particularly inaccurate at the low frequencies of the incident pulse in Fig. 5, where there is considerable energy.

C. Phenomenological Studies

The principal goal of developing theoretical scattering models for surface and subsurface UXO is to gain a better understanding of the underlying wave phenomenology. We consider such phenomenological investigations below. Despite the fact that the PO model results in Sections IV-A–B looked promising, when performing phenomenological studies it is best to use a rigorous model, to be assured unexpected phenomena are witnessed, if present. Therefore, in the examples below, all data are computed via MoM and MLFMA.

The SAR imagery presented in the previous section indicates that the measured and computed SAR image for the target and soil considered is a function of the target orientation relative to the linear synthetic aperture. This is a very important phenomenon, for it is such aspect-dependent UXO scattering that will help discriminate targets from clutter. To address this issue more completely, in Fig. 8 we plot the time-domain signature of the UXO target due to a pulsed plane wave of excitation in Fig. 5. The target is a 155-mm shell buried as in Fig. 7 and here we plot the time-domain backscattered waveform as a function of target-sensor orientation. In particular, in Fig. 8, we plot the VV and HH time-domain scattered fields, for: 1) incidence normal to the target broad side; 2) at 45° relative to broad side; and 3) along the target axis. The waveforms in 1)–3) are related to Fig. 7(a)–(c), respectively. These data were computed via the MoM model discussed in Section II-B. As expected, the time-domain signature is a strong function of orientation. The far-zone time-domain scattered waveforms in Fig. 8 are normalized by the distance r between the target and observer and, therefore, the fields have units meters (analogous to square meters for RCS).

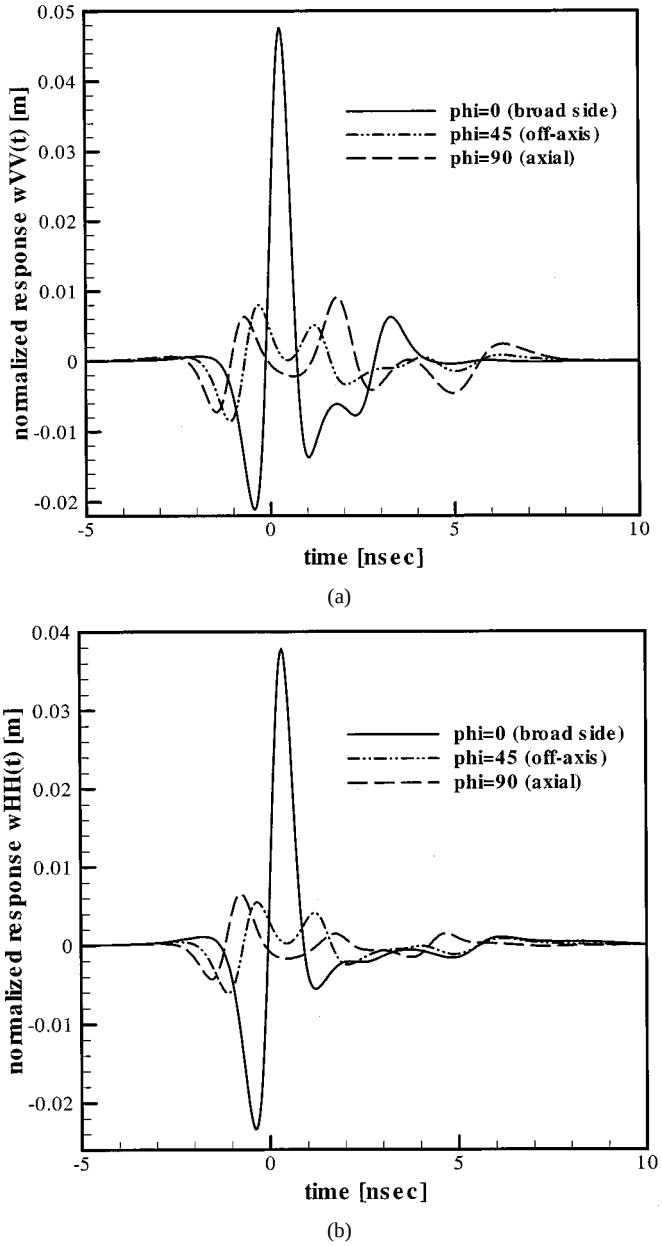


Fig. 8. Time-domain backscattered waveforms due to pulsed plane wave excitation at 30° from grazing, with an incident pulse as in Fig. 5. The 155-mm shell (Fig. 1) is buried 2.54 cm in 5% Yuma soil [15], with its axis parallel to the air–soil interface. Considered are plane wave incidence normal to the target axis, 45° from the target axis, and parallel to the target axis. (a) VV polarization. (b) HH polarization.

As expected, the broadside scattered waveform ($\phi = 0°$) in Fig. 8 is significantly stronger than the other two angles, highlighting the distinction between primarily specular ($\phi = 0°$) and edge-diffraction ($\phi = 45°$ and $\phi = 90°$) scattering. Considering first the case $\phi = 0°$, we note the large specular return is followed by a considerably smaller second pulse (although that second pulse is nearly as large in amplitude as the scattered fields at $\phi = 45°$ and $\phi = 90°$). The second waveform for $\phi = 0°$ is attributed to interaction between the target and the air–ground interface since it is not seen for the PO results [see Fig. 7(a)]. The bandwidth is not sufficiently large to distinguish the specular return from diffraction from the two ends of the UXO, these two scattering mechanisms arriving nearly simultaneously. Considering the other two incident angles ($\phi = 45°$ and $\phi = 90°$), the scattered waveforms are characterized at early times by two consecutive scattered waveforms with, as expected, the duration between these waveforms larger for the case of $\phi = 90°$ than for $\phi = 45°$.

Note that the $\phi = 0°$ case is relevant for Fig. 6(a), for which the linear aperture yields scattering primarily from the target broadside. The fact that the $\phi = 0°$ signature is characterized principally by a single strong response explains the fact that Fig. 6(a) is also represented by a single strong return (in the image domain). Similarly, the two diffracted components for the $\phi = 45°$ case in Fig. 8 give rise to the two diffracted components characteristic of the image in Fig. 6(b), these apparently

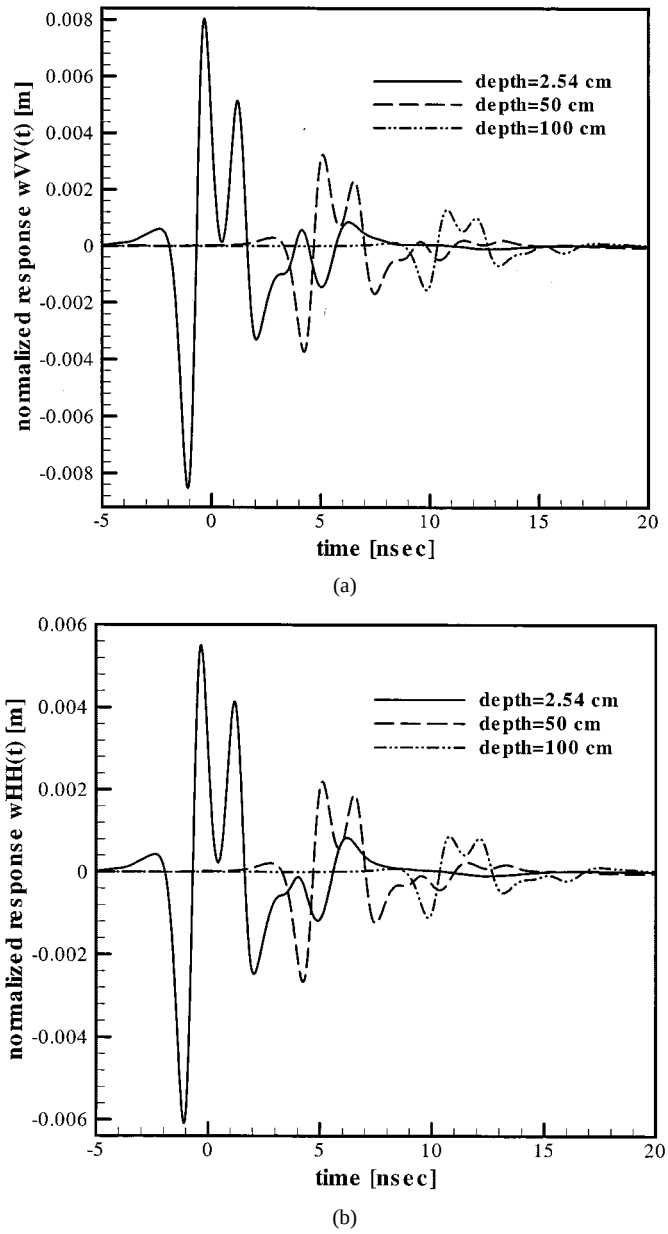


Fig. 9. Time-domain backscattered waveforms due to pulsed plane wave excitation 30° from grazing, with an incident pulse as in Fig. 5. Three burial depths are considered for the 155-mm shell (Fig. 1): 2.54 cm, 50 cm, and 100 cm, with the target axis parallel to the air–soil interface (5% Yuma soil [15]). The incident plane wave propagates 45° from the target axis. (a) VV polarization. (b) HH polarization.

merging in the measured image into a single diagonal response. Moreover, the energy in the $\phi = 0°$ signature is significantly stronger than for $\phi = 45°$, this explaining why Fig. 6(b) is more contaminated by clutter than does Fig. 6(a).

In our next set of results, we again consider the 155-mm shell buried in 5% Yuma soil, but now we consider variable target depth. In particular, for simplicity, again assume that the UXO axis is parallel to the soil interface (the model is applicable to arbitrary target orientation [17], [22], [23]). We consider three target depths, all measured from the top of the target to the interface: 1) 2.54 cm as in Figs. 2 and 3; 2) 50 cm; and 3) 100 cm. We again consider time-domain plane-wave incidence, characterized by the pulse in Fig. 5 and all results are computed via the MLFMA. In Fig. 9, we present results for 30° incidence relative to grazing, oriented 45° relative to the target axis [as in Fig. 7(b)]. As expected, there is a significant reduction in the scattered-field amplitude with target depth. This phenomenon indicates that SAR-based sensing of UXO is of most utility for targets near the air–ground interface.

IV. Conclusion

This paper has considered UWB scattering from UXO buried in lossy dispersive soil. Three scattering models were considered, and the numerical test considered here demonstrated that the MLFMA results are in close agreement with MoM for the buried UXO considered. Moreover, the agreement of the PO with the MoM/MLFMA solutions has been good, viewed both in the backscatter domain as well as in the SAR-image domain.

In addition to these theoretical investigations, we have performed SAR measurements via an experimental UWB sensor [15]. The agreement between theoretical and computed SAR images was good, despite uncertainty in the aspect-dependent radiated polarization and pulse shape. However, the clutter proved to be a significant issue, despite the fact that the UXO considered were shallow buried. The clutter was particularly problematic when the target orientation relative to the linear SAR aperture was such that no specular return was excited. In this case, the scattered fields are generated primarily by edge diffraction, the associated small scattered fields approaching the clutter level. We also demonstrated theoretically that this issue is exacerbated as the target depth increases. These theoretical and measured results indicate that SAR-based detection is most viable for detection of former bombing *ranges*, rather than for each individual (possibly deeply buried) UXO. This is motivated by the fact that most former ranges are littered by much surface and shallow-buried UXO. Once a former range is detected, alternative technology can be employed to attempt detection of each individual UXO [31].

The results presented here also indicate directions for future work. A principal concern in this context is gaining a better understanding of the mitigating clutter sources. For the UXO, the target shapes are generally well known (although the target is often deformed, due to impact) and, therefore, can be modeled. The clutter, on the other hand, is poorly understood. For example, the measured results presented here considered flat and weakly foliated terrain, but the SAR image still had clutter returns on the order of the *shallow* UXO, for the nonspecular case. The clutter source is conjectured to be subsurface soil inhomogeneties (e.g., spatially varying water content). Future work requires a combination of soil science and EM modeling to examine such issues in greater detail.

References

[1] R. W. P. King and C. W. Harrison, "The transmission of electromagnetic waves and pulses into the earth," *J. Appl. Phys.*, vol. 39, pp. 4444–4452, Aug. 1968.

[2] J. A. Fuller and J. R. Wait, "Electromagnetic pulse transmission in homogeneous dispersive rock," *IEEE Trans. Antennas Propagat.*, vol. AP-20, pp. 530–533, July 1972.

[3] D. L. Moffat and R. J. Puskar, "A subsurface electromagnetic pulse radar," *Geophys.*, vol. 41, pp. 506–518, June 1976.

[4] L. Peters and J. D. Young, "Applications of subsurface transient radars," in *Time-Domain Measurements in Electromagnetics*, E. K. Miller, Ed. New York: Van Nostrand Reinhold, 1986.
[5] G. S. Smith and W. R. Scott, "A scale model for studying ground penetrating radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 27, pp. 358–363, July 1989.
[6] C. Liu and C. Shen, "Numerical simulation of subsurface radar for detecting buried pipes," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 795–798, Sept. 1991.
[7] N. Osumi and K. Ueno, "Microwave holographic imaging of underground objects," *IEEE Trans. Antennas Propagat.*, vol. AP-33, pp. 152–159, Feb. 1985.
[8] L. Peters, J. J. Daniels, and J. D. Young, "Ground penetrating radar as an environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.
[9] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of ground penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–48, Jan. 1996.
[10] S. F. Mahmoud and J. R. Wait, "Scattering from a buried insulated loop excited by an incident electromagnetic plane wave," *J. Electromagn. Waves Applicat.*, vol. 12, pp. 1297–1311, Oct. 1998.
[11] B. A. Baertlein, J. R. Wait, and D. G. Dudley, "Scattering by a conducting strip over a lossy half space," *Radio Sci.*, vol. 24, pp. 485–497, July/Aug. 1989.
[12] S. Vitebskiy, L. Carin, M. Ressler, and F. Le, "Ultrawide-band, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.
[13] L. Carin, R. Kapoor, and C. Baum, "Polarimetric SAR imaging of buried land mines," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1985–1988, Nov. 1998.
[14] L. Nguyen, J. Sichina, K. Kappra, D. Wong, and R. Kapoor, "Minefield detection algorithm utilizing data from an ultra wideband wide-area surveillance radar," in *Proc. 1998 SPIE Conf.*, Orlando, FL, Apr. 1998, pp. 627–643.
[15] L. Carin, N. Geng, M. McClure, J. Sichina, and L. Nguyen, "Ultra-wide-band synthetic-aperture radar for mine-field detection," *IEEE Antennas Propagat. Mag.*, vol. 41, pp. 18–33, Feb. 1999.
[16] J. K. Jao, "Performance analysis for synthetic aperture radar detection of underground targets," MIT Lincoln Laboratory, Lexington, MA, Project Rep. GPR-4, Oct. 1995.
[17] N. Geng, A. Sullivan, and L. Carin, "Multi-level fast-multipole algorithm for scattering from conducting targets above or embedded in a lossy half space," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1567–1579, July 2000.
[18] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*. New York: McGraw-Hill, 1961.
[19] M. A. Ressler and J. W. McCorkle, "Evolution of the Army Research Laboratory ultra-wideband test bed," in *Ultra-Wideband Short-Pulse Electromagnetics 2*, L. Carin and L. B. Felsen, Eds. New York: Plenum, 1995, pp. 109–123.
[20] N. Geng, M. Ressler, and L. Carin, "Wideband VHF scattering from a trihedral reflector situated above a lossy dispersive halfspace," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2609–2617, Sept. 1999.
[21] N. Geng and L. Carin, "Wideband electromagnetic scattering form a dielectric BOR buried in a layered, dispersive medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.
[22] N. Geng, A. Sullivan, and L. Carin, "Fast multipole method for scattering from 3-D PEC targets situated in a half-space environment," *Microwave Opt. Tech. Lett.*, vol. 21, pp. 399–405, June 20, 1999.
[23] ——, "Fast multipole method analysis of half-space scattering problems," IEEE Trans. Antennas Propagat., to be published.
[24] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas Propagat. Mag.*, vol. 35, pp. 7–12, June 1993.
[25] J. M. Song and W. C. Chew, "Fast multipole method solution using parametric geometry," *Microwave Opt. Tech. Lett.*, vol. 7, pp. 760–765, Nov. 1994.
[26] J. M. Song, C. C. Lu, and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 1488–1493, Oct. 1997.
[27] A. F. Peterson, S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*. Piscataway, NJ: IEEE Press, 1998.
[28] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering from surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 409–418, May 1982.
[29] J. J. Yang, Y. L. Chow, and D. G. Fang, "Discrete complex images of a three-dimensional dipole above and within a lossy ground," *Proc. Inst. Elect. Eng.*, pt. H, vol. 138, no. 4, pp. 319–326, Aug. 1991.
[30] T. K. Sarkar and E. Arvas, "On a class of finite step interative methods (conjugate directions) for the solution of an operator equation arising in electromagnetics," *IEEE Trans. Antennas Propagat.*, vol. AP-33, pp. 1058–1066, Oct. 1985.
[31] A. C. Dubey and R. L. Barnard, Eds., *Detection and Remediation Technologies for Mines and Minelike Targets*. Orlando, FL: Proc. SPIE, 1997, vol. 3079.

Anders Sullivan (M'93) [redacted] in Staten Island, NY. He received the B.S. and M.S. degrees in aerospace engineering from the Georgia Institute of Technology, Atlanta, GA, in 1985 and 1987, respectively, and the Ph.D. degree in electromagnetics from Polytechnic University, Brooklyn, NY, in 1997.

From 1988 through 1998, he was with the Air Force Research Laboratory, Elgin Air Force Base, FL. From 1998 to 1999, he was with the Electrical Engineering Department, Duke University, Durham, NC, as a Research Associate. Since September 1999, he has been with the Army Research Laboratory, Adelphi, MD. His current research interests include modeling complex targets and short-pulse scattering.

Dr. Sullivan is a member of the Tau Beta Pi and Sigma Gamma Tau honor societies.

Raju Damarla (M'99) received the B.Sc.(mathematics), B.Tech. (electronics and electrical communication engineering), and M.Tech. (controls and computation engineering) degrees from the Indian Institute of Technology, Kharagpur, India, and the Ph.D. degree in engineering from Boston University, Boston, MA.

He was an Assistant Professor in the Electrical Engineering Department, University of Kentucky, Lexington, until 1994. He is currently working as an Electronics Engineer at the U.S. Army Research Laboratory, Adelphi, MD. His interests include signal processing, automatic target detection and recognition algorithm development, and neural networks.

Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl. Ing. and Dr. Ing. degrees in electrical engineering from the University of Karlsruhe, Germany, in 1991 and 1996, respectively.

From 1991 to 1996, he was with the Institute for Microwaves and Electronics at the University of Karlsruhe working on full-wave propagation modeling for radio communication systems. In January 1997, he joined the Department of Electrical and Computer Engineering, Duke University, Durham, NC, as a Visiting Postdoctor for 18 months. In July 1998, he returned to the University of Karlsruhe and, in 1999, he joined the Siemens Corporation, Munich, Germany. His current research interests include computational methods in electromagnetics and wave propagation modeling.

Dr. Geng received the Mannesmann Innovation Award in 1997.

Yanting Dong (S'00) received the B.S. and M.E. degrees in electrical engineering from Tianjin University, China, in 1993 and 1996, respectively. She is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, Duke University, Durham, NC.

From 1996 to 1998, she was with the Department of Automation, Tianjin University, China, working on process instrumentation design. Since 1999, she has been working as a Research Assistant at Duke University. Her main research interests are in signal processing, detection and estimation theory, applied electromagnetics, and instrumentation design.

Lawrence Carin (SM'96) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, where he is an Associate Professor. His current research interests include short-pulse scattering, propagation, and signal processing. He is the Principal Investigator on an Army Research Office funded Multidisciplinary University Research Initiative (MURI) on demining.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.

A Comparison of the Performance of Statistical and Fuzzy Algorithms for Unexploded Ordnance Detection

Leslie M. Collins, *Member, IEEE*, Yan Zhang, Jing Li, *Student Member, IEEE*, Hua Wang, Lawrence Carin, *Fellow, IEEE*, Sean J. Hart, Susan L. Rose-Pehrsson, Herbert H. Nelson, and Jim R. McDonald

***Abstract*—In most field environments, unexploded ordnance (UXO) items are found among extensive surface and subsurface clutter and shrapnel from ordnance. Traditional algorithms for UXO remediation experience severe difficulty distinguishing buried targets from anthropic clutter. Furthermore, naturally occurring magnetic geologic noise often adds to the complexity of the discrimination task. These problems render site remediation a very slow, labor-intensive, and inefficient process. While sensors have improved significantly over the past several years in their ability to *detect* conducting and/or permeable targets, reduction of the false alarm rate has proven to be a significantly more challenging problem. Our work has focused on the development of signal processing algorithms that incorporate the underlying physics characteristic of the sensor and of the anticipated UXO target, in order to address the false alarm issue. In this paper, we describe several algorithms for discriminating targets from clutter that have been applied to data obtained with the multisensor towed array detection system (*MTADS*). This sensor suite has been developed by the U.S. Naval Research Laboratory (NRL), and includes both electromagnetic induction (EMI) and magnetometer sensors. We describe four signal processing techniques that incorporate features derived from simple physics-based sensor models: a generalized likelihood ratio technique, a maximum likelihood estimation-based clustering algorithm, a probabilistic neural network, and a subtractive fuzzy clustering technique. These algorithms have been applied to the data measured by *MTADS* in a magnetically clean test pit and at a field demonstration. We show that overall the subtractive fuzzy technique performs better than the alternative techniques when the training and testing data sets are separate. The results also allow us to quantify the utility of fusing the magnetometer and the EMI data, and we show that performance is improved when both EMI and magnetometer features are utilized. The results indicate that the application of advanced signal processing algorithms could provide up to a factor of two reduction in false alarm probability for the UXO detection problem.**

***Index Terms*—Electromagnetic induction, magnetometer, sensor fusion, statistical detection, unexploded ordnance detection.**

Manuscript received June 23, 2000; revised November 11, 2000 and November 27, 2000. This work was supported by the Strategic Environmental Research and Development Program (SERDP) by Grant DACA72-99-C-0012-CU-1123 and CU-1092.

L. M. Collins, Y. Zhang, J. Li, H. Wang, and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA (e-mail: lcollins@ee.duke.edu).

S. J. Hart, S. L. Rose-Pehrsson, H. H. Nelson, and J. R. McDonald are with the Naval Research Laboratory, Chemistry Division, Code 6110, Washington, DC 20375 USA (e-mail: j.mcdonald@nrl.navy.mil).

Publisher Item Identifier S 1063-6706(01)01505-3.

I. Introduction

THERE are many areas in the United States and many more outside the continental United States that are contaminated or potentially contaminated with unexploded ordnance (UXO). Several million acres of Department of Defense (DOD) ranges have been returned to civilian control without adequate UXO clearance, and several hundred thousand additional acres on ranges are returned to civilian control each year. In the United States alone there are 1900 formerly used defense sites (FUDS) and 130 base realignment and closure (BRAC) installations that need to be cleared of UXO. Using current technologies, the cost of identifying and disposing of UXO in the United States is estimated to range up to $500 billion. Site specific clearance costs vary from $400/acre for surface UXO to $1.4 million/acre for subsurface UXO. Thus, there is a clear need to effectively and cost-efficiently remediate UXO contaminated lands, rendering them safe for their current or intended uses. Development of new UXO detection technologies with improved data analysis has been identified as a high priority requirement for over a decade.

Several sensor modalities are currently being explored for the detection and identification of surface and buried UXO. These include electromagnetic induction (EMI), magnetometers, radar, and seismic sensors. These sensors generally experience little difficulty detecting the UXO, thus detection does not create the bottleneck that results in the high cost of remediating sites. The primary contributor to the costs and time associated with remediating a UXO-contaminated site is the high false-alarm rate associated with each of the sensors when operated individually. This high cost results since the false alarms require digging numerous nonordnance targets for each intact ordnance item recovered. Recently, the joint use of towed arrays of magnetometer and time-domain electromagnetic sensors in conjunction with GPS navigation has generated high quality, digital georeferenced mapped data. With the availability of such high quality data, advanced physics-based computer data analysis tools have been developed and have demonstrated the ability to detect essentially all DOD ordnance up to its maximum self-burial depths on a wide variety of military ranges [1]–[5]. The false-alarm rate has, however, continued to be an impediment to efficient range clearance. Other research efforts have indicated that the application of statistical, fuzzy, and neural approaches to the detection of landmines and other ordnance-like objects has shown considerable promise [6]–[35]. Performance improvements have been shown in both

1063–6706/01$10.00 © 2001 IEEE

a theoretical context as well as on field-collected data. The goal of this effort is to apply an optimized set of algorithms to data collected in the field with a modern multisensor system in order to investigate to what extent the false-alarm rate could be mitigated.

The Jefferson Proving Ground (JPG) site was a 55 000 acre live fire testing range used from World War II until 1994 [1], [36]. Various estimates exist as to the amount of unexploded ordnance on the site, but a minimum of more than 1 million rounds is thought to still exist. Given the cleanup needs of this and other sites, the development and demonstration of emerging technologies for the detection, identification, and cleanup of sites contaminated by UXO has been pursued over the last several years. Several test sites have been prepared at JPG to evaluate sensor performance. These sites have been seeded with ordnance placed at particular positions, depths, and orientations. There have been four demonstrations of technology at JPGs controlled test sites to establish a baseline for performance.

Based on the success of the detection technologies demonstrated in the first three JPG tests, the goal of the fourth JPG demonstration (JPG IV) was not to detect ordnance, but to determine individual contractor's ability to discriminate ordnance items from nonordnance items [36]. In this test, demonstrators knew the location of each test item, but not its identity, depth, or orientation. After the completion of the demonstration, ground truth for the test site along with raw data for some of the demonstrators was released to the public, allowing the research community to develop and test discrimination algorithms. In this paper, we consider the application of a variety of statistical and fuzzy signal processing techniques to data collected during the JPG IV demonstration with a fielded system, the multisensor towed array detection system (*MTADS*) [1]–[5], [30]–[35]. Specifically, we have investigated the performance of four algorithms whose performance has been optimized for the task of discriminating UXO objects from the clutter objects emplaced during the JPG IV test. These algorithms include a generalized likelihood ratio approach, a maximum likelihood estimation-based clustering algorithm, a probabilistic neural network, and a subtractive fuzzy clustering technique. The *MTADS* sensor suite has been developed by the Naval Research Laboratory (NRL), and includes both EMI and magnetometer sensors. The implementation of the UXO detection algorithms is described as well as their robustness and performance. We also consider the impact of fusing the EMI and magnetometer data on performance. Furthermore, the impact of the specific training set used on overall performance is investigated.

II. The *MTADS* System and the JPG IV Data Set

The *MTADS* technology has been described in detail previously [1]–[5], however the most salient features are summarized here. The system hardware includes a low-magnetic-signature vehicle that is used to tow linear arrays of magnetic and EMI sensors to conduct surveys of large areas. The *MTADS* tow vehicle is a custom-built off-road vehicle specifically modified to have an extremely low magnetic self-signature. Most ferrous components have been removed from the body, drive train, and engine, and have been replaced by nonferrous alloys.

The *MTADS* magnetic sensors are Cs vapor full-field magnetometers (a variant of the Geometrics 822 sensor, designated as the Model 822ROV). An array of eight sensors is deployed as a magnetometer array. The time-dependence of Earth's background field is measured by a ninth sensor, deployed at a static site during survey operations. The EMI sensors are deployed as an array of three pulsed induction sensors and are a variant of the Geonics EM-61 instrument. These sensors, configured as an overlapping horizontal array, transmit a tailored electromagnetic pulse into the earth. The received EMI signal is time-sampled by six detection coils that are colocated with and above the three transmission coils. The Geonics EM-61 sensors have been extensively modified. These modifications include changing the time position and time width of the sampling window monitoring the return signal. The power of the transmitted pulse has been increased, as has the pulse repetition rate. The amplifier gain of the detectors has been increased and the time constant applied to the signal has been significantly reduced. The overall detection signal strength has been increased by a factor of 3–6 depending upon the composition, size, and depth of the target [37].

The sensor positions on the surface of Earth (latitude, longitude, and height) are determined using GPS navigation. This technology provides a real-time position update (at 5 Hz) with an accuracy of about 5 cm. GPS satellite clock-derived time is used to time-stamp both position and sensor data information for subsequent correlation. All navigation and sensor data are provided through electronic interfaces to a data acquisition computer in the tow vehicle.

In the JPG IV demonstration, a group of inert ordnance items and a second group of nonordnance items were buried in known positions at unknown depths and orientations. Demonstrators were provided with the positions of the flagged targets and asked to determine the identity of each object by scoring it in terms of its probability of being ordnance. Demonstrators were given the identities, though not the locations, of the ten types of inert ordnance (20, 57, 76, 90, 105, 152, and 155 mm projectiles, and 60-, 81-, and 108-mm mortars) and some were provided with examples of the nonordnance items included. The nonordnance objects (scrap simulants) consisted of steel items welded primarily from half-inch sheet, rods, I-beams, pipes, channels, and assorted scrap items. These items included about two dozen different plates, cubes, rods, cylinders, and esoteric welded objects. The dimensions, shapes, sizes, and weights of many of the nonordnance items closely approximated the ordnance items emplaced for the test. No information was provided regarding the relative proportions of the ordnance and nonordnance items, or the number of items of each type that were emplaced.

NRL conducted the *MTADS* surveys during the JPG IV demonstration using the magnetometer array in a single continuous mapping survey and using the EMI array to conduct two mapping surveys in orthogonal directions. NRL was not provided access to the UXO and non-UXO items in advance. Their classification performance on the demonstration was based on prior experience with ordnance signatures as opposed to training *in situ*. After the demonstration was completed, ground truth for the site was released. The number of targets and clutter items listed in the JPG IV Ground Truth is 50 and 110, respectively. Five of each of the ten types of ordnance items were emplaced.

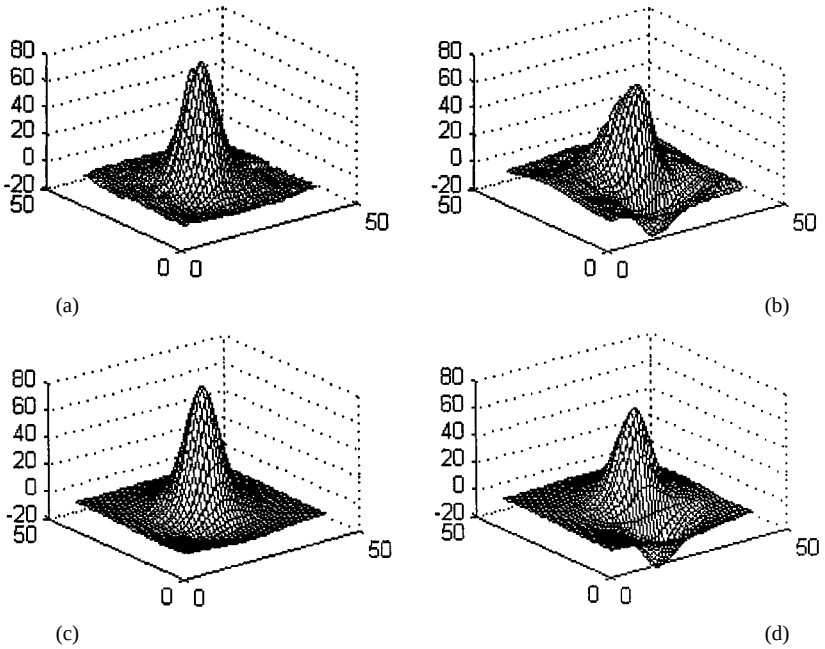


Fig. 1. Measured (a), (b) and modeled (c), (d) magnetometer responses for a 152 mm projectile buried at a depth of 0.57 m (a), (c) and a 4.1 kg nonordnance, or scrap, item (b), (d) buried at a depth of 0.48 m. The vertical axis plots sensor output and the two horizontal axes indicate distance in the north and east directions.

III. Phenomenological Model and Feature Extraction

As the name implies, the *MTADS* was designed as a detection system. In the early testing of the system [1], it was discovered that several features of the *MTADS* survey data allowed the development of some classification ability. Because of the low noise and precise sensor positioning associated with *MTADS* magnetometer data, the depths and sizes obtained from fits of magnetic anomalies to a dipole model are precise enough to allow these parameters to be used as features for discrimination. A survey was conducted at the Badlands Bombing Range [2], a site where the targets consist of practice bombs, rockets, and fins and clutter associated with the bombs and rockets. The results of this survey indicated that an impressive degree of classification was possible based on these parameters from the magnetometer model and the judgement of a skilled analyst concerning the anomaly shape and model goodness of fit.

In the same data set, it was also observed that there was significant target shape-dependence of the observed EMI anomalies [33]. Methods have been developed to take advantage of this shape-dependence of the EMI response, and these methods have recently been applied at a live test site [34]. Although these methods showed some promise, there was still progress needed to reach an acceptable classification performance. These features, among others, have been investigated in the context of this paper.

An example of the magnetometer data measured by the *MTADS* system at the JPG IV demonstration for an ordnance and a nonordnance item is shown in Fig. 1 in panels A and B. Sensor output (z axis) is plotted as a function of distance in the North (x axis) and East (y axis) directions. The data in panel A was measured from a 152-mm projectile buried at a depth of 0.57 m and the data in panel B was measured from a 4.1-kg nonordnance item buried at a depth of 0.48 m. Fig. 2 plots the output of the EMI sensor for these same two objects. Again, the data in panel A was measured from the 152-mm projectile and the data in panel B was measured from the nonordnance item.

Features are extracted from magnetometer data by fitting the data to a dipole magnetic model [33]. This approach to model-based feature extraction is consistent across all of the algorithms considered. The features that can be extracted from a magnetic dipole model include

1) Object position.
2) Object depth.
3) Object size.
4) Magnetic moment.
5) Dipole azimuth.
6) Dipole inclination.
7) Fit quality.

Each algorithm described in the next section employed a particular subset of these parameters. Each feature set was chosen using statistical performance criteria that are different for each algorithm. In this analysis, a particular feature set was not imposed on all of the algorithms as our goal was to optimize the performance capability for each algorithm and compare their relative performance. Panels C and D of Fig. 1 plot the magnetometer model fits obtained for the 152-mm projectile and the 4.1-kg nonordnance item, respectively.

A *baseline* EMI model has been developed previously for use with the *MTADS* system [33]. This modeler utilizes a dipole model for the transmitter coil and a point approximation for the receive coil. The target is represented as a magnetic polarizability dyadic $\{M_t, M_z\}$ [38], [39], and the secondary field is approximated as target dipole moment. Based on these assumptions the modeler reports the following features:

1) Object position.
2) Object depth.
3) Object size.
4) Object orientations.

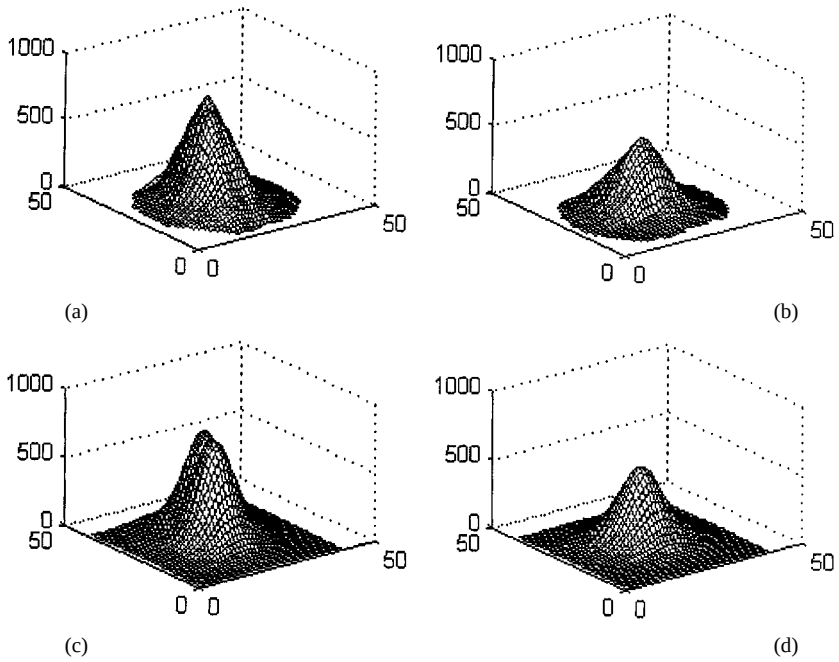


Fig. 2. Measured (a), (b) and modeled (c), (d) EMI responses for a 152 mm projectile buried at a depth of 0.57 m (a), (c) and a 4.1 kg nonordnance, or scrap, item (b), (d) buried at a depth of 0.48 m. The vertical axis plots sensor output and the two horizontal axes indicate distance in the north and east directions.

5) Magnetostatic model coefficients $\{M_x, M_t\}$ (polarizabilities).
6) Fit quality.

Panels C and D of Fig. 2 plot the EMI model fits obtained for the 152-mm projectile and the 4.1-kg nonordnance item, respectively.

The approximations associated with a dipole model for the excitation coil, the point model for the receive coil, and the colocated dipole model for the target may be tenuous, particularly for shallow objects. Two modifications were made to this model in order to improve feature estimation. First, the excitation field has been calculated accurately for the known transmit coil shape via the Biot–Savart law [40]. Second, the model has been modified so that different effective centers for the transverse and axial target dipoles $\{M_x, M_t\}$ are included, and the center locations are determined during the fitting procedure. A modification to the model to incorporate the shape of the receive coil was considered, but the slight changes in the parameter estimates were not reflected by a change in algorithm performance. This modified model was used for parameter estimation for all of the algorithms discussed below with the exception of the probabilistic neural network. Utilization of the features obtained with the modified EMI model did not substantially affect performance, so the trends in the results discussed in Section V are valid regardless of the model that was utilized. The modified EMI model did provide substantially better estimates of the depth of the object, especially for the shallow-buried objects. However, as object depth is not a useful discriminator of targets and clutter for this particular data set, performance was not affected by utilization of the features from the modified model.

The EMI model fitting is implemented by the Levenberg–Marquardt method [41]. The model parameters are fit to the data multiple times with randomly selected initial conditions, and the fit quality is monitored. In general, multiple local minima are obtained. Only the fitted model estimates with $M_z/M_t > 1$ are retained, as, in general, UXO objects have an aspect ratio that is greater than unity. Often, the fitted depths of the multiple local minima are different. To select between the various local minima, the depth information determined from the magnetometer data is utilized. When multiple local minima with $M_z/M_t > 1$ are obtained in the training phase, a minimum-within-class-distance criteria is used, i.e., the fit that minimizes the distance of the fitted parameters between different targets within the same class is selected. In the testing phase, each local minimum is processed and the final output is the maximum of the processed values. The features generated by this *modified* EMI model are the same as those listed for the baseline model, with the addition of the location of the two dipole centers. These additional features were not considered in this effort, however, but will be considered in future work.

IV. Signal Processing Algorithms

Several signal processing algorithms are considered in this analysis. First, we consider a statistical signal processing approach based on the generalized likelihood ratio test [42]. Traditionally, such statistical processors are less robust when limited training data are available, or when the statistics of the parameters that affect the signal cannot be adequately characterized, as was hypothesized to be the case for the JPG IV demonstration. Therefore, suboptimal but potentially more robust approaches were considered. These include a maximum likelihood estimation-based clustering algorithm [41], [42] whose clusters are initially determined using the ISODATA algorithm [43], [44], a probabilistic neural network [47], and a subtractive fuzzy clustering technique [48]. The various approaches are described below along with the method by which each of the algorithms was trained.

As aforementioned, the feature set that provided the optimum performance on a training data set was chosen independently for each algorithm. Each algorithm initially used the same feature set, but performed a down-select to obtain the subset that provided the best performance. Thus, the final feature set was not necessarily consistent across algorithms. This approach was taken since the eventual goal of this effort is to develop algorithms that provide the best performance in a field environment, and because this approach mimics the mechanisms by which a variety of algorithms would be developed and tested for this system.

A. Generalized Likelihood Ratio Test Algorithm

The two hypotheses to be tested are H_1: UXO, or target; and H_0: clutter, or nonordnance. The observation $\mathbf{x}$ is an N-dimensional feature vector and contains a subset of the features described above. Assuming each UXO item is equally likely, the likelihood ratio (LR) or Bayesian processor is the ratio between the probability density functions (pdfs) describing the data under the two hypotheses [42]. In the JPG IV data set there were ten targets which fell into two classes: mortars and projectiles. An analysis of the data indicated that the pdfs for the features for the objects within a class were fairly similar, however the pdfs across classes differed, thus we assumed that uncertainty regarding target class would affect the underlying pdf. With this assumption, the LR is given by (1), shown at the bottom of the page, where $p(\,)$ denotes a probability density function. Since the pdfs are not known *a priori*, each of the pdfs must be estimated from the training data. The use of an estimate of the parameters in the likelihood ratio formulation results in a generalized likelihood ratio (GLR). To execute the generalized likelihood ratio test (GLRT), the GLR is compared to a threshold, and if the GLR is greater than a threshold, an ordnance object is declared, otherwise a nonordnance object is declared. This approach is not guaranteed to provide the optimal decision performance, but does provide a statistical means of treating the data. In some cases the performance of a GLR can approach that of the LR, particularly when the signal to noise ratio is high enough to provide accurate estimates of the parameters of interest [42]. The GLRT can then be performed for each feature vector in the test set, and then the GLRT can be compared to a sequence of thresholds to evaluate the True Positive and False Positive rates.

Based on an analysis of the feature data, it was assumed that $p(\mathbf{x}\,|\,H_1)$ and $p(\mathbf{x}\,|\,H_0)$ follow multivariate Gaussian distributions. This assumption was not necessarily valid in all cases, but allowed the correlation between features to be incorporated into the processing. When more accurate pdfs of the features were utilized, and independence was assumed in order to allow the formulation of the joint pdf as the product of the individual feature pdfs, performance was degraded. Thus, in the remainder of this paper it is assumed that $p(\mathbf{x}\,|\,H_1) = N_N(\mathbf{u_1}, \Sigma_1)$ and $p(\mathbf{x}\,|\,H_0) = N_N(\mathbf{u_0}, \Sigma_0)$, where $\mathbf{u}_1$, $\mathbf{u}_0$, Σ_1, Σ_0 are the mean vectors and covariance matrices of $\mathbf{x}$ under H_1 and H_0 and $N_N(\,)$ represents an N-dimensional normal, or Gaussian, pdf. There were ten types of UXO buried in the JPG IV test, and we assumed each type was equally likely. Three of the UXO items belong to the mortar class and the remaining items belong to the projectile class so $p(\text{mortar})$ and $p(\text{projectile})$ were defined appropriately. The estimates, $\hat{\mathbf{u}}_\mathbf{1}$, $\hat{\mathbf{u}}_\mathbf{0}$, $\hat{\Sigma}_1$, $\hat{\Sigma}_0$, are obtained from the training data. Without considering the different target classes (mortar and projectile), the GLRT becomes:

$$\lambda(x) = \frac{p(\mathbf{x}\,|\,H_1)}{p(\mathbf{x}\,|\,H_0)} = \frac{\dfrac{1}{(2\pi)^{N/2}|\Sigma_1|^{1/2}}\exp\left(-\dfrac{1}{2}(\mathbf{x}-\mathbf{u}_1)^T\Sigma_1^{-1}(\mathbf{x}-\mathbf{u}_1)\right)}{\dfrac{1}{(2\pi)^{N/2}|\Sigma_0|^{1/2}}\exp\left(-\dfrac{1}{2}(\mathbf{x}-\mathbf{u}_0)^T\Sigma_0^{-1}(\mathbf{x}-\mathbf{u}_0)\right)}. \tag{2}$$

Taking the natural logarithm and incorporating the constant term into the threshold, (2) becomes

$$\ln(\lambda(\mathbf{x})) = (\mathbf{x}-\mathbf{u}_0)^T\Sigma_0^{-1}(\mathbf{x}-\mathbf{u}_0) - (\mathbf{x}-\mathbf{u}_1)^T\Sigma_1^{-1}(\mathbf{x}-\mathbf{u}_1). \tag{3}$$

It is a simple extension of the GLRT to incorporate the different target classes, although the succinct form of the previous equation is not obtained in this case.

B. Maximum Likelihood Estimation-Based Clustering Algorithm

A clustering algorithm was investigated since it is well documented that this type of algorithm performs well for cases involving limited training data in which the goal is to decide to which class a particular object belongs. The difference between the standard approach and the approach we have pursued lies in both how the clusters are determined and how the iteration that is used to assign objects to clusters is performed. The clusters for the maximum likelihood estimation-based clustering algorithm are initialized using the ISODATA algorithm [43], [44]. This algorithm is a modified K-means approach that was utilized to initially divide the training data into clusters associated with UXO and nonordnance. A traditional K-means algorithm assumes that the data can be modeled as K clusters. The modification considered here assumes that K is unknown, and is then estimated from the data. The training data is a set of training feature vectors as described above. In the algorithm, K' cluster centroids are initialized, and then each training fea-

$$\lambda(x) = \frac{p(\mathbf{x}\,|\,H_1)}{p(\mathbf{x}\,|\,H_0)} = \frac{\sum_{\text{target class}} p(\mathbf{x}\,|\,H_1, \text{target class})p(\text{target class})}{p(\mathbf{x}\,|\,H_0)} = \frac{p(\mathbf{x}\,|\,H_1,\,\text{mortar})p(\text{mortar}) + p(\mathbf{x}\,|\,H_1, \text{projectile})p(\text{projectile})}{p(\mathbf{x}\,|\,H_0)} \tag{1}$$

ture vector is associated with the closest of the K' centroids. The K' centroids are then reestimated, and clusters are merged if their centroids are within a predefined distance measure, or split if enough members are present and the cluster has a high average distance between its mean and its members. The process of re-associating the feature vectors with the various centroids and re-estimating the centroids is continued until all feature vectors do not change their association from iteration to iteration. A detailed discussion of this algorithm can be found in [43]–[46].

The ISODATA algorithm assumes that the clusters associated with the data do not overlap. Furthermore, it assumes that the variance of the data within each of the clusters is the same. These two assumptions are violated for the problem of discriminating UXO from nonordnance. To mitigate this problem, a maximum likelihood classification rule is implemented. The conditional pdfs $p(i/X_j)$, which is the probability that the feature vector X_j belongs to cluster i, and new estimates of the cluster centers V_i, are determined iteratively using the training data. Briefly, we define $p(i/X_j)$ as above, P_i to be the *a priori* probability of selecting the ith cluster, Σ_i as the covariance of the ith cluster, V_i to be the center of the ith cluster, M to be the number of training vectors, and K to be the number of clusters. Prior to estimating the conditional densities, the parameters are initialized as

$$p(i/X_j) = 1, \quad i = \arg\min_{i=1\ldots K} d(X_j, V_i)$$
$$P_i = \frac{\sum_{j=1}^{M} p(i/X_j)}{M}$$
$$\Sigma_i = \frac{\sum_{j=1}^{M} p(i/X_j)(X_j - V_i)(X_j - V_i)^T}{\sum_{j=1}^{M} p(i/X_j)} \tag{4}$$

where d is an exponential distance metric. The updated versions of the conditional pdfs, $\hat{p}(i/X_j)$ are calculated using the iterative procedure outlined below:

Step 1)

$$\hat{p}(i/X_j) = p(i/X_j)$$
$$V_i = \frac{\sum_{i=1}^{M} p(i/X_j) X_j}{\sum_{i=1}^{M} p(i/X_j)}$$

Step 2)

$$p(i/X_j) = \frac{d_e(V_i, X_j)}{\sum_{k=1}^{K} d_e(V_k, X_j)} \tag{5}$$
$$d_e(V_i, X_j) = \frac{P_i}{\det(\Sigma_i)^{0.5}} \exp\left[-(X_j - V_i)^T \Sigma_i^{-1} (X_j - V_i)\right]$$
$$P_i = \frac{\sum_{j=1}^{M} p(i/X_j)}{M} \tag{6}$$

Step 3)

$$\Sigma_i = \frac{\sum_{j=1}^{M} p(i/X_j)(X_j - V_i)(X_j - V_i)^T}{\sum_{j=1}^{M} p(i/X_j)} \tag{7}$$

Step 4)

$$\max_{i,j} |p(i/X_j) - \hat{p}(i/X_j)| < \varepsilon,$$
go to step 1 if false, else terminate. (8)

The exponential distance metric may render this algorithm unstable when the initial conditions are not properly chosen, thus the need to implement the ISODATA algorithm as the initial cluster estimation procedure.

To classify a particular test vector y, the probability that y belongs to the ith cluster is calculated as

$$p(i/y) = \frac{d_e(V_i, y)}{\sum_{k=1}^{K} d_e(V_k, y)}$$
$$d_e(V_i, y) = \frac{P_i}{\det(\Sigma_i)^{0.5}} \exp\left[-(y - V_i)^T \Sigma_i^{-1} (y - V_i)\right] \tag{9}$$

and the decision rule utilized is

$$\max_i p(i/y) > \gamma \tag{10}$$

where the threshold γ determines the relationship between the detection and false-alarm probabilities.

C. Probabilistic Neural Network (PNN)

The probabilistic neural network (PNN) is based upon Bayes' classification method [47], [49]–[51]. The basis of the classification method is given in the following equation, where P_i and P_j are the prior probabilities for class i and j, c_i and c_j are the costs of misclassification, and $p_i(x)$ and $p_j(x)$ are the true probability density functions:

$$P_i c_i p_i(x) > P_j c_j p_j(x). \tag{11}$$

The difficulty with this relationship is that the prior probabilities (the probability that a sample will come from a given population distribution) are generally unknown and must be estimated from training data. This is done using Parzen's method of pdf estimation [52]. Bayes' classification will be more likely to group a new sample $\mathbf{x}$ into class i if the prior probability or the cost of misclassification is high. This is especially important for classifications where the cost of misclassification is not equal among the classes. In our application, UXO detection, the potential cost of mis-classifying UXO is much more serious than the cost for nonordnance. In the PNN method, different costs can be set for each class, thus producing a better classification for those classes that demand higher performance. In the analysis performed here, however, these costs were not adjusted. Finally, if the probability density of a given class is large in the region of the new sample $\mathbf{x}$ then that class is favored. This allows for multimodal distributions to be dealt with appropriately when a

nearest neighbor-based classifier might fail. It is has proven convenient and practical to implement the Parzen pdf estimator in a neural network format [52].

The PNN is a nonlinear, nonparametric pattern recognition algorithm that operates by defining a pdf for each data class based on the training set data and the optimized kernel width parameter. For ordnance discrimination, the inputs are the physics-based modeler outputs or pattern vectors. The outputs of the PNN are the Bayesian a posteriori probabilities (i.e., measures of confidence in the classification) that the input pattern vector is a member of one of the possible output classes, for example, UXO or non-UXO.

The hidden layer of the PNN is the core of the algorithm. During the training phase, the pattern vectors in the training set are simply copied to the hidden layer of the PNN. Thus, there are the same number of feature sets at the hidden level as there are training vectors. Unlike other types of artificial neural networks, the basic PNN only has a single adjustable parameter. This parameter, termed sigma (σ), or kernel width, along with the members of the training set, define the pdf for each data class. In a PNN, each pdf is composed of Gaussian-shaped kernels of width σ located at each pattern vector. The pdf essentially determines the boundaries for classification. The kernel width is critical because it determines the amount of interpolation that occurs between adjacent pattern vectors. As the kernel width approaches zero, the PNN essentially reduces to a nearest neighbor classifier. A large kernel width has the advantage of producing a smooth pdf, which exhibits good interpolation properties for predicting new pattern vectors. Small kernel widths reduce the amount of overlap between adjacent data classes. The optimized kernel width is a compromise between an overly small or large σ.

Prediction of new targets using a PNN is more complicated than the training step. Each member of the training set of pattern vectors (i.e., the patterns stored in the hidden layer of the PNN and their respective classifications), and the optimized kernel width are used during each prediction. As new pattern vectors are presented to the PNN for classification, they are serially propagated through the hidden layer by computing the Euclidean distance d between the new pattern and each pattern stored in the hidden layer. The Euclidean distance scores are then processed through a nonlinear transfer function (the Gaussian kernel) given in the following equation:

$$\text{Hidden_Neuron_Output} = \exp\left(\frac{-d}{\sigma^2}\right). \tag{12}$$

Because each pattern in the hidden layer is used during each prediction, the execution speed of the PNN is considerably slower than some other algorithms. The mass data storage requirements can also be quite large since every pattern in the hidden layer is needed for prediction. Several researchers have developed modified PNN algorithms to overcome this limitation, but they were not deemed necessary for this application [53], [54].

The summation layer consists of one neuron for each output class and simply collects the outputs from all hidden neurons of each respective class. The products of the summation layer are forwarded to the output layer, where we compute the estimated probability of the new pattern being a member of each class. In the PNN, the sum of the output probabilities equals 100%. The pattern being predicted is then assigned to the class with the highest output probability. One of the main advantages associated with using a PNN is the ability to output a probability for each of its classifications. For critical applications, such as ordnance detection and remediation, such an indicator of confidence is extremely useful in assisting the decision making process and reducing the likelihood that individual ordnance items are missed by reducing the detection probability.

As discussed earlier, the calculation of the optimum kernel width σ is imperative for high classification rates to be achieved. For the work described herein, an iterative algorithm for σ optimization was employed [55]. The approach by which the training data was used to determine this parameter is described in Section IV-E.

D. Subtractive Fuzzy Clustering Algorithm

In contrast to the maximum likelihood classification algorithm described above, the subtractive fuzzy clustering algorithm allows a feature vector of interest to be associated with multiple clusters, i.e., it is not restricted to have membership in only one cluster. In addition, once a feature vector's level of membership in each of the possible clusters is determined, a set of fuzzy logic rules can be imposed to determine whether a feature vector should be associated with an ordnance or nonordnance class.

In the UXO discrimination problem, the input is the N-dimensional feature vector $\mathbf{x}$ and the system output is related to the likelihood that the data is a target, although it is not a probability in the mathematical sense. In the training data, the output z is set to 1 for an ordnance item or 0 for a nonordnance item. The training of the clusters is implemented via a modification of the mountain clustering method described by Yager and Filev [56]. One shortcoming of this procedure is that the computation time grows exponentially with the dimensionality of the feature space. The subtractive clustering method used here [48] is a modification to the mountain clustering method that is computationally more efficient. As implemented in this application, the N features are normalized in each dimension so that their coordinate ranges in each dimension are equal, and a standard subtractive clustering algorithm is applied. The clusters are determined by applying this approach to the training data.

The cluster validation problem, or the criteria for accepting and rejecting cluster centers is important for any clustering method, and the problem is essentially unsolved [57]. In the subtractive clustering algorithm, the rule that was adopted is to set an upper threshold $\bar{\varepsilon}$ and a lower threshold $\underline{\varepsilon}$. If $m_{c_k} > \bar{\varepsilon}$, the cluster is accepted definitely, if $m_{c_k} < \underline{\varepsilon}$, the cluster is rejected definitely, where m_{c_k} is the mountain function value of the cluster candidate. Otherwise, if the mountain function value falls between the two thresholds, the shortest distance to the existing cluster centers is calculated. This value is used to determine whether the candidate cluster center provides a good tradeoff between having a large mountain function value and being far away from existing cluster centers. Although the number of clusters is determined by the algorithm automatically, the user-specified parameters strongly affect the number of clusters. In our implementation, the default values from the original algorithm were used.

Instead of determining the absolute membership of a test vector in one of the clusters, a fuzzy approach is implemented where each cluster center defines a rule. Thus, the subtractive

clustering training algorithm establishes the number of fuzzy rules (number of clusters) as well as the rule premises (cluster centers). Given a test data point $\mathbf{x}$ in the N-dimensional feature space ($\mathbf{x}$ is a $1 \times N$ vector), the degree to which the ith rule is fired, i.e., its firing strength, is defined as

$$u_i = \exp\left(-\frac{\|\mathbf{x} - \mathbf{V}_i\|^2}{2\sigma^2}\right) \tag{13}$$

where $\mathbf{V}_i$ is the ith cluster center and σ is the application-specific constant that determines the height and smoothness of the mountain function. The output z is computed via

$$z = \frac{\sum_{i=1}^{K} u_i z_{c_i}}{\sum_{i=1}^{K} u_i} \tag{14}$$

where z_{c_i} is the consequence of the ith rule and K is the number of rules.

The rule structure adopted was a first-order Sugeno rule [58], since it has been established that this model can accurately represent complex behavior with only a few rules [59]. The rules thus have the form

$$z_{c_i} = \mathbf{g}_i \mathbf{x}^T + h_i \tag{15}$$

where $\mathbf{g}_i$ a $1 \times N$ vector and h_i is a constant. Given that we have a set of rules with fixed premises, optimizing the parameters in the consequent equations with training data reduces to a least squares problem [58]. Letting

$$\rho_i = \frac{u_i}{\sum_{j=1}^{K} u_j} \tag{16}$$

the equation for the output of the system can be written

$$z = \sum_{i=1}^{K} \rho_i z_{c_i} = \sum_{i=1}^{K} \rho_i (\mathbf{g}_i \mathbf{x}^T + h_i). \tag{17}$$

Given M training vectors, the set of outputs can be written

$$\begin{bmatrix} z_1 \\ \vdots \\ z_M \end{bmatrix} = \begin{bmatrix} \rho_{1,1}\mathbf{x}_1^T & \rho_{1,1} & \cdots & \rho_{K,1}\mathbf{x}_1^T & \rho_{K,1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{1,M}\mathbf{x}_M^T & \rho_{1,M} & \cdots & \rho_{K,M}\mathbf{x}_M^T & \rho_{K,M} \end{bmatrix} \times \begin{bmatrix} \mathbf{g}_1^T \\ h_1 \\ \vdots \\ \mathbf{g}_K^T \\ h_K \end{bmatrix} \tag{18}$$

where $\rho_{i,j}$ denotes ρ_i evaluated at $\mathbf{x}_j$. Again, an output z is set to 1 or 0 depending on whether the corresponding input feature vector $\mathbf{x}$ is associated with an ordnance or nonordnance item. Given the set of training vectors, the first matrix on the right-hand side (RHS) of the above equation is constant while the second one consists of the parameter vectors to be estimated. Thus, it takes the form of a standard linear problem with the format of $\mathbf{AP} = \mathbf{B}$ that can be solved by least squares. To solve for the parameters of interest, the recursive least squares (RLS) estimation algorithm is used in order to avoid both the computational complexity associated with calculating the inverse of $\mathbf{A}$, as well as the numerical stability issues. RLS has been fully described [60], and determines the parameters $\mathbf{P}$ via the following iterative formula

$$\begin{aligned} \mathbf{P}_{i+1} &= \mathbf{P}_i + \Sigma_{i+1}\mathbf{a}_{i+1}\left(\mathbf{b}_{i+1}^T - \mathbf{a}_{i+1}^T\mathbf{P}_i\right) \\ \Sigma_{i+1} &= \Sigma_i - \frac{\Sigma_i \mathbf{a}_{i+1}\mathbf{a}_{i+1}^T \Sigma_i}{\mathbf{1} + \mathbf{a}_{i+1}^T \Sigma_i \mathbf{a}_{i+1}} \quad i = 0, 1, \ldots, n-1 \end{aligned} \tag{19}$$

where

- $\mathbf{P}_i$ parameter estimate at the ith iteration;
- Σ_i covariance matrix;
- $\mathbf{a}_i^T$ ith row vector of $\mathbf{A}$;
- $\mathbf{b}_i^T$ ith row vector of $\mathbf{B}$;
- $\mathbf{P}_n$ final least squares estimate.

The initial condition of the iteration can be set to $\mathbf{P_0} = \mathbf{0}$ and $\Sigma = \mu\mathbf{I}$, where $\mathbf{I}$ is an identity matrix and μ is a large positive number.

Next, the (first-order) Sugeno type fuzzy inference system is constructed. There are K fuzzy IF-THEN rules; each rule follows the format

$$\begin{aligned} &\text{If } x(1) \text{ is } A_{i1} \text{ and } x(2) \text{ is } A_{i2} \text{ and } \ldots x(n) \text{ is } A_{in} \\ &\text{then } z = \mathbf{g}_i \mathbf{x}^T + h_i \end{aligned} \tag{20}$$

where

- $x(j)$ jth element of the vector under test, $\mathbf{x}$;
- z system output;
- $\mathbf{g}_i$, h_i defined above.

A_{ij} is an exponential, or Gaussian, membership function given by

$$A_{ij}(q) = \exp\left(-\frac{(q - x_{c_i}(j))^2}{2\sigma^2}\right) \tag{21}$$

where $x_{c_i}(j)$ is the jth element of $\mathbf{x}_{c_i}$ (i.e., the ith rule premise), and σ is defined in the firing strength expression [(13)]. The membership functions are completely defined by the cluster centers. In effect, aggregation is performed by adding the outputs after multiplication by the firing strength and defuzzification is implemented by weighting the average of the outputs. A Gaussian membership function was chosen since it is completely specified by two parameters, it is invariant under multiplication and Fourier transformation, it is smooth, and the notation is concise. Multiplication is used as the AND operator. The fuzzy inference system weights the output of each rule by the rule's firing strength, and computes the output value as a weighted average of the output of each rule.

E. Training of the Algorithms

There was no separate training area available to NRL prior to or during the JPG IV demonstration, so a round-robin training method, also known as a leave-one-out method, was one of the approaches used to determine the parameters required by the various algorithms. This approach attempted to provide a baseline as to the best possible performance attainable by these algorithms for the particular data set. The method uses the data set circularly, so that the object being tested is not used to develop the statistics that are used to test it. For example, when attempting to determine whether the object associated with measurement $\hat{\mathbf{x}}$ is a target or clutter item for the GLRT processor, we

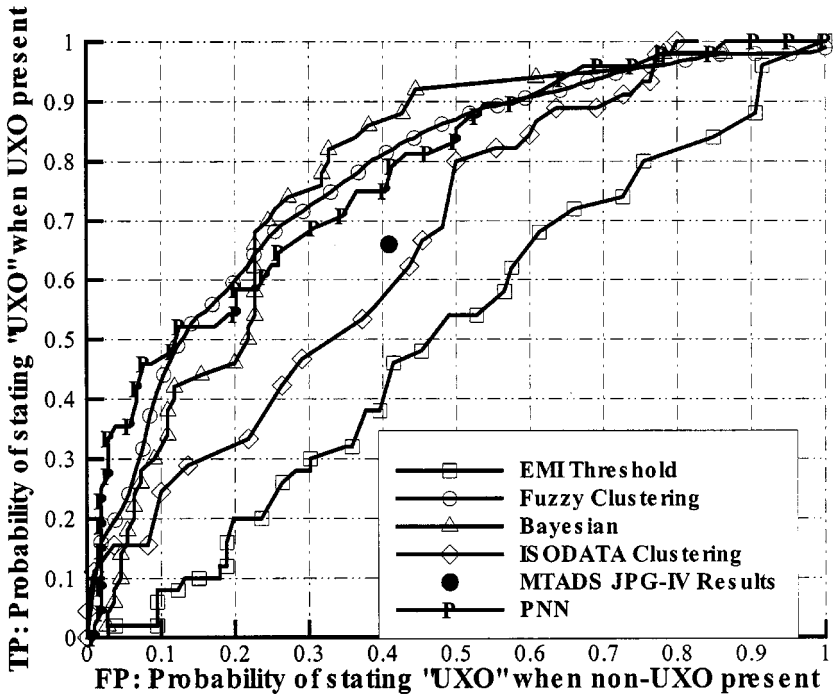


Fig. 3. Performance for the four algorithms and a baseline measure (threshold on the EMI data) for cross-validation testing on the JPG IV data. *MTADS* system performance at the JPG IV demonstration is also shown.

train the model pdfs, $p(\mathbf{x} \mid H_1, \bullet)$ and $p(\mathbf{x} \mid H_0)$, with the entire data set with the exception of $\hat{\mathbf{x}}$. In this way the training and detection data are separate, and the largest possible training set is used. This approach results, for example, in 109 feature vectors at the hidden level of the PNN.

After the ground truth was released, NRL acquired the nonordnance items used at JPG IV. Measurements at several depths and orientations of 20 of these items, as well as 9 inert ordnance items were made in the test pit located at the NRL test facility in Blossom Point, MD, USA using the magnetometer array and the EMI arrays to conduct orthogonal surveys. The test pit data thus formed a comprehensive, but separate, "training data set." This test pit data was used to train a subset of the algorithms described earlier, and then the JPG IV data was used as a test set. In this scenario, no training was performed on the JPG IV data. Such an approach is often preferred to evaluate the robustness of the candidate algorithms. Other data collected with the *MTADS* system was not utilized for training as the nonordnance items emplaced for the JPG IV test were substantially different than nonordnance items commonly encountered in field exercises.

In the leave-one-out training approach, 110 different training sets were used each composed of 109 training vectors, and thus the discrimination algorithm for each of the four discrimination approaches change slightly as the training sets change. For example, the cluster centers for the subtractive fuzzy clustering algorithm change slightly across each of the 110 training sets; however, the changes are not dramatic. Although this is not an approach that would be used in practice, it is a standard mechanism by which to evaluate performance when the amount of data available is limited. When the test pit data was used for training, each algorithm was trained once and the resulting algorithm was applied to the JPG IV data. For the fuzzy-clustering algorithm, the number of cluster centers, which is the same as the number of fuzzy rules, was four in the leave-one-out training approach. This relatively small number of cluster centers attests to the particular advantage of using the Sugeno type fuzzy rules in that complex behavior can be represented with only a few rules. Three clusters were generated when the test pit data was used for training the algorithm.

V. Results

The ROC curves for the four algorithms described in the previous section are shown in Fig. 3 for the case where leave-one-out training was applied in conjunction with the JPG IV data. ROC performance is plotted as the "Probability of stating UXO is present when UXO is present," or True Positives, as a function of the "Probability of stating UXO is present when UXO is not present," or False Positives. All processors utilize both magnetometer and EMI data, thus the results are essentially "sensor fusion" results. For the GLRT, or Bayesian, processor, the feature set that provides the best performance consists of magnetometer-moment azimuth, transverse-EMI-polarizability, and the ratio between the two EMI polarizability moments. The remaining processors utilized all of the features—performance was not dramatically affected by exclusion of the features not included in the GLRT processor. Fig. 3 plots the ROC performance of each of the four processors considered, as well as a ROC obtained for a simple threshold processor operating on the EM data alone. The EM threshold is indicative of the performance obtained by a cueing processor that might be used as a preprocessor. In addition, the performance of the *MTADS* system using "man in the loop" processing as scored for the JPG IV demonstration is shown.

Clearly, a simple threshold on the EMI data is not an effective discriminator of UXO from non- UXO items. The maximum-

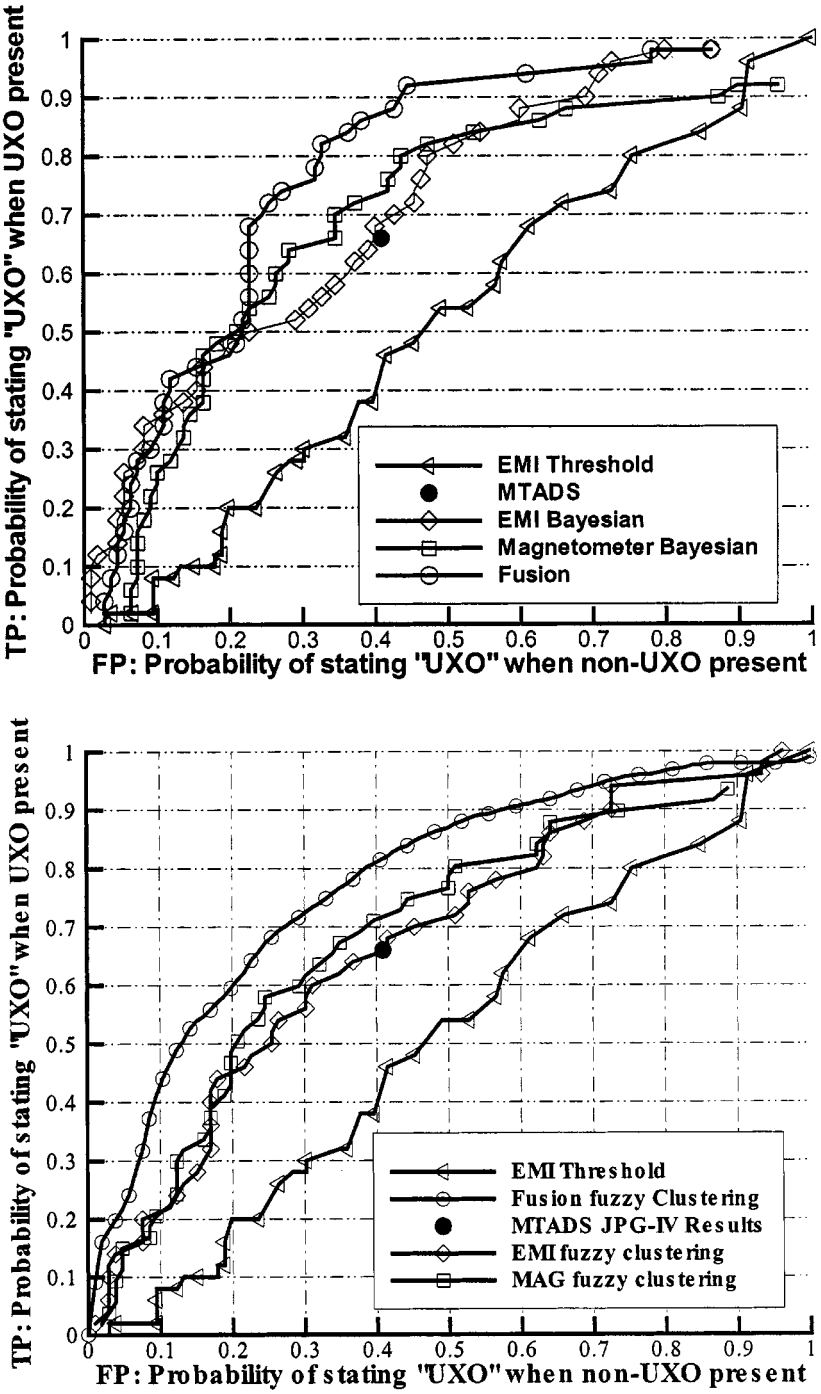


Fig. 4. Comparison of single sensor performance versus fused performance for the Bayesian/GLRT algorithm (top panel) and the subtractive fuzzy clustering algorithm (bottom panel).

likelihood clustering algorithm improves performance somewhat, and its performance is near that achieved by the NRL "man-in-the-loop" approach. The PNN, subtractive fuzzy clustering technique, and the Bayesian approach provide substantial performance improvements over that obtained by the maximum likelihood clustering technique. This may be a result of the additional information incorporated by the fuzzy rule base and the PNN, or the correlation structure incorporated into the Bayesian processor. Although the limited amount of data does not allow strict statistical comparisons across algorithms, it appears that the PNN performance is better than the other algorithms at the lower true positive rates and that the GLRT performance is better than the other algorithms at the higher true positive rates. The subtractive fuzzy technique tends to perform at the average of the PNN and GLRT algorithms. No attempt has been made to date to fuse the processor outputs using a technique such as algorithm ensembles [61] (i.e., joint decisions made using many diverse classification algorithms), although this and other techniques will be considered in future work.

ROC curves for the GLRT processor and the subtractive fuzzy processor, where the EMI and magnetometer features have been considered separately are shown in Fig. 4. These

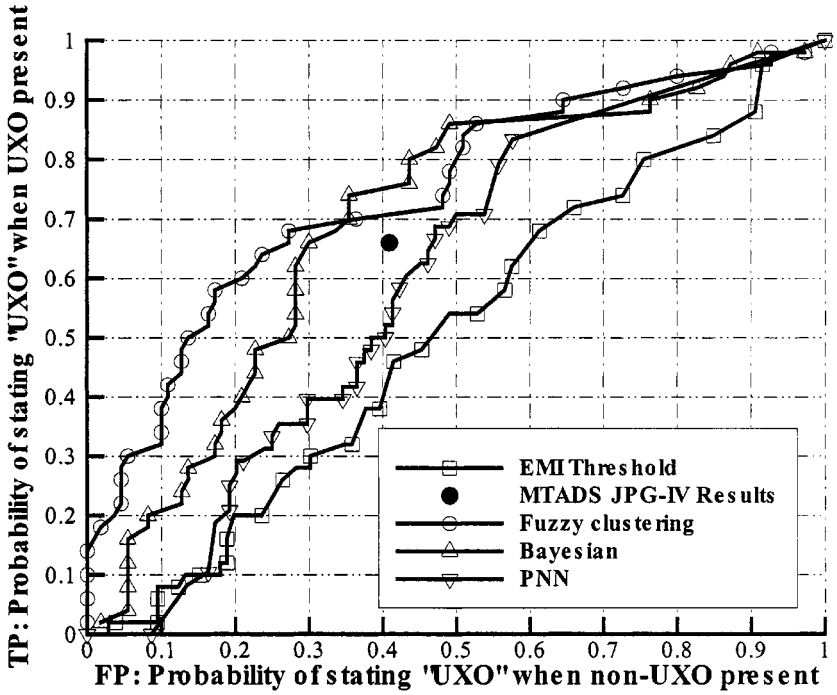


Fig. 5. Cross-site testing and training results.

processors were considered so that the relative performance associated with each sensor could be gauged. Again, the performance of a simple EMI thresholding algorithm and the NRL "man in the loop" processing results are shown for comparison. The GLRT (top panel) operating on the magnetometer features alone provides results comparable to NRLs man-in-the loop processing. The features used for the GLRT/magnetometer-only processor consisted of dipole-azimuth and dipole-orientation. Utilization of these features has been suggested and validated in previous studies that suggested that discriminating UXO based on induced versus remnant magnetization was possible [62]. The GLRT operating on the EMI features alone, which consisted of the depth estimate, transverse-EMI-polarizability, and the ratio between the two EMI polarizability moments, results in better performance than that obtained with the magnetometer alone. Fusion of the data features between the EMI and the magnetometer provides an additional performance improvement over that obtained with the EMI features alone, especially at the higher true positive probabilities. The fuzzy technique (bottom panel) operating on the EMI features alone provide results comparable to NRLs man-in-the loop processing, and the fuzzy technique applied to the magnetometer features results in slightly better performance. As observed with the GLRT, fusion of the data features across sensors provides an additional performance improvement over either sensor alone. This performance improvement is observed for all true-positive levels.

Fig. 5 shows ROC curves with the sensor features fused for the Bayesian, fuzzy clustering, and PNN for the case where the algorithms were trained on the test pit data collected at the Blossom Point site and tested on the data collected during the JPG IV exercise. In this case, the training data set and the test data set were completely separate. As in the previous figures, the performance of an algorithm which simply thresholds the EMI data as well as the NRL "man in the loop" performance is shown for comparison. As expected, the performance for each of the algorithms degrades over that obtained with the leave-one-out training on the JPG IV data set. The PNN performance is degraded the most, although its performance continues to exceed that of the simple EMI threshold. The subtractive fuzzy technique provides the best performance overall, although the GLRT performance is similar at the higher true positive rates. In the UXO classification problem, the 90%–95% true positive region or higher on the ROC curve is considered the ultimate discrimination goal, given the risks associated with lower UXO clearance rates. Comparing the results at a 90% true positive rate shows that the PNN and GLRT results are equivalent (approximately 76% false positives) while the fuzzy clustering algorithm has higher performance (approximately 64% false positives).

One possible reason that the performance is degraded so substantially for the cross-site training results is that the EMI data at JPG IV were collected using an older version of the data acquisition system than was used to obtain the training data at the Blossom Point site. In the JPG IV collection with the older data acquisition system, there was a variable and drifting offset between the three EMI sensors. As a result, what should have been symmetric signals from vertical objects often were not as symmetric as theory would predict, and this lack of symmetry affected the estimation of the features. Thus, the statistical properties of the physics-based features associated with the Blossom Point training set could be different from those encountered at JPG IV. The magnitude of this effect will be considered in future work using data from other sites. An additional source of error may arise from the presence of naturally occurring magnetic noise at JPG IV which was not present in the Blossom Point training data.

VI. Conclusion

The performance of a variety of signal processing techniques has been evaluated on a set of data collected by the *MTADS* sensor suite with the goal of decreasing the false-alarm rate associated with discriminating UXO from clutter. Prior to this analysis, data was collected with the *MTADS* sensor during the JPG IV technology demonstration exercise. This exercise was designed to test the ability to discriminate between ordnance and nonordnance. The techniques that were applied to this data included a maximum likelihood clustering technique, a probabilistic neural network, a generalized likelihood ratio test, and a subtractive fuzzy clustering technique. These approaches were selected based on their hypothesized robustness to uncertainty or their robustness when faced with a limited training set. A set of physics-based features obtained from the simple magnetometer and EMI models was available to each algorithm. The algorithms were compared based on their ability to discriminate nonordnance, or clutter, items from ordnance items. Performance of the various algorithms was evaluated under two training scenarios: a leave-one-out approach where the JPG IV data was used for both testing and training, and a cross-site approach where data from a collection separate from the JPG IV exercise was utilized. This approach allowed both an assessment of performance bounds and an estimate of the robustness of the various algorithms.

All of the signal processing algorithms considered, with the exception of the maximum likelihood clustering technique, appear to hold promise for processing EMI and magnetometer data for UXO discrimination when limited training data is available. The *MTADS* false-alarm probability was 42% during the demonstration, and the detection probability was 66%. The PNN, GLRT, and subtractive fuzzy technique achieved false-alarm probabilities (respectively) of 27%, 22%, and 24% in the leave one out training approach and 47%, 30%, and 27% in the cross-site training approach at a detection probability of 66%. Thus, the false alarm rate can be reduced by as much as a factor of 2 using these approaches on this data set. Alternatively, at a fixed false-alarm probability, the improvement in the detection rate can also be considered. The PNN, GLRT, and subtractive fuzzy technique achieved detection probabilities (respectively) of 75%, 87%, and 81% in the leave one out training approach and 50%, 76%, and 71% in the cross-site training approach at a false alarm probability of 42%. Again, the algorithms considered here provided substantial improvements in performance.

The success of the GLRT and the PNN may be at least partially due to the fact that these approaches model the decision process statistically, and thus deal effectively with problem uncertainty. The success of the subtractive fuzzy technique is probably also due, at least partially, to the ability of such techniques to robustly handle uncertainty in the data. Such approaches must be tested on more extensive data sets prior to implementing them for a system such as *MTADS*.

The analysis of the JPG IV data is predicated on the assumption that a model can be used to extract features from data that has been collected in an area where either a target or clutter item is located. The features obtained from this region of interest (ROI) are then processed by one of the algorithms to determine whether the object is ordnance or nonordnance. During field operations, ground truth is not available, and an algorithm or operator must cue the algorithms to process specific data segments. Clearly, it is important that the cueing algorithm maintain a very high detection rate. During the third JPG demonstration, the *MTADS* was reported to achieve 95% detection at a false-alarm rate of approximately 42 per hectare [63]. Currently, the *MTADS* uses three cueing approaches: an operator (man-in-the-loop), an EM-based cue-er, and a magnetometer-based cue-er. The performance of such cueing algorithms without the subsequent signal processing algorithms was simulated in this work by analyzing the performance of a simple EM threshold algorithm. The cross-site training and testing results indicate that the false-alarm rate was reduced by about 30% by the best performing algorithm for most detection rates over this simple cuing algorithm.

The performance of all of the algorithms considered exceeded that obtained by a trained human operator processing the JPG IV data "in the loop," although this operator was not allowed access to training data from this particular site. When the data was re-evaluated by a separate operator, the "man in the loop" performance improved [31]. However, the performance of the automated algorithms presented here continues to out-perform the man in the loop approach. The performance also exceeded that of a simple EMI threshold. It is important to acknowledge that in addition to the classification performance improvement, all of the algorithms remove the human expert from the detection and remediation crew while at the same time significantly reducing the time required for analysis, thus reducing the associated clean-up costs. Future work will consider the utilization of features from more sophisticated EMI and magnetometer models.

Acknowledgment

The authors would like to thank Dr. C. Baum for the helpful discussions regarding the modifications to the dipole EMI model.

References

[1] H. Nelson, J. McDonald, and R. Robertson, "MTADS TECHEVAL Demonstration, October 1996," NRL/PU/6110-97-348, 1997.

[2] J. McDonald, H. Nelson, J. Neece, R. Robertson, and R. Jeffries, "MTADS unexploded ordnance operations at the Badlands Bombing Range on the Pine Ridge Reservation, Cuny Table, SD," NRL/PU/6110-98-353, 1998.

[3] J. McDonald, H. Nelson, and R. Robertson, "MTADS Live Site Survey, Bombing Target #2 at the Former Buckley Field, Arapahoe County, CO," NRL/PU/6110-99-379, 1999.

[4] J. McDonald, H. Nelson, R. Robertson, and R. Jeffries, "MTADS Live Site Demonstration at the Pueblo of Laguna, NM," NRL/PU/6110-00-398, July 6–Aug. 7 2000.

[5] J. McDonald, H. Nelson, R. Jeffries, and R. Robertson, "Results of the MTADS technology demonstration #3," NRL Tech. Rep., NRL/PU/6110-99-375, 1999.

[6] S. Tantum and L. Collins, "A comparison of algorithms for subsurface object detection and identification using time domain electromagnetic induction data," IEEE Trans. Geosc. Remote Sensing, to be published.

[7] L. Collins, S. Tantum, P. Gao, J. Moulton, L. Makowsky, D. Reidy, and D. Weaver, "Improving detection of low metallic content landmines using EMI data," in *Proc. SPIE*, Apr. 2000, pp. 14–24.

[8] T. Dogaru, L. Collins, and L. Carin, "Optimal time-domain detection of a deterministic target buried under a randomly rough interface," IEEE Trans. Antennas Propagat., to be published.

[9] P. Gao and L. Collins, "A 2-dimensional generalized likelihood ratio test for landmine detection," *Signal Processing*, vol. 80, pp. 1669–1686, July 2000.

[10] ——, "A theoretical performance analysis and simulation of time-domain EMI sensor data for land mine detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 2042–2055, July 2000.

[11] P. Gao, L. Collins, P. Garber, N. Geng, and L. Carin, "Classification of landmine-like metal targets using wide-band electromagnetic induction," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1352–1361, May 2000.

[12] P. Gao and L. Collins, "A comparison of optimal and sub-optimal processors for classification of buried metal objects," *IEEE Signal Processing Lett.*, vol. 6, Aug. 1999.

[13] L. Collins, P. Gao, and L. Carin, "An improved Bayesian decision theoretic approach for land mine detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 811–819, Mar. 1999.

[14] P. Gader, M. Mystkowski, and Y. Zhao, "Application of hidden Markov models to landmine detection using ground penetrating radar," IEEE Trans. Geosci. Remote Sensing, to be published.

[15] P. Gader, J. Keller, and B. Nelson, "Recognition technology for the detection of buried land mines," *IEEE Trans. Fuzzy Syst.*, vol. 9, pp. 31–43, Feb. 2001.

[16] P. Gader, B. Nelson, H. Frigui, G. Vaillette, and J. Keller, "Landmine detection in ground penetrating radar using fuzzy logic," *Signal Processing—Special Issue Fuzzy Logic Signal Processing*, vol. 80, pp. 1069–1084, June 2000.

[17] P. Gader, B. Nelson, A. Hocaoglu, S. Auephanwiriyakul, and M. Khabou, "Neural versus heuristic development of choquet fuzzy integral fusion algorithms for land mine detection," in Neuro-Fuzzy Pattern Recognition, H. Bunke and A. Kandel, Eds. Singapore: World Scientific Co., submitted for publication.

[18] P. Gader, Y. Zhao, M. Mystkowski, M. Khabou, Y. Zhang, and A. Hocoaglu, "Hidden Markov models and morphological neural networks for GPR-based landmine detection," in *Proc. SPIE Conf. Detection and Remediation Technology Mines and Minelike Targets IV*, Orlando, FL, to be published.

[19] J. Keller, S. Auephanwiriyakul, and P. Gader, "New fuzzy set tools to aid in predictive sensor fusion," in *Proc. SPIE Conf. Detection and Remediation Technology for Mines and Minelike Targets IV*, to be published.

[20] A. Hocaoglu, P. Gader, E. Gelenbe, and T. Kocak, "Optimal linear combination of order statistics filters and their relationship to the delta-operator," in *Proc. SPIE Conf. Detection and Remediation Technology Mines and Minelike Targets IV*, Apr. 1999, pp. 1323–1329.

[21] M. McClure and L Carin, "Matched pursuits with a wave-based dictionary," *IEEE Trans. Signal Processing*, vol. 45, pp. 2912–2927, Dec. 1997.

[22] T. Dogaru and L. Carin, "Time-domain sensing of targets buried under a rough air-ground interface," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 360–372, Mar. 1998.

[23] M. McClure and L. Carin, "Wave-based matched-pursuits detection of submerged elastic targets," *J. Acoustical Soc. Amer.*, vol. 104, pp. 937–946, Aug. 1998.

[24] L. Carin, R. Kapoor, and C. E. Baum, "Polarimetric SAR imaging of buried landmines," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1985–1988, Nov. 1998.

[25] D. Wong and L. Carin, "Analysis and processing of ultra-wideband SAR imagery for buried landmine detection," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 1747–1748, Nov. 1998.

[26] L. Carin, N. Geng, M. McClure, J. Sichina, and L. Nguyen, "Ultra-wide-band synthetic aperture radar for mine field detection," *IEEE Antennas Propagat. Mag.*, vol. 41, pp. 18–33, Feb. 1999.

[27] P. Runkle, P. Bharadwaj, and L. Carin, "Hidden Markov model multi-aspect target classification," *IEEE Trans. Signal Processing*, vol. 47, pp. 2035–2040, July 1999.

[28] P. K. Bharadwaj, P. R. Runkle, and L. Carin, "Target identification with wave-based matching pursuits and hidden Markov models," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 1543–1554, Oct. 1999.

[29] A. Sullivan, R. Damarla, N. Geng, Y. Dong, and L. Carin, "Ultrawide-band synthetic aperture radar for detection of unexploded ordnance: Modeling and measurements," IEEE Trans. Antennas Propagat., to be published.

[30] S. Hart, H. Nelson, R. Grimm, S. Rose-Pehrsson, and J. McDonald, "Probabilistic neural networks for unexploded ordnance (UXO) classification using data fusion of magnetometry and EM physics-derived parameters," in *Proc. UXO Forum*, Anaheim, CA, May 2000.

[31] B. Barrow and H. Nelson, "Model-based characterization of EM induction signatures for UXO/clutter discrimination using the *MTADS* platform," in *Proc. UXO Forum*, Atlanta, GA, May 1999.

[32] H. Nelson and J. McDonald, "Pulsed induction and magnetic sensor fusion results using the *MTADS*," in *UXO Forum Proceedings*, Atlanta, GA, May 1999.

[33] B. Barrow and H. Nelson, "Collection and analysis of multi-sensor ordnance signatures," *J. Environ. Eng. Geophys.*, vol. 3, no. 2, pp. 71–79, 1998.

[34] ——, "Model-based characterization of EM induction signatures obtained with the *MTADS* EM array," IEEE Trans. Geosci. Remote Sensing, to be published.

[35] S. Hart, R. Shaffer, S. Rose-Pehrsson, and J. McDonald, "A feasibility study on using physics-based modeler outputs to train probabilistic neural networks for UXO classification," NRL/PU/6110-99-385, 1999.

[36] "UXO Techology demonstration program and jefferson proving ground: Phase IV," U.S. Army Environmental Ctr., Rep. SFIM-AEC-ET-CR-99 051, May 1999.

[37] B. Barrow, "Testing of the Modified Geonics EM61 for the MTADS System," AETC Tech. Rep. VA-074-063-TR, 1996.

[38] C. E. Baum, Ed., *Detection and Identification of Visually Obscured Targets*: Taylor Francis, 1998.

[39] L. Carin, H. Yu, Y. Dalichaouch, and C. Baum, "On the wide-band EMI response of a rotationally symmetric permeable and conducting target," IEEE Trans. Geosci. Remote Sensing, to be published.

[40] C. Balanis, *Advanced Engineering Electromagnetics*. New York: Wiley, 1989.

[41] P. Gill, W. Murray, and M. Wright, *Practical Optimization*. New York: Academic, 1981.

[42] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.

[43] G. Ball and D. Hall, "Isodata—A novel method of data analysis and pattern classification," Tech. Rep., (NTIS AD699 616), Stanford Res. Inst., Stanford, CA, 1965.

[44] ——, "ISODATA, an iterative method of multivariate data analysis and pattern classification," *IEEE Trans. Int. Comm. Conf.*, June 1966.

[45] A. Jain and R. Dubes, *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[46] C. Therrien, *Decision, Estimation and Classification—An Introduction to Pattern Recognition and Related Topics*. New York: Wiley, 1989.

[47] D. Specht, *Neural Networks*, 1990, vol. 3, pp. 109–118.

[48] S. L. Chiu, "Fuzzy model identification based on cluster estimation," *J. Intell. Fuzzy Syst.*, vol. 2, no. 3, pp. 267–278, 1994.

[49] R. Shaffer, S. Rose-Pehrsson, and R. McGill, "A comparison study of chemical sensor array pattern recognition algorithms," *Anal. Chim. Acta.*, vol. 384, pp. 305–317, 1999.

[50] ——, "Probabilistic neural networks for chemical sensor array pattern recognition: Comparison studies, improvements and automated outlier rejection," Naval Research Lab., Formal Rep. 6110-98-9879, 1998.

[51] R. Shaffer and S. Rose-Pehrsson, "Improved probabilistic neural network algorithm for chemical sensor array pattern recognition," Analytical Chem., to be published.

[52] T. Masters, *Advanced Algorithms for Neural Networks a C++ Sourcebook*. New York: Academic, 1995.

[53] Y. Chtioui, D. Bertrand, and D. Barba, "Reduction of the size of the learning data in a probabilistic neural network by hierarchical clustering—Application to the discrimination of seeds by artificial vision," *Chemom. Int. Lab. Syst.*, vol. 35, pp. 175–186, 1996.

[54] Y. Chtioui, D. Bertrand, M. F. Devaux, and D. Barba, "Comparison of multivariate perceptron and probabilistic neural networks in artificial vision. Application to the discrimination of seeds," *J. Chem.*, vol. 11, pp. 111–129, 1997.

[55] D. F. Specht, *Computational Intelligence: A Dynamic System Perspective*, M. Palaniswami, Y. Attikiouzel, R. J. Marks, D. Fogel, and T. Fokuda, Eds. New York: IEEE Press, 1995.

[56] R. R. Yager and D. P. Filev, *Essentials of Fuzzy Modeling and Control*. New York: Wiley, 1995.

[57] D. Langan, J. Modestino, and J. Zhang, "Cluster validation for unsupervised stochastic model-based image segmentation," *Proc. IEEE Trans. Image*, vol. 7, pp. 180–195, 1998.

[58] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, pp. 116–132, 1985.

[59] M. Sugeno and K. Tanaka, "Successive identification of a fuzzy model and its applications to prediction of a complex system," *Fuzzy Sets Syst.*, vol. 42, pp. 315–334, 1991.

[60] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[61] D. Opitz and R. Maclin, "Popular ensemble methods: An empirical study," *J. Artificial Intell. Res.*, vol. 11, pp. 169–198, 1999.
[62] H. Nelson, T. Altshuler, E. Rosen, J. McDonald, B. Barrow, and N. Khadr, "Magnetic modeling of UXO and UXO-like targets and comparison with signatures measured by *MTADS*," in *Proc. UXO Forum*, Anaheim, CA, May 1998.
[63] S. Park and M. Mander, "Evaluation of the multi-sensor towed array detection system (*MTADS*) performance at Jefferson proving ground, January 14–24, 1997," Institute for Defense Analysis, IDA Document D-2174, Feb. 1999.

Leslie M. Collins (M'96) She received the B.S.E.E. degree from the University of Kentucky, Lexington, in 1985 and the M.S.E.E. and Ph.D. degrees in electrical engineering, both from the University of Michigan, Ann Arbor, in 1986 and 1995, respectively.

She was a Senior Engineer with the Westinghouse Research and Development Center, Pittsburgh, PA, from 1986 to 1990. In 1995, she became an Assistant Professor with the Electrical and Computer Engineering Department, Duke University, Durham, NC. Her current research interests include incorporating physics-based models into statistical signal processing algorithms. She is pursuing applications in subsurface sensing, as well as enhancing speech understanding by hearing imparied individuals.

Dr. Collins is a member of the Tau Beta Pi, Eta Kappa Num, and Sigma Xi honor societies.

Yan Zhang He received the B.S., M.S., and Ph.D. degrees in electrical engineering all from Jilin University of Technology, Changchun, China, in 1993, 1996, and 1998, respectively.

Since January 1999, he has been a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, Duke University, Durham, NC. His present research interests include statistical signal processing, sensor fusion, and related applications.

Jing Li (S'97) He received the B.S. and M.S. degrees in automatic control and system engineering from Xian Jiao-Tong University, China, in 1993 and 1996, respectively. He is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, Duke University, Durham, NC.

His research interests include fuzzy control, LPV control, signal processing, sensor fusion, and pattern recognition. His particular interest is in advanced control design for industrial systems with high performance and reliability.

Mr. Li was a recipient of the Sunshine Fellowship in 1996. He is a member of SIAM.

Hua Wang received the Ph.D. degree from the University of Maryland, College Park.

He is currently an Assistant Professor of Electrical and Computer Engineering, University of Maryland. He is an Associate Editor of the conference editorial board of the IEEE Control Systems Society. His interests include nonlinear control systems, nonlinear dynamics, intelligent systems and control, sensor fusion, control network, biomedical systems, robotics and applications.

Lawrence Carin (S'85–M'89–SM'96–F'00) He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, Durham, NC, where he is an Associate Professor. He is the Principal Investigator on a Multidisciplinary University Research Initiative (MURI) on demining. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies. He is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.

Sean J. Hart received the Ph.D. degree in analytical chemistry in 1998 from Tufts University, Medford, MA.

He then accepted a position as a Senior Staff Scientist with Nova Research, Inc. working on-site at the Naval Research Laboratory (NRL). In June 2000, he accepted a permanent position as a Research Chemist at NRL. He has been involved with the analysis of multivariate data from a variety of chemical and physical sources and the development of optically based chemical sensor systems.

Susan L. Rose-Pehrsson received the Ph.D. degree in analytical chemistry from Pennsylvania State University, University Park, in 1988.

She is a Research Chemist in the Chemistry Division, Naval Research Laboratory. During the past 17 years, she has worked on chemical sensors and multivariate analysis for the detection of hazardous species in air and water.

Herbert H. Nelson He received the B.S. degree in 1975 from Tulane University, New Orleans, LA, and the Ph.D. degree in physical chemistry from the University of California, Berkeley, in 1980.

From 1980 to 1983, he was an NRC Postdoctoral Associate at the Naval Research Laboratory (NRL), Washington, DC, where he is now Head of the Molecular Dynamics Section. For the past ten years, he has been involved in the area of UXO detection, most recently as a Member of the MTADS team. His research at NRL involves investigation of gas-phase reaction dynamics focusing on chemistry important in combustion systems and the atmosphere.

Dr. Nelson is a member of Sigma Xi, the American Chemical Society, and the American Physical Society.

Jim R. McDonald He received the B.S. degree from Southwestern University, Georgetown, TX, and the Ph.D. degree from Louisiana State University, Baton Rouge, in 1964 and 1969, repsectively.

After a Postdoctoral Fellowship, he joined the Naval Research Laboratory in 1970, where he is now the head of the Chemical Dynamics and Diagnostics Branch in the Chemistry Division. Since the mid 1980s, he has headed programs sponsored by both the Department of Defense and other government agencies to develop automated technologies for detection, location, and characterization of buried unexploded ordnance associated with military ranges and associated adjacent lands. These research and development efforts have resulted in the transition of the *MTADS* technology to the commercial sector, which is now available to perform commercial UXO survey services for the Department of Defense and other agencies.

Optimal Time-Domain Detection of a Deterministic Target Buried Under a Randomly Rough Interface

Traian Dogaru, Leslie Collins, *Member, IEEE,* and Lawrence Carin, *Senior Member, IEEE*

***Abstract*—We consider pulsed plane-wave scattering from targets buried under a rough air–ground interface. The properties of the interface are parametrized as a random process with known statistics, and therefore the fields scattered from a particular surface constitute one realization of an ensemble, characterized by corresponding statistics. Moreover, since the fields incident upon a *buried* target must first penetrate the rough interface, they and the subsequent scattered fields are random processes as well. Based on this understanding, an optimal detector is formulated, accounting for the clutter and target-signature statistics (the former due to scattering at the rough surface, and the latter due to transmission); the statistics of these two processes are in general different. Detector performance is compared to that of a matched filter, which assumes the target signature is known exactly (i.e., nonrandom). The results presented here, as a function of angle and polarization, demonstrate that there is often a significant gain in detector performance if the target signature is properly treated as a random process.**

***Index Terms*—Optimal detection, rough surface scattering, time-domain analysis.**

I. Introduction

THE scattering of waves at a rough surface has motivated considerable research [1]–[9]. These studies have been performed primarily in the frequency domain, although there have been some recent time-domain investigations [7], [9]. The rough surface is usually parametrized as a random process (i.e., while the details of a *particular* realization of the surface are not known exactly, each surface is assumed to represent one realization of an *ensemble*, characterized by known statistics). Since the rough surface is treated statistically, the fields scattered and transmitted (for penetrable surfaces) at such a surface must be parametrized statistically as well. In the frequency domain, the fields scattered at a given frequency are parametrized as a random variable [1]–[8], usually by the mean and variance (under the assumption that the random variable has a Gaussian distribution). For time-domain scattering, the fields must be parametrized as a random *process* [9]. In our previous investigation of transient scattering from a Gaussian rough surface [9], we have found the backscattered fields to be wide-sense-stationary and Gaussian; however, the data was correlated (i.e., it was *not*-white).

Manuscript received October 27, 1997; revised April 19, 2000. This work was supported in part by the Army Research Office under Grant DAAH04-96-1-0448 (Demining MURI), the Army Research Laboratory (Adelphi, MD), and by the Office of Naval Research under Grant N00014-96-1-0861.

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0018-926X(01)01270-4.

The statistics of fields scattered by a rough surface are of interest for many applications. For example, such statistics can be used to infer properties of the rough surface, of interest for remote sensing of soil or vegetation [4]–[6]. In the work presented here, we are interested in detection of a target situated near a rough surface, for which the rough-surface-induced scattered fields are usually characterized as clutter. In detection applications, the clutter statistics are often used to prewhiten the data [9]–[11], representing the initial stage in most detectors.

The most commonly used scheme for detection of transient signals is the matched filter [10], in which the target signature is assumed known exactly. However, as discussed above, the fields transmitted through a penetrable rough surface constitute a random process (like the surface-scattered fields), and therefore the fields that impinge upon a buried target are, in turn, random (further randomness can incur if the medium under the surface is itself random [6], although this is not considered here). Therefore, even if the buried target is known exactly (but the rough surface is treated statistically), its scattered fields must be treated as a random process. In a previous paper [9], we have demonstrated that, in many cases, the random quality of the target signature results in matched-filter detector performance—quantified in terms of probability of detection versus probability of false alarm, termed the receiver operating characteristic (ROC)—which is significantly inferior to the expectations of ideal matched-filter theory (which assumes all underlying assumptions are valid). In this paper, we consider implementation of an optimal detector which properly accounts for the random nature of the target's scattered signature.

The development of an optimal detector for a random signal in noise (clutter) requires *a priori* knowledge of both the target and clutter statistics. As discussed above, a whitening filter is generally used to convert the clutter into a white process. The whitening filter is implemented using the clutter correlation matrix [10], [11], computed here via Monte Carlo simulations for scattering from the rough surface alone, in the absence of buried targets (as would be done for the experimental collection of clutter statistics). Therefore, while frequency-domain fields scattered from a rough surface are parametrized by their mean and variance, here time-domain scattering (a random process) is characterized via its correlation matrix. Having quantified the statistics of time-domain rough-surface scattering, it remains to account for the statistics of the stochastic buried-target signature.

As discussed subsequently, an optimal scheme for detection of a random signature requires integration over the signature's density function [10]. Unfortunately, such a density function is difficult to obtain in general. Therefore, we implement the

0018–926X/01$10.00 © 2001 IEEE

optimal detector approximately via Monte Carlo integration, through consideration of multiple realizations of the target signature (each for a particular rough surface, from an ensemble of such). Thus, instead of requiring the statistics of the target response, *per se*, we only require access to a set of target signatures representative of such. Multiple waveforms are calculated here using a forward-scattering algorithm. Recall from above that, in principle, clutter statistics can be measured and therefore a forward algorithm is not essential for such. However, to build an optimal detector, one requires multiple realizations of the target signature, for different manifestations of the rough surface. A fast forward-scattering algorithm is therefore essential for optimal detection of such targets, with random scattering signatures (the statistics of which are not easily measured for variable targets and target positions). While fast algorithms are well known to be requisite elements in inverse-scattering schemes (generally for deterministic scattering data) [12], [13], here we introduce the use of such in optimal detectors (for stochastic scattering data).

The optimal detector is of general utility, with results presented here for the special case of two-dimensional scattering. All scattering data are computed via a finite-difference time-domain (FDTD) algorithm [14]–[30], in which we consider plane-wave incidence and far-zone scattering, using an appropriate near-to-far-zone transformation [9], [30]. Moreover, since this algorithm must be run many times—to compute clutter statistics, multiple realizations of the stochastic target signature, as well as for generation of ROC curves—it is essential that the FDTD be as fast as possible. To reduce the computational domain, we have utilized the perfectly-matched-layer (PML) absorbing-boundary condition [23]–[25], with appropriate modifications for handling lossy soil [26], [27].

The remainder of the paper is organized as follows. In Section II we develop the optimal detector for targets with random scattering signatures. Since the fast forward-scattering algorithm is an integral element of such, it is discussed in this section. Results are presented in Section III, for both TE and TM polarization, wherein comparisons are performed with the idealized matched filter. Conclusions are discussed in Section IV, as are directions for future research.

II. Optimal Detector

A. Target-Signature Model

We consider development of an optimal detector for known targets buried under a rough interface. In a previous paper [9], we investigated the matched filter and evaluated its performance. The matched filter is optimal if the target signature is known exactly and the noise (clutter) is Gaussian and wide-sense stationary. In our previous work [9], we have found that a Gaussian rough surface yields clutter which is Gaussian and wide-sense stationary, although it has been found to be correlated. Therefore, for the problem considered here, the matched filter is preceded by a whitening filter [10].

Having verified that the *clutter* meets the requirements of a matched filter, it remains to investigate the characteristics of the target signature. The assumption of a deterministic target response is investigated by comparing the receiver operating characteristics (ROCs) obtained via Monte Carlo simulation with *idealized* ROC curves, computed presupposing that all underlying assumptions are valid. In such computations, a natural choice for the "known" target signature is its response when buried under a flat interface, represented in discretized form by the N-dimensional vector $\boldsymbol{s} = [s_1 \quad s_2 \quad \dots \quad s_N]^T$; the discretized measured field and clutter are similarly expressed as $\boldsymbol{e} = [e_1 \quad e_2 \quad \dots \quad e_N]^T$ and $\boldsymbol{c} = [c_1 \quad c_2 \quad \dots \quad c_N]^T$, respectively. Therefore, the matched filter assumes that under hypothesis H_1 (target plus clutter) $\boldsymbol{e} = \boldsymbol{s} + \boldsymbol{c}$, while under hypothesis H_0 (clutter only) $\boldsymbol{e} = \boldsymbol{c}$. In light of the fact that the clutter as measured is not white, we assume $\boldsymbol{e}, \boldsymbol{s}$, and $\boldsymbol{c}$ have been passed through a whitening filter. In the Bayesian approach [10], we partition the observation space into two regions, corresponding to the two hypotheses, such that the cost of the decision is minimized. This is done by transforming the N-dimensional problem into a one-dimensional likelihood ratio test

$$\Lambda(e) = \frac{p_c(\boldsymbol{e} - \boldsymbol{s})}{p_c(\boldsymbol{e})} \tag{1}$$

where $p_c(\cdot)$ is the probability distribution function of the clutter $\boldsymbol{c}$. The likelihood ratio is compared against a threshold T [10], and we decide that the signal is present if $\Lambda(e) > T$ and it is not present if $\Lambda(\boldsymbol{e}) < T$. After prewhitening the Gaussian clutter, the natural log of (1) yields the conventional matched-filter sufficient statistic [10] $l = \boldsymbol{e}^T \boldsymbol{s}$.

The ROCs give a measure of the detector performance, through consideration of the probability of detection (P_D) versus the probability of false alarm (P_F). By definition,

$$P_D = \int_T^\infty p(\Lambda|H_1)\, d\Lambda$$
$$P_F = \int_T^\infty p(\Lambda|H_o)\, d\Lambda \tag{2}$$

where $p(\Lambda|H_i)$ is the probability of Λ for hypothesis H_i. It can be shown that, for the case of additive white Gaussian noise (after prewhitening), the probabilities of detection and false alarm depend only on the parameter $d = E/\sigma^2$, where E is the energy of the target response and σ^2 is the noise variance [10].

Therefore, from E and σ^2, one can readily compute the *theoretical* ROC, representative of idealized detector performance, if all the assumptions mentioned earlier are valid. By comparing these theoretical characteristics with the ones obtained by Monte Carlo simulation (considering a large number of surface realizations), we have determined in [9] that the actual detector performance is often well below that of the ideal, especially for near-grazing incidence. We attributed this to an inappropriateness of the deterministic-target-response assumption.

B. Optimal Detector

One can model the randomness in the target response by introducing a generalized stochastic parameter $\boldsymbol{\theta}$. Instead of assuming the target signature $\boldsymbol{s}$ is known exactly, we introduce an uncertainty which we symbolize by introducing the notation $\boldsymbol{s}(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ represents a vector of stochastic parameters that

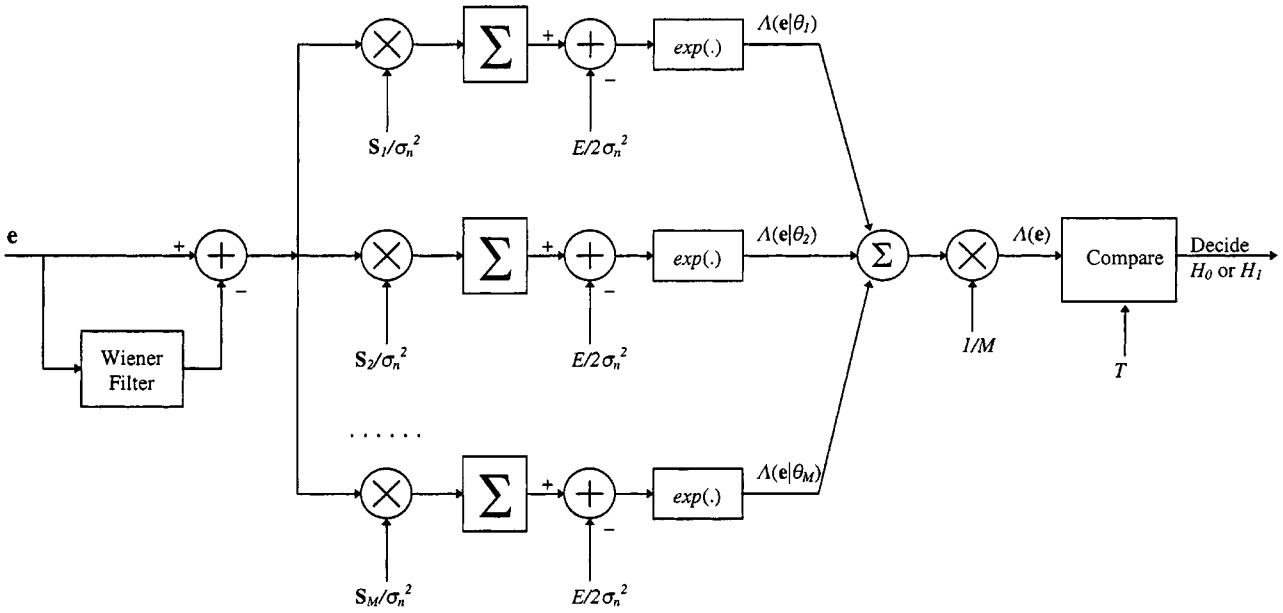


Fig. 1. Optimal detector for a random signal in additive clutter, using a Wiener whitening filter and M realizations of the random signature.

are responsible for the random nature of $\boldsymbol{s}$. Thus, under the hypothesis H_1, the received signal becomes $\boldsymbol{e} = \boldsymbol{s}(\boldsymbol{\theta}) + \boldsymbol{c}$. The previous, simple likelihood ratio is now generalized as

$$\Lambda(\boldsymbol{e}) = \frac{\int p_c[\boldsymbol{e} - \boldsymbol{s}(\boldsymbol{\theta})] p_{\boldsymbol{\theta}}(\boldsymbol{\theta})\, d\boldsymbol{\theta}}{p_c(\boldsymbol{e})} = \int \Lambda(\boldsymbol{e}|\boldsymbol{\theta}) p_{\boldsymbol{\theta}}(\boldsymbol{\theta})\, d\boldsymbol{\theta} \quad (3)$$

where $\Lambda(\boldsymbol{e}|\boldsymbol{\theta})$ represents the likelihood ratio for a particular $\boldsymbol{s}(\boldsymbol{\theta})$ and $p_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ is the probability density function of the vector $\boldsymbol{\theta}$.

Note that we have not specified $\boldsymbol{\theta}$, nor do we quantify its distribution. Even though there are ways to model $\boldsymbol{\theta}$ as a physical quantity, we avoid this and consider an alternative manner of computing the likelihood ratio. Considering (3), note that the expression on the right side is simply the ensemble average of $\Lambda(\boldsymbol{e}|\boldsymbol{\theta})$, computed for all possible values of $\boldsymbol{\theta}$. We can approximate this quantity numerically by performing Monte Carlo integration. Thus, we consider M realizations of the random vector $\boldsymbol{\theta}$, the mth of which is represented by $\boldsymbol{\theta}_m$ (physically corresponding to M realizations of the rough surface), and compute the approximate likelihood ratio as

$$\Lambda(\boldsymbol{e}) = \int \Lambda(\boldsymbol{e}|\boldsymbol{\theta}) p_{\boldsymbol{\theta}}(\boldsymbol{\theta})\, d\boldsymbol{\theta} \approx \frac{1}{M} \sum_{m=1}^{M} \Lambda(\boldsymbol{e}|\boldsymbol{\theta}_m). \quad (4)$$

Consequently, we obtain the structure of the optimal receiver described in Fig. 1.

The whitening filter is implemented as a forward linear prediction-error filter [31]. A linear prediction filter is a Wiener filter, with tap weights represented by the vector $\boldsymbol{w} = [w_0, w_1 \ldots w_P]^T$, ($P$th-order filter) satisfying the Yule–Walker equation $\boldsymbol{R}\boldsymbol{w} = \boldsymbol{r}$, where $r_k = E[c_n c_{n+k}]$, $\boldsymbol{r} = [r_1 \;\; r_2 \;\; \ldots \;\; r_P]^T$, and $\boldsymbol{R}$ represents the correlation matrix $E[\boldsymbol{c}\boldsymbol{c}^T]$, which, for a wide-sense stationary process, has Toeplitz symmetry.

C. Numerical Model

The data used in our study are obtained by numerical simulation using a two-dimensional finite-difference time-domain code (with no variation in the z direction, see Fig. 2). We use the standard Yee cell [14], for both TE and TM polarization. Plane-wave excitation is considered, employing a (total field)–(scattered field) formulation [9], [19], [20]. The backscattered fields to be processed are observed in the far zone, necessitating a near-to-far zone transformation [9], [30], with separate Huygens surfaces used to enclose the target and the rough surface (Fig. 2). Finally, with regard to the rough air–ground interface, it has been demonstrated that the staircase approximation to such (inherent to the Yee scheme) is accurate, as long as the grid size is very small compared with the wavelength, the surface correlation length, and the surface variance [7].

To implement the detector, we must first characterize $p_c(\boldsymbol{c})$, which requires hundreds of rough-surface realizations (N_c), in the absence of a target. Subsequent implementation of the detector in (4), for a given $\boldsymbol{e}$, requires computation of the likelihood ratios $\Lambda(\boldsymbol{e}|\boldsymbol{\theta}_m)$, therefore requiring M realizations of $\boldsymbol{s}(\boldsymbol{\theta}_m)$; in the subsequent examples, we demonstrate that, for the data considered here, M must be on the order of forty. Finally, $\boldsymbol{e}$ is a random process (and therefore $\Lambda(\boldsymbol{e}|\boldsymbol{\theta}_m)$ a random variable). Therefore, the *statistical* detector characterization requires consideration of N_e realizations of the scattered signal $\boldsymbol{e}$, where here we have considered N_e on the order of several hundred. Therefore, detector implementation and characterization requires $2M + N_e + N_c$ scattered-field computations (the factor two in $2M$ is clarified below). Therefore, it is critical that the modeling algorithm be as computationally efficient as possible.

To achieve computational efficiency, it is essential that the FDTD computational domain be as small as possible. Therefore, we have employed a PML absorbing medium for the outwardly

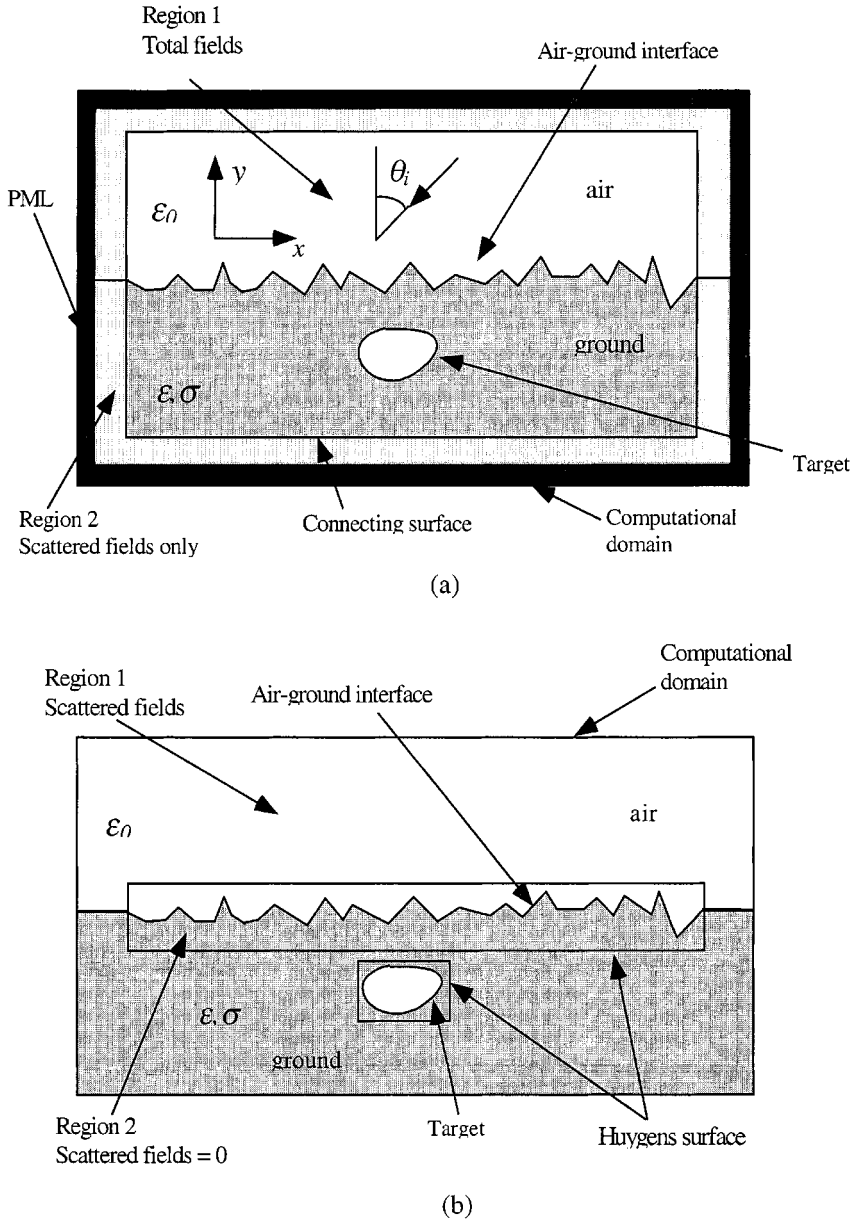


Fig. 2. Schematization of the buried-target problem and the finite-difference time-domain algorithm used for its modeling. (a) Target buried under a rough surface. (b) Summary of FDTD algorithm.

propagating waves leaving the computational domain. The PML used here is designed to absorb waves in lossy media [26], [27], of interest for the lossy half space. For the problems studied, we have employed the incident pulse shown in Fig. 3, representative of a Rayleigh wavelet [32], and have utilized 12 FDTD spatial samples per wavelength, for the smallest wavelength of interest in the problem. The computational size of each FDTD calculation was $9\lambda_c \times 1.5\lambda_c$ (width $\times$ depth), where λ_c is the center wavelength of the incident pulse (Fig. 3) in free space. For this problem size and spatial discretization numerical dispersion [17], [18] was not found to present a problem.

Before proceeding to the results, we reiterate that we have considered plane-wave incidence, motivated by ground-penetrating systems with a large stand-off distance. For example, ultrawideband synthetic aperture radar (SAR) systems [33], [34] employ sensors that are quite distant from the target, for which plane-wave incidence appears most appropriate. However, this introduces a problem. In particular, diffraction is induced at the ends of the numerical rough surface (see Fig. 2), with such *not* representative of the statistics of the rough surface itself. To mitigate this problem, several authors have considered beam excitation [1], [2], [7], which removes the edge effects but also solves a problem different from the plane-wave case of interest. For beam excitation, one could make the beamwidth large enough such that, at least paraxially, a reasonable plane-wave approximation could be made; however, this results in a significant increase in the computational domain. Moreover, we note that previous beam-excitation studies (e.g., [1], [2], [7]) have considered narrow-band problems, while here we are addressing ultrawideband fields (see Fig. 3). For such, one has the added complexity of requiring the beamwidth to be wide relative to *all* wavelengths in the incident pulse, which would require pro-

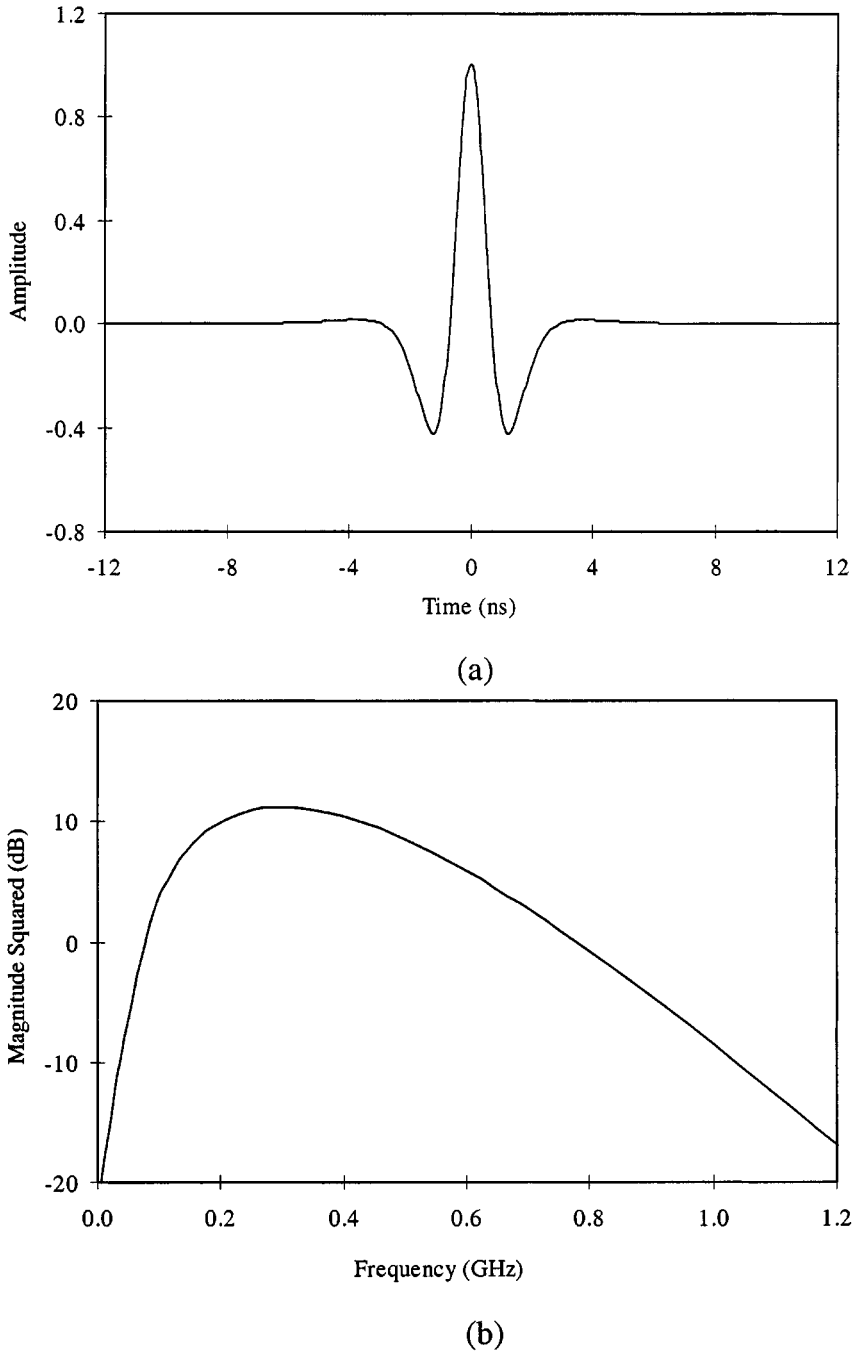


Fig. 3. Incident pulse and corresponding spectrum used in the numerical computations. The energy is peaked at a center frequency of 300 MHz ($\lambda_c = 1$ m). (a) Time-domain waveform. (b) Spectrum.

hibitively wide beams (and hence very large computational domains) at the low frequencies of interest in the pulses used for ground-penetration applications.

Fortunately, for the detector problem of interest here, these difficulties are avoided to a large extent. In particular, for the white Gaussian clutter (after prewhitening) $\boldsymbol{c}$, the optimal detector in (4) reduces to projecting the measured signal $\boldsymbol{e}$ onto M realizations of $\boldsymbol{s}(\boldsymbol{\theta}_m)$. For the targets considered here, $\boldsymbol{s}(\boldsymbol{\theta}_m)$ is nonzero over a relatively limited temporal support, and it is only over this time that the values of $\boldsymbol{e}$ are important to detector performance. For the geometrical parameters considered here, the edge-diffraction-induced effects in $\boldsymbol{e}$ are well outside the temporal support of $\boldsymbol{s}(\boldsymbol{\theta}_m)$, and therefore have no impact on detector performance. We also note that, for the same reasons, the only portion of the clutter $\boldsymbol{c}$ meaningful to detector performance is that in the vicinity of the nonzero support of $\boldsymbol{s}(\boldsymbol{\theta}_m)$; therefore, it was also this portion of $\boldsymbol{c}$ that was used to compute the clutter correlation matrix $\boldsymbol{R}$.

Although the above discussion addresses concerns of edge diffractions induced at the end of the numerical rough surface, we have not accounted for the fact that such diffractions may launch surface waves that would not be present in an actual rough surface (without such endpoints). We have carefully examined the backscattered clutter $\boldsymbol{c}$, to see if such surface waves

are induced (as witnessed in [1] for a highly conducting interface). Such effects, if present, were within the noise of the numerical results and were therefore deemed unimportant for the relatively low-loss medium considered here. Therefore, while plane-wave incidence presents unavoidable problems at the ends of the numerical rough surface, for the ultrawideband applications of interest, beam excitation is not a viable alternative; moreover, for the reasons discussed above, plane-wave excitation does not appear to produce significant problems for the detection problems and material properties considered here.

III. Results

In the examples considered here, we assume a Gaussian rough surface, with zero mean, standard deviation 3.95 cm, and correlation length 18.75 cm. This distribution is felt reasonable for many air–soil interfaces, but it is *not* based on the measured properties from any particular interface. Other parameters and distributions are clearly possible (and, possibly, in some cases more realistic). However, Gaussian interfaces have been well studied [1]–[9] and constitute a good starting point. The soil is represented by a lossy dielectric with $\epsilon_r = 6$ and conductivity $\sigma = 0.005$S/m, which is characteristic of many soil types [35]. Finally, the target is placed 25 cm under the mean air–ground interface position, and consists of a lossless dielectric with $\epsilon_r = 2$, a width of 37.5 cm ("parallel" to the interface), and a thickness of 12.5 cm. We consider both TE and TM polarizations and plane-wave incidence at 30° and 70°, with respect to the normal (see Fig. 2). Finally, in all cases the incident pulse shape corresponds to the Rayleigh wavelet in Fig. 3, with bandwidth representative of current ground penetrating radar systems.

A. Signature of Target Buried Under a Flat Interface

As discussed in Section II-A, the matched-filter detector is effected by projecting the measured data $\boldsymbol{c}$ onto a canonical waveform, here the response of the target when buried under a flat surface, thereby yielding the test statistics $l = \boldsymbol{c}^T\boldsymbol{s}$ ($\boldsymbol{s}$ is the target response for a *flat* interface). It is therefore of interest to examine the characteristics of $\boldsymbol{s}$, as well as the statistics of the clutter $\boldsymbol{c}$, which are fundamental to detector performance.

The properties of $\boldsymbol{s}$ are addressed in Fig. 4(a) and (b), in which the energy spectral density (ESD) of $\boldsymbol{s}$ is investigated for TE and TM excitation, considering angles of incidence $\theta_i = 30°$ and $\theta_i = 70°$, respectively. For the soil considered here, the Brewster angle is 67.8° (neglecting the effects of the conductivity σ). We see that, over much of the spectrum, for $\theta_i = 70°$ the far-zone backscattered fields are stronger for TM excitation than for TE, consistent with the enhanced penetration expected for TM polarization near the Brewster angle. For incidence angle $\theta_i = 30°$ (well away from the Brewster angle), the energy spectral densities for TE and TM excitation are very similar. We note that the polarization and incidence-angle dependence of the target itself, apart from the Brewster-angle effects at the interface, also play a significant role concerning the results in Fig. 4.

B. Clutter Statistics

From Fig. 4, one might expect that, for near-Brewster incidence, a matched filter would perform better for TM polarization than for TE, assuming the underlying matched-filter assumptions are valid: that the target signature is deterministic and similar to that for a target under a flat interface. However, detector performance is strongly influenced by the cutter characteristics, the power spectral density (PSD) for which are shown in Fig. 5, using the surface statistics, polarizations, and incidence angles considered above. We see from Fig. 5 that the PSD is stronger for TM excitation than for TE, appreciably so for $\theta_i = 70°$. Therefore, it is possible that detector performance will actually be worse for near-Brewster TM excitation, as compared to TE incidence at the same angle, despite the fact that the flat-surface target response is larger for TM incidence. Note that the clutter is nonwhite (is correlated), and the detector performance will ultimately be determined by the properties of the whitened clutter. Nevertheless, the results in Figs. 4 and 5 indicate that the assumption of optimal detector performance for near-Brewster TM excitation may be undermined by the properties of the rough-surface-generated clutter.

C. Detector Performance

In the next series of figures, we demonstrate ROC performance for the detectors discussed in Section II, considering the geometrical properties and operating conditions addressed in Figs. 4 and 5. In each figure, three results are presented: idealized matched-filter performance, assuming all underlying assumptions are valid; actual matched filter performance; and detector performance for the optimal detector in (4), using $M = 50$ (for the examples considered here, the results stabilized for $M > 40$). The idealized matched-filter results were computed by applying the matched filter to $N_e = 300$ *synthesized* waveforms, constructed by adding the flat-surface target response to 300 realizations of the fields scattered from the rough surface, in the absence of a buried target (these were *not* the same $N_c = 300$ clutter realizations used to design the whitening filter). Actual matched-filter performance was computed using $N_e = 300$ realizations of fields scattered from a target buried under a randomly rough surface (making no assumptions that the target signature is deterministic). This same scattering data was used to characterize the optimal detector in (4).

To effect the optimal detector, we require M realizations of the target signature $\boldsymbol{s}(\boldsymbol{\theta}_m)$, which must characterize the statistical variation of $\boldsymbol{s}(\boldsymbol{\theta})$, dictated by $p_\theta(\boldsymbol{\theta})$. To compute $\boldsymbol{s}(\boldsymbol{\theta}_m)$, we consider M realizations of the rough surface, and for each we run the FDTD code twice: once with the target and once without, obtaining $\boldsymbol{c}_m + \boldsymbol{s}(\boldsymbol{\theta}_m)$ and $\boldsymbol{c}_m$, respectively. The difference between these two waveforms is defined to be $\boldsymbol{s}(\boldsymbol{\theta}_m)$, for use in (4). Thus, we parametrize the target signature as random and explicitly enforce (by definition) the additivity of the target signature and clutter, both of which are random.

In Figs. 6 and 7, we plot ROC curves for TE and TM polarization, respectively, for incidence angle $\theta_i = 30°$. From Figs. 4(a) and 5(a), for $\theta_i = 30°$, the target ESD and clutter PSD are very similar for TE and TM polarization. In fact, defining E_{TE} and E_{TM} as the energies in the flat-surface target signatures for TE and TM polarization, respectively, we have found that $E_{\mathrm{TM}}/E_{\mathrm{TE}} = 1.416$ for $\theta_i = 30°$. Moreover, defining σ^2_{TE} and σ^2_{TM} as the clutter variance for TE and TM polarization, respectively, we have found that, for $\theta_i = 30°$, $\sigma^2_{\mathrm{TM}}/\sigma^2_{\mathrm{TE}} = 1.584$.

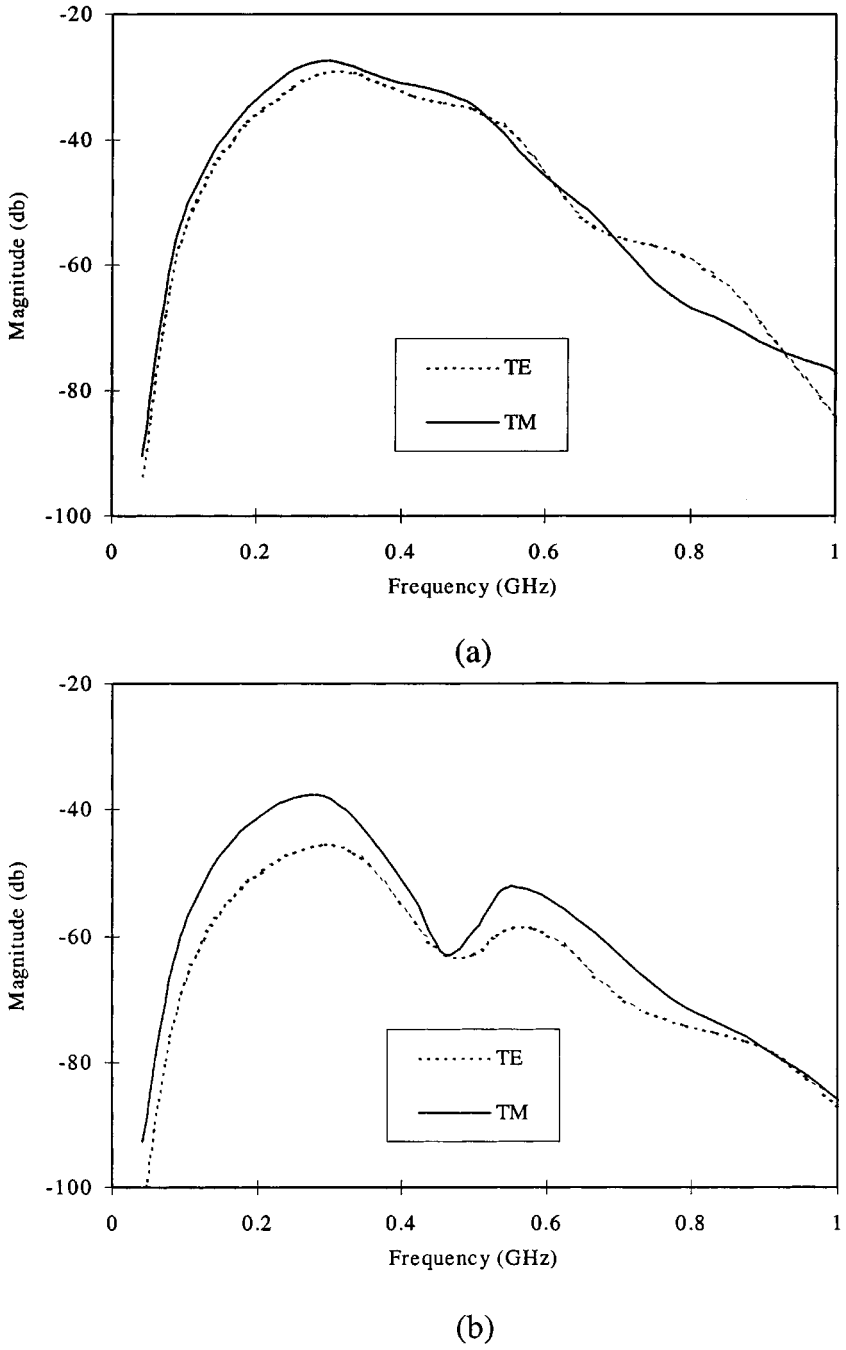


Fig. 4. Energy spectral density of a 37.5 cm × 12.5 cm dielectric target ($\epsilon_r = 2$) buried 25 cm beneath a flat air–ground interface, with the lossy soil characterized by $\epsilon_r = 6$ and $\sigma = 0.005$S/m. Results are plotted for both TE (horizontal) and TM (vertical) polarization. (a) Incidence angle $\theta_i = 30°$. (b) Incidence angle $\theta_i = 70°$.

The similarity of the target energies and clutter variances for TE and TM polarization forecasts the similarity in detector performance manifested in Figs. 6 and 7, for $\theta_i = 30°$. The other significant observation from these figures. is that, for $\theta_i = 30°$, the optimal detector performs only slightly better than the simple matched filter. This issue will be addressed after first examining performance for incidence angle $\theta_i = 70°$.

In Figs. 8 and 9 we consider detector performance for TE and TM polarizations, respectively, for an incidence angle of $\theta_i = 70°$. For this example, using the same notation as above, $E_{\rm TM}/E_{\rm TE} = 5.671$ and $\sigma^2_{\rm TM}/\sigma^2_{\rm TE} = 11.018$. Therefore, it is not surprising that, for $\theta_i = 70°$, the detectors perform markedly better for TE polarization than for TM (reduced false alarms required to achieve a given detection probability), despite the fact that the angle of incidence is very near the Brewster angle. While the ratios $E_{\rm TM}/E_{\rm TE}$ and $\sigma^2_{\rm TM}/\sigma^2_{\rm TE}$ predict that the detector will perform better for TE polarization than for TM, they do not explain the other principal feature of these results: for TE polarization, the actual matched-filter performance is significantly degraded relative to idealized results, and the optimal detector yields a significant performance enhancement, while, for TM incidence, the actual matched-filter

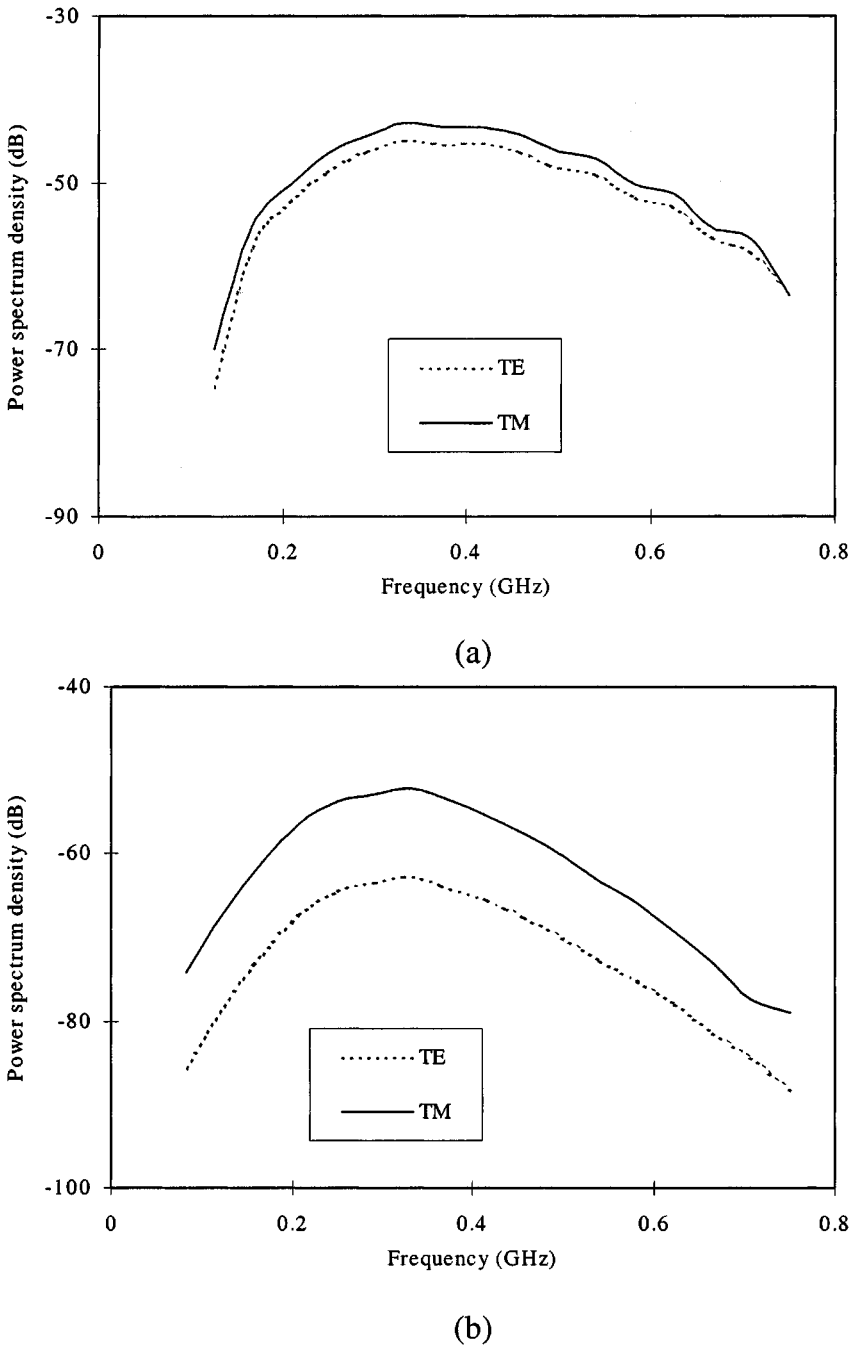


Fig. 5. Power spectral density for backscattered clutter induced by a Gaussian rough surface with mean height zero, variance 3.95 cm, and correlation length 18.75 cm, with results presented for TE and TM polarizations. (a) Incidence angle $\theta_i = 30°$. (b) Incidence angle $\theta_i = 70°$.

performance is close to those of the idealized results, and the optimal detector provides only a slight improvement. To explain this phenomenon, consider Fig. 10, in which we have plotted the relative power transmitted into the soil ($\epsilon_r = 6$ and $\sigma = 0.005\text{S/m}$) when the air–soil interface is flat, as a function of incidence angle, for both TE and TM polarization (i.e., we plot $1 - |\Gamma|^2$, where Γ is the flat-surface reflection coefficient). For incidence near $\theta_i = 70°$, there is very little variation in $1 - |\Gamma|^2$ for TM polarization (due to the Brewster angle, and the corresponding stationary point in $1 - |\Gamma|^2$), while the variation for TE polarization is relatively strong. At the rough surface, the incident wave impinges the interface at angles that vary randomly about θ_i. For the relatively modest roughness considered here, the random variation about θ_i is small. Therefore, from Fig. 10, one would expect that, for $\theta_i = 70°$, the fields transmitted into the soil will be more randomized for TE polarization than for TM. Recall that the optimal detector is designed for problems in which the target signature is random, due to the stochastic nature of the fields transmitted through a random interface. Therefore, for $\theta_i = 70°$ and relatively modest surface roughness, one would expect that a matched filter would be sufficient for TM polarization while the optimal detector would be most beneficial for TE excitation, consistent with the results in Figs. 8 and 9.

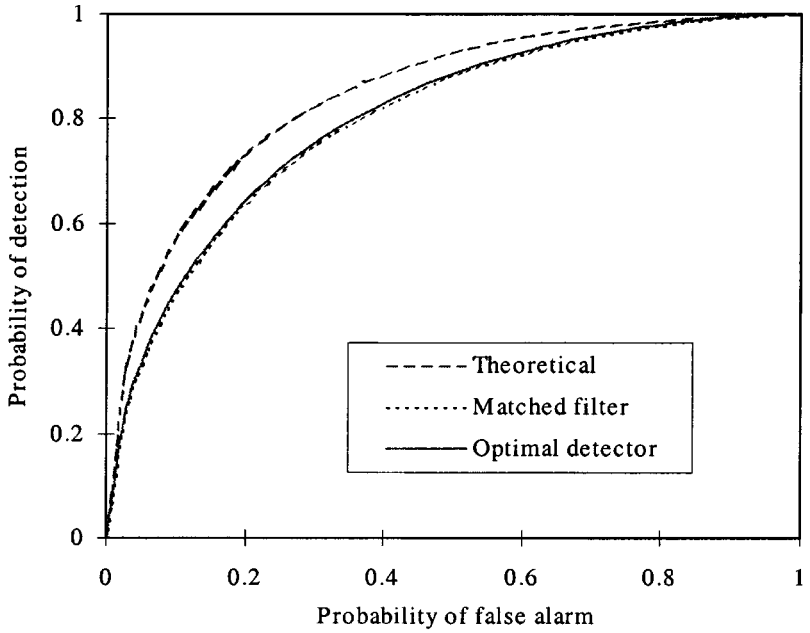


Fig. 6. Receiver operating characteristic for the target in Fig. 4 buried under the random rough surface in Fig. 5, for TE excitation at $\theta_i = 30°$. Results are presented for idealized matched-filter performance (if the target signature were deterministic and known exactly), actual matched-filter performance, and for the optimal detector in Fig. 1.

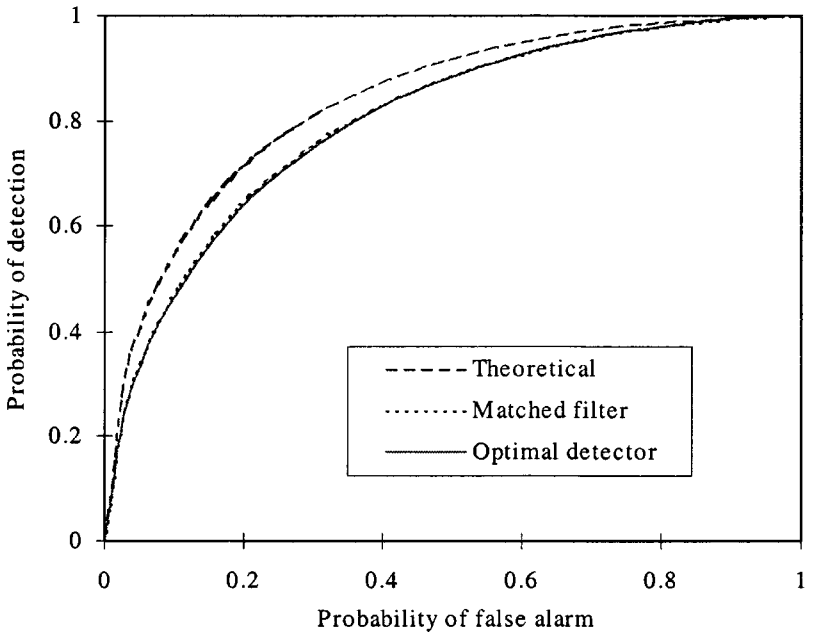


Fig. 7. As in Fig. 6, but for TM excitation.

With this insight, we return to the results for incidence angle $\theta_i = 30°$. From Fig. 10, we see that the variation in $1 - |\Gamma|^2$ is relatively modest in the vicinity of $\theta_i = 30°$, with similar variability seen for TE and TM excitation. Therefore, it is expected that the matched filter will do relatively well for such incidence angles (the transmitted fields only being weakly perturbed due to the random surface) and that modest improvements are expected from the optimal detector, consistent with Figs. 6 and 7.

Summarizing, if the variability of $\boldsymbol{s}(\boldsymbol{\theta}_m)$ is large, the optimal detector will yield significant improvements relative to the matched filter. However, if the rough surface is such that the fields which penetrate the interface are only weakly perturbed relative to the flat-surface response $\boldsymbol{s}$, the optimal detector reduces to a matched filter. To quantify such, we compute the mean and standard deviation of the correlation between $\boldsymbol{s}(\boldsymbol{\theta}_m)$ and $\boldsymbol{s}$ as

$$c_m = \frac{\boldsymbol{s}^T \boldsymbol{s}(\boldsymbol{\theta}_m)}{\sqrt{\boldsymbol{s}^T \boldsymbol{s}} \sqrt{E[\boldsymbol{s}(\boldsymbol{\theta}_m)^T \boldsymbol{s}(\boldsymbol{\theta}_m)]}} \tag{5}$$

the results for which are tabulated in Table I for the examples in Figs. 6–9. From Table I, we see that the relative variation in $c, \sigma_c/m_c$, is small for TM polarization, at both $\theta_i = 30°$ and

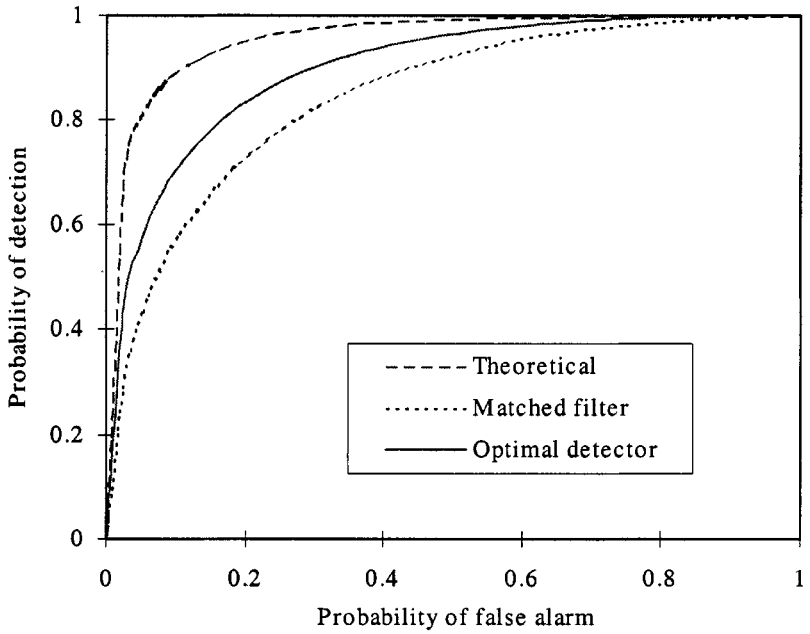


Fig. 8. Receiver operating characteristic for the target in Fig. 4 buried under the random rough surface in Fig. 5, for TE excitation at $\theta_i = 70°$.

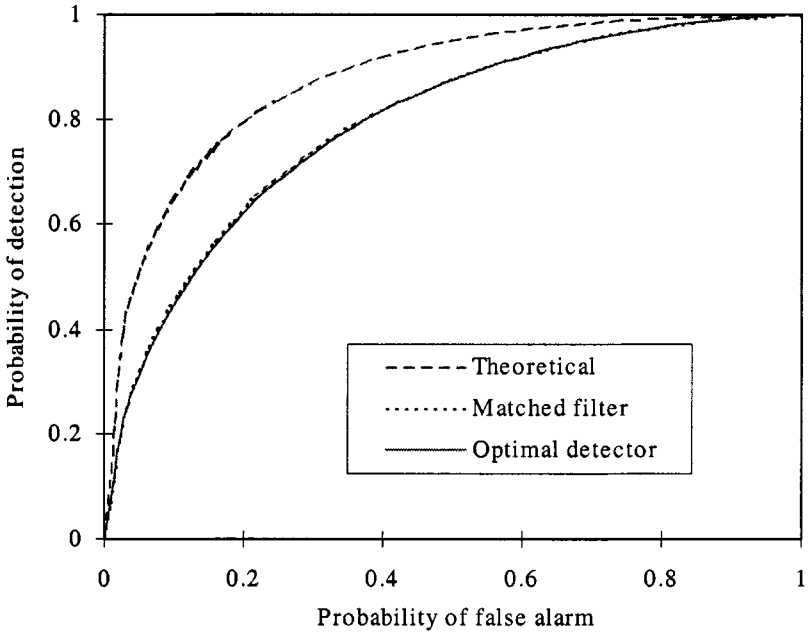


Fig. 9. As in Fig. 8, but for TM excitation.

TABLE I
MEAN AND STANDARD DEVIATION OF THE CORRELATION IN (5), FOR THE EXAMPLES IN FIGS. 6–9

	mean c, m_c	standard dev. of c, σ_c	σ_c/m_c
TE, θ_i=30°	0.6704	0.3703	0.5522
TM, θ_i=30°	0.7678	0.2943	0.3828
TE, θ_i=70°	0.5392	0.3970	0.7379
TM. θ_i=70°	0.6969	0.2593	0.3716

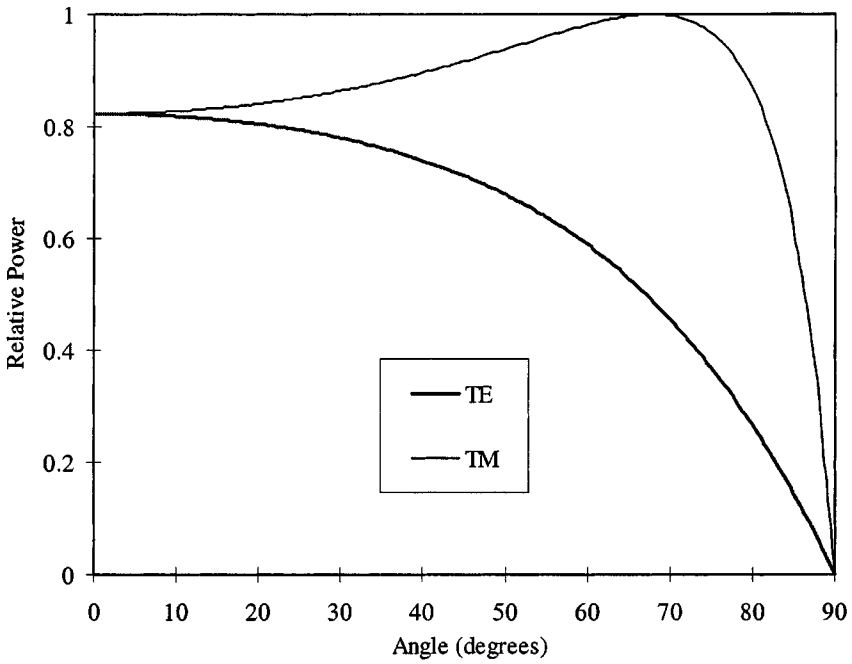


Fig. 10. Relative power transmitted into soil with $\epsilon_r = 6$ and $\sigma = 0.005$, as a function of incidence angle (flat air–ground interface).

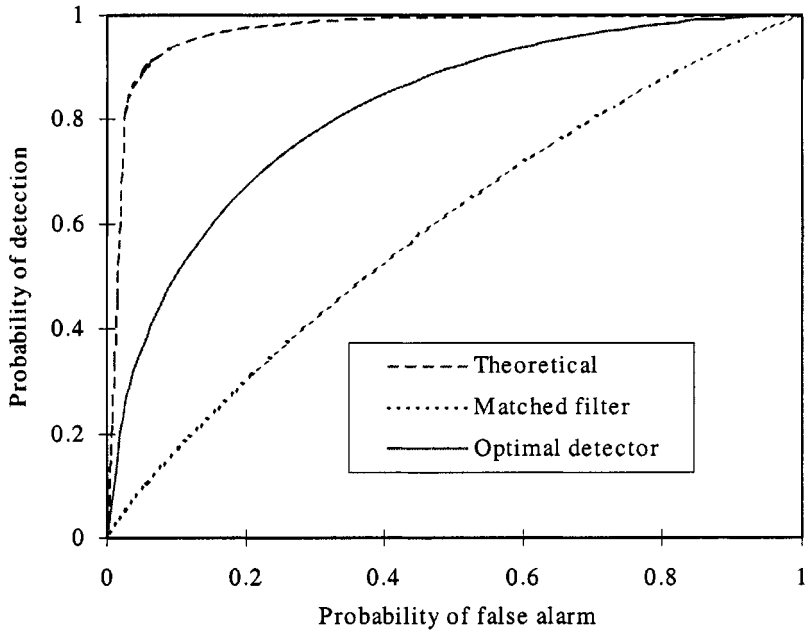


Fig. 11. Receiver operating characteristic as in Fig. 8 (TE polarization and $\theta_i = 70°$), but considering a rough surface with variance 8.84 cm and correlation length 37.5 cm.

$\theta_i = 70°$. The largest such variability occurs for TE excitation at $\theta_i = 70°$, for which we saw the most dramatic performance gain manifested by the optimal detector.

D. Increased Surface Roughness

To demonstrate the effect of the optimal detector for a rougher surface, we consider the same soil and target as above, but now the Gaussian surface has a variance of 8.84 cm and a correlation length of 37.5 cm (we are allowing larger surface variance than before, but have also increased the correlation length, such that the surface is still relatively smoothly varying). We restrict ourselves to $\theta_i = 70°$, to examine if the increased surface variance produces enough randomness in the transmitted fields (and hence the target signature) such that the optimal detector yields gains, even for TM excitation. In Figs. 11 and 12 are plotted ROC curves for TE and TM polarization, respectively, in the same format as before. We see from these Figs. that the optimal detector yields significant performance improvement for TE excitation, as found in Fig. 8. Of more interest, in Fig. 12 we note that the increased level of surface variability has yielded optimal detector performance for TM polarization which is significantly superior to that of the matched filter. This underscores the

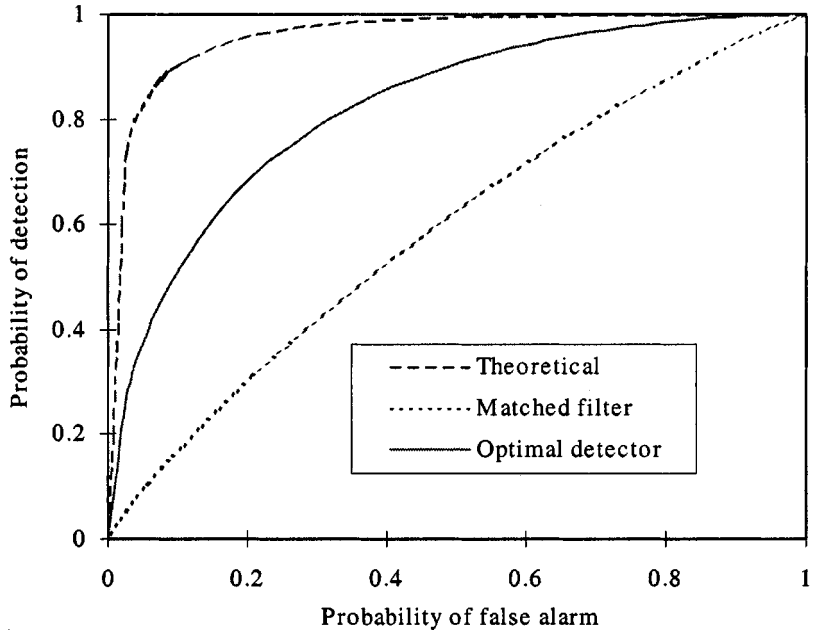


Fig. 12. As in Fig. 11, but for TM polarization.

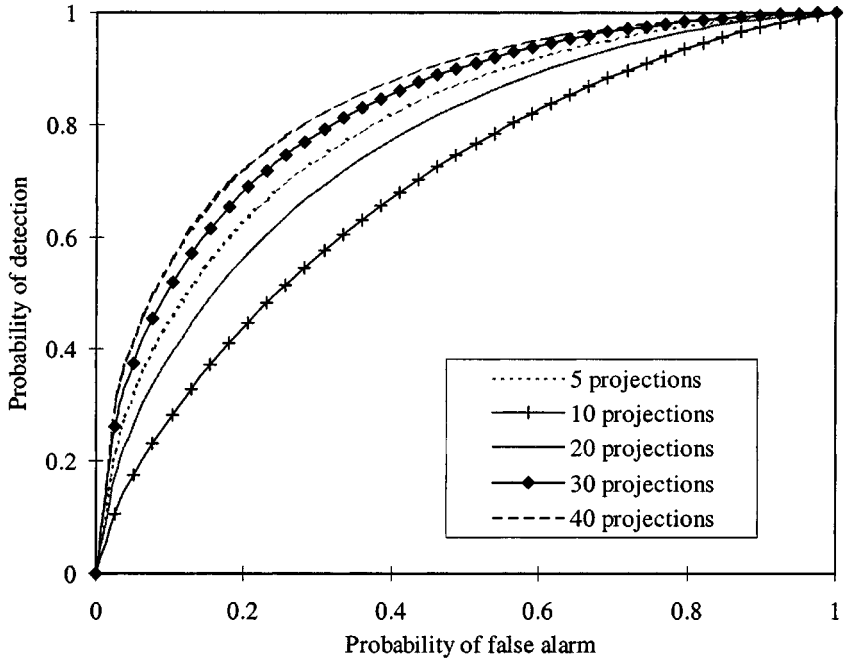


Fig. 13. Convergence of the optimal detector as a function of M [see (4)], for the example considered in Fig. 12.

fact that the performance gain accrued by the optimal detector is more significant as the transmitted fields (and, therefore, the target signature) becomes more random.

Because the surface is more random in these examples, one might expect the value of M required for convergence of the optimal detector will in turn be larger (relative to the results in Figs. 6–9). This was found to be the case; for the results in Figs. 11 and 12, we have found that M must be greater than about 40 to achieve convergence of (4). To examine the convergence of the optimal detector, in Fig. 13, we plot ROC curves as a function of M, considering TM excitation, $\theta_i = 70°$, and the surface roughness in Figs. 11 and 12. For this relatively rough surface, the optimal detector converges for $M > 40$ (recall from Fig. 8 that the less-rough surface only required $M > 20$). We note that one will generally not know *a priori* which value of M is required for a given problem, but the appropriate M can be determined adaptively by simply considering further projections $\boldsymbol{s}(\boldsymbol{\theta}_m)$ until convergence is achieved in (4).

E. General Observations

In the above examples, we have considered one soil type, one target, two angles of incidence, and two distributions for the rough-surface statistics. Therefore, it is difficult to draw general conclusions. However, we have demonstrated that, when

the surface roughness is sufficient to introduce randomization of the transmitted fields, the optimal detector yields improved performance relative to the matched filter. In many examples the performance enhancement was significant. To address the complexity of the optimal detector, note that it requires M projections onto waveforms $\boldsymbol{s}(\boldsymbol{\theta}_m)$ rather than a single projection onto the deterministic waveform $\boldsymbol{s}$, as per a matched filter. However, for a given target and surface roughness, the M waveforms $\boldsymbol{s}(\boldsymbol{\theta}_m)$ are computed once and stored, and can be used to test all $\boldsymbol{e}$ characteristic of said target and clutter. For the problems considered here, M was a relatively small number (<50) and therefore the attendant significant performance gain in several cases appears to justify the associated modest escalation in complexity.

IV. Conclusion

An optimal detector has been presented for the time-domain detection of deterministic targets buried under a randomly rough air–ground interface. The random character of the fields scattered from and penetrating through the random interface requires the parametrization of the clutter and buried-target signature as random processes, generally with different statistics. For the Gaussian surfaces considered here, we have found the clutter characterized as a correlated (nonwhite) Gaussian random process. Therefore, the optimal detector invokes a whitening filter. While in principle the optimal detector requires the statistics of the target signature, such are not easily quantified in general. Therefore, we have implemented the optimal detector approximately via Monte Carlo integration, utilizing M realizations of the random target signature, generated from M realizations of the rough surface.

For the examples considered here, designed to be of interest for buried-target detection, we have found that M should be greater than approximately forty to achieve convergence. However, this number is a function of the degree of randomness in the target signature and was found to increase as the surface became more rough. The enhancement in detector performance yielded by the optimal detector depends as well on the degree to which the target is stochastic. Interestingly, we found that for relatively modest roughness the signature was minimally randomized for vertical polarization near the Brewster angle, at which the transmitted power has a stationary point as a function of incidence angle. However, this phenomenon was vitiated when the roughness became more severe, and the optimal detector then provided dramatic improvements in detector performance.

In the examples considered here, the soil properties, target depth, and target orientation were assumed known, and all randomness was induced by the rough surface. In practice, these parameters will not be known and must be treated statistically. For example, for the detection of buried mines or unexploded ordnance, the target depth will be unknown, but there is likely to be *a priori* knowledge as to its statistical distribution, with similar issues holding for the soil properties and target orientation. Therefore, in the context of the optimal detector in (4), one must perform Monte Carlo integration over these parameters as well. With the large number of random parameters that one may encounter in practice, and the need to consider a sufficient number of waveforms M to span the statistical space of same, it is essential that the forward algorithm used to compute each waveform (from a statistical ensemble) be as fast as possible. This will be an area of future research, constituting the synergy of fast forward algorithms and target detection for random scattering, much as fast forward algorithms have played a critical role in the development of inverse-scattering algorithms for deterministic scattering.

References

[1] M. Saillard and D. Maystre, "Scattering from metallic and dielectric rough surfaces," *J. Opt. Soc. Amer. A.*, vol. 7, pp. 982–990, June 1990.

[2] ——, "Scattering from random rough surfaces: A beam simulation method," *J. Opt.*, vol. 19, pp. 173–176, 1988.

[3] E. Bahar and Y. Zhang, "A new unified full-wave approach to evaluate the scatter cross sections of composite random rough surfaces," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, pp. 973–980, July 1996.

[4] A. Nashahibi, F. T. Ulaby, and K. Sarabandi, "Measurement and modeling of the millimeter-wave backscatter response of soil surface," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, pp. 561–572, Mar. 1996.

[5] R. M. Narayanan, R. Pardipuram, and D. C. Rudquist, "Statistical characteristics of simulated radar imagery from bare soil surfaces: Effects of surface roughness and soil moisture variability," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, pp. 159–168, Jan. 1994.

[6] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. New York: Academic, 1978.

[7] F. D. Hastings, J. B. Schneider, and S. L. Broschat, "A Monte-Carlo FDTD technique for rough surface scattering," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 1183–1191, Nov. 1995.

[8] G. S. Brown, "A new approach to the analysis of rough surface scattering," *IEEE Trans. Antennas Propagat.*, vol. 39, pp. 943–948, July 1991.

[9] T. Dogaru and L. Carin, "Time-domain sensing of a target buried under a rough air–ground interface," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 360–372, Mar. 1998.

[10] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.

[11] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.

[12] M. Moghaddam and W. C. Chew, "Study of some practical issues in inversion with the Born iterative method using time-domain data," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 177–184, Feb. 1993.

[13] C.-C. Chiu and Y.-W. Kiang, "Electromagnetic inverse scattering of a conducting cylinder buried in a lossy half-space," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 1562–1565, Aug. 1992.

[14] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, vol. AP-14, pp. 302–307, 1966.

[15] K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*. Boca Raton, FL: CRC Press, 1993.

[16] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech, 1995.

[17] P. G. Petropoulos, "Stability and phase error analysis of FDTD in dispersive dielectrics," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 62–69, Jan. 1994.

[18] ——, "Phase error control for FDTD methods of second and fourth order accuracy," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 859–862, June 1994.

[19] D. E. Merewether, R. Fisher, and F. W. Smith, "On implementing a numerical Huygens surface in a finite difference program to illuminate scattering bodies," *IEEE Trans. Nuclear Sci.*, vol. NS-27, pp. 1829–1833, Dec. 1980.

[20] T.-T. Hsu and L. Carin, "FDTD analysis of plane-wave diffraction from microwave devices on an infinite dielectric slab," *IEEE Microwave Guided Wave Lett.*, vol. 6, pp. 16–18, Jan. 1996.

[21] G. Mur, "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic-field equations," *IEEE Trans. Electromagn. Compat.*, vol. EMC-23, pp. 377–382, Nov. 1981.

[22] R. L. Higdon, "Absorbing boundary conditions for difference approximations to the multi-dimensional wave equations," *Math. Comput.*, vol. 47, no. 176, pp. 437–459, Oct. 1986.

[23] J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, vol. 114, pp. 185–200, Oct. 1994.

[24] D. Katz, E. Thiele, and A. Taflove, "Validation and extension to three dimensions of the Berenger PML absorbing boundary condition for FD-TD meshes," *IEEE Microwave Guided Wave Lett.*, vol. 4, pp. 268–270, Aug. 1994.
[25] W. C. Chew and W. Weedon, "A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates," *Microwave Opt. Technol. Lett.*, vol. 7, pp. 599–604, Sept. 1994.
[26] J. Fang and Z. Wu, "Generalized perfectly matched layer—An extension of Berenger's perfectly matched layer boundary condition," *IEEE Microwave Guided Wave Lett.*, vol. 5, pp. 451–453, Dec. 1995.
[27] Z. Wu and J. Fang, "Numerical implementation and performance of perfectly matched layer boundary condition for waveguide structures," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 2676–2683, Dec. 1995.
[28] K. Umashankar and A. Taflove, "A novel method to analyze electromagnetic scattering of complex objects," *IEEE Trans. Electromagn. Compat.*, vol. EMC-24, pp. 397–405, Nov. 1982.
[29] A. Taflove and K. Umashankar, "Radar cross section of general three-dimensional scatterers," *IEEE Trans. Electromagn. Compat.*, vol. EMC-25, pp. 433–441, Nov. 1983.
[30] K. Demarest, Z. Huang, and R. Plumb, "An FDTD near-to-far-zone transformation for scatterers buried in stratified grounds," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1150–1157, Aug. 1996.
[31] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.
[32] P. Hubral and M. Tygel, "Analysis of the Rayleigh pulse," *Geophys.*, vol. 54, pp. 654–658, 1989.
[33] M. A. Ressler and J. W. McCorkle, "Evolution of the Army Research Laboratory ultra-wideband test bed," in *Ultra-Wideband Short-Pulse Electromagnetics 2*, L. Carin and L. B. Felsen, Eds. New York: Plenum Press, 1995, pp. 109–123.
[34] S. Vitebskiy, L. Carin, M. Ressler, and F. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Theory and measurement," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, pp. 762–772, May 1997.
[35] J. E. Hipp, "Soil electromagnetic parameters as functions of frequency, soil density, and soil moisture," *Proc. IEEE*, vol. 62, pp. 98–103, Jan. 1974.



Traian Dogaru [redacted] He received the engineering degree from the Polytechnic University of Bucharest in 1990. He received the M.S. degree in electrical engineering and the Ph.D. degree from Duke University, Durham, NC. in 1997 and 1999, respectively.

From 1992 to 1995, he held different engineering positions in the magnetic recording industry. Currently, he is a Research Associate with Duke University. His research interests are in electromagnetic wave theory, computational electromagnetics, rough surface scattering and radar-related signal processing.

Leslie Collins (M'96) [redacted] She received the B.S.E.E. degree from the University of Kentucky, Lexington, the M.S.E.E. and the Ph.D. degree both in electrical engineering, from the University of Michigan, Ann Arbor, in 1985, 1986 and 1995, respectively.

From 1986 to 1990, she was a Senior Engineer with the Westinghouse Research and Development Center, Pittsburgh, PA. In 1995, she joined the Electrical and Computer Engineering Department at Duke University, Durham, NC., as an Assistant Professor. Her current research interests include incorporating physics-based models into statistical signal processing algorithms and she is pursuing applications in subsurface sensing as well as enhancing speech understanding by hearing impaired individuals.

Dr. Collins is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi honor societies.

Lawrence Carin (S'86–M'89–SM'96) [redacted] . He received the B.S., M.S., and Ph.D. degrees in electrical engineering, from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department at Polytechnic University, Brooklyn, NY., as an Assistant Professor, and in 1994, became an Associate Professor. In 1995, he joined the Electrical Engineering Department at Duke University, Durham, NC., where he is an Associate Professor. Currently, he is principal investigator on a Multidisciplinary University Research Initiative (MURI) on demining. He is an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies.

Multiresolution Time-Domain Using CDF Biorthogonal Wavelets

Traian Dogaru, *Member, IEEE,* and Lawrence Carin, *Fellow, IEEE*

***Abstract*—A new approach to the multiresolution time-domain (MRTD) algorithm is presented in this paper by introducing a field expansion in terms of biorthogonal scaling and wavelet functions. Particular focus is placed on the Cohen–Daubechies–Feauveau (CDF) biorthogonal-wavelet class, although the methodology is appropriate for general biorthogonal wavelets. The computational efficiency and numerical dispersion of the MRTD algorithm are addressed, considering several CDF biothogonal wavelets, as well as other wavelet families. The advantages of the biorthogonal MRTD method are presented, with emphasis on numerical issues.**

***Index Terms*—Biorthogonal wavelets, multiresolution analysis, time-domain methods.**

I. Introduction

THE multiresolution time-domain (MRTD) method has been applied successfully to various electromagnetic-field problems, such as microwave cavities and circuits [1], [2], as well as scattering by general targets [4]–[6]. It has been shown that MRTD often yields important computational savings, *vis-à-vis* the traditional finite-difference time-domain (FDTD) method [7]–[9] without sacrificing solution accuracy.

Higher order finite-difference schemes have been devised within the FDTD framework [10], [11]. However, significant issues are associated with their implementation for general scattering problems, requiring hybrid approaches for treatment of interfaces between different media, as well as for absorbing boundary conditions [11]. The MRTD achieves a multigrid structure by introducing denser discretization (scaling functions and wavelets) in zones with relatively fast spatial field variation, while keeping a lower resolution representation (scaling functions alone) in slowly varying regions. Multigrid schemes have been studied in the context of FDTD [12]–[14], but in that paper, there were difficulties in formulating a general stability criterion. By contrast, the multiresolution wavelet expansion employed by MRTD provides a natural multigrid formulation, for which rigorous stability criteria can be established. Another feature of the MRTD algorithm is that it incorporates a sub-cell methodology for treating the boundary between two dielectric media, without having to introduce a staircase approximation. In this context, an analogous formulation could be devised for the FDTD algorithm, in which the dielectric/magnetic properties of the media are averaged inside the Yee cells [9]. Summarizing, for each of the principal issues addressed by MRTD (higher order approximation of the fields, multigrid structure, and accurate treatment of the interface between different media), there exists a more or less *ad hoc* solution within the traditional FDTD framework. However, the MRTD simultaneously addresses all of these issues, within a rigorous and stable construct.

In the MRTD algorithm, the fields are expanded in a wavelet basis [15], and Maxwell's curl equations are discretized using a method-of-moments procedure [16]. In the existing literature on MRTD, the basis functions of choice have been orthonormal wavelets from the Battle–Lemarie wavelet family [1]–[3], the Haar wavelet family [4]–[6], as well as, more recently, the Daubechies scaling functions [17], [18]. The Battle–Lemarie wavelets display good smoothness properties, but they have the disadvantage of infinite (though exponentially decaying) support. This means that many neighboring terms must be included in the update equation at each node, and that there may be numerical issues associated with the truncation of the MRTD coefficient sequence (as explained in Section III). The Haar wavelet family yields a simple algorithm. In fact, it can be shown that an n-level Haar wavelet expansion is equivalent to the Yee FDTD scheme applied to a grid of 2^n higher resolution. Unfortunately, the Haar wavelets lack smoothness. In this paper, we seek a balance between smoothness and reduced spatial support. The former addresses numerical dispersion, and the latter yields algorithmic simplicity while avoiding truncation errors.

The remainder of this paper is organized as follows. The choice of the biorthogonal wavelet family is discussed in Section II. We present the formulation of the biorthogonal MRTD scheme in Section III, using scaling functions alone, with consideration of the stability criterion and dispersion curves. Section IV performs a similar analysis of the biorthogonal MRTD scheme, including a first level of wavelet functions. In Section V, we present a simple numerical example using various versions of the MRTD algorithm. Conclusions are drawn in Section VI.

II. Choosing the Wavelet Family for Field Expansion

The main objective of the MRTD method is a minimization of the computational resources required for a given accuracy of the electromagnetic solution. In this context, we would like to reduce the number of unknowns, by decreasing the number of discretization points per wavelength, while simultaneously keeping the numerical dispersion under control. A second and distinct issue involves reduction of the total number of computations required by the algorithm (i.e., while it is generally desirable to reduce the discretization rate per wavelength, this salu-

Manuscript received September 21, 2000.

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0018-9480(01)03315-4.

0018–9480/01$10.00 © 2001 IEEE

TABLE I
COURANT NUMBER AT THE STABILITY LIMIT FOR THE MRTD ALGORITHM IN ONE DIMENSION

	CDF (2,2)	CDF (2,4)	CDF (2,6)	CDF (3,3)	Cubic spline Battle-Lemarie
Scaling only	0.7500	0.6844	0.6585	0.6844	0.6371
Scaling + One Level Wavelet	0.6046	0.4831	0.4221	0.4815	0.2625

tary goal is mitigated if the associated algorithmic complexity is simultaneously increased).

The first problem can be reformulated as follows: find a wavelet basis that minimizes the numerical dispersion error of the MRTD algorithm for a given discretization rate. This is clearly related to how "good" an approximation of the fields can be obtained by considering a wavelet expansion limited to a certain number of levels. This is a typical problem encountered in data-compression techniques [15]. Let us consider the following general expansion of the function f:

$$f = \sum_{l,m} \langle f, \psi_{l,m} \rangle \tilde{\psi}_{l,m}. \tag{1}$$

Here, $\psi_{l,m}$ represents the wavelet function at level l, shifted by $m/2^l$ units, whereas $\tilde{\psi}_{l,m}$ is the dual of this function. They must satisfy the following orthogonality relationship:

$$\left\langle \psi_{l,m}, \tilde{\psi}_{l,m} \right\rangle = \delta(l-l')\,\delta(m-m'). \tag{2}$$

If $\psi_{l,m} = \tilde{\psi}_{l,m}$, then the wavelets are orthonormal. If the basis functions are different, then the family of wavelets is biorthogonal. Two common requirements used in designing wavelet systems are regularity and vanishing moments [15]. Regularity refers to the degree of differentiability of a function, whereas the nth moment of the wavelet function is defined as $m_1(n) = \int x^n \psi(x)\,dx$. Suppose the function f is smooth and we want to approximate it by truncating the expansion in (1) to a finite number of levels. It can then be shown [15] that, in order to obtain a good approximation, we require that $\psi_{l,m}$ have as many vanishing moments as possible, and $\tilde{\psi}_{l,m}$ be as regular as possible.

An additional requirement for the wavelet system is that the expansion functions should have minimal support. As will become more evident in Sections III and IV, this condition addresses the algorithmic computational complexity. If we restrict ourselves to orthonormal wavelet systems, it can be shown that regularity and minimal support are two conflicting requirements [15]. The Battle–Lemarie family of wavelets, which are derived from B-spline functions [1], [15], have good regularity properties (depending on the order of the spline functions used in design), but they have infinite support. This results, theoretically, in an infinite number of MRTD terms in each update equation. Since the Battle–Lemarie functions display exponential decay, the higher order MRTD coefficients also decay fast. Nevertheless, truncating the sequence of MRTD coefficients [1] to a reasonable number (usually 8–12 on each side) poses problems in terms of arithmetic precision by vitiating the properties of the wavelet functions imposed by design. Reference [19] draws attention to implementation of wavelet systems in finite precision arithmetic, which can result in nonzero moments and eventually nonorthogonal systems.

Compactly supported orthonormal wavelets have been obtained by Daubechies [15] by maximizing the number of vanishing moments. The Cohen–Daubechies–Feauveau (CDF) family of biorthogonal wavelets [15] uses spline functions as dual wavelets $\tilde{\psi}_{l,m}$ and maximizes the vanishing moments of $\psi_{l,m}$ for a given extent of their support. The main difference, compared to Battle–Lemarie wavelets, is that the biorthogonal wavelets yield compact support. This means that the sequence of MRTD coefficients required in the update equations is rigorously finite (no truncation is needed).

In conclusion, the biorthogonal wavelet systems strike a good balance between regularity and reduced support. While the Battle–Lemarie wavelets are symmetric and smooth, they must be truncated to achieve finite support.

III. BIORTHOGONAL MRTD USING SCALING FUNCTIONS

A. MRTD Formulation

The derivation of the MRTD equations for a biorthogonal-wavelet field expansion is similar to the derivation for the orthogonal case [1], [6]. To keep the presentation simple, we consider one-dimensional (1-D) wave propagation with components E_z and H_y propagating in the x-direction. The full wavelet expansion of the component E_z can be written as

$$E_z(x,t) = \sum_{k,m=-\infty}^{\infty} \left[E^{\Phi}_{k,m} \tilde{\Phi}_m(x) + \sum_{l=0}^{\infty} E^{\Psi^l}_{k,m}\, \tilde{\Psi}^l_m(x) \right] h_k(t). \tag{3}$$

Here, we denote by $\tilde{\Phi}_m$ the dual scaling function shifted by m units, and by $\tilde{\Psi}^l_m$ the lth level dual-wavelet function displaced by $m/2^l$ units. For time discretization, we use rectangular pulses $h_k(t)$, where k represents the shift in time units. A similar equation holds for H_y, only the supports of the scaling/wavelet functions are displaced half a unit relative to E_z. In this section, we consider expansion in terms of scaling functions alone as follows:

$$E_z(x,t) = \sum_{k,m=-\infty}^{\infty} E^{\Phi}_{k,m} \tilde{\Phi}_m(x) h_k(t). \tag{4}$$

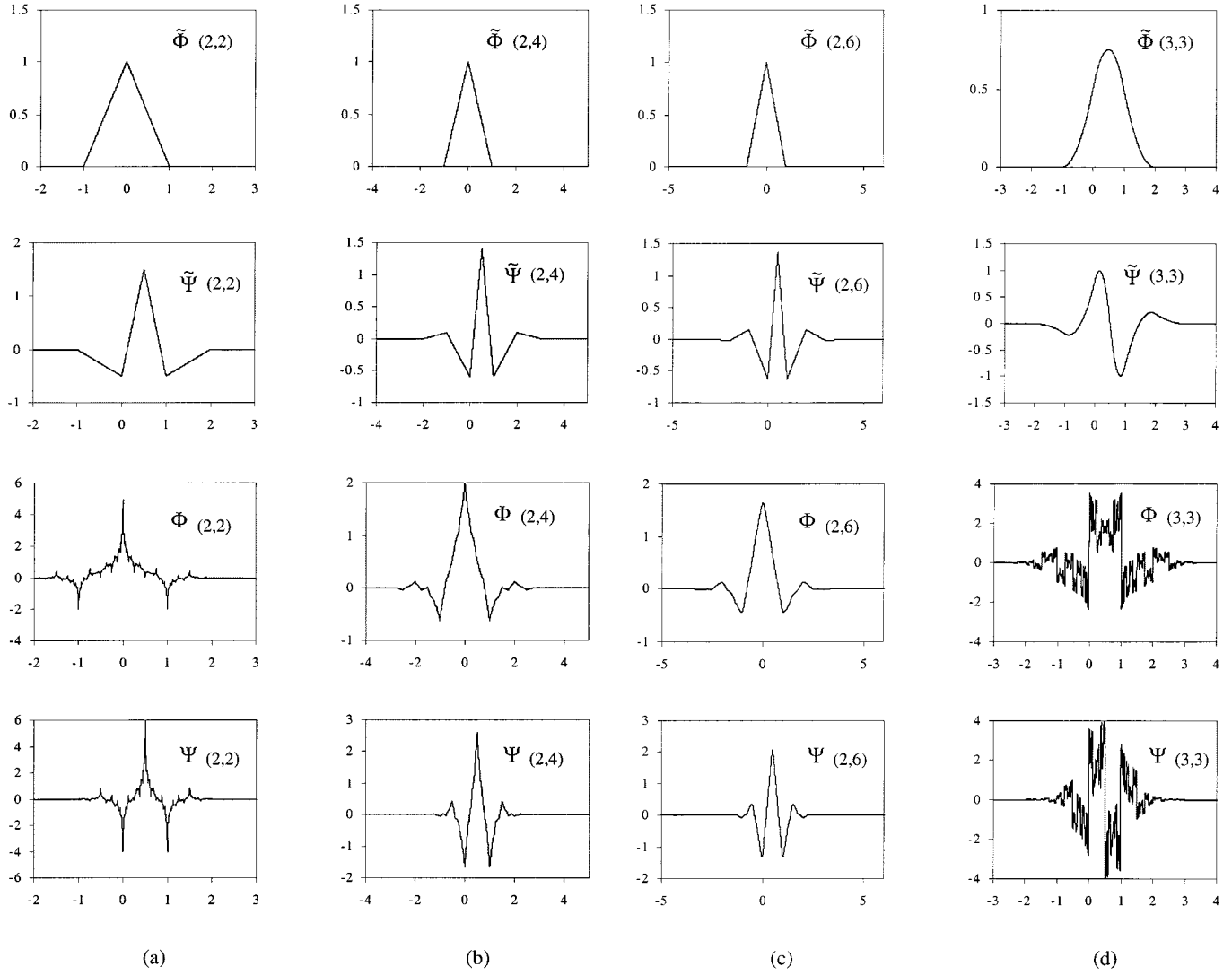


Fig. 1. Scaling and wavelet functions and their duals for several families of CDF biorthogonal wavelet systems. In all cases, the graphs are approximations obtained by the cascade algorithm after eight iterations. (a) CDF $(2, 2)$. (b) CDF $(2, 4)$.

Fig. 1. *(Continued)*. Scaling and wavelet functions and their duals for several families of CDF biorthogonal wavelet systems. In all cases, the graphs are approximations obtained by the cascade algorithm after eight iterations. (c) CDF $(2, 6)$. (d) CDF $(3, 3)$.

The method-of-moments procedure [16] is affected by testing Maxwell's equations with the scaling functions Φ_m. The resulting equations are

$$H^{\Phi}_{k,\,m} = H^{\Phi}_{k-1,\,m} + \frac{\Delta t}{\mu \Delta x} \sum_{i=1}^{n_a} a(i)\left(E^{\Phi}_{k,\,m+i} - E^{\Phi}_{k,\,m-i+1}\right) \tag{5a}$$

$$E^{\Phi}_{k+1,\,m} = E^{\Phi}_{k,\,m} + \frac{\Delta t}{\varepsilon \Delta x} \sum_{i=1}^{n_a} a(i)\left(H^{\Phi}_{k,\,m+i-1} - H^{\Phi}_{k,\,m-i}\right) \tag{5b}$$

where $a(i) = \int \left((\partial \tilde{\Phi}_{m+i}(x))/(\partial x)\right) \Phi_{m+1/2}(x)\, dx$; Δx and Δt are the spatial and temporal steps, respectively. The number n_a, also called the "stencil size," equals half the number of the nonzero coefficients in the MRTD scheme. Note that we differentiate with respect to the dual function and, therefore, smoothness (differentiability) is imposed on such. The inner products are formed with the nondual functions, on which we, therefore, impose moment suppression.

The stability criterion for this algorithm was established in [3] and can be written as

$$\frac{c\Delta t}{\Delta x} \leq \frac{1}{\sum_{i=1}^{n_a} |a(i)|} \tag{6}$$

where c is the phase velocity. The maximum values of $(c\Delta t)/(\Delta x)$ required by a stable algorithm (in one dimension) are given in Table I for different choices of the scaling functions. In two or three dimensions, these values must be adjusted by factors of $1/\sqrt{2}$ and $1/\sqrt{3}$, respectively [9].

B. Numerical Dispersion

In analyzing the numerical dispersion of the MRTD schemes, we consider a discretized monochromatic plane wave propagating in accord with the MRTD equations and study the phase

TABLE II
MRTD COEFFICIENTS FOR VARIOUS WAVELET BASES. (a) COEFFICIENTS $a(i)$. (b) COEFFICIENTS $b(i)$

i	CDF (2,2)	CDF (2,4)	CDF (2,6)	CDF (3,3)	Cubic spline Battle-Lemarie coefficients truncated at 10^{-3}
1	1.2291667	1.2918134	1.3110317	1.2918129	1.2918462
2	-0.0937500	-0.1371348	-0.1560124	-0.1371343	-0.1560761
3	0.0104167	0.0287617	0.0419962	0.0287618	0.0596391
4	0	-0.0034701	-0.0086543	-0.0034701	-0.0293099
5	0	0.0000080	0.0008308	0.0000080	0.0153716
6	0	0	0.0000109	0	-0.0081892
7	0	0	-0.0000000	0	0.0043788
8	0	0	0	0	-0.0023433
9	0	0	0	0	0.0012542

(a)

i	CDF (2,2)	CDF (2,4)	CDF (2,6)	CDF (3,3)	Cubic spline Battle-Lemarie coefficients truncated at 10^{-3}
1	1.4375000	1.6890090	1.8610039	1.6890090	2.4725388
2	0.0937500	0.1955818	0.2903309	0.1955818	0.9562282
3	-0.0104167	-0.0287622	-0.0426845	-0.0287618	0.1660587
4	0	0.0034701	0.0086545	0.0034701	0.0939244
5	0	-0.0000080	-0.0008308	-0.0000080	0.0031413
6	0	0	-0.0000109	0	0.0134936
7	0	0	0.0000000	0	-0.0028589
8	0	0	0	0	0.0027788
9	0	0	0	0	-0.0011295

(b)

error as compared with the continuous case. Following the procedure in [3], we obtain the following equation (for the 1-D case):

$$\frac{1}{q}\sin\frac{\pi q}{n_l} = \sum_{i=1}^{n_a} a(i)\sin\left[\frac{\pi u}{n_l}(2i-1)\right] \tag{7}$$

where $q = (c\Delta t)/(\Delta x)$ (the Courant number), $u = (\lambda_{\text{continuous}})/(\lambda_{\text{discrete}})$ (the ratio between the theoretical and numerical wavelength), and n_l are the number of sampling points per wavelength. Ideally, we should have $u = 1$. We measure the phase error (in degrees per wavelength) as $360(u-1)$, where u represents the solution of the nonlinear equation (7) for a given discretization rate n_l.

TABLE II *(Continued).*
MRTD COEFFICIENTS FOR VARIOUS WAVELET BASES (c) COEFFICIENTS $c(i)$. (d) COEFFICIENTS $d(i)$

i	CDF (2,2)	CDF (2,4)	CDF (2,6)	CDF (3,3)	Cubic spline Battle-Lemarie coefficients truncated at 10^{-3}
1	-0.1510417	-0.1569410	-0.1390804	0.2422510	-0.0465973
2	0.0833333	0.1363198	0.1482517	-0.3855128	0.0545394
3	-0.0052083	-0.0443966	-0.0742472	0.1597666	-0.0369996
4	0	0.0047841	0.0192157	-0.0178193	0.0205745
5	0	-0.0003301	-0.0026345	0.0013174	-0.0111530
6	0	0.0000008	0.0002887	-0.0000030	0.0059769
7	0	0	-0.0000146	0	-0.0032026
8	0	0	-0.0000002	0	0.0017141
9	0	0	0.0000000	0	-

(c)

i	CDF (2,2)	CDF (2,4)	CDF (2,6)	CDF (3,3)	Cubic spline Battle-Lemarie coefficients truncated at 10^{-3}
1	-0.0416667	-0.0726900	-0.0954514	-0.0181721	-0.0465973
2	0.0208333	0.0364721	0.0468862	0.0272902	0.0545394
3	0	-0.0000856	0.0005564	-0.0091396	-0.0369996
4	0	0	-0.0000002	0.0000214	0.0205745
5	0	0	0	0	-0.0111530
6	0	0	0	0	0.0059769
7	0	0	0	0	-0.0032026
8	0	0	0	0	0.0017141
9	0	0	0	0	-

(d)

In the following, we consider a few families of CDF biorthogonal wavelet functions and study the dispersion errors associated with their use in the MRTD algorithm. The results are also compared to those obtained with cubic-spline Battle–Lemarie wavelets and with the traditional Yee FDTD algorithm. For the CDF wavelet families, we use the notations CDF (α, β) [15], where the integers α and β are related to the lengths of the reconstruction and decomposition filters of the family [15], respectively (they are also related to the support extent of the scaling function Φ_m and the dual scaling function $\tilde{\Phi}_m$, respectively). Plots of the CDF families considered in this paper (including scaling and first-level wavelet functions) are given in Fig. 1. The scaling/wavelet functions in Fig. 1 were computed using the cascade algorithm [15]. The coefficients $a(i)$ were computed by

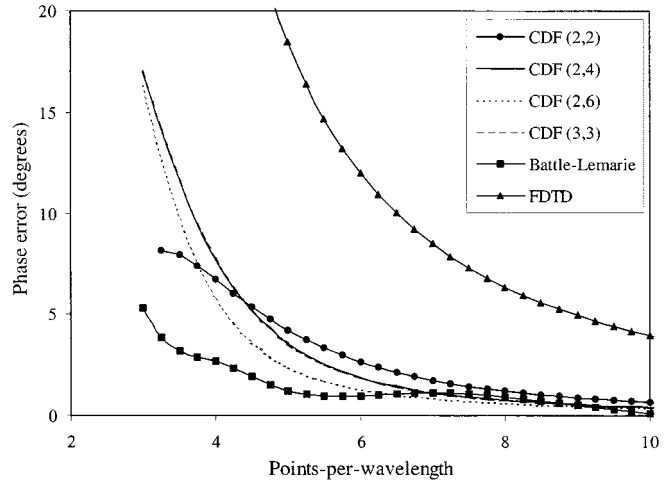


Fig. 2. Phase error (in degrees per wavelength) versus discretization rate for MRTD schemes using expansion of the fields in terms of scaling function only (at low resolution). For the MRTD schemes, the Courant number is $q = 0.25$, for FDTD, the Courant number is $q = 0.6$. 1-D propagation. The curves for CDF $(2, 4)$ and CDF $(3, 3)$ are on top of one another.

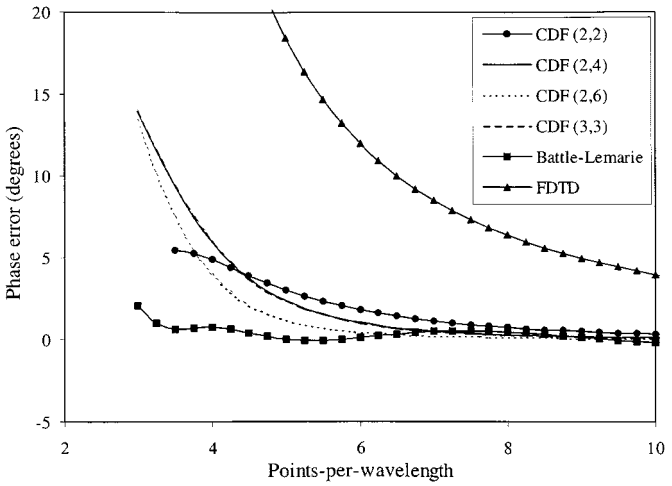


Fig. 3. Phase error (in degrees per wavelength) versus discretization rate for MRTD schemes using expansion of the fields in terms of scaling function only (at low resolution). For the MRTD schemes, the Courant number is $q = 0.1$, for FDTD, the Courant number is $q = 0.6$. 1-D propagation. The curves for CDF $(2, 4)$ and CDF $(3, 3)$ are on top of one another.

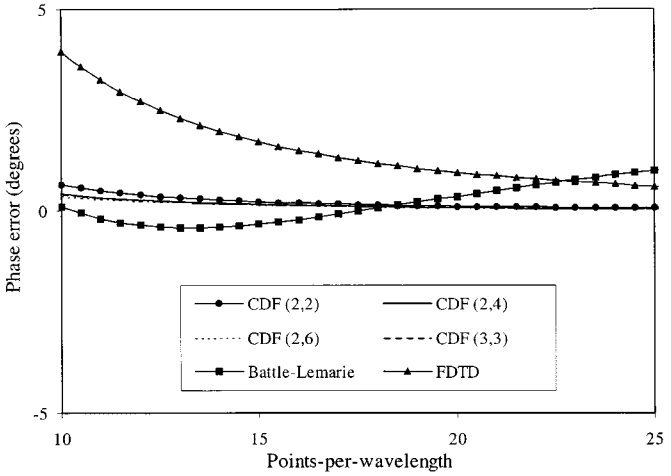


Fig. 4. Phase error (in degrees per wavelength) versus discretization rate for MRTD schemes using expansion of the fields in terms of scaling function only (at high resolution). For the MRTD schemes, the Courant number is $q = 0.25$, for FDTD, the Courant number is $q = 0.6$. 1-D propagation. The curves for CDF $(2, 4)$ and CDF $(3, 3)$ are on top of one another.

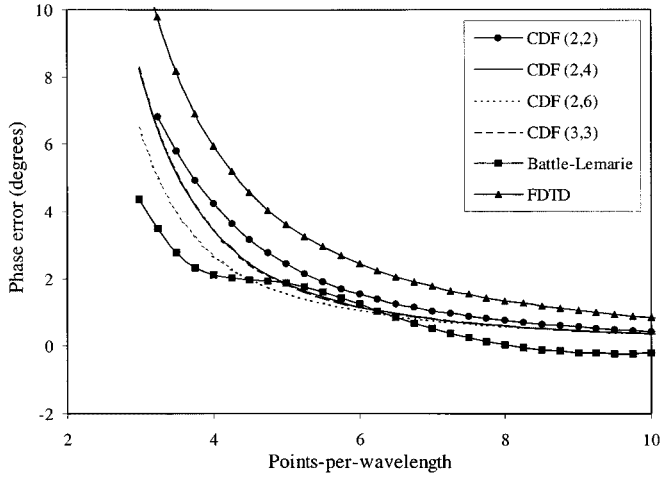


Fig. 5. Phase error (in degrees per wavelength) versus discretization rate for MRTD schemes using expansion of the fields in terms of scaling function only (at low resolution). For the MRTD schemes, the Courant number is $q = 0.25$, for FDTD, the Courant number is $q = 0.6$. 2-D propagation at 45°. The curves for CDF $(2, 4)$ and CDF $(3, 3)$ are on top of one another.

numerical differentiation/integration and are given in Table II. Notice that, in some cases, the functions Φ_m can be very rough (we do not impose any smoothness condition on them, only moment suppression) and, therefore, they must be discretized very finely in order to obtain sufficient precision in the numerical integration (up to 20 levels of the cascade algorithm were used for six to seven decimal places). On the other hand, we notice that the dual scaling/wavelet functions are piecewise smooth and, therefore, the numerical differentiation is not a problem (actually it can be done in closed form for spline functions).

The dispersion curves for scaling-function field expansion are given in Figs. 2–5, in which the phase error is depicted as a function of the number of discretization points per wavelength. As a control, we plot the dispersion curves for the FDTD algorithm on the same graphs. It is important to mention that the errors are sensitive to the Courant number q. For all MRTD schemes considered, we noticed that the phase error decreases when q is decreased (consistent with the results reported in [3]). However, for the FDTD algorithm, the opposite is true. In the following examples, we keep $q = 0.6$ for all the FDTD dispersion curves. More details on the influence of the Courant number on the FDTD dispersion curves can be found in [3]. In Fig. 2, we considered $q = 0.25$ for the MRTD schemes (since the stability limits are comparable for all MRTD schemes using scaling function expansions, we can directly compare them at the same Courant number). In Fig. 3, we lower the time step such that $q = 0.1$ for the same schemes, and in Fig. 4, we plot the same curves as in Fig. 2, but display curves for higher values of n_l (number of sample points per wavelength). The most striking aspect of the dispersion curves for the Battle–Lemarie system is that the error reaches a minimum and then increases as n_l is increased (we used a stencil size 9 for this MRTD scheme). This suggests that the error should increase at higher discretization

rates and that we would never reach the exact solution, no matter how finely we discretize the fields. However, we attribute this anomaly to the truncation of the MRTD coefficient sequence $a(i)$. Actually, it is easy to see that, in order to converge to the exact solution as $\Delta x \to 0$, we require

$$\sum_{i=1}^{n_a} a(i)(2i-1) = 1. \tag{8}$$

The finite precision arithmetic used in computing the coefficients $a(i)$ leads to an error in satisfying this relationship, and it can be shown that this error is directly related to the phase error in the limit $\Delta x \to 0$. If we use a larger stencil size for the Battle–Lemarie MRTD scheme, this error can be reduced and the shape of the curves looks more similar to the ones obtained for the other MRTD schemes [3]. In fact, we noticed that the phase error in the case of the MRTD algorithms using biorthogonal wavelet expansion also does not converge exactly to zero in the limit $\Delta x \to 0$ (due to finite numerical precision), however, the errors are much smaller than in the Battle–Lemarie case. For comparison, the error associated with the coefficient computation is on the order of 10^{-6}, whereas the error associated with the coefficient sequence truncation (for the Battle–Lemarie scheme with stencil size 9) is on the order of 10^{-3}.

The behavior of the dispersion curves in the limit $\Delta x \to 0$ has theoretical value rather than a practical one since we are mainly interested in operating the MRTD algorithm at a discretization rate as low as possible. If we accept the phase error obtained for the FDTD scheme at ten points-per-wavelength, then we can use as few as three to five points-per-wavelength in the MRTD algorithm, depending on the choice of the basis and on the Courant number. The best performance appears to be provided by the cubic-spline Battle–Lemarie functions, at least at this level of discretization. However, these curves were obtained considering 1-D wave propagation. If we consider the two-dimensional (2-D) case, the dispersion error depends on the angle of propagation. In Fig. 5, we look at the phase errors for propagation at 45° for the same parameters as in Fig. 1. One can notice that now, at about five points per wavelength, the Battle–Lemarie scheme does not provide better accuracy than the biorthogonal MRTD schemes considered. Another interesting fact about these curves is the reduction of the error for the FDTD scheme at 45°, when a Courant number of $q = 0.6$ is employed. In fact, it is well known that, for a Courant number close to the stability limit ($q = 1/\sqrt{2}$ in two dimensions), the dispersion error of the Yee algorithm vanishes for propagation at 45° [9]. This discussion suggests that one set of curves at one angle and one Courant number do not completely characterize the error obtained in a complex scattering configuration. In the case of the Battle–Lemarie MRTD scheme, we expect spurious errors caused by the truncation of the coefficient sequence. On the other hand, the performance of the biorthogonal MRTD schemes should be more predictable since the precision errors involved in arithmetic computations are much smaller. Finally, we note that all the above dispersion analyses were performed at Courant numbers that guarantee stability in 2-D configurations for all MRTD and FDTD schemes (in fact, the 1-D cases correspond to 2-D propagation along one of the Cartesian axes).

C. Computational Complexity

Another issue to be addressed is the number of floating-point calculations involved in each update equation. If the stencil size of the MRTD algorithm is n_a, using only scaling functions, then in order to achieve a similar number of computations with the Yee FDTD scheme for the same problem, we must reduce the number of cells in one direction by a factor of n_a in one dimension, $\sqrt{n_a}$ in two dimensions, and $\sqrt[3]{n_a}$ in three dimensions. It is easy to see that the savings of the MRTD scheme become more important as the number of spatial dimensions is increased. In the following, we make the comparisons with the Yee scheme employing ten points per wavelength. Since the Battle–Lemarie scheme needs at least $n_a = 9$, for this wavelet choice, the MRTD algorithm becomes more efficient than FDTD (in terms of computational time) only in a three-dimensional (3-D) configuration. Increasing the stencil size for the Battle–Lemarie scheme does not reduce the dispersion error significantly at low discretization rates [3], but it does increase the computation time. We also tried to decrease the stencil size of the Battle–Lemarie scheme to four, but the results in terms of numerical dispersion were very poor. The most economical MRTD scheme from this point-of-view appears to be the CDF $(2, 2)$ basis, which has a stencil size of $n_a = 3$ and offers phase accuracy comparable with the other MRTD schemes. It can also be noticed that increasing the support of the scaling and wavelet functions in the CDF families does not lead to significant improvements in terms of phase error, but could increase the computation time significantly. The size of the time step is also important in evaluating the total CPU time. For all MRTD schemes, the accuracy is better as the time step is decreased. However, this increases the computational time necessary to obtain the numerical solution. Therefore, a tradeoff must be achieved between the accuracy of the scheme and total number of time steps (notice that, in the FDTD case, this tradeoff is not necessary because choosing the Courant number close to the stability limit is nearly optimal).

IV. Biorthogonal MRTD Using the First Level of Wavelet Functions

In this section, we perform a similar error analysis of the MRTD algorithm, considering an expansion of the fields in terms of scaling functions and a single level of wavelet functions. The formulation depends on whether the wavelet functions (and their duals) are symmetric or antisymmetric about one-half (see Fig. 1). If the wavelets are symmetric (e.g., CDF $(2, 2)$, $(2, 4)$, $(2, 6)$, and cubic-spline Battle–Lemarie), then, at cell m (in one dimension), the component H^{Ψ} occupies the same position as the component E^{Φ}, whereas E^{Ψ} occupies the same position as H^{Φ} (we call the "position" of a field component, the coordinate where the scaling or wavelet function associated with it peaks out). If the wavelets are antisymmetric [e.g., CDF $(3, 3)$], then, at cell m, the component E^{Ψ} aligns with the component E^{Φ}, whereas H^{Ψ} aligns with H^{Φ}. The

MRTD update equations in the symmetric case are (for 1-D propagation)

$$H^{\Phi}_{k,m} = H^{\Phi}_{k-1,m} + \frac{\Delta t}{\mu \Delta x} \cdot \left(\sum_{i=1}^{n_a} a(i) \left(E^{\Phi}_{k,m+i} - E^{\Phi}_{k,m-i+1} \right) + \sum_{i=1}^{n_c} c(i) \left(E^{\Psi}_{k,m+i} - E^{\Psi}_{k,m-i} \right) \right) \tag{9a}$$

$$H^{\Psi}_{k,m} = H^{\Psi}_{k-1,m} + \frac{\Delta t}{\mu \Delta x} \cdot \left(\sum_{i=1}^{n_d} d(i) \left(E^{\Phi}_{k,m+i} - E^{\Phi}_{k,m-i} \right) + \sum_{i=1}^{n_b} b(i) \left(E^{\Psi}_{k,m+i-1} - E^{\Psi}_{k,m-i} \right) \right) \tag{9b}$$

$$E^{\Phi}_{k+1,m} = E^{\Phi}_{k,m} + \frac{\Delta t}{\varepsilon \Delta x} \cdot \left(\sum_{i=1}^{n_a} a(i) \left(H^{\Phi}_{k,m+i-1} - H^{\Phi}_{k,m-i} \right) + \sum_{i=1}^{n_c} c(i) \left(H^{\Psi}_{k,m+i} - H^{\Psi}_{k,m-i} \right) \right) \tag{9c}$$

$$E^{\Psi}_{k+1,m} = E^{\Psi}_{k,m} + \frac{\Delta t}{\varepsilon \Delta x} \cdot \left(\sum_{i=1}^{n_d} d(i) \left(H^{\Phi}_{k,m+i} - H^{\Phi}_{k,m-i} \right) + \sum_{i=1}^{n_b} b(i) \left(H^{\Psi}_{k,m+i} - H^{\Psi}_{k,m-i+1} \right) \right) \tag{9d}$$

where $b(i) = \int ((\partial \tilde{\Psi}_{m+i}(x))/(\partial x)) \Psi_{m+1/2}(x)\, dx$, $c(i) = \int ((\partial \tilde{\Psi}_{m+i}(x))/(\partial x)) \Phi_{m+1/2}(x)\, dx$ (we have $c(0) = 0$) and $d(i) = \int ((\partial \tilde{\Phi}_{m+i}(x))/(\partial x)) \Psi_{m-1/2}(x)\, dx$ (we have $d(0) = 0$). Notice that we have dropped the index zero indicating the wavelet level. For the biorthogonal MRTD schemes, the $c(i)$ and $d(i)$ coefficient have different values, unlike the Battle-Lemarie scheme [1], which uses orthonormal basis functions. For the antisymmetric case, the corresponding equations are

$$H^{\Phi}_{k,m} = H^{\Phi}_{k-1,m} + \frac{\Delta t}{\mu \Delta x} \cdot \left(\sum_{i=1}^{n_a} a(i) \left(E^{\Phi}_{k,m+i} - E^{\Phi}_{k,m-i+1} \right) + \sum_{i=1}^{n_c} c(i) \left(E^{\Psi}_{k,m+i} + E^{\Psi}_{k,m-i+1} \right) \right) \tag{10a}$$

$$H^{\Psi}_{k,m} = H^{\Psi}_{k-1,m} + \frac{\Delta t}{\mu \Delta x} \cdot \left(\sum_{i=1}^{n_d} d(i) \left(E^{\Phi}_{k,m+i} + E^{\Phi}_{k,m-i+1} \right) + \sum_{i=1}^{n_b} b(i) \left(E^{\Psi}_{k,m+i} - E^{\Psi}_{k,m-i+1} \right) \right) \tag{10b}$$

$$E^{\Phi}_{k+1,m} = E^{\Phi}_{k,m} + \frac{\Delta t}{\varepsilon \Delta x} \cdot \left(\sum_{i=1}^{n_a} a(i) \left(H^{\Phi}_{k,m+i-1} - H^{\Phi}_{k,m-i} \right) + \sum_{i=1}^{n_c} c(i) \left(H^{\Psi}_{k,m+i-1} + H^{\Psi}_{k,m-i} \right) \right) \tag{10c}$$

$$E^{\Psi}_{k+1,m} = E^{\Psi}_{k,m} + \frac{\Delta t}{\varepsilon \Delta x} \cdot \left(\sum_{i=1}^{n_d} d(i) \left(H^{\Phi}_{k,m+i-1} + H^{\Phi}_{k,m-i} \right) + \sum_{i=1}^{n_b} b(i) \left(H^{\Psi}_{k,m+i-1} - H^{\Psi}_{k,m-i} \right) \right) \tag{10d}$$

where the coefficients $b(i)$, $c(i)$, and $d(i)$ have the same expressions as above. Their values are listed in Table II for all the wavelet families considered above.

The stability criterion can be written as (11), shown at the bottom of this page [3], in which $n = \max(n_a, n_b)$. The maximum values of $(c\Delta t)/(\Delta x)$ required by a stable algorithm (in one dimension) are given in Table I.

To analyze the dispersion properties of the MRTD scheme using one level of wavelets, we plot similar phase error graphs

$$\frac{c\Delta t}{\Delta x} \leq \frac{\sqrt{2}}{\sqrt{\left(\sum_{i=1}^{n_a} |a(i)| \right)^2 + \left(\sum_{i=1}^{n_b} |b(i)| \right)^2 + 2 \sum_{i=1}^{n_c} |c(i)| \sum_{i=1}^{n_d} |d(i)| + \left(\sum_{i=1}^{n} |a(i) + b(i)| \right) \sqrt{\left(\sum_{i=1}^{n} |a(i) - b(i)| \right)^2 + 4 \sum_{i=1}^{n_c} |c(i)| \sum_{i=1}^{n_d} |d(i)|}}} \tag{11}$$

versus the discretization rate, as discussed in the previous section. The dispersion relationship for the numerical MRTD solution of Maxwell's equations can be written as

$$1-\left(s_a^2+s_b^2+s_cs_d\right)+\left(s_as_b-s_cs_d\right)^2=0 \tag{12}$$

where

$$s_a=q\frac{\sum_{i=1}^{n_a}a(i)\sin\left[\frac{\pi u}{n_l}(2i-1)\right]}{\sin\frac{\pi q}{n_l}} \tag{13a}$$

$$s_b=q\frac{\sum_{i=1}^{n_b}b(i)\sin\left[\frac{\pi u}{n_l}(2i-1)\right]}{\sin\frac{\pi q}{n_l}} \tag{13b}$$

$$s_c=q\frac{\sum_{i=1}^{n_c}c(i)\begin{matrix}\sin\\ \cos\end{matrix}\left(\frac{\pi u}{n_l}2i\right)}{\sin\frac{\pi q}{n_l}} \tag{13c}$$

$$s_d=q\frac{\sum_{i=1}^{n_d}d(i)\begin{matrix}\sin\\ \cos\end{matrix}\left(\frac{\pi u}{n_l}2i\right)}{\sin\frac{\pi q}{n_l}}. \tag{13d}$$

In the numerators of s_c and s_d, we take sine for symmetric wavelet systems and cosine for antisymmetric wavelet systems. The graphs in Fig. 6 were obtained for $q=0.25$ (which insures that all schemes are stable) and the same families of wavelets as in the previous sections. In theory, by introducing a level of wavelets, the resolution of the MRTD schemes should double. However, the curves in Fig. 6 show a modest reduction in the phase error for the Battle-Lemarie MRTD scheme (stencil size 9) when wavelets are considered, as compared to the case where the expansion was provided by scaling functions only. One explanation for this result can be again provided by the truncation of the MRTD coefficient sequence. However, we conjecture that the principal reason for the small difference between the two Battle–Lemarie MRTD schemes (one using scaling functions only and the other using scaling plus first-level wavelet functions), when they both use the same Courant number, is the fact that, in the second case, we operate much closer to the stability limit (see Table I). The curves in Fig. 6 indicate, against expectations, that the performance decreases as the order of the wavelet system increases. However, we can explain this by taking into account that the numerical dispersion error is also a function of the Courant number. One can notice that, e.g., in the Battle–Lemarie scheme, $q=0.25$ is very close to the stability limit, whereas for CDF $(2,2)$, this value is well below the stability limit of the scheme. As mentioned in Section III, the errors decrease with q. In Fig. 7, we consider the same dispersion curves, but operate all MRTD schemes at Courant numbers equal to 95% of their stability limits. Now the dispersion performances of the schemes are as expected (i.e., the higher order schemes offer better performance). However, the comparison we make in Fig. 6 is meaningful because, in this case, all MRTD schemes have the same time step. The graphs in this figure clearly show that, for the same time step, the biorthogonal schemes have lower numerical dispersion than the Battle–Lemarie scheme.

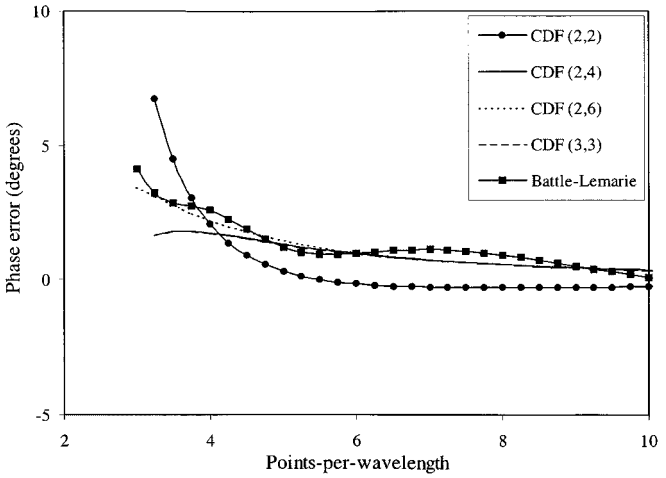


Fig. 6. Phase error (in degrees per wavelength) versus discretization rate for MRTD schemes using expansion of the fields in terms of scaling and one level wavelet functions. For all schemes, the Courant number is $q=0.25$. 1-D propagation. The curves for CDF $(2,4)$ and CDF $(3,3)$ are on top of one another.

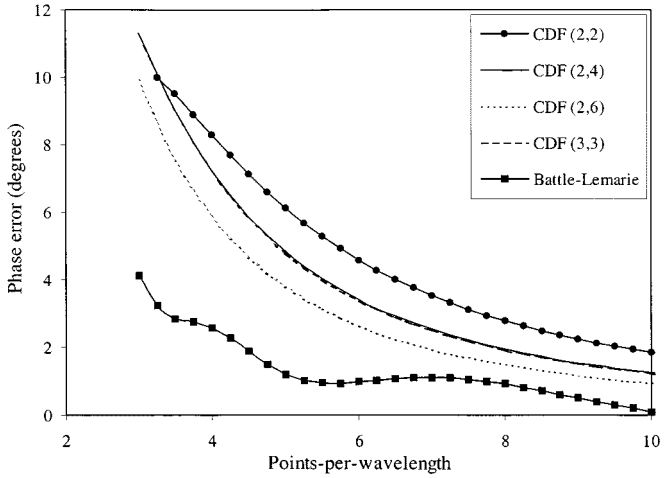


Fig. 7. Phase error (in degrees per wavelength) versus discretization rate for MRTD schemes using expansion of the fields in terms of scaling and one level wavelet functions. For all schemes, the Courant number is 95% of the stability limit. 1-D propagation. The curves for CDF $(2,4)$ and CDF $(3,3)$ are on top of one another.

The stencil sizes for the coefficient sequences $a(i)$, $b(i)$, $c(i)$, and $d(i)$ are given in Table III. Once again, the biorthogonal MRTD schemes with low-order support, e.g., CDF $(2,2)$ and CDF $(2,4)$, are the most efficient in terms of the number of computations (while providing better dispersion performance than the Battle–Lemarie basis for a similar Courant number). In this case, where one scaling and one level wavelet functions are employed, the number of calculations in the update equation for one field component is about $n_a+n_b+n_c+n_d$ times larger than the ones required by FDTD. Again, the savings of the MRTD algorithm increase with the number of spatial dimensions. We should also mention that the differences in the Courant number for the schemes using scaling plus one level wavelet functions

TABLE III
STENCIL SIZES FOR DIFFERENT MRTD SCHEMES

	CDF (2,2)	CDF (2,4)	CDF (2,6)	CDF (3,3)	Cubic spline Battle-Lemarie coefficients truncated at 10^{-3}
n_a	3	5	7	5	9
n_b	3	5	7	5	9
n_c	3	6	9	6	8
n_d	2	3	4	4	8

TABLE IV
RESONANT FREQUENCIES FOR AN AIR-FILLED METALLIC CAVITY (ALL FREQUENCIES IN MEGAHERTZ)

Analytic	S-MRTD				W-MRTD				FDTD
	CDF (2,2)	CDF (2,4)	CDF (2,6)	Battle-Lemarie	CDF (2,2)	CDF (2,4)	CDF (2,6)	Battle-Lemarie	
150.00	150.11	150.10	150.11	150.22	150.01	150.01	150.01	150.14	149.60
167.71	167.87	167.85	167.86	168.14	167.72	167.71	167.73	168.01	167.43
270.42	271.11	271.05	271.03	270.79	270.50	270.51	270.51	270.25	269.63
300.00	301.12	300.92	300.88	300.66	300.05	300.15	300.15	299.92	296.81

are significant, as one compares, e.g., the CDF $(2, 2)$ scheme with the Battle–Lemarie scheme (see Table I). This means that, in the former algorithm, we can choose a time step that is more than double the one used by the latter scheme, and still maintain stability. However, increasing the time step adversely affects the accuracy of the algorithm (compare Figs. 6 and 7). Finally, we note that, when using scaling functions and wavelets (here, a single wavelet level), the MRTD also has the added benefit of multiresolution. In this context, scaling functions alone can be used in regions in which the fields vary relatively slowly, with wavelets added locally in regions of fast field variation.

V. NUMERICAL EXAMPLE

To address implementation of the MRTD with CDF biorthogonal wavelets, we consider a simple electromagnetic problem, i.e., studying the resonant frequencies of an air-filled metallic cavity, in a manner similar to the example in [1]. We considered a 2-D cavity of 1 m $\times$ 2 m (with perfectly conducting walls) discretized with a grid step of $\Delta x = 0.1$ m for all versions of FDTD/MRTD models. We performed simulations using the MRTD algorithm based on scaling functions alone (S-MRTD with the notations in [1]) and the same algorithm based on scaling and first level wavelet functions (W-MRTD). In the latter case, we employed the full expansion in both directions, uniformly throughout the whole computational domain. The Battle–Lemarie MRTD scheme uses a stencil size 9 (as given in Table III). We also implemented the same problem with the Yee FDTD algorithm. The Courant numbers were $q = 0.42$ for S-MRTD, $q = 0.17$ for W-MRTD (same for all wavelet bases), and $q = 0.6$ for FDTD. We considered four resonant frequencies, which were obtained either in TE or TM polarizations. We chose a relatively high discretization rate in order to have a number of cells larger than the stencil size of any scheme in each direction. In this manner, only one image of the cavity is required at each metallic wall in order to enforce the boundary conditions. The numerical results are given in Table IV.

Based on these results, we can conclude that all MRTD schemes provide good accuracy in measuring the resonant frequencies of the cavity, and they all perform better than FDTD. The differences between the various schemes are small—some model the higher frequencies more accurately, whereas others do better on the lower frequencies. As expected, adding one level of wavelets resulted in improved precision. However, one should be aware that the Courant number for W-MRTD is different from that used by the S-MRTD (the time steps are different). Also, based on the discussion in the previous section, the numerical results confirm that, for the same Courant number, the CDF $(2, 2)$ W-MRTD scheme provides the most accurate results. We emphasize that the point of this

numerical experiment is simply to underscore that the MRTD is implemented easily, using either CDF biorthogonal wavelets or the orthogonal Battle-Lemarie wavelets. In particular, in these computations, we have not focused on the reduced spatial discretization afforded by the various MRTD schemes, this issue having been addressed in the dispersion studies.

VI. Conclusions

In this paper, we have formulated the MRTD algorithm using biorthogonal wavelet families. We chose the CDF biorthogonal systems because they fit well with the requirements for an efficient MRTD scheme: maximum number of vanishing moments for the wavelet function (for a given support), good regularity for the dual-wavelet function, and compact support. We analyzed the dispersion properties of the MRTD scheme for a few families of CDF biorthogonal wavelets and compared them with the cubic-spline Battle–Lemarie scheme. Although, in theory, the latter should perform better in terms of numerical dispersion, the error associated with truncating the MRTD coefficient sequence vitiates the accuracy of the scheme, when finite arithmetic precision is considered. The effect is more evident when we consider field expansion in terms of scaling and first-level wavelet functions. The biorthogonal wavelet systems do not present this problem, due to their compact support, although high precision is required in the MRTD coefficient calculation in order to minimize arithmetic inaccuracies. Also, the compact support of the biorthogonal wavelets reduces the total CPU time relative to the Battle-Lemarie wavelets for similar levels of phase error.

References

[1] M. Krumpholz and L. Katehi, "MRTD: New time-domain schemes based on multiresolution analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 555–571, Apr. 1996.

[2] E. M. Tentzeris, R. L. Robertson, and L. Katehi, "PML absorbing boundary conditions for the characterization of open microwave circuit components using multiresolution time domain techniques (MRTD)," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 1709–1716, Nov. 1999.

[3] E. M. Tentzeris, R. L. Robertson, J. Harvey, and L. Katehi, "Stability and dispersion analysis of Battle–Lemarie based MRTD schemes," *IEEE Trans. Microwave Theory Tech.*, vol. 47, pp. 1004–1013, July 1999.

[4] T. Dogaru and L. Carin, "Application of multiresolution time-domain schemes to two-dimensional electromagnetic scattering problems," *IEEE Trans. Antennas Propagat.*, submitted for publication.

[5] ——, "Multiresolution time-domain analysis of scattering from a rough dielectric surface," *Radio Sci.*, vol. 35, pp. 1279–1292, Nov.–Dec. 2000.

[6] T. Dogaru, "Modeling and signal processing for electromagnetic subsurface sensing," Ph.D. dissertation, Dept. Elect. Comput. Eng., Duke Univ., Durham, NC, 1999.

[7] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, vol. AP-14, pp. 302–307, May 1966.

[8] K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*. Boca Raton, FL: CRC Press, 1993.

[9] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 1995.

[10] P. G. Petropoulos, "Phase error control for FDTD methods of second and fourth order accuracy," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 859–862, June 1994.

[11] M. F. Hadi and M. Piket-May, "A modified FDTD (2,4) scheme for modeling electrically large structures with high-phase accuracy," *IEEE Trans. Antenna Propagat.*, vol. 45, pp. 254–264, Feb. 1997.

[12] M. W. Chevalier, R. J. Luebbers, and V. P. Cable, "FDTD local grid with material traverse," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 411–421, Mar. 1997.

[13] M. Okoniewski, E. Okoniewska, and M. A. Stuchly, "Three-dimensional subgridding algorithm for FDTD," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 422–429, Mar. 1997.

[14] M. J. White, M. F. Iskander, and Z. Huang, "Development of a multigrid FDTD Code for three-dimensional applications," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 1512–1517, Oct. 1997.

[15] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[16] R. F. Harrington, *Field Computation by Moment Methods*. Melbourne, FL: Krieger, 1968.

[17] Y. W. Cheong, Y. M. Lee, K. H. Ra, J. G. Kang, and C. C. Shin, "Wavelet–Galerkin scheme of time-dependent inhomogeneous electromagnetic problems," *IEEE Microwave Guided Wave Lett.*, vol. 9, pp. 297–299, Aug. 1999.

[18] M. Fujii and W. J. R. Hoefer, "Dispersion of time domain wavelet Galerkin method based on Daubechies' compactly supported scaling functions with three and four vanishing moments," *IEEE Microwave Guided Wave Lett.*, vol. 10, pp. 125–127, Apr. 2000.

[19] J. Gotze, J. E. Odegard, P. Rieder, and C. S. Burrus, "Approximate moments and regularity of efficiently implemented orthogonal wavelet transforms," in *Proc. IEEE Int. Circuits Syst. Symp. Dig.*, May 1996, pp. 405–408.

Traian Dogaru (S'96–M'99) [redacted] He received the Engineering degree from the Polytechnic University of Bucharest, Bucharest, Romania, in 1990, the M.S. degree in electrical engineering and the Ph.D. degree from Duke University, Durham, NC, in 1997 and 1999, respectively.

From 1992 to 1995, he has held different engineering positions in the magnetic recording industry. He is currently a Research Associate at Duke University, where his main research interests are in electromagnetic wave theory, computational electromagnetics, rough surface scattering and radar-related signal processing.

Lawrence Carin (S'85–M'85–SM'96–F'01) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland at College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor and, in 1994, became an Associate Professor. In September 1995, he joined the Electrical Engineering Department, Duke University, Durham, NC, where he is currently a Professor. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is a member of Tau Beta Pi and Eta Kappa Nu. He is the principal investigator of a Multidisciplinary University Research Initiative (MURI) on demining. He is currently an associate editor for the IEEE Transactions on Antennas and Propagation.

On the Wideband EMI Response of a Rotationally Symmetric Permeable and Conducting Target

Lawrence Carin, *Fellow, IEEE*, Haitao Yu, Yacine Dalichaouch, Alexander R. Perry, *Senior Member, IEEE*, Peter V. Czipott, and Carl E. Baum, *Fellow, IEEE*

***Abstract*—A simple and accurate model is presented for computation of the electromagnetic induction (EMI) resonant frequencies of canonical conducting and ferrous targets, in particular, finite-length cylinders and rings. The imaginary resonant frequencies correspond to the well known exponential decay constants of interest for time-domain EMI interaction with conducting and ferrous targets. The results of the simple model are compared to data computed numerically, via method-of-moments (MoM) and finite-element models. Moreover, the simple model is used to fit measured wideband EMI data from ferrous cylindrical targets (in terms of a small number of parameters). It is also demonstrated that the general model for the magnetic-dipole magnetization, in terms of a frequency-dependent diagonal dyadic, is applicable to general rotationally symmetric targets (not just cylinders and rings).**

***Index Terms*—Induction, magnetization, resonance, sensing.**

I. Introduction

ELECTROMAGNETIC induction (EMI) is widely used for the detection and discrimination of conducting and ferrous targets. In the context of subsurface sensing, EMI is used to sense buried metal mines as well as unexploded ordnance (UXO) [1], [2]. Since the conductivity of such targets is typically many orders of magnitude larger than the conductivity of soil, the target can usually be analyzed in free space. We consider such an analysis here. Most land mines and UXO exhibit rotational symmetry, with this symmetry yielding a simplification of the EMI phenomenology. In particular, as discussed further in Section II, the magnetization tensor of such targets becomes diagonal (with regard to a coordinate system with an axis along the target axis of symmetry). Moreover, we demonstrate that each element of the diagonal corresponds to a magnetic dipole moment, oriented either parallel to or orthogonal to the target axis.

The wideband frequency-domain EMI response of a conducting and ferrous target is characterized by EMI resonant frequencies, which exist along the imaginary axis of the complex-frequency plane [3]. Note that here we use the ω plane for complex frequencies, instead of the $s = \Omega + j\omega$ plane as in the two-sided Laplace transform [4]. Assuming an $\exp(j\omega t)$ time dependence, a purely imaginary resonant frequency corresponds to the exponential decay characteristic of time-domain EMI interaction with conducting and ferrous targets. As is discussed in Section II, the imaginary resonant frequencies corresponding to magnetic-dipole modes play an important role in characterizing the frequency and time-domain properties of the magnetization tensor.

The magnetization tensor and EMI resonances have been discussed previously [4]. The principal contribution of this paper is a recognition that the EMI resonant modes of high-conductivity targets, characterized by purely imaginary resonant frequencies, are analogous to the resonant modes of high-permittivity, low-loss microwave resonators [5], which resonate at nearly purely real resonant frequencies. Once this duality is understood, all of the many techniques developed in the microwave community for characterization of high-permittivity dielectric resonators [6]–[8] (operating at nearly real frequencies) can be directly transitioned to the EMI high-conductivity, high-permeability problem (with resonant frequencies at imaginary frequencies). Moreover, this understanding yields an explicit interpretation of the EMI resonant modes in terms of dipole and higher order magnetic moments, from which the physical significance of particular EMI modes can be accrued.

Many metal land mines and UXO are composed of parts that resemble finite-length cylinders and rings and therefore, the EMI resonant modes of such targets are of interest. Such targets, as well as more-general shapes, can be analyzed via numerical algorithms, such as the method of moments (MoM) [3]. Unfortunately, such calculations are very expensive computationally, since the size of the MoM basis functions typically must be small relative to the skin depth [3]. This requirement significantly limits the size of the targets that can be considered, and results in significant computer run times, and computer memory requirements. However, for simple but important shapes, such as finite-length cylinders and rings, one can make use of the duality between EMI resonances and microwave resonances to yield very simple but highly accurate approximate models. In particular, as detailed in Section II, we have employed these simple models to calculate the (imaginary) resonant frequencies of conducting cylinders and rings, with the computed data compared favorably to data simulated numerically. In those examples we restrict ourselves to relatively small targets, for which the numerical analysis is tractable. The approximate model is applicable to cylinders and

Manuscript received June 18, 2000; revised October 27, 2000. This work was supported in part by Strategic Environmental Research and Development Program Contract DACA72-99-C-0012-CU-1123, a U.S. Army Research Office Demining MURI, and by the Air Force Office of Scientific Research.

L. Carin and H. Yu are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Y. Dalichaouch, A. R. Perry, and P. V. Czipott are with Quantum Magnetics, San Diego, CA 92192 USA.

C. E. Baum is with the Air Force Research Laboratory, Directed Energy Directorate, Kirtland Air Force Base, Albuquerque, NM 87108 USA.

Publisher Item Identifier S 0196-2892(01)04840-9.

0196–2892/01$10.00 © 2001 IEEE

rings of arbitrary size. A long conducting and ferrous cylinder, for example, provides a good representation of a UXO, and therefore we can use our simple model to predict the imaginary resonant frequencies (time-domain decay constants) of such targets, as a function of material parameters (conductivity and permeability).

The remainder of the paper is organized as follows. In Section II, we explain the duality between conducting and ferrous targets at EMI frequencies and high-permittivity resonators at microwave frequencies. We also elucidate how to transition approximate models from the microwaves community, for calculation of the EMI resonant frequencies of finite-length cylinders and rings. In Section III, we compare the resonant frequencies computed via the approximate model with those computed by rigorous numerical algorithms. A demonstration of how the model is used to characterize measured data is discussed in Section IV, in which we investigate the measured wideband EMI response from ferrous cylinders. Conclusions from this work are discussed in Section V.

II. Fundamental Theory

A. Resonant EMI Model

Consider an arbitrarily shaped highly (but not perfectly) conducting and possibly permeable target, residing in free space. At induction frequencies [1], [2], the displacement current can be neglected and the wavenumber k inside the target satisfies $k^2 = -j\omega\mu_o\mu_r\sigma$, where μ_o is the free-space permeability, μ_r is the relative permeability, σ is the conductivity, and ω is the angular frequency. The source-free fields inside the target satisfy

$$\nabla^2\mathbf{\Phi} - j\omega\mu_o\mu_r\sigma\mathbf{\Phi} = 0 \tag{1}$$

where $\mathbf{\Phi}$ denotes either the electric or magnetic fields. This source-free wave equation can be solved for the resonant modes of the target, from which the target resonant frequencies and fields can be computed. A similar analysis is commonly performed for microwave low-loss dielectric resonators, for which k^2 is nearly purely real (resonators with a high quality factor [5]), as is the resonant frequency. For the EMI frequencies of interest here, we note that k^2 is positive real when the frequency ω is positive imaginary. For the $\exp(j\omega t)$ time dependence assumed, this implies that the EMI resonant modes decay exponentially with time. This is consistent with the well known late-time characteristics of transient EMI excitation of conducting and permeable targets [1].

The computation of the real resonant frequencies of microwave dielectric resonators is of significant interest for filter design, and therefore there is a vast literature on this subject [5]–[8]. Moreover, van Bladel [5] has developed an elegant theory for the modes of microwave resonators of very high permittivity, with this directly transferable to the EMI resonances of highly (but not perfectly) conducting and permeable targets. Of more importance from a practical standpoint, simple algorithms have been developed for calculation of the resonant modes of high-permittivity rotationally symmetric dielectric resonators [6]–[8], with that previous work transitioned here for characterization of the EMI modes of rotationally symmetric

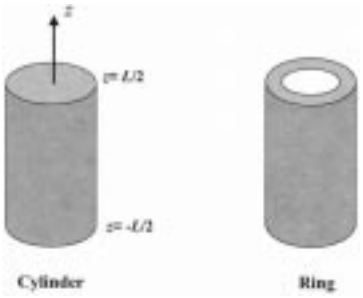


Fig. 1. Cylinder and ring targets for which the approximate model in Section II-B is developed, each highly conducting and possibly ferrous.

highly conducting and permeable targets. While the restriction to rotationally symmetric targets is clearly a limitation, such are of interest for many conducting and permeable land mines and unexploded ordnance (UXO) [1], [2].

B. Finite-Length Cylinder or Ring Target

We develop below a simple model for the EMI response of general conducting and permeable rotationally symmetric targets. However, we initially limit the discussion to the case of a finite-length cylinder or ring (Fig. 1), for which the model parameters can be computed explicitly. To analyze the resonant characteristics (with purely imaginary resonant frequencies) of finite-length cylinders or rings, we first consider such targets of infinite length, followed by imposition of appropriate boundary conditions at $z = \pm L/2$ (see Fig. 1). As discussed above, for purely imaginary frequencies, the infinite-length conducting and permeable cylinder or ring behaves analogous to an infinite-length low-loss dielectric waveguide (when the low-loss waveguide operates at real frequencies). If the target axis is oriented along the z axis of a cylindrical coordinate system, the rotational symmetry of the targets dictates a $\cos(m\phi)$ azimuthal variation for the waveguide modes, where m is an integer greater than or equal to zero. The lowest-order modes have no azimuthal variation, corresponding to $m = 0$. The $m = 0$ modes reduce to two classes [9]. Those with electric fields transverse to the z direction (TE) and those with magnetic fields transverse to z (TM). The resonant modes corresponding to the TM modes can be found by replacing the boundary of the target by a magnetic wall [5], [6] and letting β^{TM} represent the longitudinal wavenumber for a TM mode (infinite waveguide). The nth corresponding resonant mode, for a cylinder or ring target of length L, resonates at the imaginary frequency for which the wavenumber β^{TM} satisfies

$$\beta^{\mathrm{TM}}L = 2\pi n \qquad n = 1, 2, 3, \ldots \tag{2}$$

For the TE modes, the simple magnetic-wall approximation at the ends of the target is not sufficiently accurate, and therefore we utilize a higher-order approximation [6]. In particular, we note in Fig. 2 that the cylindrical or ring target has a symmetry plane at $z = 0$, with this represented by a short-circuit or open-circuit plane, depending on the mode of interest. The mode that resonates at the lowest frequency corresponds to a short circuit [6], so we consider that case. Let β^{TE} denote the longitudinal wavenumber for the TE modes supported by the infinite-length cylinder or ring. The impedance seen from $z = L/2$, looking

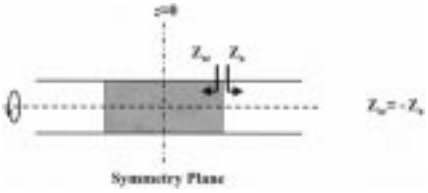


Fig. 2. Schematic of the resonance condition applied for computation of the resonant frequencies of the TE modes supported by the finite-length cylinders and rings considered.

toward the short-circuit symmetry plane is defined as Z_{sc}, while the impedance seen at $z = L/2$ looking toward the air region is denoted Z_a, where (using magnetic walls)

$$Z_{sc}(\omega = j\omega_i) = -\frac{\omega_i \mu_o \mu_r}{\beta^{\mathrm{TE}}} \tan(\beta^{\mathrm{TE}} L/2)$$

$$Z_a(\omega = j\omega_i) = -\omega_i \mu_o / \sqrt{k_c^2 + \omega_i^2 \mu_o \varepsilon_o} \tag{3}$$

and k_c is the cutoff wavenumber for the TE mode under consideration. The nth purely imaginary resonant frequency occurs under the condition $Z_{sc} = -Z_a$ [6], and there are an infinite number of solutions to this transcendental equation (corresponding to the infinite number of such modes with short-circuit symmetry).

A multipole expansion of the fields associated with the above TE and TM modes yields important insight into the fields radiated by these source-free modes. It is straightforward to demonstrate that the lowest-order moment associated with the $m = 0$ (no azimuthal variation) TE resonant modes corresponds to a magnetic dipole oriented along the z direction. By contrast, the lowest-order moment associated with the $m = 0$ TM resonant modes corresponds to a magnetic quadrupole (equivalently, an electric dipole [5]). Consequently, to a sensor that measures the magnetic field, the $m = 0$ TM resonant modes contribute negligibly *vis-a-vis* the TE modes.

To account for the resonant modes with a dipole magnetic moment oriented orthogonal to the z axis, we must consider the $m = 1$ waveguide modes, with the associated physical significance discussed subsequently. Unlike the $m = 0$ waveguide modes, for infinite cylindrical or ring waveguide, the $m = 1$ modes are not simply TE or TM, but are hybrid modes corresponding to a superposition of these two simpler mode types [9]. Consequently, to determine the resonant frequencies of the $m = 1$ modes, we first numerically compute the longitudinal wavenumber for the corresponding infinite-length waveguide, denoted β^{EM} for the hybrid modes [9]. The imaginary resonant frequencies are then computed using a relationship analogous to (2). For the hybrid modes one cannot easily define a modal impedance, and therefore we cannot employ a more-sophisticated resonant condition, as in (3). Nevertheless, as demonstrated below, the accuracy of the resonant frequencies computed via this simple model are in very good agreement with data computed via a rigorous numerical model. Finally, through consideration of the $m = 1$ modal fields, it is easily shown that the lowest-order multipole term corresponds to a magnetic dipole oriented perpendicular to the axis of rotation.

Summarizing, we have demonstrated that the lowest-order resonant modes of a finite-length highly conducting and permeable cylindrical or ring target correspond to magnetic dipoles oriented along and perpendicular to the target axis. Each of these resonant modes resonates with a purely imaginary resonant frequency. The modes with magnetic dipole oriented along the target axis correspond to the $m = 0$ TE modes, while the magnetic moments oriented orthogonal to the axis correspond to the $m = 1$ hybrid modes. The explicit imaginary resonant frequencies can be computed easily using the simple expressions reviewed above, these borrowed from the microwave community [5]–[8]. In Fig. 3, we schematically plot the currents characteristic of the two mode classes. Note that the dipole moment perpendicular to the target axis corresponds to a current density for which $\boldsymbol{J}(\phi) = -\boldsymbol{J}(\phi + \pi)$, explaining the $\cos(\phi)$ variation.

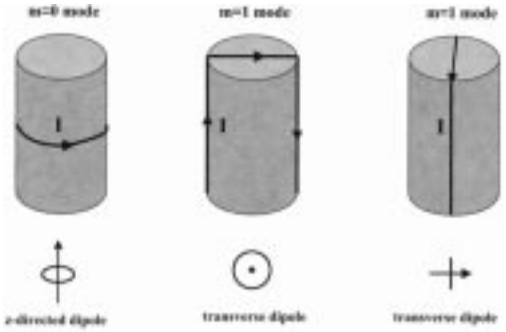


Fig. 3. Schematic of the currents associated with the magnetic dipoles EMI-excited on a conducting and possibly ferrous target. The dipole oriented along the z axis corresponds to the $m = 0$ TE modes, and the transverse magnetic dipole moments correspond to the $m = 1$ hybrid modes.

C. Target Transfer Function

The resonant frequencies corresponding to each EMI resonant mode are manifested as poles in the complex frequency plane [3], [4], where as discussed above all EMI poles reside on the imaginary frequency axis. Recalling that the $m = 0$ TE modes correspond to dipoles oriented in the z direction, and the $m = 1$ hybrid modes correspond to dipoles perpendicular to z, the frequency-dependent dyadic dipole magnetization is expressed as

$$\boldsymbol{M}(\omega) = \boldsymbol{zz}\left[m_z(0) + \sum_k \frac{\omega m_{zk}}{\omega - j\omega_{zk}}\right] + (\boldsymbol{xx} + \boldsymbol{yy})\left[m_p(0) + \sum_i \frac{\omega m_{pi}}{\omega - j\omega_{pi}}\right]. \tag{4}$$

The frequency-independent parameters $m_z(0)$ and $m_p(0)$ account for the dipole moments contributed by ferrous targets, with these constants equal to zero for nonpermeable targets. In (4) we sum over the infinite number of resonant $m = 0$ TE and $m = 1$ hybrid modes, although in practice, a small number of modes are sufficient for representation of the frequency-dependent dyadic dipole magnetization. The quantities ω_{zk} and ω_{pi} are real, and therefore, each of the summed terms has a pole on the imaginary frequency axis. Concerning the frequency dependence of each of the summed terms, we note that at high EMI frequencies the resonant mode asymptotically approaches the inductive high-frequency limit [10], for which the real part of the response approaches a constant, while the imaginary part goes

to zero. Moreover, as the frequency goes to zero, the response of each EMI resonant mode will also go to zero [10] [leaving only the constants $m_z(0)$ and $m_p(0)$]. These two conditions motivate the ω in the numerator of each summed term. We also note that the *relative* strengths m_{zk} of the z-directed dipole modes are fixed by the target geometry and material properties, with this also true for the *relative* weights of m_{pi} [11] Finally, the impulse response of the dipole magnetization is easily found via a Fourier analysis of (4), and is given as

$$\hat{\boldsymbol{M}}(t) = \boldsymbol{zz}\left[m_z(0)\delta(t)+\frac{\partial}{\partial t}\sum_k u(t)m_{zk}\exp(-\omega_{zk}t)\right] + (\boldsymbol{xx}+\boldsymbol{yy})\left[m_p(0)\delta(t)+\frac{\partial}{\partial t}\sum_i u(t)m_{pi}\exp(-\omega_{pi}t)\right]. \tag{5}$$

Such that real time-domain waveforms are realized, (5) implies that $m_z(0)$, $m_p(0)$, m_{zk}, and m_{pi} must be real quantities.

If we consider an incident, frequency-dependent magnetic field $\boldsymbol{H}^{inc}(\omega)$, then the magnetic fields induced by a target described by the $\boldsymbol{M}(\omega)$ in (4) are given as [9]

$$\boldsymbol{H}(\omega) = -\frac{1}{4\pi}\nabla\times\left[\boldsymbol{M}(\omega)\cdot\boldsymbol{H}^{inc}(\omega)\times\nabla\frac{1}{r}\right] \tag{6}$$

where r is the distance from the target dipoles to the observation point. A similar, convolutional relationship holds for time-domain excitation.

D. General Rotationally Symmetric Targets

In (4) and (5), we have presented the frequency and time-dependent dipole magnetization for a conducting and permeable finite-length cylinder or ring. In particular, we have demonstrated that these expressions are characterized by the purely imaginary resonant frequencies of the targets, with explicit tools for their approximate calculation presented in Section II-B. We note, however, from the discussion of Section II-A, that *any* finite size conducting and permeable target will be characterized by resonant modes with purely imaginary resonant frequencies. For simple shapes like the cylinder and ring considered above, the imaginary resonant frequencies can be calculated approximately via simple models, while more-general shapes require numerical algorithms. In [3] we discussed a MoM model for general body-of-revolution (BoR) conducting and permeable targets, and in Section III, we discuss a finite-element method (FEM) applicable to general target shapes. Moreover, due to symmetry [4], the dipole magnetization dyadic for a general BoR will be in the same form as (4) and (5), although for general targets one must estimate the model parameters via a numerical model or from measured data. Many targets of interest (mines and ordnance) are bodies of revolution, and therefore the simple models in (4) and (5) are particularly useful.

While beyond the scope of the present BoR computations, we note that the longitudinal/transverse decomposition in (4) and (5) applies to more general metal targets as well. In particular, the same results apply to such bodies with an N-fold rotation axis [12] for $N \geq 3$ (i.e., C_N symmetry with $N \geq 3$). Axial symmetry planes are not required (see [4]). Our BoR corresponds to $N = \infty$ with axial symmetry planes (O_2 symmetry).

III. Comparison of Approximate and Numerical Data

A. Numerical Models Considered

Below we compare the accuracy of the EMI resonances computed using the simple models in Section II-B with data computed via rigorous numerical models. We summarize the numerical models, for completeness. For BoR targets a MoM analysis is particularly useful, in which the currents are expanded in terms of azimuthal harmonics with variation $\cos(m\phi)$. For such problems the currents are expanded only on the surface of the generating arc [3]. This approach has two advantages for our applications: 1) the problem complexity simplifies significantly, since the unknowns, for each mode m, are only on a two-dimensional (2-D) arc (rather than in three dimensions), and 2) the azimuthal BoR harmonics with index m are directly applicable to the azimuthal modes discussed in Section II-B, allowing a direct comparison of numerical and approximate EMI resonant frequencies. The details of our BoR MoM model are discussed in [3].

While the BoR MoM model is useful and computationally efficient, it is restricted to a limited class of problems. If the target is not a BoR [15], or if it is composed of inhomogeneous material parameters, a more general model is preferable. In this context we have developed a general hybrid model that combines a volumetric finite-element method (FEM) solution with a boundary element method (BEM) [16], [17]. In the FEM-BEM formulation, the fields inside the target are expanded in terms of general volumetric basis functions, and the surface region is represented in terms of generalized surface basis functions. The unknowns therefore exist inside and on the surface of the target, and there is no need to truncate the computational domain in an absorbing boundary region. The details of the FEM-BEM model have been published elsewhere. In particular, the general FEM-BEM formulation is discussed in [16], [17], with the type of surface basis functions employed discussed in [18]. The FEM-BEM and BoR-MoM models are distinct, and their respective accuracy was assured by demonstrating that these two disparate models yielded the same results, when considering problems for which they are both applicable.

B. Finite-Length Cylinder

Our initial example is for a stainless-steel cylinder, of length 5.08 cm and 1.27 cm diameter. We perform a comparison between resonant frequencies computed via the simple approximate model (Section II-B) and via the rigorous BoR-MoM solution. Note that the target size considered is small, especially relative to large ordnance that may be of interest. This is a limitation of any numerical model, including the BoR-MoM. In particular, the currents on the surface of the BoR are expanded in a basis set [3], and the size of the basis functions must be small relative to the skin depth. Since for large targets, the skin depth may be very small relative to other target dimensions (e.g., length), a large number of basis functions may be required to obtain accurate results. Therefore, as the target size increases,

one realizes a commensurate increase in the required computer memory and run times. In particular, if N basis functions are required to expand the surface currents, the required memory is order N^2 and the required computer time is also order N^2. The FEM-BEM solution has similar restrictions. By contrast, the simple models developed in Section II-A have the significant advantage of being applicable independent of the target size. The computational burden of the numerical algorithms is exacerbated when one considers highly permeable targets, and therefore all comparisons to numerically generated data is for nonferrous targets. We address ferrous targets through consideration of measured data, in Section IV.

The imaginary resonant frequencies correspond to zeros of the MoM or FEM determinant [3] when computed along the imaginary frequency axis. In practice, the determinant does not go to zero at the resonant frequencies, due to numerical truncation errors, but the resonant modes occur at sharp dips in the MoM/FEM determinant when viewed as a function of frequency (see Figs. 4 and 5). In Fig. 4 we plot the determinant of the MoM impedance matrix, assuming the stainless-steel conductivity $\sigma = 1.3 \times 10^6$ S/m, for the $m = 0$ and $m = 1$ modes. Recall that the $m = 0$ modes are in the form of TE and TM modes. The TM modes correspond to magnetic quadrupoles, and therefore are of little interest. However, the MoM solution does not distinguish between TE and TM modes, it calculating the imaginary resonant frequencies of *all* modes. Consequently, an advantage of the simple EMI solution is its close connection to the underlying physics. Similar information can be obtained from the MoM solution, but one must examine the modal currents to ascertain the mode identity [3]. In addition to plotting in Fig. 4 the determinant of the MoM matrix, we tabulate the imaginary resonant frequencies computed via the approximate model, for the first few magnetic-dipole modes.

In Fig. 4(a) the MoM and approximate model agree to less than 5% error while in Fig. 4(b) the error is typically less than 2%. Given the simplicity and computational efficiency of the approximate model, such accuracy is good. The approximate results in Fig. 4(b) are a little better than those in Fig. 4(a) because the $m = 0$ modes are computed using a magnetic wall for the sides of the TE-mode waveguide, while for the hybrid modes characteristic of the $m = 1$ modes, we numerically evaluate the modal propagation constant [9] without any such approximation. The $m = 0$ approximate solution in Fig. 4(a) are improved slightly if one computes the TE waveguide propagation constant numerically (without a magnetic wall), as was done for the hybrid modes.

Note that, for this example, the $m = 1$ mode has a lower fundamental resonant frequency than the $m = 0$ mode; we have consistently found this to be true for the nonferrous cylinders examined. The MoM results, for computation of the determinant over the range of imaginary frequencies considered, requires on the order of one hour of CPU time on a 500 MHz Pentium III personal computer. On the same computer, the approximate model requires approximately 20 s.

C. Finite-Length Ring

We next consider a ring of 2.54 cm length, outer radius 1.27 cm and a wall thickness of 3.175 mm, composed of aluminum ($\sigma = 3 \times 10^7$ S/m). The numerical results are computed via the

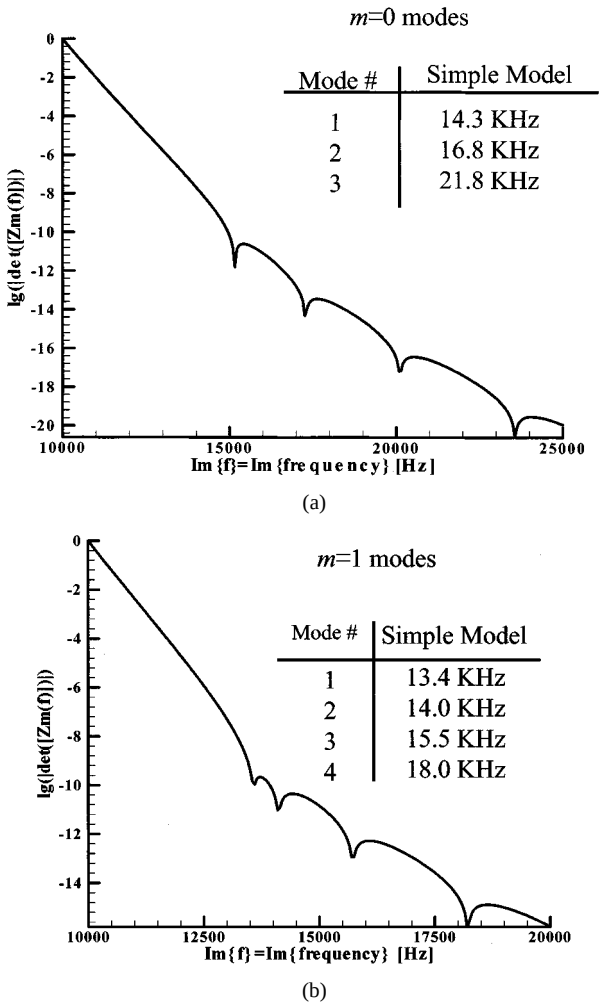


Fig. 4. Characterization of the magnetic dipoles associated with a cylindrical target of length 5.08 cm and 1.27 cm diameter, and conductivity $\sigma = 1.3 \times 10^6$ S/m. The curves correspond to the determinant of the MoM impedance matrix. In the inserted table is shown the results from the approximate model in Section II-B. In this context, for the $m = 0$ modes we only tabulate results for the magnetic-dipole TE modes; the highest-frequency MoM dip in Fig. 4(a) corresponds to a TM mode (magnetic quadrupole). (a) Modes with magnetic dipole oriented along the target axis and (b) modes with magnetic dipole orthogonal to target axis.

FEM-BEM model, this model applicable to targets of general composition and shape. The disadvantage of this generality is that the FEM-BEM model does not decompose the fields into $\cos(m\phi)$ harmonics, and therefore, a comparison between the approximate and FEM-BEM results is less direct. We therefore compare the accuracy of the approximate solution for the ring mode of lowest resonant frequency, this corresponding to the fundamental $m = 1$ mode in the approximate model. As in the BoR-MoM solution, for computation of the numerical resonant frequency, we compute the determinant of an FEM-BEM matrix along the imaginary frequency axis, with the dips corresponding to resonant frequencies. We see in Fig. 5 that the approximate and FEM-BEM solutions agree to within 1.6%. The FEM-BEM solutions, being three-dimensional (3-D), are even more costly computationally than the BoR-MoM solution. While the FEM-BEM solution requires hours of CPU time, on the personal computer discussed above, the approximate solution again requires approximately 20 s.

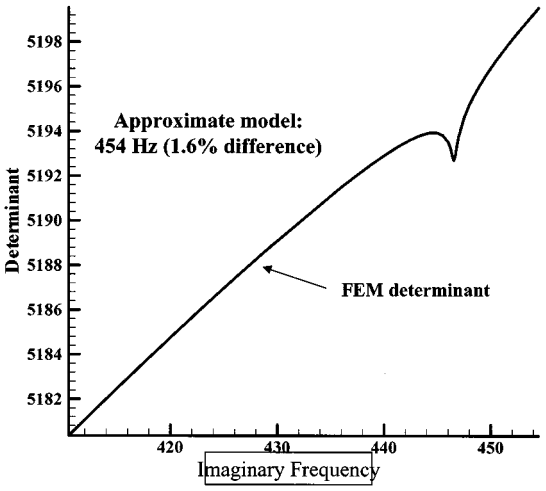


Fig. 5. Characterization of the lowest-order (fundamental) magnetic dipole for a ring target of 2.54 cm length, outer radius 1.27 cm and a wall thickness of 3.175 mm, and of conductivity $\sigma = 3 \times 10^7$ S/m. The curve is the determinant of the FEM-BEM impedance matrix, while the approximate model predicted an imaginary resonant frequency of j454 Hz.

IV. Comparisons to Measured Data

A. Sensor Details

The above comparisons between the approximate model of Section II-B and numerical data computed via the BoR-MoM and FEM-BEM models are encouraging. However, as discussed above, the numerical models have difficulties when considering ferrous targets. Therefore, to consider the accuracy of the model in Section II-B for such targets, we perform a comparison to measured data. The data were taken using a magnetic sensing probe developed by Quantum Magnetics. This probe uses a wide bandwidth room temperature magnetoresistive (MR) sensor operated using a high-performance electronics developed by QM. Unlike induction receiver coils, which measure the rate of change of magnetic field, MR sensors measure the magnetic field itself. As a result, they operate with a frequency-independent response from dc to several MHz. The current sensor operates over the frequency band 0.1 and 400 KHz.

B. Comparison of Measured and Theoretical Data

When considering ferrous targets, there are two parameters that can be varied: the conductivity and permeability. In practice, for a given metal type, there is often a wide range of variability in these parameters, and therefore, it is somewhat misleading to use particular values of μ_r and σ in the approximate model (for the resonant frequencies) and compare with measured data. In particular, one can choose μ_r and σ such that there is a match, but this does not mean that the μ_r and σ used in the model actually correspond to the parameters of that target that was measured. Consequently, we therefore perform an alternative comparison of modeled and measured data. In particular, from (4) and (6) it is clear that the frequency-dependent magnetic field can be expressed in the form

$$H_\alpha(\omega) = a + \sum_k \frac{\omega b_k}{\omega - j\omega_k} \tag{7}$$

where

- H_α particular component of magnetic field;
- a constant offset due to the ferrous target (zero for a nonferrous target);
- $j\omega_k$ imaginary resonant modes excited;
- b_k excitation strengths of the modes, with these dictated by the details of the target and the incident magnetic field.

For the ferrous targets considered below, we consider the optimal match of (7) to the measured data, using just a one-pole model. We therefore perform a least-squared match of (7) to the measured data, where the frequency-independent (constant) free parameters are a, b_1 and ω_1.

We consider a carbon steel pin of 6.858 mm length and 1.5748 mm diameter. This target is characteristic of a small firing pin used in plastic antipersonnel land mines. The EMI response was measured over the 0.1–400 KHz band, with the pin axis aligned with the sensor axis (the magnetic dipole moment along the target axis is excited preferentially with this target-sensor orientation, motivating the single-pole model). In Fig. 6, we compare the measured real and imaginary EMI signature (curves) to the EMI model in (7), where the model parameters are shown with points. One notices an excellent match to the computed and measured EMI signature, where here the model parameters employed were $a = -0.1$, $b_1 = 0.1$ and $f_1 = 8.5$ KHz ($\omega_1 = 2\pi f_1$).

In practice, while the targets of interest may be rotationally symmetric, the exact conductivity and permeability are generally known only approximately (the same metal type can have substantial variability in these parameters). Therefore, in the context of target classification, it may be best to compute the EMI model parameters from measured data, as above. The extracted model parameters can then be employed in a statistical target-classification algorithm [19], in which the statistical variation of the target-model parameters is accounted for. The comparisons to numerical data in Section III therefore serve to demonstrate the approximate-model accuracy, while in practice the model will typically be applied directly to measured data (as in Fig. 6). The approximate but highly accurate models discussed in Sections II-B and II-C can be used to quickly determine the range of variability in the imaginary resonant frequencies of finite-length cylinder and ring targets, as a function of real-world variability of the target material parameters.

Before closing this section, it is of interest to examine the characteristics of the target in Fig. 6, for the case in which the cylinder axis is orthogonal to the sensor axis. This corresponds to the case for which the orthogonal (to the target axis) magnetic dipoles will be excited preferentially. We employ the same sensor as discussed above, with the comparison between the measured data and the single-pole model shown in Fig. 7. For this case the fitted model parameters in (7) correspond to $a = -0.0114$, $b_1 = 0.0111$, and $f_1 = 7.6$ KHz. Note that for this case, the general fit between the measured and model data is good, although not as good as in Fig. 6. The orthogonal magnetic dipole is apparently not excited as strongly, for the EMI response (measured magnetic field orthogonal to the sensor loop) is now an order of magnitude smaller than the results in Fig. 6

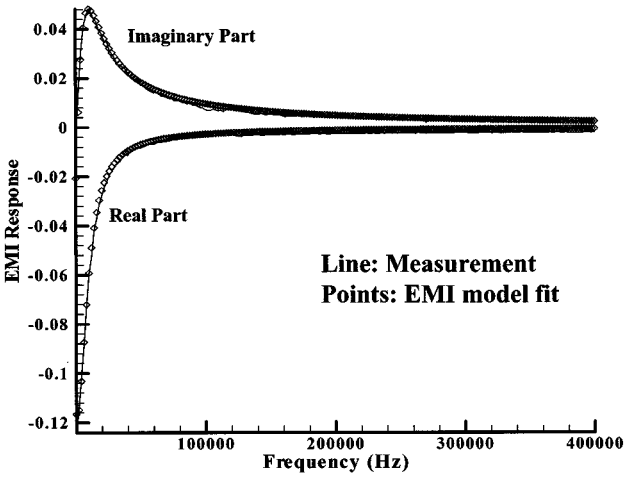


Fig. 6. Magnetic fields measured by an EMI sensor, with the excitation-coil axis aligned with the target axis. The curve corresponds to the measured data, and the points correspond to the fit to the data using a one-pole model, as in (7).

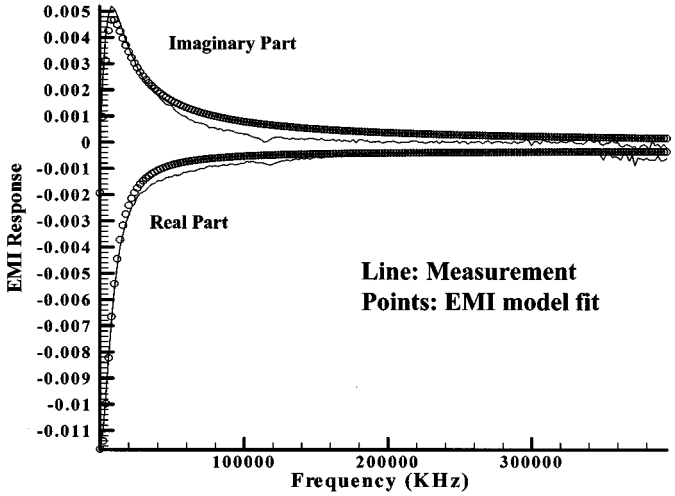


Fig. 7. As in Fig. 6, but now the sensor coil is placed orthogonal to the cylinder axis, aligned with the middle of the cylinder.

(for the same target-sensor distance). The small differences between the measured and model data may therefore be attributed to sensor noise, due to the relatively weak measured signal at this target-sensor orientation. A better fit can also be obtained using a two-pole model.

With regard to the measured data in Figs. 6 and 7, we reiterate that it appears that the magnetic dipole oriented along the target axis is excited more strongly than the dipole orthogonal to this axis. Our theory from Section II can help explain this phenomenon. In particular, the magnetic dipole oriented along the target axis corresponds to an $m = 0$ TE mode, while the magnetic dipole orthogonal to this axis corresponds to a higher-order $m = 1$ hybrid mode. We therefore attribute the differences in the excitation strengths in Figs. 6 and 7 to relatively stronger excitation of lower order (simpler) modes as compared to weaker excitation of higher order (more complex) modes.

V. Conclusions

The principal contribution of this paper is an explanation of the duality between the response of high-permittivity dielectric resonators at microwave frequencies and high-conductivity, high-permeability targets at EMI frequencies. The dielectric resonators resonate at nearly purely real microwave frequencies, while the conducting and ferrous targets have purely imaginary EMI frequencies. An understanding of this duality allows exploitation of the wide range of work that has been done in the microwave community for analyzing high-permittivity resonators. In particular, high-conductivity and high-permeability cylinders and rings are of interest in many targets interrogated via an EMI sensor, and we have transitioned models originally developed in the microwave community to perform an EMI analysis of such targets. The approximate-model EMI results were found to be in close agreement with data computed via rigorous numerical algorithms.

The EMI model, in terms of a diagonal magnetization tensor, in which the EMI resonant frequencies are used explicitly, is applicable to general rotationally symmetric targets, not simply cylinders and rings. This model characterizes the wideband EMI response of such targets in terms of a small number of parameters. This is important in the analysis of measured EMI data which, by exploitation of the simple model, can be fit in terms of a small number of parameters, with these compact parameters employed in a subsequent signal-processing algorithm. We have demonstrated the fit of wideband EMI data, over a 0.1–400 KHz bandwidth, in terms of only three parameters. The target considered was a small ferrous cylinder, characteristic of the firing pin in a plastic land mine. Such fits can be performed with EMI data from more general rotationally symmetric targets.

Future work is required on consideration of more-complex targets, such as actual UXO, that may be composed of multiple parts (e.g., a long cylinder, rings, and possibly a nose composed of a different metal type). Each of these individual components is characterized in terms of its own resonant EMI modes, and it is of interest to examine the degree to which such modes couple in the composite. Moreover, the magnetization tensor is representative of a superposition of various magnetic dipole modes. For a composite target, the different magnetic dipoles from each of the components will in general not be colocated (i.e., the magnetic dipoles associated with a UXO ring may not be located at the same position as the dipoles associated with the larger body). In this context, the dipole magnetization model, in terms of a finite number of magnetic dipoles, must be augmented to consider the location of the dipoles associated with respective subcomponents.

References

[1] Y. Das, J. E. McFee, and R. H. Chesney, "Time domain response of a sphere in the field of a coil: Theory and experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. GE–22, pp. 360–367, July 1984.

[2] Y. Das, J. E. McFee, J. Toews, and G. C. Stuart, "Analysis of an electromagnetic induction detector for real-time location of buried objects," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 278–287, May 1990.

[3] N. Geng, C. E. Baum, and L. Carin, "On the low-frequency natural response of conducting and permeable targets," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 347–359, Jan. 1999.

[4] C. E. Baum, Ed., *Detection and Identification of Visually Obscured Targets*. London, U.K.: Taylor & Francis, 1998.

[5] J. van Bladel, "On the resonances of a dielectric resonator of very high permittivity," *IEEE Trans. Microwave Theory Tech.*, vol. MTT–23, pp. 199–208, Feb. 1975.

[6] K. K. Chow, "On the solution and field pattern of cylindrical dielectric resonators," *IEEE Trans. Microwave Theory Tech.*, vol. MTT–14, p. 439, Feb. 1966.
[7] M. Verplanken and J. van Bladel, "The electric-dipole resonances of ring resonators of very high permittivity," *IEEE Trans. Microwave Theory Tech.*, vol. MTT–24, pp. 108–112, Feb. 1976.
[8] ——, "The magnetic-dipole resonances of ring resonators of very high permittivity," *IEEE Trans. Microwave Theory Tech.*, vol. MTT–27, pp. 328–333, Apr. 1979.
[9] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*: McGraw Hill, 1961.
[10] F. Grant and G. West, *Interpretation Theory in Applied Geophysics*. New York: McGraw Hill, 1965.
[11] C. E. Baum, "Use of residue and constant-dyadic information in magnetic singularity identification," Air Force Res. Lab., Kirtland AFB, NM, Interaction Note 547, Aug. 1988.
[12] C. E. Baum and H. N. Kritikos, "Symmetry in electromagnetics," in *Electromagnetic Symmetry*, C. E. Baum and H. N. Kritikos, Eds. London, U.K.: Taylor & Francis, 1995, ch. 1, pp. 1–90.
[13] C. E. Baum, "Symmetry in electromagnetic scattering as a target discriminant," Air Force Res. Lab., Kirtland AFB, NM, Interaction Note 523, Oct. 1996.
[14] ——, "Application of symmetry to magnetic-singularity identification of buried targets," Air Force Res. Lab., Kirtland AFB, NM, Interaction Note 543, June 1998.
[15] C. E. Baum, N. Geng, and L. Carin, "Integral equations and polarizability for magnetic singularity identification," Air Force Res. Lab., Kirtland AFB, NM, Interaction Note 524, Mar. 1997.
[16] S. Wakao and T. Onuki, "Electromagnetic field computations by the hybrid FE-BE method using edge elements," *IEEE Trans. Magn.*, vol. 29, pp. 1487–1490, Feb. 1993.
[17] H. T. Yu and K. R. Shao, "H method for solving 3D eddy current problems," *IEEE Trans. Magn.*, vol. 31, May 1995.
[18] A. Bossavit and I. D. Mayergoyz, "Edge elements for scattering problems," *IEEE Trans. Magn.*, vol. 25, pp. 2816–2821, Apr. 1989.
[19] L. Collins, P. Gao, and L. Carin, "An improved Bayesian decision theoretic approach for land mine detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 811–819, Mar. 1999.

Lawrence Carin (SM'96–F'01) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor, and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department at Duke University, Durham, NC, where he is now a Professor. He is the Principal Investigator on a Multidisciplinary University Research Initiative (MURI) on demining. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is a member of Tau Beta Pi and Eta Kappa Nu and is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.



Haitao Yu [redacted] He received the Ph.D. degree in electrical engineering from Huazhong University of Science and Technology (HUST), China, in 1995. From 1995 to 1997, he worked as postdoctoral fellow in the Department of Electrical Engineering, HUST. In 1997, he joined the same department as an Associate professor. From 1998, he was with Duke University, Durham, NC. His current research interests include wideband electromagnetic computation and numerical techniques for high and low frequencies.

Yacine Dalichaouch received the B.S. in physics from the University of Constantine, Algeria, and the Ph.D. in physics from the University of California, San Diego, in 1988, where he fabricated and investigated the magnetic and superconducting properties of highly correlated electron materials.

Since 1996, his work has concentrated on the development of sensitive magnetic sensors and their applications in nondestructive evaluation of metal structures and detection/characterization of magnetic targets. He is currently a Senior Scientist and Magnetics Group Leader with Quantum Magnetics, Inc., San Diego, CA. He has close to 100 publications.

Dr. Dalichaouch is a member of the SPIE and APS.

Alexander R. Perry (M'96–SM'00) received the B.A. and Ph.D. degrees in engineering from Cambridge University, Cambridge, U.K., in 1988 and 1992, respectively, where he used cryogenic magnetometers for an investigation of the flux creep properties of high temperature superconducting materials.

Since 1992, his career has focused on applications of magnetic-field sensing using innovative configurations of both superconducting and room-temperature sensors. Currently, he serves as Group Leader for Advanced Systems with Quantum Magnetics, Inc., San Diego, CA, on research and development programs for data acquisition systems for a variety of magnetic sensing technologies.

Peter V. Czipott received the B.A. degree in physics and the Ph.D. degree from the University of California, San Diego, in 1975 and 1983, respectively, where he used superconducting gravimeters for geophysical studies and for an investigation of possible non-Newtonian gravitation at short distance scales.

Since 1993, his career has focused on applications of magnetic-field sensing, using innovative configurations of both superconducting and room-temperature sensors. Currently, he serves as Research Applications Manager with Quantum Magnetics, Inc., San Diego, CA, where he oversees the company's research and development programs outside the area of chemically specific explosives detection.

Dr. Czipott is a member of the APS, AAAS, AGU and SPIE.



Carl E. Baum (S'62–M'63–SM'78–F'84) [redacted] He received the B.S. (with honors), M.S., and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1962, 1963, and 1969, respectively.

He was stationed at the Air Force Research Laboratory, Directed Energy Directorate (formerly Phillips Laboratory, formerly Air Force Weapons Laboratory), Kirtland AFB, Albuquerque, NM, from 1963 to 1967 and from 1968 to 1971. Since 1971, has served as a civil servant with a position as Senior Scientist, Air Force Research Laboratory. He is editor of several interagency note series on EMP (electromagnetic pulse) and related subjects and has published four books: *Transient Lens Synthesis: Differential Geometry in Electromagnetic Theory, Electromagnetic Symmetry, Ultra-Wideband, Short-Pulse Electromagnetics 3*, and *Detection and Identification of Visually Obscured Targets*. He has led an EMP short course and HPE workshops at numerous locations around the globe.

Dr. Baum has been awarded the Air Force Research and Development Award (1970), the AFSC Harold Brown Award (1990), and Air Force Research Laboratory Fellow (1996). He has received the Richard R. Stoddart award (1984) of the IEEE EMC Society. He is a recipient of the 1987 Harry Diamond Memorial Award, one of the IEEE Field Awards, with citation "for outstanding contributions to the knowledge of transient phenomena in electromagnetics." He is a member of Commissions A, B, and E of the U.S. National Committee of the International Union of Radio Science (URSI) and is Founder and President of SUMMA Foundation, which sponsors various electromagnetics-related activities including scientific conferences, publications, short courses, fellowships, and awards.

Multi-Aspect Detection of Surface and Shallow-Buried Unexploded Ordnance via Ultra-Wideband Synthetic Aperture Radar

Yanting Dong, *Student Member, IEEE*, Paul R. Runkle, Lawrence Carin, *Fellow, IEEE*, Raju Damarla, Anders Sullivan, *Member, IEEE*, Marc A. Ressler, and Jeffrey Sichina

***Abstract*—An ultra-wideband (UWB) synthetic aperture radar (SAR) system is investigated for the detection of former bombing ranges, littered by unexploded ordnance (UXO). The objective is detection of a high enough percentage of surface and shallow-buried UXO, with a low enough false-alarm rate, such that a former range can be detected. The physics of UWB SAR scattering is exploited in the context of a hidden Markov model (HMM), which explicitly accounts for the multiple aspects at which a SAR system views a given target. The HMM is trained on computed data, using SAR imagery synthesized via a validated physical-optics solution. The performance of the HMM is demonstrated by performing testing on measured UWB SAR data for many surface and shallow UXO buried in soil in the vicinity of naturally occurring clutter.**

***Index Terms*—Detection, scattering, synthetic aperture radar (SAR), wideband.**

I. Introduction

In many former war zones and bombing test ranges, there are a large number of unexploded ordnance (UXO) that must be removed before the terrain can be safely returned to civilian use. There are many sensors that have been developed to address this problem, including electromagnetic induction (EMI) and magnetometers [1]. While such sensors are often effective in detecting and discriminating UXO, they generally require that the sensor be close to the ground, implying that they are typically employed in a ground-based system. In many cases, the former bombing ranges have been unused for many years, and their location is poorly known. In such situations, wide-area surveillance is often required to circumscribe local regions in which UXO may reside, with such regions sensed subsequently via ground-based sensors.

A wide-area surveillance sensor requires rapid coverage with significant standoff. An airborne synthetic aperture radar (SAR) can meet these requirements, but SAR has significant limitations for sensing UXO. In particular, it is well known that radar waves typically suffer significant attenuation upon propagation in soil [2]–[4], severely limiting the utility of SAR for detecting deeply buried UXO. However, at many former bombing ranges there are a large number of surface and shallow-buried UXO. It is of interest to quantify whether UWB SAR can detect a high-enough percentage of surface and shallow UXO, with a low-enough false-alarm rate, such that such a system can be utilized for detection of former bombing ranges (*not* for detection of each, possibly deeply buried, individual UXO). The definitions of "high-enough" and "low-enough" are dictated by mission requirements and public policy. Rather than attempting to deal with these inchoate terms, we investigate performance of a particular SAR system at a particular test site (discussed further below), incorporating the underlying wave physics into the signal-processing construct. The site considered is relatively favorable for this mission (good soil penetration and relatively benign clutter), and therefore, the results from this study help quantify performance under propitious operating characteristics.

Having restricted ourselves to surface and shallow-buried UXO, we must now address the wave phenomenology one might exploit to reliably detect such targets. For UXO that have one or more dimension large relative to wavelength, the backscattered waveform is typically a strong function of target-sensor orientation. While this aspect dependence may complicate target detection, because there is not a single target signature representative of all target orientations, the aspect dependence of the target signature is often different than that of clutter. A large SAR aperture, which manifests looks at the target from multiple orientations, provides a natural vehicle for exploiting the aspect dependence of the target signature. We have previously developed hidden Markov models (HMMs) for performing such multi-aspect processing [5]–[7], with those tools applied here. Since the basic HMM construct has been described by us previously [5]–[7], for target detection and classification, here we only provide a concise summary, for completeness.

Although the scattered waveform (or SAR image) from a given target is often a strong function of target-sensor orientation, there are generally contiguous sets of orientations over which the scattering physics is slowly varying (or stationary). Such angular sectors are called "states," with a given target described by M states, with M dictated by the target complexity and sensor bandwidth. When one performs a sequence of N measurements, from N consecutive target-sensor orientations, we implicitly sample waveforms (or images) from N states of

Manuscript received June 27, 2000; revised February 12, 2001. This work was supported by the Strategic Environmental Research and Development Program (SERDP) under Contract DACA72-99-C-0012-CU-1123.

Y. Dong, P. R. Runkle, and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA (e-mail: lcarin@ee.duke.edu).

R. Damarla, A. Sullivan, M. A. Ressler, and J. Sichina are with the Army Research Laboratory, Adelphi, MD 20783 USA.

Publisher Item Identifier S 0196-2892(01)04845-8.

0196–2892/01$10.00 © 2001 IEEE

the target. Some states may be sampled more than once, and some not all, depending on the target orientation and sensor motion. We have demonstrated that, to a good approximation, this transitioning from state to state can be modeled as a Markov process, which implies that the probability of transitioning into any given state is approximately dictated entirely by the current state being occupied. Moreover, since the target is typically distant or concealed, the actual states being sampled are "hidden," motivating an HMM. The HMM paradigm has been applied successfully to several detection and classification problems, based on processing multi-aspect scattering data [5]–[7].

A limitation of many target-classification algorithms, including the HMM, is that they often require substantial training data, and in practice, this training data is typically similar to the measured data on which the algorithm is tested. An advantage of our problem is that electromagnetic models have been developed to simulate UWB scattering from targets such as UXO, in the presence of realistic lossy, dispersive media (soil). In particular, method-of-moments (MoM) [8], fast-multiple method (FMM) [9] and finite-difference time-domain (FDTD) [10] models have been developed over the last several years. These models provide the capability of simulating the UWB SAR imagery of a surface or subsurface UXO, with such simulated data applicable to training a classifier. In this manner, the classifier is trained and tested on independent data, yielding a meaningful assessment of classifier performance.

As indicated, substantial training data is desirable, with which the target can be viewed from all possible target-sensor orientations. Consequently, it is required that the forward solver used to generate the scattering data be as efficient as possible, while retaining accuracy. Although well designed MoM, FMM, and FDTD codes are efficient, such numerical algorithms are still relatively computationally intensive for the UXO problem. It is therefore desirable to employ an approximate solver, such as physical optics (PO) [11]. We validate here that the PO approximation accurately generates UWB SAR imagery for surface and buried UXO, in comparison to a rigorous MoM solution, and therefore employ the PO solutions to train the HMM.

The measured UWB SAR data is collected at the Yuma Proving Ground, Yuma, AZ, using the U.S. Army Research Laboratory (ARL) BoomSAR system, which has been described elsewhere [12], [13]. The soil at Yuma is relatively dry, providing favorable electromagnetic soil penetration over the 50–1200 MHz bandwidth of the ARL system. Moreover, the naturally occurring clutter at this site provides a significant challenge, but is relatively benign. In particular, there is little large foliage, with most clutter manifested from small shrub trees, rocks, subsurface inhomogeneities, and animal burrows. As demonstrated in Section III, even this relatively benign environment presents a significant challenge to UWB SAR sensing, necessitating a sophisticated processing algorithm.

The remainder of the text is organized as follows. In Section II, we review the HMM and the key components employed for its implementation with UWB SAR data. We address HMM training, with a validation of the PO forward solver employed in this context. Performance of the HMM on measured UWB SAR data is discussed in Section III, in which we also address the use of prescreeners to determine the points of interest subsequently interrogated by the HMM. Conclusions from this work are discussed in Section IV.

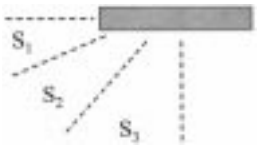


Fig. 1. Schematic of the three states used to characterize a UXO. State one (S_1) is characterized principally by end scattering S_2 by diffractive scattering from the two ends, and S_3 by broadside scattering. Due to symmetry assumptions, these three states describe the scattering from any target-sensor orientation.

II. Summary of HMM Classifier and Associated Tools

A. Hidden Markov Model (HMM) Basics

Assume that a given target is described by M states, each state characteristic of a generally contiguous set of target-sensor orientations over which the scattering physics is approximately stationary as a function of orientation. For the UXO problem considered here, we employ an $M = 3$ state model. As indicated in Fig. 1, state one (S_1) corresponds primarily to end-on excitation, S_2 corresponds primarily to oblique excitation, and S_3 corresponds primarily to excitation of the UXO broad side. Note that we are assuming that the UXO has 90° symmetry. Although this may not be exactly true (a UXO typically has 180° symmetry in reality), this is a good approximation over the frequencies of interest. The initial state decomposition for the UXO targets is performed as indicated in Fig. 1, with the details of the angles used to define the states not critical, since the entire HMM is subsequently optimized, as discussed further below. The decomposition in Fig. 1 is therefore simply a starting point.

Assume that a sequence of N measurements are performed. For the SAR problem of interest here, the N measurements correspond to N subaperture images. In particular, assume that the SAR aperture defines an angular extent Θ relative to the target center. The total aperture can be divided into N subapertures, each spanning an angle Θ/N (here we assume the subapertures do not overlap, although this restriction can be loosened). The N subaperture SAR images implicitly "view" the target from N distinct orientations, and it is these subaperture images that are processed via the HMM. The N images implicitly sample physics from the M target states, with the particular states sampled dictated by the details of the target and SAR aperture under consideration.

The HMM characterizes a doubly stochastic process. The sequence of N measurements sample N states of the target, denoted $\{q_1, q_2, \ldots, q_N\}$, where the state sequence is treated statistically because the absolute target-sensor orientation is "hidden." Each state is characterized by an ensemble of scattered waveforms (images), with the selection of a particular signal from the ensemble also treated statistically. For characterization of the HMM, we first define the probability that the initial subaperture image in the sequence will be in state S_m, or $p(q_1 = S_m) \equiv \pi_m$. For subsequent subaperture images in the sequence, we define the probability of transitioning from state k to state i, $p(q_{n+1} = S_i | q_n = S_k) \equiv A_{ki}$. The π_m constitute an M-dimensional vector π, and the A_{ki} an $M \times M$ matrix

$\boldsymbol{A}$. The initial values of these quantities are easily defined in terms of the initial state decomposition [5]–[7], with these also refined via a subsequent optimization processed. We have discussed the initial computation of π_m and A_{ki} previously, with the reader referred to [5]–[7] for details.

The HMM is based on the understanding that the scattering physics is often a strong function of the target-sensor orientation, but within each of the M states, the physics is relatively stationary as a function of orientation. This is addressed here, in the context of the UXO problem, in the following manner. We consider a UXO on the surface of the soil, and produce K subaperture images, from K angles uniformly distributed over the extent of the M states (over 90°, see Fig. 1). These K images constitute what is termed a codebook $\boldsymbol{C}$, with $\boldsymbol{C}$ composed of subaperture images $\boldsymbol{I}_k$, $k = 1, \ldots, K$. The K elements of the codebook $\boldsymbol{C}$ are termed codes. For each of the N subaperture images, characteristic of the subaperture sequence under consideration, we perform a correlation with each of the K images in $\boldsymbol{C}$. We associate a given image in the sequence with $\boldsymbol{I}_k$ if, for all the characteristic images in $\boldsymbol{C}$, the correlation is greatest with $\boldsymbol{I}_k$. By this process, each of the original subaperture images in the sequence under test is mapped to one of the codes in $\boldsymbol{C}$, with this process constituting a generalization of vector quantization [14].

The final HMM element that must be defined is the probability that a given code will be associated with a given target state, with b_{km} denoting the probability that code k is associated with state S_m. These probabilities b_{km} constitute a $K \times M$ matrix $\boldsymbol{B}$. Simple training procedures exist for defining $\boldsymbol{B}$, with the interested reader referred to [5]–[7] for details.

For a given target under consideration, the HMM classifier is implemented as follows. Assume that a sequence of N subaperture images are considered, with each mapped into one of the discrete codes in $\boldsymbol{C}$. We therefore now have an observation sequence $\boldsymbol{O} = \{O_1, O_2, \ldots, O_N\}$, with each of the $O_n \in \boldsymbol{C}$. The likelihood that $\boldsymbol{O}$ is characteristic of target T is given, within the context of the HMM, as

$$\begin{aligned} p(O_1, O_2, \ldots, O_N|T) &= \sum_{m}^{M}\sum_{n}^{M}\cdots\sum_{l}^{M}\sum_{q}^{M} \\ &\cdot [p(O_1|q_1 = s_m, T)p(O_2|q_2 = s_n, T)\cdots \\ &\quad p(O_{N-1}|q_{N-1} = s_l, T)p(O_N|q_N = s_q, T)] \\ &\cdot [\pi_m a_{m,n}\cdots a_{l,q}]. \end{aligned} \tag{1}$$

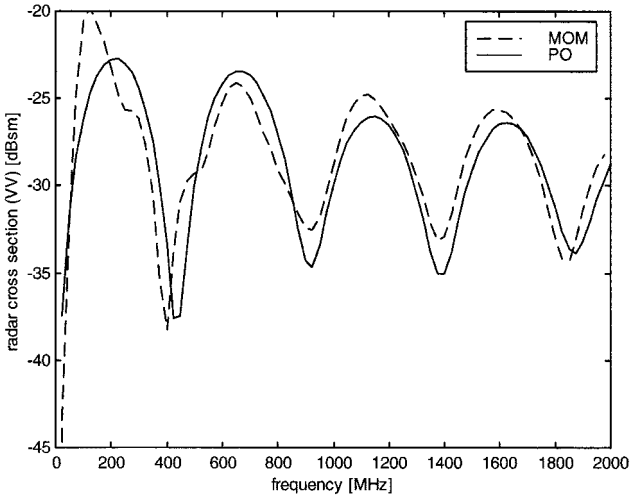
The conditional probabilities for the codes O_n are defined by the matrix $\boldsymbol{B}$. In (1), each sum is over all possible state sequences for the set of N subaperture images (M^N possible state sequences), with each sequence weighted by its likelihood of occurrence (quantified through π and $\boldsymbol{A}$). Although the expression in (1) appears computationally expensive to implement, there is much redundancy in the required summations, and the forward–backward (FB) algorithm [15], [16] has been developed to exploit this fact, making computation of (1) highly efficient. The expression in (1) is also often treated in an alternative manner, through exploitation of the Viterbi algorithm [17]. We have found that HMM classification performance is generally insensitive to whether the FB or Viterbi algorithm is utilized.

In the training phase of the HMM classifier, as many N-dimensional subaperture sequences as possible are utilized, each mapped into an observation sequence $\boldsymbol{O}$ via vector quantization, as discussed above. In the training phase, the expression in (1) is computed for all training sequences, and the HMM parameters π, $\boldsymbol{A}$, and $\boldsymbol{B}$ are iteratively optimized to maximize the average likelihood from (1). In this process, the state decompositions themselves are also optimized (refined), underscoring our earlier comment that the initial state decomposition (Fig. 1) need not be perfect to yield good HMM performance after training. In the testing phase, a given sequence of subaperture images is mapped to an observation sequence, as above, and the likelihood (1) is computed. The sequence of images is deemed associated with the target representative of the HMM if the likelihood is above a prescribed threshold. By varying the threshold, one can generate the receiver operating characteristic [18]. If there are multiple targets of interest, an HMM is designed for each, and the sequence of images is deemed associated with that target for which the respective HMM yields the highest likelihood. If the likelihood of all HMMs is below a given threshold, the sequence of images is deemed characteristic of clutter (i.e., none of the targets).

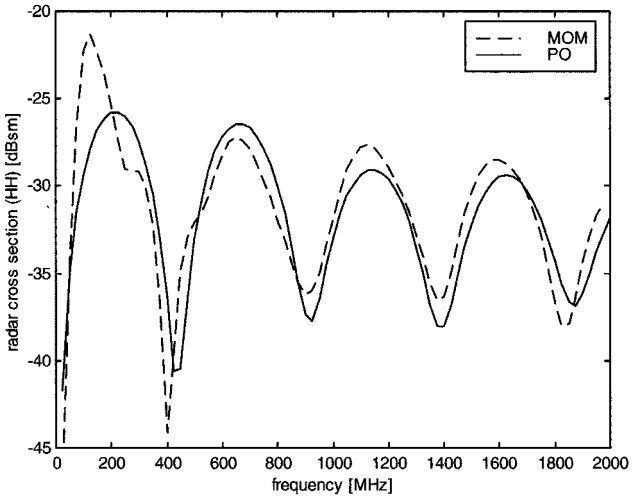
B. Physical-Optics Approximation

A key aspect of the HMM, as implemented here for the UXO problem, is that all training is performed on computed data (for both the HMM and prescreeners, see Section IV). All such computations have been performed via the PO approximation, due to its computational efficiency and accuracy (for this problem). To address the latter, we plot in Fig. 3 the backscattered RCS of the PO approximation, relative to a rigorous MoM solution [8], with the meshed target (155 mm shell). This target has a maximum diameter of 155 mm and a length of 58.4 cm. Results are shown for the backscattered RCS from this target, with incidence 30° from grazing, and 45° with respect to the target broadside. In the results of Fig. 2, the target is buried parallel to the flat air–ground interface, with the top of the target 2.54 cm below the interface (recall that we are focusing on surface and shallow-buried UXO). The soil properties considered are characteristic of Yuma soil [12], [13] with 5% water content ($\epsilon_r = 3.6 - j\ 0.12$ with $\sigma = 0.005$ S/m). Results are shown over the bandwidth of the ARL system [12], [13], and it is demonstrated that the PO approximation generally matches the MoM solution well, for both VV and HH polarizations. We note that, in addition to not modeling the target currents in the shadow region, the PO approximation has the related shortcoming of not modeling multiple wave interaction between the target and interface. This explains why the PO solutions are still not perfect, even at the high-frequency end.

In Fig. 3, we plot SAR imagery computed using scattered fields from the PO approximation, assuming plane-wave incidence, VV polarization, and an incident pulse characteristic of the ARL BoomSAR [12], [13]. Results are shown for the target situated as in Fig. 2, but from three different target-sensor orientations, representative of the three states sketched in Fig. 1.



(a)



(b)

Fig. 2. Comparison of the backscattered RCS computed via a physical optics (PO) and method of moments (MoM) solution, for a 155 mm UXO shell buried parallel to the air–ground interface, at a depth of 2.54 cm. The assumed plane-wave excitation is incident 30° from grazing and 45° with respect to broad side. The target mesh used for these computations is shown in Fig. 3. (a) VV polarization and (b) HH polarization.

These images correspond to a 60° SAR aperture, VV polarization, and again are for incidence 30° from grazing (details of the imaging procedure are discussed in [12]). These images underscore the strong aspect dependence to the SAR imagery from a UXO, with this exploited explicitly via the HMM. Space limitations preclude showing comparisons of these images with similar images computed using MoM-computed scattered fields. However, we have found that the PO and MoM imagery are, on average, correlated to better than 0.95 for the UXO cases considered here. Before closing this section, we note from Fig. 3(a) that the image, when the UXO is parallel to the linear SAR aperture, is not absolutely symmetrical in the cross-range direction. This is due to the fact that the end and nose of the UXO are not identical. This asymmetry vitiates the simple three-state model used here and depicted in Fig. 1. However, it is felt that the slight asymmetry in Fig. 3 does not warrant additional HMM complexity (more states), given that small variations in the target signature are likely to be undermined by clutter.

C. Subaperture Images

The HMM processes a sequence of subaperture SAR images (each subaperture image using a subset of the total available SAR aperture). The total aperture from the BoomSAR is $\Theta = 60°$ [12], [13], dictated by the system antennas, which only provide approximate polarization purity over this angular range. Each of the N subaperture images could be formed by using the UWB scattered waveforms (the "raw" scattering data) measured by antennas within the associated subaperture angular sector. However, this implies that one would have to store all the UWB scattered waveforms for each antenna position along the SAR aperture. Alternatively, we can use all of the available scattered-field data to form a single, full-aperture image. A sequence of two-dimensional (2-D) filters (one filter for each of the desired N subaperture images) is then utilized on this full-aperture imagery to yield the desired sequence of images applied to the HMM. These filters are simple to design and implement, and have been described in detail elsewhere [7]. This latter approach is desirable, since it implies that we need only store the full-aperture image, after which, the original scattered waveforms can be discarded for our purposes.

It is of interest to examine a typical sequence of subaperture images, manifested by the sequence of N filters applied to the original full-aperture image. In Fig. 4 is shown the full-aperture VV-polarized image of a 155 mm UXO on the surface of the ground, oriented at 45° with respect to the linear SAR aperture (computed using PO-generated scattering data, assuming an incident pulse characteristic of the ARL BoomSAR [12], [13]). Also in this figure is shown a sequence of $N = 7$ overlapping subaperture images of this target, with the center angle of the subaperture images corresponding to $-22.5°$, $-15°$, $-7.5°$, 0°, 7.5°, 15°, and 22.5° (0° is pointing orthogonal to the linear SAR aperture, through the target center). Each subaperture image has an associated aperture of 15° ($\pm 7.5°$ about the center angle of the subaperture).

We note from Fig. 4 that each subaperture image is characterized by a loss of cross-range resolution relative to the original full-aperture image. This is the price one pays for exploiting the aspect dependence of the target scattering in the context of subaperture images, since the cross-range resolution is tied to the size of the SAR aperture. Nevertheless, one notes that the sequence of imagery embodies significant information. It is this phenomenology we seek to exploit, in the context of the HMM, with the goal of distinguishing UXO from clutter.

Before proceeding, it is of interest to address the utility of partitioning the original full-aperture image into a sequence of subaperture images, given that all information in the sequence of images is implicitly contained in the original full-aperture image (and we lose cross-range resolution in the subaperture images). Clutter is likely to undermine some subaperture images more than others, depending on the clutter position relative to portions of the full aperture. The HMM is a statistical algorithm that does not require that all subaperture images be well matched to the target in question [5]–[7], as long as a subset is so matched, and the sequence order of measured images is consistent with the HMM. By contrast, if one were to build a processor based only on the full-aperture image, the latter will be corrupted by

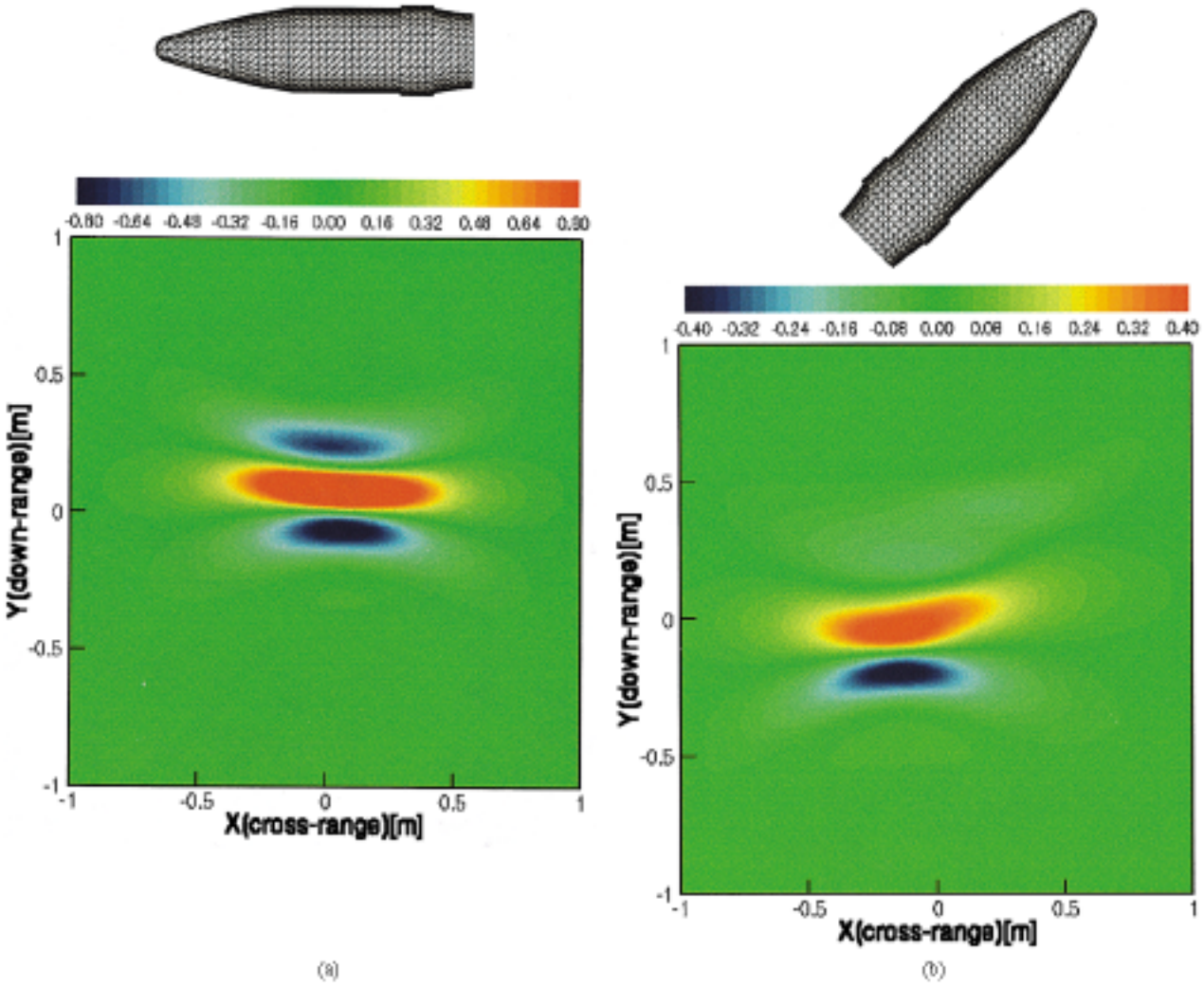


Fig. 3. Computed SAR images for a 155 mm shell buried 2.54 cm in Yuma soil [12], [13], with the target axis parallel to the air–soil interface. The computations use pulsed plane-wave excitation at 30° from grazing, characterized by the incident pulse described in [12] and [13]. The images use an aperture length that yields a 60° angle between the target center and aperture. The images are produced by scattered fields from a physical optics (PO) forward solver. All results are for VV polarization. (a) Target axis parallel to linear SAR aperture, and (b) target axis 45° with respect to linear SAR aperture.

all clutter that appears in the image region under consideration, even if the clutter only contributes from a few subapertures of the total measurement.

III. HMM Performance on Measured UWB SAR Data

A. Measurements

The UWB SAR measurements were performed by the U.S. Army Research Laboratory at Yuma Proving Ground, Yuma, AZ. The majority of the sensor hardware is mounted to the basket of a telescoping boom lift capable of moving at approximately 1 km/h, while the basket is elevated to 45 m. For a typical collection geometries, down-look angles to the target vary from 45° to approximately 10° (from grazing), depending on the range to the target and the height of the boom. For all measurements reported here, the depression angle was approximately 40° and therefore, over the extent of the linear SAR aperture, to a good approximation, only the azimuthal angle is varying with respect to a given target (as assumed in the HMM for simplicity). The UWB SAR system employed is operated directly in the time domain, covering a usable bandwidth of approximately 50–1200 MHz. Since the sensor is typically distant from the targets under investigation, the target excitation is approximately a plane wave (it is so modeled in the simulations discussed in Section III). The pulse emitted by the radar, as well as other details of the sensor, are discussed in [12], [13].

At the Yuma test site, a large set of inert ordnance were buried in the ground at various depths and orientations with respect to the linear SAR aperture. Although there were many deeply buried UXO (at 1 m depth or more), only the surface and shallow-buried UXO were consistently visible in the SAR imagery, underscoring the difficulty of radar-based UXO detection and motivating our goal of UXO-range detection, rather than detection of each individual UXO. As indicated in the introduction, in UXO-range detection, one hopes to detect a large percentage of the surface and flush-buried UXO while realizing a low false-alarm rate (FAR). With regard to the latter, the FAR is strongly dictated by the environment in which the measurements are performed. Yuma is relatively favorable in this context, since the test site is characterized primarily by

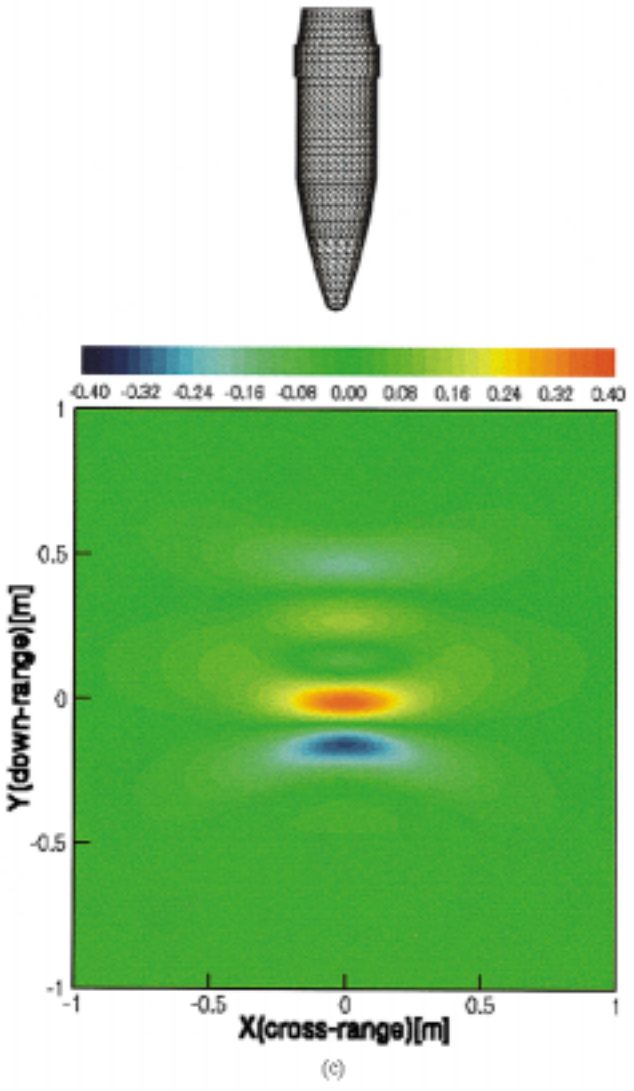


Fig. 3. (*Continued.*) (c) Target axis perpendicular to linear SAR aperture.

desert conditions in which there are rocks, animal burrows, and relatively benign foliage. The soil conditions at Yuma are detailed in [12], [13]. However, as indicated below, even this relatively propitious environment poses a significant challenge to radar-based sensors. For example, general soil inhomogeneities constitute clutter to a radar system, with such poorly understood, since such inhomogeneities are often destroyed upon excavation. Consequently, while new electromagnetic models are yielding increased insight into the phenomenology of scattering from man-made targets, such as UXO, scattering from naturally occurring clutter is far less understood. This is not due to a shortcoming in the electromagnetic modeling tools, but rather to a poor understanding of the subsurface environment.

B. Prescreeners

A SAR system can quickly collect a large amount of imagery. It is therefore of interest to develop simple algorithms that prescreen the imagery with the goal of eliminating for subsequent consideration those regions at which it is relatively clear no target exists. The requirements of a prescreener are that it be computationally efficient, while yielding a high probability of detection (any targets lost to the prescreener cannot be recovered subsequently). It is acceptable if the FAR of a prescreener is relatively high, since these false alarms will (ideally) be pruned subsequently by more sophisticated algorithms. In the work considered here, we have considered three prescreeners classes.

Our first prescreener is based on our ability to model the SAR images of surface and buried UXO. We consider the SAR image of a flush-buried UXO, from several target-sensor orientations. In particular, we consider the full-aperture image $(\Theta = 60°)$ of the UXO, with the center of the SAR aperture ranging azimuthally from 0° to 180°. In this way, we produce $L = 181$ full-aperture SAR images of the UXO (using a 1° azimuthal shift between consecutive images). Each of the L SAR images is arranged as a one-dimensional (1-D) column vector $\boldsymbol{v}_l$, from which we computed the average correlation between images

$$\boldsymbol{R} = \frac{1}{L}\sum_{l=1}^{L} \boldsymbol{v}_l \boldsymbol{v}_l^T. \tag{2}$$

An eigenvector decomposition of (2) yields the principal eigenvectors characteristic of the space spanned by $\boldsymbol{v}_l$, with these eigenvectors associated with the dominant eigenvalues. This is the Karhunen–Loève transform (KLT) [19]. In Fig. 5, we plot the eigenvalues associated with the 155 mm UXO shell considered in Figs. 3–5. The first 19 eigenvalues contain 99% of the energy in the sum of all eigenvalues of $\boldsymbol{R}$, and we use the associated 19 eigenvectors in our first prescreener.

One could apply these 19 characteristic images in the context of a matched filter. However, it is well known that matched filters are only optimal for signals in white Gaussian noise [19]. The clutter in the SAR imagery mitigates the utility of a matched filter. Alternatively, one can design a Wiener filter [18] to mitigate the effects of clutter and to produce better discrimination than a simple matched filter. This extension of the traditional matched filtering has been termed expansion matching (EXM filters) in the signal processing community [20]. Such EXM filters, based on the aforementioned eigenvectors, have been applied here in the context of our initial prescreener. The EXM filters must be implemented for all 2-D shifts across the SAR aperture, with this implemented efficiently via a 2-D FFT (on both the EXM filters and the SAR image). At each pixel value, we consider the output from the 19 EXM filters that are maximum (using full-aperture imagery) and a threshold is set, with all pixels below a particular threshold discarded for subsequent consideration. The selection of the threshold is based on a compromise between retaining as many targets as possible, while eliminating as much of the image that is clearly devoid of targets.

While the EXM filters (based on the reduced set of modeling-generated eigenvectors) produce relatively good results, such filters have limitations. In particular, any large-amplitude portion of the SAR image will produce a relatively large output from an EXM filter, even if the correlation between the filter and image is low. We therefore consider a subsequent class of prescreeners, based on the spatial extent of energy in the image. As seen from Fig. 3, the spatial extent of energy in a UXO image

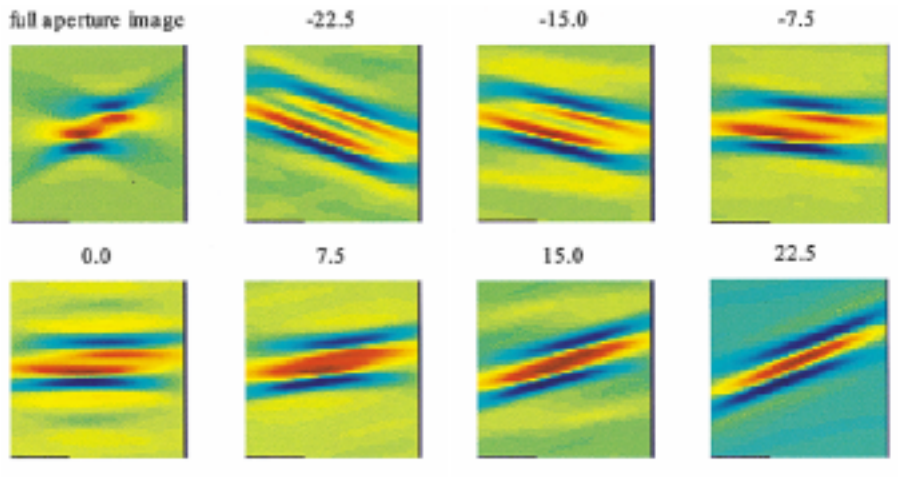


Fig. 4. Original (top left) full-aperture VV-polarized image of a 155 mm UXO on the surface of the ground, oriented at 45° with respect to the linear SAR aperture (computed using PO-generated scattering data, assuming an incident pulse characteristic in [12], [13]). Also shown are a set of overlapping subaperture images of this target, with the center angle of the subaperture images indicated. Each subaperture image has an associated aperture of 15°.

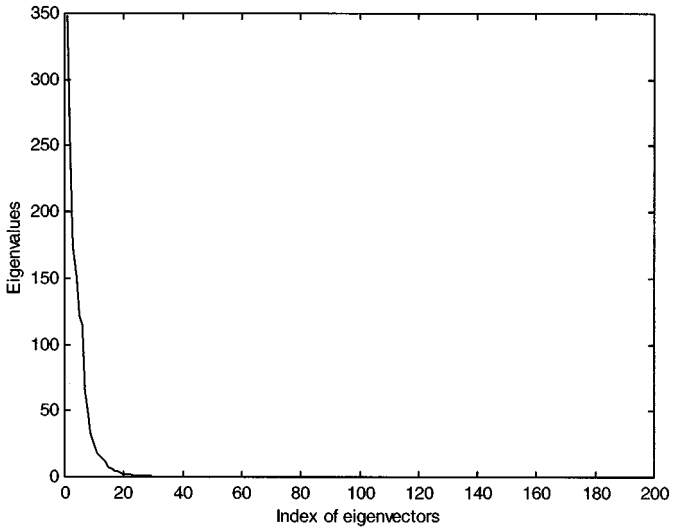


Fig. 5. Eigenvalues corresponding to the surface 155 mm shell atop Yuma soil [12], [13], as computed from the correlation matrix of the full-aperture VV polarized SAR imagery, considering center angles of the SAR image ranging from 0° to 180°, in 1° increments.

is relatively small, dictated by the relatively small target size (relative to system wavelength). Consequently, for each region in the SAR image that passed through the first prescreener, we compute the standard deviation of the image strength, about the mean energy position in a prescribed area (this area made consistent with the size of the target images, as in Fig. 3). This simple prescreener efficiently eliminates any portions of the SAR image that have energy extent too large to be consistent with a UXO. All portions of the image that pass through this stage are then sent to another prescreener of the same genre. In particular, we consider templates as shown in Fig. 6. The total energy from the SAR image inside the circle ("inner" energy) is compared to the energy in the area between the circle and ellipse ("outer" energy). The feature applied from this template is the ratio of the inner to outer energy, with a threshold again set on the output from this parameter. A proper threshold is again determined by consideration of the variability of this feature for

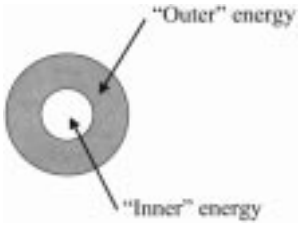


Fig. 6. Template applied in one of the prescreeners, used to eliminate clutter that occupy a spatial extent in the SAR imagery that is too large to be consistent with a UXO. The sizes of the template components are dictated by the expected extent of UXO in the SAR imagery (see Fig. 3). The diameter of the inner circle is 0.6 m and the outer-circle diameter is 2.25 m.

computed data of the form in Fig. 3. The size of the inner circle is chosen such that it encloses the target image (Fig. 3).

The three prescreeners discussed above are computationally very efficient, and therefore easily implemented. A subsequent more computationally expensive prescreener involves calculating correlations between a library of characteristic (full-aperture) image templates and the measured SAR image (this requires more than a simple 2-D FFT). However, after having implemented the three prescreeners above, the space that need be interrogated with the more sophisticated correlation filter is much reduced. We therefore consider the modeled full-aperture image of the UXO, with center aperture position oriented at 0°, 10°, 30°, ...180° with respect to the target (this number of templates is consistent with the 19 eigenvalues discussed above for the 155 mm shell). We compute the correlation of the remaining portions of the measured SAR image (after the initial prescreeners) with these 19 templates, yielding 19 correlation values for each point of interest. If the maximum correlation of this set is above a prescribed threshold, this region of the image passes on to the more sophisticated HMM test.

C. HMM Implementation

The correlation matrix used to implement the first prescreener is implemented using computed imagery for the 155 mm shell,

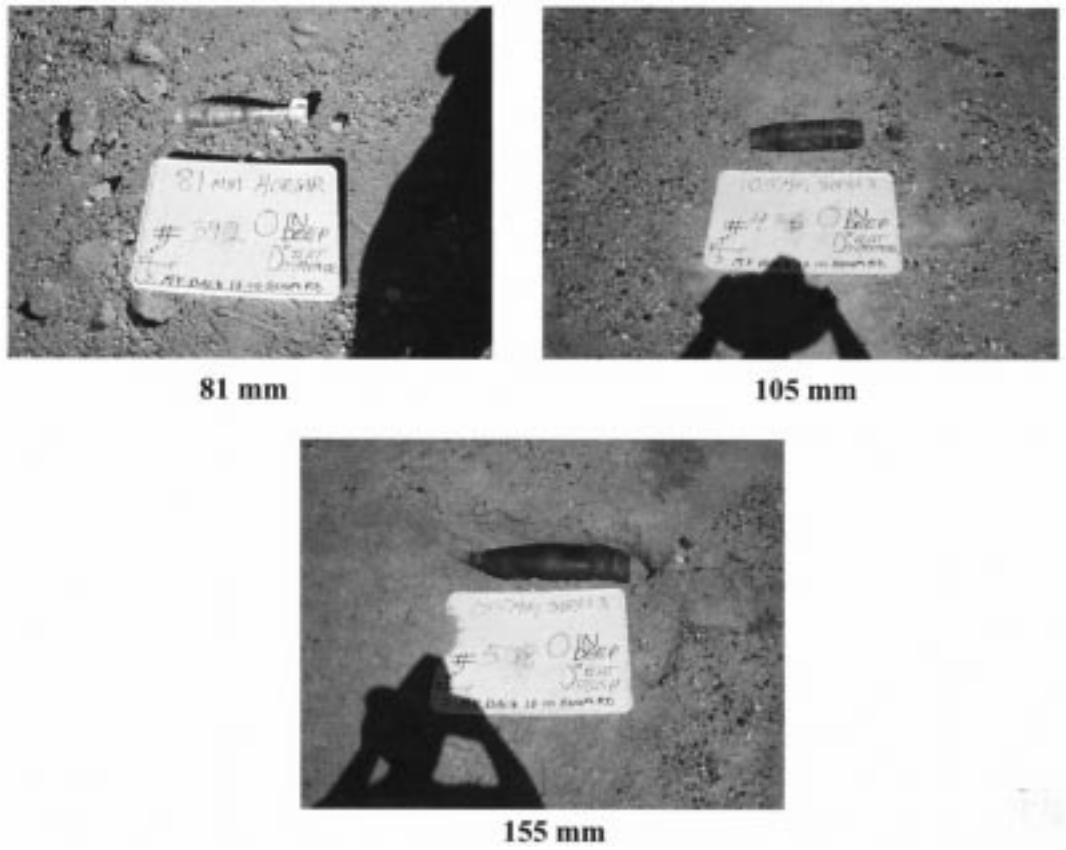


Fig. 7. Photographs of the three UXO in the measured SAR imagery used to test the performance of the classifier.

for VV polarization. The correlation filters, for the last prescreener, were also implemented based on computed imagery for the same scattering scenario. The HMM is likewise trained entirely on computed data from the 155 mm shell, with VV polarization. In particular, the codebook discussed in Section III-A was implemented with $K = 10$ overlapping subaperture images, with center angles at 0°, 10°, 20°, ... 90°, each with a 15° angular aperture. Concerning training with the computed 155 mm shell data, we considered all possible $N = 7$ length subaperture sequences (as in Fig. 4), with the center of the first subaperture image in the sequence ranging over the angles 0°, 1°, 2°,...179°, 180°. The HMM parameters π, $\boldsymbol{A}$, and $\boldsymbol{B}$ were optimized using the above training data, through application of the Viterbi algorithm [17]. After the HMM for the 155 mm shell was so trained, its performance is tested on measured data, as discussed below.

D. Algorithm Performance

We address performance of the detector through consideration of three distinct UXO, with each shown in Fig. 7. The 155 mm shell, 105 mm shell, and 81 mm shell are all of approximately the same length, with distinct maximum-diameter sizes. These targets look very similar to a radar system operating at the bandwidth considered. Therefore, as discussed above, we build a single HMM (and a single set of prescreeners), based on computed training data from the 155 mm shell on the soil surface. The full classifier (prescreeners and HMM) is tested on measured SAR data from Yuma, and if the HMM correctly declares a particular region occupied by one of the *three* targets, this is termed a detection. At regions for which the classifier declares a UXO, but none exists, a false alarm is declared. It is necessary to place a "hallow" around a given declaration, to define the spatial region in which a target is deemed present. In the work reported here, we have used circular regions of diameter 5 m. All results reported here are for VV polarization.

In the following, we present receiver operating characteristic (ROC) performance for subsets of the target classes considered, because we have found that the algorithms detect some types of targets better than others. Such performance characteristics yield insight into the types of scattering physics for which the algorithms work particularly well.

In our first set of results, we consider UXO that are in the ground at a 45° angle, with the top of the target flush buried with the interface (nose down). There were 36 such targets. The algorithms performed best for this target class, with this discussed further in Section III-E. In Fig. 8, we plot the ROC for the last three prescreeners discussed above. The ROC quantifies the probability of detection as a function of the FAR, the latter quantified in terms of false alarms/km^2 (the total region considered for this test encompassed 0.23 km^2). We see from Fig. 8 that the performance of the prescreeners improves as the prescreener complexity increases. In particular, we see that the worst performance occurs for the standard-deviation test, with significant improvement demonstrated by the energy ratio summarized in Fig. 6. Finally, the most-complex, correlation-based prescreener, which uses computed full-aperture imagery, yields

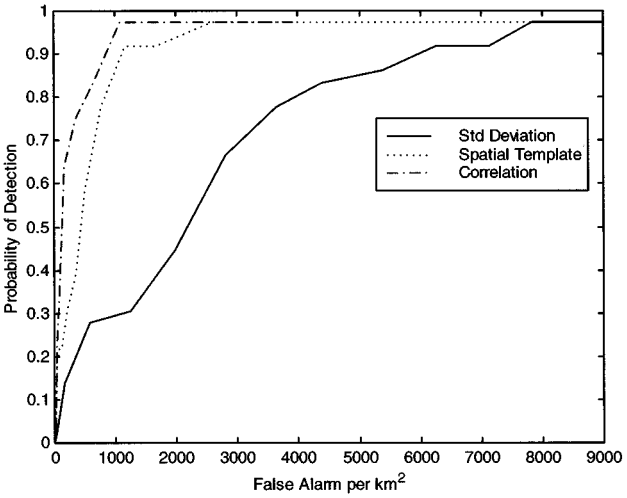


Fig. 8. Receiver operating characteristic (ROC) for three of the prescreeners, as scored by only considering the angled-entry UXO with the top of the target flush buried.

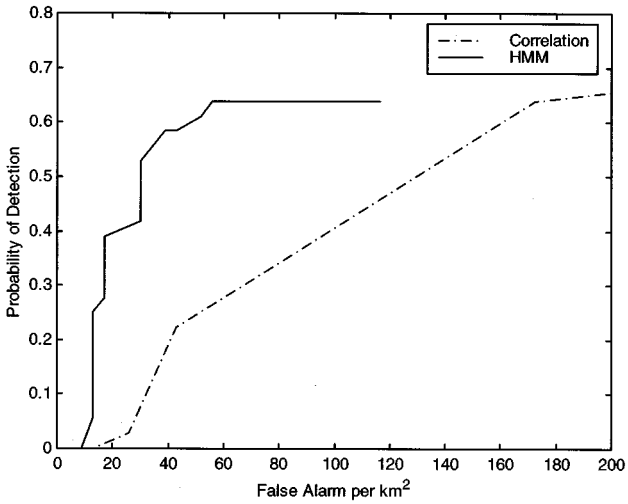


Fig. 9. As in Fig. 8, but here we consider the performance of the HMM, compared to the correlation-based prescreener.

the best performance. The advantage of pruning via a sequence of prescreeners is realized by the fact that the most complex prescreener need only be performed over a small subset of the original data.

One issue that must be discussed is that, although the ROCs in Fig. 8 demonstrate performance as a function of the detector threshold, when one transitions from one prescreener to the next, a particular threshold must be selected for the former, this dictating which parts of the image are discarded, and which go on to the next phase. This implies that the final performance of subsequent detectors may not achieve a probability of detection of unity, since targets may not be passed forward from an earlier prescreener. This explains why the correlation-based prescreener in Fig. 8 does not reach a detection probability of unity. Care must be take in selection of these prescreener thresholds, with the latter dictated by the clutter statistics.

The composite performance of the prescreeners summarized in Fig. 8 is relatively good. However, note that we are achieving a probability of detection of 0.97 with approximately 1000 false alarms/km^2. This FAR is likely too high to reliably perform UXO-range detection, especially given the relatively benign clutter environment considered here (likely to be worse for other cases). This motivates the HMM, the performance of which we now address. In Fig. 9, we plot the ROC for the HMM as scored with the angle-entry targets. For comparison, in Fig. 9, we also plot the results of the correlation detector from Fig. 8. We note that the HMM yields a substantial improvement in classifier performance, especially at the low FAR values.

The next set of ROC curves are as in Figs. 8 and 9, but now, in addition to considering the angle-entry UXO, we consider UXO parallel to the air–ground interface (both on the surface and flush-buried). With regard to the latter, we only consider those interface-parallel targets for which, for some portion of the SAR aperture, the target axis is parallel to or orthogonal to the linear sensor path. For this target class, we have a total of 54 targets. Again, we discuss in Section III-E the reasons for considering these special cases. In Figs. 10 and 11, we show the

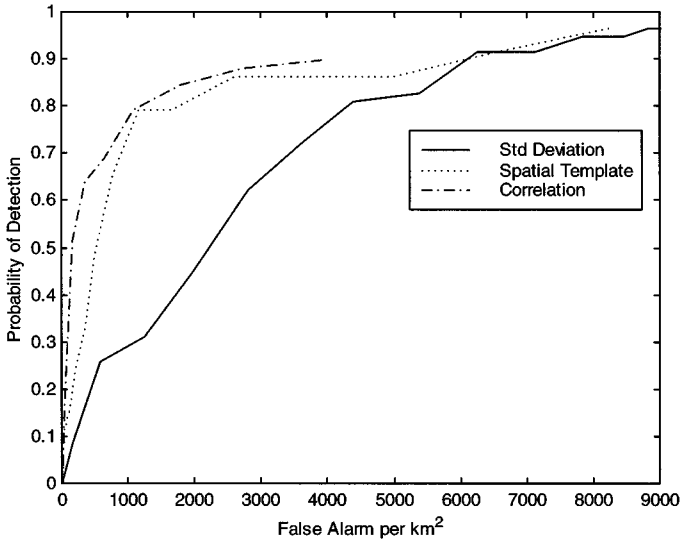


Fig. 10. Receiver operating characteristic (ROC) for three of the prescreeners, as scored by only considering the angled-entry UXO with the top of the target flush buried, as well as the UXO for which the target was either parallel or perpendicular to some part of the linear SAR aperture.

results of the prescreeners and of the HMM, as in Figs. 8 and 9, respectively. We notice that, by inclusion of this additional target class, the overall ROC performance deteriorates, although the relative trends between different algorithm types are as in Figs. 8 and 9. In particular, note in Fig. 11 that at an FAR of 40/km^2, the HMM correctly detects over 40% of the targets, with the performance of the correlation detector significantly lower (about 10%).

In Figs. 12 and 13, we plot the results of the prescreeners and of the subsequent incorporation of the HMM, as applied to all surface and flush-buried UXO, constituting a total of 84 targets. The principal difference between Figs. 12 and 13 and Figs. 10 and 11 is the inclusion of surface and flush UXO that are not at some point oriented parallel to or orthogonal to the linear SAR aperture. The performance of the classifiers is now further

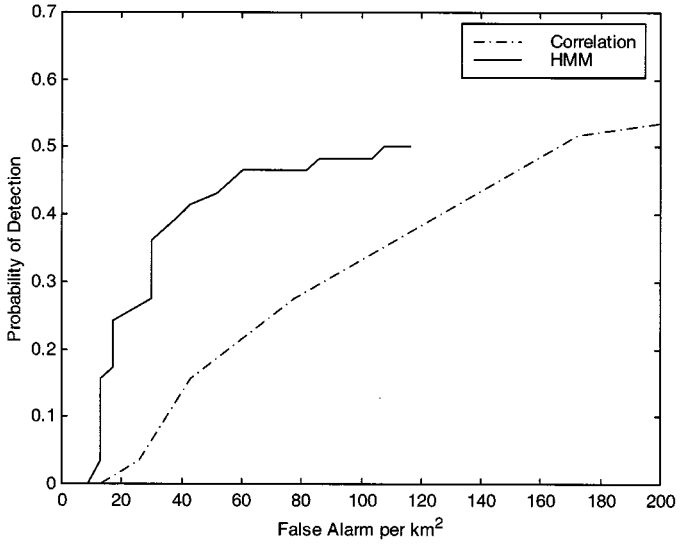


Fig. 11. As in Fig. 10, but here we consider the performance of the HMM compared to the correlation-based prescreener.

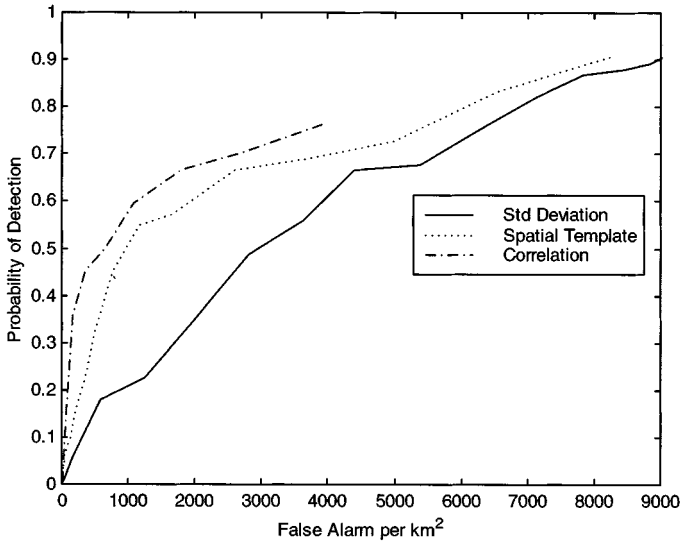


Fig. 12. Receiver operating characteristic (ROC) for three of the prescreeners, as scored by considering all surface and flush-buried UXO.

degraded, although the HMM is still correctly identifying over a fourth of the targets at a FAR of 40/km^2.

E. Discussion

We have broken the above presentation into the three portions, with the goal of delineating the types of targets for which the SAR system is particularly effective. In particular, we found that best performance was achieved for angle-entry targets, with the top of the target flush buried. We explain this phenomenon as follows. The portion of the target nearest the surface produces an end-diffracted component, that looks approximately the same from all antenna positions along the linear SAR aperture. This implies that the response from this component is approximately the same for each of the antenna positions used to form the full and subaperture images, yielding a strong signature that is relatively less sensitive to competing clutter.

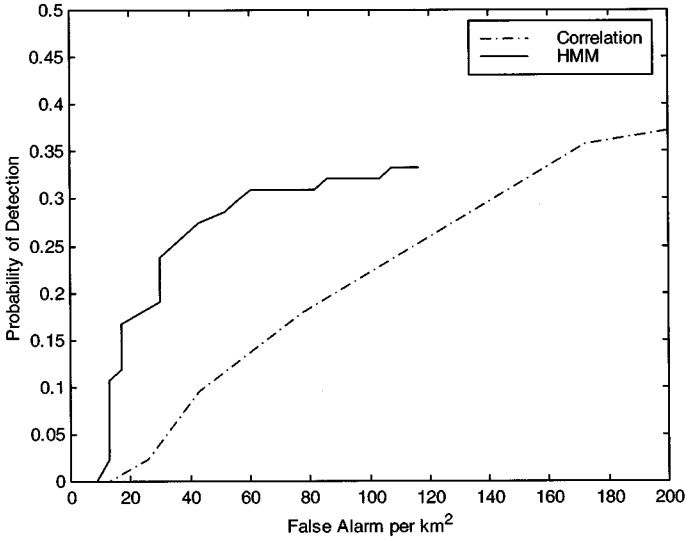


Fig. 13. As in Fig. 12, but here we consider the performance of the HMM compared to the correlation-based prescreener.

In this same context, if the target broadside is parallel to the linear SAR aperture for some antenna positions, at these positions the UXO signature will be strong, and less sensitive to clutter. Moreover, for such targets, as one moves along the sequence of subaperture images, the signature will increase as broadside excitation is approach, at which point the signature will be maximized in strength, and then the image strength will diminish as one moves away from this cardinal angle. This sequence produces a very distinctive signature, exploited effectively via the HMM. Similar issues hold for the case in which the target is orthogonal to the portion of the linear SAR aperture.

IV. Conclusions

The utility of employing UWB SAR for detection of UXO has been investigated. We have emphasized that a SAR-based system is unlikely to find all UXO, especially those buried deeply. However, at a former bombing range, there is often a large amount of surface and flush-buried UXO. The goal of the work reported here was investigation of whether a UWB SAR system can detect a "large enough" percentage of such UXO, at a "low enough" false-alarm rate, such that a former bombing range can be detected. The answer to this question involves both engineering and policy issues. We have sought to focus on the former, with the goal of developing algorithms that provide an optimal utilization of the wave physics, explicitly utilizing models that simulate electromagnetic scattering from such targets.

We have considered several prescreeners and have implemented an HMM, with all of these algorithms trained entirely on computed data. We have demonstrated that the performance is a strong function of the UXO position in the ground or on the surface. This underscores that it is salutary to view the UXO from as many target-sensor orientations as possible. Here we have considered a single linear aperture. However, the work presented here suggests that it may be desirable to consider multiple linear sensor paths, or perhaps a circular SAR, to

improve the likelihood that one or more of the favorable target-sensor orientations discussed in Section III-E is realized.

The work reported here constitutes only a beginning of more research that should be directed at this problem. For example, we have only considered imagery from a single polarization (VV). Although not presented here, we have also addressed fusing the performance of prescreeners based on both VV and HH polarized fields, and have witnessed further improvements in performance. We have not discussed such results further, because a significant issue in UWB SAR sensors involves accurate polarization calibration. This is an ongoing topic of research, but it is expected that the fusion of multiple polarizations will further improve the detection results.

The SAR-based detection of UXO, with the goal of bombing range detection, is a clutter-limited problem. In particular, the success of such a system is predicated on realizing measured data from UXO with signature amplitudes and features (as a function of target-sensor orientation) that allow one to separate targets from clutter. This implies that performance is strongly linked to the clutter characteristics. The test conditions at Yuma Proving Ground were relatively favorable in this connection, the soil being relatively low-loss and the clutter generally benign. Nevertheless, desirable performance was only achieved after implementing sophisticated detectors that explicitly accounted for the wave physics. To further quantify the utility of such a system, more challenging clutter scenarios must be considered. Moreover, since the number of different environments that can be considered empirically is limited, further insight is required into the clutter sources, with this insight transitioned to the electromagnetic models.

REFERENCES

[1] Y. Das, J. E. McFee, J. Toews, and G. C. Stuart, "Analysis of an electromagnetic induction detector for real-time location of buried objects," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 278–287, May 1990.

[2] R. W. P. King and C. W. Harrison, Jr., "The transmission of electromagnetic waves and pulses into the earth," *J. Appl. Phys.*, vol. 39, pp. 4444–4452, Aug. 1968.

[3] J. A. Fuller and J. R. Wait, "Electromagnetic pulse transmission in homogeneous dispersive rock," *IEEE Trans. Antennas Propagat.*, vol. AP-20, pp. 530–533, July 1972.

[4] D. L. Moffat and R. J. Puskar, "A subsurface electromagnetic pulse radar," *Geophysics*, vol. 41, pp. 506–518, June 1976.

[5] P. Runkle, P. K. Bharadwaj, L. Couchman, and L. Carin, "Hidden Markov models for multiaspect target classification," *IEEE Trans. Signal Processing*, vol. 47, pp. 2035–2040, July 1999.

[6] P. Runkle, L. Carin, L. Couchman, T. Yoder, and J. Bucaro, "Multiaspect target identification with wave-based matched pursuits and continuous hidden Markov models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, pp. 1371–1378, Dec. 1999.

[7] P. Runkle, L. Nguyen, J. McClellan, and L. Carin, "Multiaspect target detection for SAR imagery using hidden Markov models," *IEEE Trans. Geosci. Remote Sensing*, to be published.

[8] J. Q. He, T. J. Yu, N. Geng, and L. Carin, "Method of moments analysis of electromagnetic scattering from a general three-dimensional dielectric target embedded in a multilayered medium," *Radio Sci.*, vol. 35, pp. 305–313, Mar.–Apr. 2000.

[9] N. Geng, A. Sullivan, and L. Carin, "Multi-level fast multipole algorithm for conducting targets embedded above or within a lossy half space," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1561–1573, July 2000.

[10] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of ground penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-34, pp. 36–28, Jan. 1996.

[11] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*. New York: McGraw Hill, 1961.

[12] L. Carin, N. Geng, M. McClure, J. Sichina, and L. Nguyen, "Ultra-wide-band synthetic-aperture radar for mine-field detection," *IEEE Antennas Propagat. Mag.*, vol. 41, pp. 18–33, Feb. 1999.

[13] S. Vitebskiy, L. Carin, M. A. Ressler, and F. H. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.

[14] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84–95, Jan. 1980.

[15] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Statist.*, vol. 41, pp. 164–171, 1970.

[16] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, pp. 257–285, Feb. 1989.

[17] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. 13, pp. 260–269, Apr. 1967.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1998, vol. 2, Detection Theory.

[19] S. S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[20] D. Nandy and J. Ben-Arie, "Generalized feature extraction using expansion matching," *IEEE Trans. Image Processing*, vol. 8, pp. 22–32, Jan. 1999.

Yanting Dong (S'00) received the B.S. degree in electrical engineering and the M.E. degree in automation from Tianjin University, China, in 1993 and 1996, respectively. She received the M.S. degree in electrical engineering from Duke University, Durham, NC, in 2000, and is now pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering, Duke University.

From 1996 to 1998, she was with the Department of Automation, Tianjin University, working on electromagnetic tomography systems. Since 1999, she has been working as a Research Assistant with Duke University. Her research interests are in statistical signal processing, information theory, detection and estimation, tomographic technique, and applied electromagnetics.

Paul R. Runkle received the B.S. degree in applied mathematics and physics from the University of Wisconsin, Madison, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, MI, in 1992 and 2000, respectively.

Since 1998, he has been a Research Associate in the Department of Electrical and Computer Engineering, Duke University. His research interests include statistical signal processing and modeling, multichannel signal processing, image processing, and related applications.

Lawrence Carin (SM'96–F'01) [redacted] in Washington, DC, and earned the B.S., M.S., and Ph.D. degrees in electrical engineering at the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor, and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, Durham, NC, where he is now a Professor. He is currently the Principal Investigator on a Multidisciplinary University Research Initiative (MURI) on demining. His current research interests include short-pulse scattering, subsurface sensing, and wavebased signal processing.

Dr. Carin is an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION and is a member of Tau Beta Pi and Eta Kappa Nu.

Raju Damarla received the B.Sc. degree (hons.) in mathematics, the B.Tech. degree (hons.) in electronics and electrical communication engineering, and the M.Tech. degree in controls and computation engineering from the Indian Institute of Technology, Kharagpur, India. He received Ph.D. in engineering from Boston University, Boston, MA.

He was an Assistant Professor in the Electrical Engineering Department, University of Kentucky, Lexington, until 1994. He is currently an Electronics Engineer with the U.S. Army Research Laboratory, Adelphi, MD. His interests include signal processing, automatic target detection and recognition algorithm development, and neural networks.

Anders Sullivan (M'93) [redacted] in Staten Island, NY. He received the B.S. and M.S. degrees in aerospace engineering from the Georgia Institute of Technology, Atlanta, GA, in 1985 and 1987, respectively, and the Ph.D. degree in electromagnetics from Polytechnic University, Brooklyn, NY, in 1997.

From 1988 through May 1998, he was with the Air Force Research Laboratory, Eglin AFB, FL. From June 1998 through September 1999, he was with the Electrical Engineering Department, Duke University, Durham, NC, as a Research Associate. Since September 1999, he has been with the Army Research Lab, Adelphi, MD. His current research interests include modeling complex targets and short-pulse scattering.

Dr. Sullivan is a member of Tau Beta Pi and Sigma Gamma Tau.

Marc A. Ressler received the B.S. degree in electrical engineering from the University of Maryland, College Park, in 1969, and the M.S. degree in electrical and computer engineering from the University of Michigan, Ann Arbor, in 1973.

He has been with the Army Research Laboratory (ARL) and one of its predecessors, the Harry Diamond Laboratories, since 1969, where he worked on a variety of programs including battlefield data processing and fusion, and low-frequency and microwave radar systems.

Mr. Ressler is currently assigned to the Sensors and Electron Devices Directorate of ARL, where his area of interest is low frequency, ultra-wideband radar for foliage and ground penetration applications.

Jeffrey Sichina received the B.S.E.E. degree in 1975 from the University of Pittsburgh, Pittsburgh, PA. He subsequently undertook graduate study at the University of Maryland, College Park.

In 1974, he joined Harry Diamond Laboratories, Adelphi, MD, a component of the U.S. Army Materiel Command. From 1974 through 1985, he was a Staff Engineer, working on a variety of electronic fuzing programs. His primary research interest included digital signal processing, FM/CW systems, and electronic counter-counter measures. In 1988, he joined the Radar Branch, and was responsible for development of a miniature moving-target-indicator radar for unmanned air vehicles. In 1992, he was named Chief of the Radar Branch and was responsible for a variety of long-term applied research undertakings within the Army. Among these was an effort aimed at determining the effectiveness of ultra-wideband (UWB) radar for detection of concealed objects such as buried mines and UXO. His current research interests include algorithm development and detection theory.

Mr. Sichina is a member of Eta Kappa Nu.

Fast Multipole Method for Scattering from an Arbitrary PEC Target Above or Buried in a Lossy Half Space

Norbert Geng, *Member, IEEE,* Anders Sullivan, *Member, IEEE,* and Lawrence Carin, *Fellow, IEEE*

***Abstract*—The fast multipole method (FMM) was originally developed for perfect electric conductors (PECs) in free space, through exploitation of spectral properties of the free-space Green's function. In the work reported here, the FMM is modified, for scattering from an arbitrary three-dimensional (3-D) PEC target above or buried in a lossy half space. The "near" terms in the FMM are handled via the original method-of-moments (MoM) analysis, wherein the half-space Green's function is evaluated efficiently and rigorously through application of the method of complex images. The "far" FMM interactions, which employ a clustering of expansion and testing functions, utilize an approximation to the Green's function dyadic via real image sources and far-field reflection dyadics. The half-space FMM algorithm is validated through comparison with results computed via a rigorous MoM analysis. Further, a detailed comparison is performed on the memory and computational requirements of the MoM and FMM algorithms for a target in the vicinity of a half-space interface.**

***Index Terms*—Method of moments, scattering.**

I. Introduction

RESEARCHERS have directed significant attention toward development of numerical scattering models, with the method of moments (MoM) representing a prominent example [1]–[7]. For an MoM problem with N unknowns, the memory requirements are of order $O(N^2)$, while the computational complexity depends on whether a direct (LU-decomposition) or iterative (conjugate gradient [2], [8]) matrix solver is applied, the former requiring $O(N^3)$ and the latter $P \cdot O(N^2)$ computations, with P the number of iterations required to achieve convergence. To reduce computational complexity, several researchers have undertaken the development of modified integral-equation solvers [9]–[17], in an effort to achieve better computational efficiency and reduced memory requirements. For example, the adaptive integral method (AIM) [12] exploits properties of the FFT, and researchers have demonstrated memory requirements and computational complexity of $O(N^{1.5})$ and $O(N^{1.5}\log N)$, respectively [12]. Further, there has been significant interest in the fast multipole method (FMM). The simplest two-level (single-stage) FMM [13], [14] has $O(N^{1.5})$ computational complexity and memory requirements, while a multilevel extension [15]–[17] reduces these numbers to $O(N \log N)$.

While the AIM and FMM algorithms represent promising tools, they have heretofore been applied primarily to free-space scattering [12], [13]–[17], two-dimensional (2-D) analysis in layered media [18], or quasi-planar three-dimensional (3-D) problems in circuit and antenna design [19]. There are many applications for which the free-space model or a quasi-planar approximation are inappropriate. For example, there has been significant interest in radar sensing of buried targets, such as mines [20], unexploded ordnance (UXO) [21], drums [22], etc. In this paper we concentrate on the two-level FMM, with the goal of extending it to the case of a target in the vicinity of a lossy half space.

The FMM clusters the MoM basis elements into M groups [13], [14], and the interactions between distant groups ("far" interactions) are handled by exploiting features of the FMM spectral propagator [13], [14], discussed further below. "Near" interactions are handled in a manner analogous to a conventional MoM analysis [3], [5], [7]. For the calculation of these "near" interactions, we rigorously evaluate the dyadic half-space Green's function [6], [25] via the complex-image technique [26]–[29].

We have developed an approximate but highly accurate method for evaluating the "far" FMM interactions within the context of a half-space problem. In particular, if the target is entirely above or below the half space, the MoM and FMM analyses only require the Green's function for source and observer in the same region. Under this circumstance, each component of the dyadic Green's function can in general be represented in terms of a direct and reflected term [30], [31]. The direct term is handled as in the conventional free-space FMM [13], [14]. Further, the reflected term corresponds to radiation from the expansion function, reflection at the half-space interface, and subsequent propagation to the testing function. If the expansion and testing functions are relatively distant, than this term can be evaluated approximately asymptotically [30], [31]. The reflected term can therefore be represented in terms of an image in real space, with an appropriate amplitude, dictated by the polarization-dependent reflection coefficient.

II. Theory

A. Integral Equation and MoM Formulation

We utilize an electric-field integral equation (EFIE) formulation [2], [3], [5], which restricts its utility somewhat, due to

Manuscript received January 5, 1999; revised September 27, 2000.

N. Geng is with the University of Karlsruhe, Institute for Microwaves and Electronics, D-76128 Karlsruhe, Germany.

A. Sullivan and L. Carin are with the Duke University, Department of Electrical and Computer Engineering, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0018-926X(01)03184-2.

0018–926X/01$10.00 © 2001 IEEE

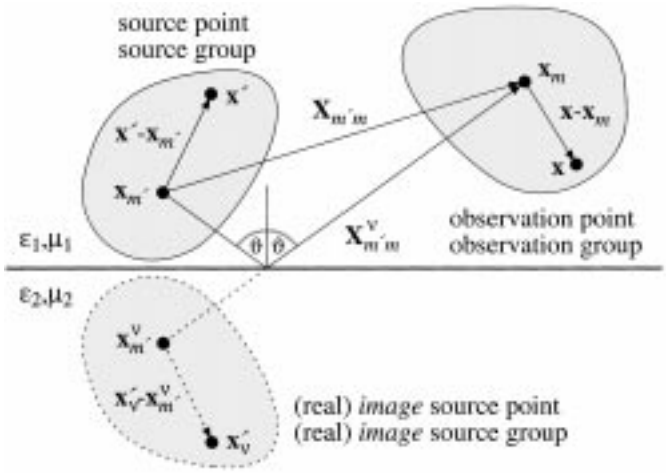


Fig. 1. Geometry for source and observation group in 3-D FMM and generalization to a half-space environment using real images to account for "far" interface interactions.

spurious internal resonances (for closed targets) [2]. However, it has been demonstrated previously [23] that when a closed target is embedded in a lossy medium (e.g., soil), these internal resonances are damped significantly and they therefore present less of a problem computationally. We note, however, that the general FMM formulation can also be applied to a combined field integral equation (CFIE) [2], [15], which is not beset by internal-resonance problems.

The half-space EFIE for a perfectly conducting 3-D target situated entirely in layer i=1 (e.g., air) or layer $i = 2$ (e.g., soil) of a half-space environment (Fig. 1) can be expressed as [6]

$$\hat{\boldsymbol{n}} \times \boldsymbol{E}^{\text{inc}}(\boldsymbol{x}) = \hat{\boldsymbol{n}} \times j\omega\mu_i \iint_{S'} \left[\boldsymbol{I} + \frac{\nabla\nabla}{k_i^2}\right] \cdot \boldsymbol{G}_{\text{Aii}}(\boldsymbol{x}, \boldsymbol{x}') \cdot \boldsymbol{J}(\boldsymbol{x}')\, dS' \tag{1a}$$

or, alternatively, in the mixed potential form of the electric field integral equation (MPIE) [6]

$$\hat{\boldsymbol{n}} \times \boldsymbol{E}^{inc}(\boldsymbol{x}) = \hat{\boldsymbol{n}} \times \left[j\omega\mu_i \iint_{S'} \boldsymbol{K}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') \cdot \boldsymbol{J}(\boldsymbol{x}')\, dS' - \frac{\nabla}{j\omega\varepsilon_i} \iint_{S'} K^{ii}_{\phi e}(\boldsymbol{x}, \boldsymbol{x}') \nabla' \cdot \boldsymbol{J}(\boldsymbol{x}')\, dS' \right] \tag{1b}$$

where $\boldsymbol{x}$ is on and $\hat{\boldsymbol{n}}$ is a unit vector normal to the target surface; $\epsilon_i = \epsilon_i' - j\sigma/\omega$, μ_i and k_i represent (in general complex) permittivity, permeability, and wavenumber of the medium in which the target resides (region i); ω is the angular frequency ($\exp(j\omega t)$ time dependence is assumed and suppressed); and $\boldsymbol{I}$ represents the unit dyadic. Details on the dyadic half-space Green's functions $\boldsymbol{G}_{Aii}$ and $\boldsymbol{K}_{Aii}$, and on the scalar Green's function $K^{ii}_{\phi e}$, have been given by Michalski and Zheng [6], where we use their formulation C. The EFIE in (1) is valid for a general layered medium [6], but here we only consider a half-space, for simplicity (see also explanation in Section II-C).

Like in a conventional MoM solution, the unknown surface current $\boldsymbol{J}$ is expanded by a set of basis (or expansion) functions $\boldsymbol{b}_{n'}(\boldsymbol{x}')$, where we use the well-known RWG vector basis functions introduced in [3], defined on adjacent triangles representing the target's surface. Testing the EFIE (1) with a set of weighting (or testing) functions $\boldsymbol{w}_n(\boldsymbol{x})$ tangential to the scatterer surface results in N linear equations $[\boldsymbol{Z}]\cdot\boldsymbol{I} = \boldsymbol{V}$ for the unknown current coefficients $\boldsymbol{I}_{n'}$, where the $N \times 1$ driving vector $\boldsymbol{V}$ represents the tangential electric field of the incident wave, tested on the target surface. The elements of the $N \times N$ impedance matrix $[\boldsymbol{Z}]$ are given by

$$Z_{nn'} = j\omega\mu_i \iint_S \iint_{S'} \boldsymbol{w}_n(\boldsymbol{x}) \cdot \left[\boldsymbol{I} + \frac{\nabla\nabla}{k_i^2}\right] \cdot \boldsymbol{G}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') \cdot \boldsymbol{b}_{n'}(\boldsymbol{x}')\, dS'\, dS \tag{2a}$$

or

$$Z_{nn'} = j\omega\mu_i \iint_S \boldsymbol{w}_n(\boldsymbol{x}) \cdot \iint_{S'} \mathbf{K_{Aii}}(\boldsymbol{x}, \boldsymbol{x}') \cdot \boldsymbol{b}_{\mathbf{n}'}(\boldsymbol{x}')\, \mathbf{dS'}\, \mathbf{dS} - \frac{1}{j\omega\varepsilon_i} \iint_S \boldsymbol{w}_n(\boldsymbol{x}) \cdot \nabla \iint_{S'} \mathbf{K}^{\mathbf{ii}}_{\boldsymbol{\phi e}}(\boldsymbol{x}, \boldsymbol{x}') \nabla' \cdot \boldsymbol{b}_{n'}(\boldsymbol{x}')\, dS'\, dS. \tag{2b}$$

Using the RWG triangular basis [3] and Galerkin testing, the last term in (2b) can be simplified using integration by parts [3].

In the FMM analysis [13], [14], we divide the computation of (expansion function)-(testing function) interactions into "near" and "far" terms. The "far" interactions are best handled via the form in (1a) and (2a), respectively, while for "near" interactions (MoM part of the FMM [13], [14]) (1b) and (2b) are utilized, rigorously accounting for the dyadic half-space Green's function [6]. The half-space Green's function is evaluated using the method of complex images [26]–[29], thereby avoiding direct numerical evaluation of Sommerfeld integrals [25]. Impedance matrix elements representing these "near" interactions are stored in a sparse matrix denoted $[\boldsymbol{Z}^{\text{near}}]$. The method of complex images for calculating the space-domain Green's function, as well as explicit equations for the impedance matrix elements, including the extraction of self-term singularities, can be found in the literature [26]–[29], [3]–[5].

"Far" interactions, applicable to the FMM framework [13], [14], are best described using (2a) and realizing that [6], [30]

$$\boldsymbol{G}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{I} g_i(\boldsymbol{x}, \boldsymbol{x}') + \boldsymbol{\Delta G}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') = \boldsymbol{I} \frac{e^{-jk_i|\boldsymbol{x}-\boldsymbol{x}'|}}{4\pi|\boldsymbol{x}-\boldsymbol{x}'|} + \boldsymbol{\Delta G}_{Aii}(\boldsymbol{x}, \boldsymbol{x}') \tag{3}$$

where the new dyadic $\boldsymbol{\Delta G}_{Aii}$ is defined as the difference between $\boldsymbol{G}_{Aii}$ and $\boldsymbol{I} g_i$ (i.e., $\boldsymbol{\Delta}$ is not an operator). The dyadic $\boldsymbol{G}_{Aii}$ is therefore split into a part which accounts for direct interactions (as in the free-space case, but using a homogeneous background medium with in general complex wave number k_i) and a dyadic accounting for interactions with the interface. Consequently, the matrix elements in (2a) for "far" interactions are also split as $Z^{\text{far}}_{nn'} = Z^{far,hom}_{nn'} + \boldsymbol{\Delta} Z^{\text{far}}_{nn'}$ with $Z_{nn'} = Z^{\text{near}}_{nn'} + Z^{\text{far}}_{nn'}$. As discussed above, $Z^{\text{far}}_{nn'}$ is computed using the form

in (2a), with $Z_{nn'}^{far,hom}$ employing the homogeneous-medium Green's function g_i and $\Delta Z_{nn'}^{\text{far}}$ employing ΔG_{Aii} [see (3)].

B. Free-Space Fast Multipole Method

As is well known from [13], [14], in the 3-D free-space FMM the scalar Green's function $g_i(\boldsymbol{x},\boldsymbol{x}')$ is expanded using the addition theorem [32] resulting in the plane wave decomposition

$$g_i(\boldsymbol{x},\boldsymbol{x}') = \frac{-jk_i}{(4\pi)^2}\iint_{4\pi} e^{-jk_i\hat{\boldsymbol{k}}\cdot(\boldsymbol{x}-\boldsymbol{x}_m)}T_L \cdot (k_iX_{m'm},\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{X}}_{m'm})e^{+jk_i\hat{\boldsymbol{k}}\cdot(\boldsymbol{x}'-\boldsymbol{x}_{m'})}\,d^2\hat{\boldsymbol{k}} \tag{4a}$$

$$T_L(k_iX_{m'm},\hat{\boldsymbol{k}}\cdot\boldsymbol{X}_{m'm}) = \sum_{l=0}^{L}(-j)^l(2l+1)h_l^{(2)}(k_iX_{m'm})P_l(\hat{\boldsymbol{k}}\cdot\boldsymbol{X}_{m'm}) \tag{4b}$$

where the distance vector $\boldsymbol{R} = |\boldsymbol{x}-\boldsymbol{x}'|$ has been subdivided into a vector from the source point $\boldsymbol{x}'$ to the center $\boldsymbol{x}_{m'}$ of a corresponding source group, a vector from an observation group center $\boldsymbol{x}_m$ to the observation point $\boldsymbol{x}$, and a vector $\boldsymbol{X}_{m'm}$ connecting the group centers (Fig. 1), i.e., $\boldsymbol{R} = \boldsymbol{x}-\boldsymbol{x}' = (\boldsymbol{x}-\boldsymbol{x}_m)+(\boldsymbol{x}_m-\boldsymbol{x}_{m'}-(\boldsymbol{x}'-\boldsymbol{x}_m) = (\boldsymbol{x}-\boldsymbol{x}_m)+\boldsymbol{X}_{m'm}-(\boldsymbol{x}'-\boldsymbol{x}_{m'})$. Therefore, the target surface is first partitioned into groups $m = 1,\cdots,M$, each of which has an average number of $A_m \approx N/M$ basis/weighting functions. Inside group m the elements are labeled as $\alpha = 1,\cdots,A_m$ [13], [14]. The group information $n(m)$, is stored in matrix form. As was shown in [13] and [14], for the two-level FMM described here, the optimal number of groups (to minimize CPU and RAM) is given by $M \propto \sqrt{N}$. Approximate equations for the number of terms L needed in (4b) are given in [13], [14], [17] for the case of a lossless environment with $k_i = k_0$ (vacuum). For a lossy background medium (e.g., soil) more terms are required, as discussed further in Section II-C. Therefore, in our implementation L is determined adaptively, depending on the FMM grouping scheme and the (complex) wavenumber. For the accurate numerical integration over the solid angle 4π in (4a), a quadrature rule with $K = 2L^2$ points is applied [13], [14]; although more efficient quadrature rules are available [33].

C. Extension to Half-Space Fast Multipole Method

In addition to the evaluation of the dyadic Green's function in the calculation of $[\mathbf{Z}^{\text{near}}]$ (here using the method of complex images [26]–[29]), it is essential to include the effects of the "far" interface interactions represented by $[\Delta\boldsymbol{Z}^{\text{far}}]$ [see (3)]. Using complex images, each term ΔG_{Aii}^{xx}, ΔG_{Aii}^{zz}, and Q_{Aii} in the space-domain half-space dyadic Green's function [6]

$$\Delta G_{Aii}(\boldsymbol{x},\boldsymbol{x}') = (\hat{\boldsymbol{x}}\hat{\boldsymbol{x}}+\hat{\boldsymbol{y}}\hat{\boldsymbol{y}})\Delta G_{Aii}^{xx} + \left(\hat{\boldsymbol{z}}\hat{\boldsymbol{x}}\frac{\partial}{\partial x}+\hat{\boldsymbol{z}}\hat{\boldsymbol{y}}\frac{\partial}{\partial y}\right)Q_{Aii} + \hat{\boldsymbol{z}}\hat{\boldsymbol{z}}\Delta G_{Aii}^{zz} \tag{5}$$

can be expressed in terms of a sum of free-space Green's functions [26]–[29] with image sources located in complex space.

The plane wave expansion (4) used in the 3-D free-space FMM remains valid for this more general case of a complex wavenumber k_i and complex source points, although the convergence is slower. This is especially true for the complex distance vectors. While complex wavenumbers $k = k_i$ can be handled by slightly increasing the number L of terms in (4) (at least for moderate losses, see Fig. 2), we have not yet found a solution to avoid the generally large upper limit L needed for general complex source points. To further illustrate this, we consider the addition theorem [13], [14], [32]

$$\frac{e^{-jkR}}{R} = \frac{e^{-jk\sqrt{(\boldsymbol{X}+\boldsymbol{d})\cdot(\boldsymbol{X}+\boldsymbol{d})}}}{\sqrt{(\boldsymbol{X}+\boldsymbol{d})\cdot(\boldsymbol{X}+\boldsymbol{d})}} = -jk\sum_{l=0}^{L=\infty}(-1)^l(2l+1)j_l(kd)h_l^{(2)}(kX)P_l(\hat{\boldsymbol{d}}\cdot\hat{\boldsymbol{X}}) \tag{6}$$

on which the plane wave expansion (4) is based. Theoretically, the expansion in (4) converges as long as $|d| < |X|$ is satisfied, wherein the distance R, wavenumber k, $X = \sqrt{\boldsymbol{X}\cdot\boldsymbol{X}}$, and $d = \sqrt{\boldsymbol{d}\cdot\boldsymbol{d}}$ can be arbitrarily complex [32]. As discussed in [13], for real arguments (i.e., real wavenumber k, real source point $\boldsymbol{x}'$, real group center $\boldsymbol{x}_{m'}$, real vectors $\boldsymbol{d}$ and $\boldsymbol{X}$), the spherical Bessel function $j_l(kd)$ and spherical Hankel function of second kind $h_l^{(2)}(kX)$ are approximately of constant magnitude for $l < kd$ and $l < kX$, respectively. The lth-order Legendre polynomial is strictly bounded according to $|P_l(\hat{\boldsymbol{d}}\cdot\hat{\boldsymbol{X}})| \le 1$. For $l > kd$ and $l > kX$, respectively, $j_l(kd)$ decays quickly and $h_l^{(2)}(kX)$ grows quickly. Consequently, L must be chosen large enough such that convergence is achieved, but it must not greatly exceed kX, or the sum becomes numerically unstable. For the free-space FMM with real arguments (e.g., $k = k_0$) it has been shown in [13], [14], [17] that, assuming a maximum cluster diameter d, approximately $L = kd+\gamma\,\ln(\pi+kd)$ terms in the series expansion are sufficient; with $\gamma \approx 1$ for a relative error less than $10^{-1} = 10\,\%$, $\gamma \approx 3$ for 10^{-3} and $\gamma \approx 5$ for 10^{-6}. If two clusters are situated such that $kX < L$, the clusters are too close for the representation in (4a) and (4b) to be valid [13] and therefore the interactions between those clusters

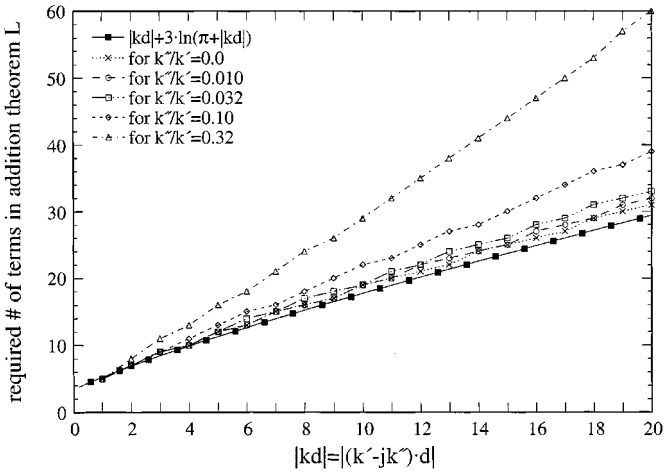


Fig. 2. Number of terms needed in (4) as a function of $|kd|$ for a relative error smaller than 10^{-3}, depending on the ratio k''/k' for a complex wavenumber $k = k'-jk''$. The empirical formula $L = kd+\gamma\,\ln(\pi+kd)$ is also plotted, with kd replaced by $|kd|$, is also plotted for comparison. Here only the direct interaction (i.e., free-space, but lossy background) is considered.

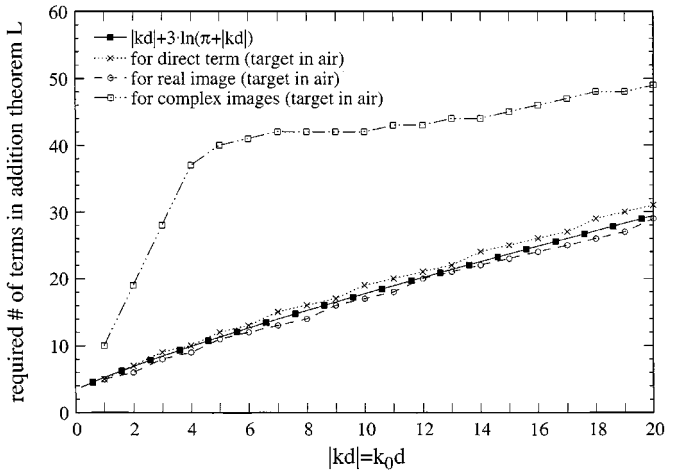


Fig. 3. Number of terms needed in (4) as a function of $|kd| = k_0 d$ for a relative error smaller than 10^{-3} for a scatterer above soil with relative permittivity $\epsilon_r = 6 - j1$. The source and observation group centers are both located at $z_{m'} = z_m = 0.4d$ and the separation was set to $k_0 X_{m'm} = L$. In addition to the number of terms required for the direct interaction, the results are given for a real image when considering all complex images.

have to be treated in the sparse part [$\mathbf{Z}^{\text{near}}$] of the original MoM impedance matrix (see Section II-A).

For image sources in complex space at least one of the vectors $\boldsymbol{X}$ or $\boldsymbol{d}$ is complex (depending on the specific choice of the image group center $\boldsymbol{x}_{m'}$) and the condition $L = kd + \gamma \ \ln\ (\pi + kd)$ is no longer valid, which is also the case for real image sources but a lossy background [i.e., complex wavenumber $k = k' - jk''$)]. Fig. 2 illustrates the increase in L for a complex wavenumber, depending on the relative losses of the surrounding medium (represented by the ratio k''/k'), where here we truncate the series expansion for a relative error less than 10^{-3}, and the distance between the group centers is set to $|kX| = L$. For moderate losses, the required number of terms in the addition theorem only increases slightly. In contrast, the (in general) significant increase of L for complex images is illustrated in Fig. 3, where we plot the required number of terms L for a target above soil of relative permittivity $\epsilon_r = 6 - j1$. The source and observation group centers are both located at $z_{m'} = z_m = 0.4d$, separated by a distance which satisfies $|k_1 X_{m'm}| = k_0 X_{m'm} = L$. Remembering that $K = 2L^2$ terms are utilized in the numerical evaluation of (4a), the computational and memory requirements increase dramatically when using complex images.

Therefore, we searched for an alternative (though approximate) formulation. Considering the fact that the dyadic half-space Green's function is rigorously accounted for in the "near" interaction matrix [$\mathbf{Z}^{\text{near}}$], using a far-field reflection coefficient approximation [30], [31] to include "far" interface interactions seems appropriate (as demonstrated in Section III). Under this assumption it is sufficient to utilize a single real image to approximately account for the reflection at the interface [30], [31]. There is only one set of real image clusters, as opposed to dyadic-component-dependent multiple complex-image locations. This property reduces the overall CPU and memory requirement. In fact, the storage and CPU are approximately twice that of a free-space FMM (see Section III), due to the extra set of image clusters. The location of the (approximate) real image at $(x', y', -z')$ (assuming the interface at $z = 0$) and the corresponding real image source group are easily determined (Fig. 1). Generalizing the free-space FMM algorithm is now straightforward. In the preprocessing stage [13], [14] we only have to include additional calculations of the translation operators between all image group and observation group centers and the spectral Fourier transforms of the image expansion functions.

To account for the polarization-dependent far-field reflection of a wave traveling from the source point $\boldsymbol{x}'$ to the observation point $\boldsymbol{x}$, we introduce the asymptotic expression for the reflection dyadic [30], [31]

$$\boldsymbol{r}^v(\boldsymbol{x}, \boldsymbol{x}') = \hat{\boldsymbol{h}}\hat{\boldsymbol{h}} R^v_{TE}(\vartheta) + [\boldsymbol{I} - \hat{\boldsymbol{h}}\hat{\boldsymbol{h}}] R^v_{TM}(\vartheta)$$

with

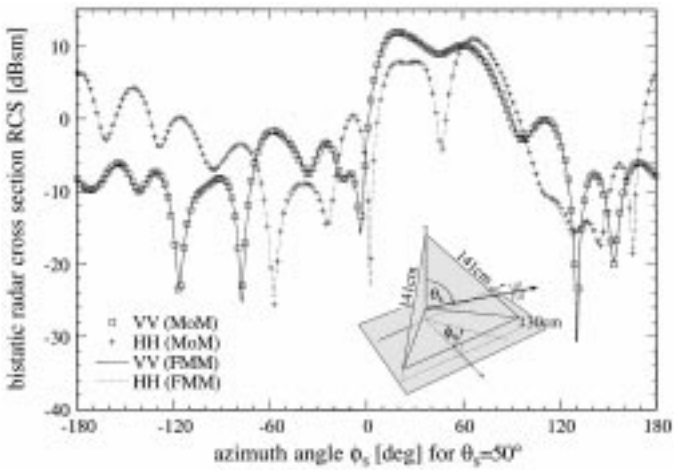
$$\hat{\boldsymbol{h}} = \frac{\hat{\boldsymbol{z}} \times (\boldsymbol{x} - \boldsymbol{x}')}{|\hat{\boldsymbol{z}} \times (\boldsymbol{x} - \boldsymbol{x}')|} \tag{7}$$

where ϑ represents the angle between the ray-optical reflection path and the unit vector $\hat{\boldsymbol{z}}$ normal to the interface (Fig. 1). The material properties and the angle ϑ determine the Fresnel reflection coefficients $R^v_{TE,TM} = R^{12}_{TE,TM}$ for a target situated in layer i=1 (e.g., air) or $R^v_{TE,TM} = R^{21}_{TE,TM}$ for a target in layer $i = 2$ (e.g., soil), respectively. The polarization is denoted as TE_z (electric field perpendicular to the plane of incidence, i.e., parallel to the unit vector $\hat{\boldsymbol{h}}$) and TM_z (magnetic field perpendicular to the plane of incidence), respectively.

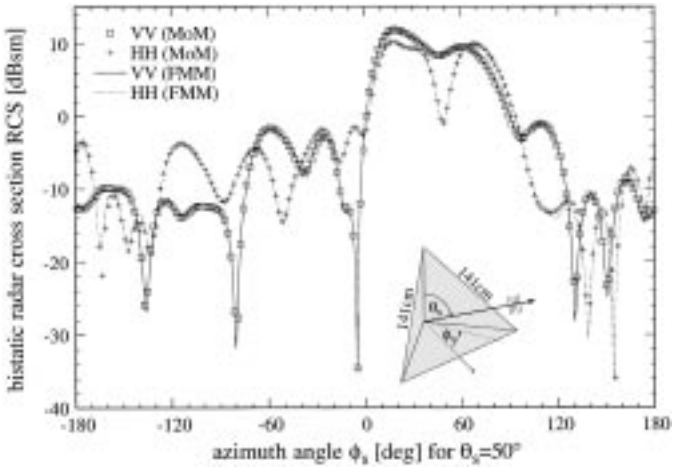
III. Results

In all examples below, the scattered fields are computed via the FMM algorithm developed here, as well as with a rigorous MoM solution [7], wherein the dyadic Green's function is characterized rigorously via the method of complex images [26]–[29]. The accuracy, computational complexity, and memory requirements of the FMM are calibrated through comparison with the MoM algorithm. The FMM model discussed here is easily modified to the free-space case [13], [14]; we therefore consider the first target (trihedral) in free space as well, such that the accuracy of the free-space FMM (vis-a-vis the MoM) can be assessed relative to the accuracy of the half-space FMM. This is particularly important because the free-space FMM makes no approximation to the Green's function beyond those in (4), while the half-space FMM code developed here makes an additional approximation (real images) to the half-space Green's function (for the "far" FMM interactions).

Concerning the FMM, recall that $M \propto \sqrt{N}$ clusters are optimal [13], [14], where N is the number of unknowns. For simplicity, we have used $M = 2^q$ FMM clusters, where q is an integer. While this approach is slightly suboptimal for the two-level FMM applied here, a similar clustering is employed in the multilevel FMM [15]–[17]. As mentioned in Section II-B, the number of terms L in (4b) is determined adaptively, where, for the problems considered here, L is typically in the range 6–15 (when truncating the series for a relative error of 10% or less, as suggested in [14], [17]). With regard to the MoM, the number of complex-image terms is also determined adaptively, and for the problems considered here 5–15 complex images are



(a)



(b)

Fig. 4. Radar cross section of a 90°-trihedral of edge length 100 cm (hypotenuse = 141 cm) with one plate parallel to the xy-plane (see inset). The incident plane wave direction is given by $\phi_i = 20°$ (relative to trihedral symmetry plane) and $\theta_i = 60°$ (30° from grazing). The figure shows the bistatic RCS (here we only plot VV and HH) for $\theta_s = 50°$ and a frequency of 1 GHz (number of unknowns $N = 5202$). (a) Trihedral 30 cm above lossy Yuma soil with 10% water content [20], [22]. (b) Trihedral in free-space.

typical. We reiterate that the "near" FMM terms employ a complex-image-computed Green's function. Finally, in the MoM solution, we have used a direct (LU-decomposition) solver of order $O(N^3)$ complexity, which could be reduced to $P \cdot O(N^2)$ if a conjugate-gradient solver [2], [8] was applied.

In all the computations, we have assumed linearly polarized plane-wave incidence. The scattered fields are computed in the far zone, via an asymptotic approximation to the half-space dyadic Green's function [30], [34].

A. Trihedral Fiducial Target

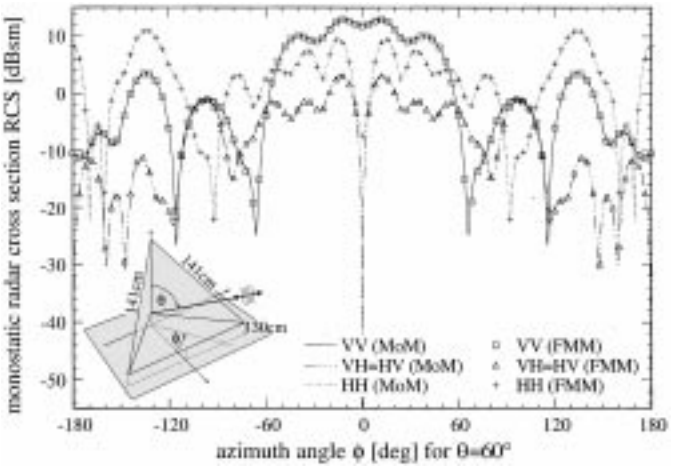
We first consider the trihedral target inset in Fig. 4(a), situated above a lossy half space. The soil is characterized as Yuma soil, with 10% water content [20], [22]. As depicted in Fig. 4(a), the trihedral is composed of three 45°-45°-90° triangles, each with a 1.41 m hypotenuse. The trihedral is placed 30 cm above the soil, with one of its triangles parallel to the air–ground interface. Such that the code is tested thoroughly, we first present bistatic results for incidence angles $\theta_i = 60°$ and $\phi_i = 20°$, and the scattered fields are observed for $\theta_s = 50°$ and $-180° \leq \phi_s \leq 180°$. The results in Fig. 4(a) present the bistatic scattered fields for VV and HH polarization, for both the FMM and MoM solutions, at a frequency of 1 GHz. At this frequency the hypotenuse is $4.7\lambda_0$, where λ_0 is the free-space wavelength. We see that the agreement between the FMM and MoM results is excellent. Similar agreement was found for the cross-polarized (VH and HV) fields, but these are not shown here, such that Fig. 4(a) is easy to read.

While the agreement between the MoM and FMM results is quite good, there are slight discrepancies in some places. While the MoM has few approximations, the FMM algorithm has three additional sources of errors. First, considering (4), the integration over the unit sphere in (4a) is approximated via a $K = 2L^2$ point quadrature rule [13]. Further, the sum in (4b) must be truncated appropriately, as discussed in Section II-C. These same approximations are present in the free-space FMM algorithm [13], [14]. The half-space problem, as modeled here, introduces a third source of error. In particular, we have approximated the half-space Green's function (for the "far" FMM terms) via real images, with appropriate polarization-dependent reflection coefficients. To examine more closely the source of errors in the FMM results of Fig. 4(a), we consider the same trihedral, situated in free space. If the agreement between the MoM and FMM algorithms is better for the free-space case than for the half-space results in Fig. 4(a), then we conclude that the half-space Green's function approximation applied here is inadequate.

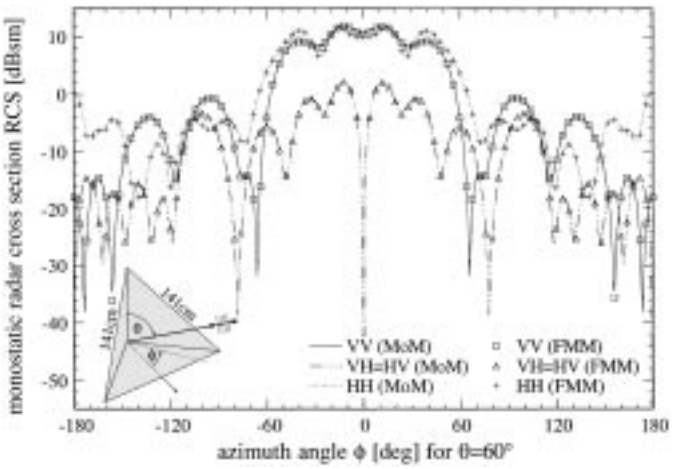
In Fig. 4(b) we consider the same case as in Fig. 4(a) but the soil half space is absent. We see by comparing Fig. 4(a) and (b) that the free-space and half-space FMM results demonstrate the same level of accuracy, relative to the MoM computations. This yields further confidence in the accuracy of the real-image approximation to the half-space Green's function, which has been examined previously in the context of the MoM [31]. With regard to the FMM results in Fig. 4(a), we also examined the solution when the reflection term was eliminated (i.e., neglecting reflections in "far" interactions), to see if the real image points were actually important. We found that the accuracy of the half-space FMM results deteriorated dramatically if the real image term was ignored, underscoring the importance and accuracy of this term.

The bistatic results in Fig. 4 provide a thorough test of the half-space FMM's accuracy, but the monostatic RCS is of more interest for practical radar systems. In Fig. 5 we consider the MoM- and FMM-computed backscattered RCS from the same trihedral as considered above, for the half-space and free-space cases, respectively. We consider a frequency of 600 MHz, with angles of incidence and scatter $\theta = 60°$ and $-180° \leq \phi \leq 180°$. We see that for these backscatter results, the agreement between the FMM and MoM results is almost perfect, for both the free-space and half-space cases.

In comparing the results in Figs. 4 and 5, we note that there is often a noticeable difference between the signature of the trihedral in free space and above soil. This is an important issue for radar calibration via such fiducial targets, since most previous results for trihedral scattering have been for the free-space case [7], [35], [36]. The importance of the FMM becomes even more evident as the size of the trihedral increases. For example,



(a)



(b)

Fig. 5. Backscattering (monostatic RCS) for the trihedral in Fig. 4 at a frequency of 600 MHz (number of unknowns $N = 1770$), for varying $\phi = \phi_i = \phi_s$ and $\theta = \theta_i = \theta_s = 60°$. (a) Trihedral 30 cm above lossy Yuma soil with 10% water content [20], [22]. (b) Trihedral in free-space.

ultra-wideband synthetic aperture radar (SAR) systems operate over frequencies from 50 to 1200 MHz [20], [22], for which very large trihedrals are required, to account for the lowest frequencies. The efficacy of applying the FMM to this problem has been examined in detail, in a separate paper [37].

Before leaving the trihedral target, we examine the memory requirements and computational speed of the FMM and MoM algorithms, for the results in Fig. 4. In Fig. 6 we compare the FMM and MoM computation speed and memory requirements, for the trihedral in free space as well as over the soil, as a function of the number of unknowns (approximately equal to 1.5 times the number of triangles [3]). The smallest number of unknowns $N = 220$ corresponds to a frequency of 200 MHz, while the largest $N = 9522$ corresponds to a frequency of 1.4 GHz. We see that, for the MoM solution, the computation time for the free-space and half-space problems coalesce after N becomes sufficiently large, representing the point at which the $O(N^3)$ LU-decomposition overwhelms the $O(N^2)$ matrix fill time. These results were computed on a HP9000/C240 workstation. The results in Fig. 6 also demonstrate the significant memory savings gained by the FMM, at $N > 2500$ for the

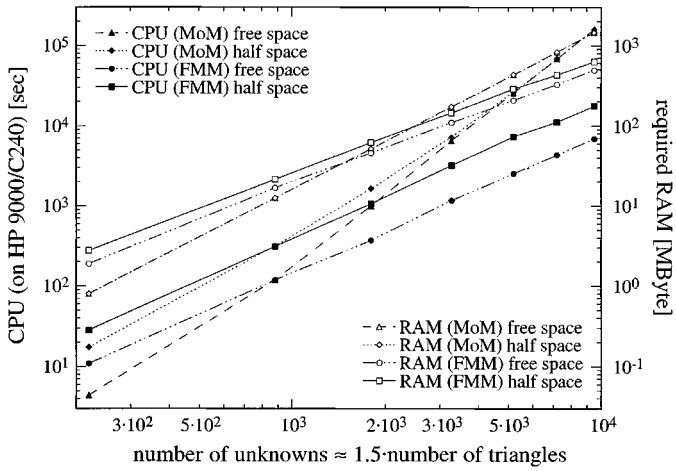


Fig. 6. Required RAM and total CPU time (on HP9000/C240) for bistatic RCS calculations in Fig. 4 (720 scattering angles), for frequencies ranging from 200 to 1400 MHz (corresponding to $N = 220$–9522 unknowns) using MoM (with LUD) and FMM. The CPU time for the MoM solution at $f > 1000$ MHz ($N > 5202$) has been extrapolated.

half-space problem. Finally, we note that the memory requirements and computation time of the half-space FMM, as formulated here, is about twice that of the free-space code. This is attributed to the extra set of image sources, at real spatial positions, which are absent for the free-space problem.

B. Buried 55-Gallon Drum

We consider a 55-gallon drum (90-cm length, 60-cm diameter) buried with its axis parallel to the air–ground interface. Further, the axis of the target is at a depth of 100 cm. For this example we consider Puerto Rico soil, as reported in [38], with 10% water content by weight. The wavelength in this soil is reduced in size by more than a factor of two (relative to free space), due to the soil's relatively high dielectric constant between about $\epsilon_r' = 5$ and $\epsilon_r' = 6.5$ [38]. We consider normal incidence relative to the air–ground interface, addressing separately the case of perpendicular (V) and parallel (H) polarization relative to the target axis. In Table I we present the copolarized VV and HH backscattered RCS of the target, as a function of frequency, from 100–1000 MHz. The MoM results are only computed up to 500 MHz, above which the MoM computation time (Fig. 7) becomes prohibitive. However, for frequencies below 600 MHz, there is excellent agreement between the rigorous MoM solution and the FMM results.

In Fig. 7 we plot the computation time and memory requirements of the MoM and FMM algorithms, for the results in Table I. All MoM computation times for frequencies over 500 MHz are estimated. These results, which were computed on a SGI Origin 2000 supercomputer (only one processor used), demonstrate the significant gains afforded by the FMM, in both computation time and memory. Upon close inspection of the FMM results in Fig. 7, one notices slight undulations in the data. This, as discussed above, is due to the fact that we utilize $M = 2^q$ FMM clusters, where q is an integer. As N is varied, $M = 2^q$ approaches and recedes the optimal $M \propto \sqrt{N}$, causing a slight undulation in the results (as the algorithm, with $M = 2^q$ clusters, approaches and recedes optimality).

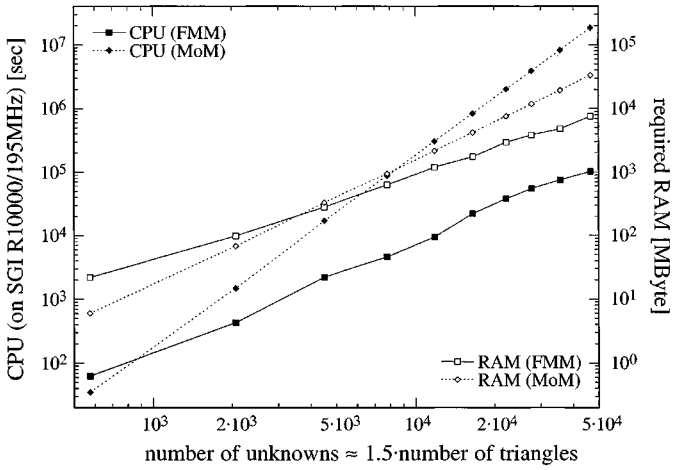


Fig. 7. Required RAM and CPU time (on SGI R10000/195 MHz) for monostatic RCS calculations of 55-gallon drum in Table I using MoM (with LUD) and FMM, i.e., $N = 576$–$46\,623$ unknowns corresponding to $f = 100$–1000 MHz. The CPU time for the MoM solution at $f > 500$ MHz ($N > 11841$) has been extrapolated.

TABLE I
BACKSCATTERING (RCS IN [dBsm]) OF A 55-GALLON DRUM (LENGTH = 90 cm, DIAMETER = 60 cm) BURIED IN PUERTO RICO CLAY (10% WATER BY WEIGHT) [38] AS A FUNCTION OF FREQUENCY FOR A NORMALLY INCIDENT PLANE WAVE ($\theta = 0°$). THE DRUM IS SITUATED WITH ITS SYMMETRY AXIS 100 cm BELOW THE HALF-SPACE BOUNDARY, PARALLEL TO THE INTERFACE. THE INCIDENT ELECTRIC FIELD IS PERPENDICULAR TO THE CYLINDER AXIS FOR VV AND PARALLEL FOR HH POLARIZATION

f [MHz]	N (unknowns)	VV (MoM)	HH (MoM)	VV (FMM)	HH (FMM)
100	576	-15.916	-13.600	-15.917	-13.598
200	2070	-13.238	-12.658	-13.247	-12.682
300	4506	-12.643	-12.468	-12.649	-12.469
400	7812	-13.276	-13.156	-13.321	-13.098
500	11841	-14.738	-14.549	-14.754	-14.526
600	16536	–	–	-16.495	-16.314
700	22200	–	–	-18.248	-18.149
800	27687	–	–	-20.119	-20.072
900	35568	–	–	-22.137	-22.102
1000	46623	–	–	-24.181	-24.127

C. Buried Model Unexploded Ordnance (UXO)

The final set of results correspond to an MK 146 MOD 2 U.S. Navy projectile, with dimensions and shape shown in the inset in Fig. 8, buried in Yuma soil with 5% water content [20], [22], the target symmetry axis in the yz-plane. The plane wave is incident at $\phi_i = -90°$ (in the yz-plane) and $\theta_i = 60°$, and the bistatic scattered fields are observed at angles $\theta_s = 60°$ and $-180° \leq \phi_s \leq 180°$, with a 500-MHz operating frequency. The MoM versus FMM computation and memory requirements are similar to those found for the buried drum (Fig. 7) and the trihedral (Fig. 6) and therefore are not shown here, for brevity.

Considering the results in Fig. 8, we see that the FMM and MoM algorithms agree almost exactly. The depth and orientation of the target in Fig. 8 is realistic (given the target size), but this is a very favorable case for the FMM. In particular, recall that the additional approximation we have introduced [beyond the usual FMM approximations in (4)] is the use of a single real image to represent the reflection at the air–ground interface. This term is needed to account for multiple interactions of

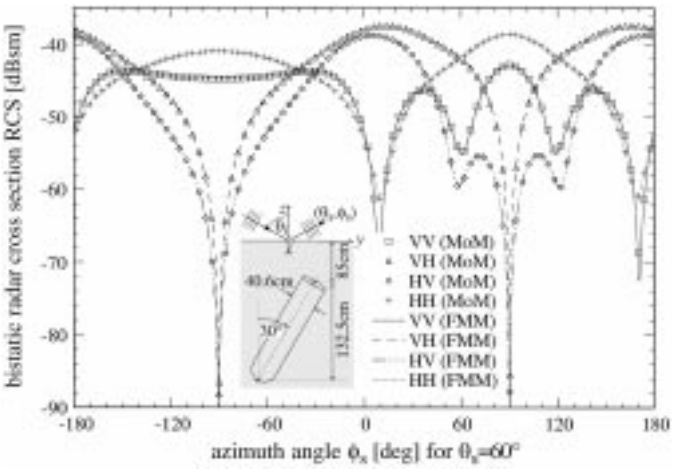


Fig. 8. Bistatic radar cross section of a model UXO (cylinder with hemispherical endcap) of length 153 cm and diameter 40.6 cm (see inset). The model UXO is buried in Yuma soil of 5% water content [20], [22] with the target symmetry axis in the yz-plane, at an angle of 30° relative to the z-axis and the nose down at a depth of $z = -217.5$ cm. The RCS is plotted for varying ϕ_s at an angle $\theta_s = 60°$ (30° from grazing) and a frequency of 500 MHz (corresponding to $fwrN = 7167$), assuming a plane wave incident at $\phi_i = -90°$ (in yz-plane) and $\theta_i = 60°$.

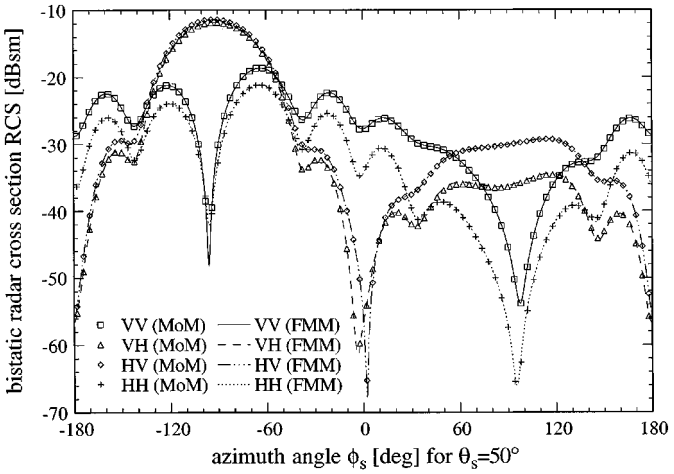


Fig. 9. Same as Fig. 8 but using an angle of 75° between target axis and z-direction (15° tilt relative to horizontal xy-plane), the nose down at a depth of $z = -89.3$ cm (the closest distance between target and interface is 30 cm), and a plane wave incident from $\phi_i = 0°$ (in xz-plane) and $\theta_i = 60°$ (compare to inset in Fig. 8). The bistatic RCS is plotted for varying ϕ_s at an angle $\theta_s = 50°$ (40° from grazing).

electromagnetic energy between the target and air–ground interface. For the lossy soil, and relatively deep target, such interactions are less important. Consequently, as a better test of the half-space FMM developed here, we consider the same ordnance considered in Fig. 8, with the angle between the target axis and z-axis (see Fig. 8) set to 75°. Moreover, the closest distance between the target and the interface is 30 cm, rather than 75 cm in Fig. 8. While such a shallow depth is unlikely for a target of this size, the target-interface interactions are considerably more important for this case [although not as important as for a target above the interface, as in Figs. 4(a) and 5(a)]. Results are shown in Fig. 9, where we consider the same frequency, but modified incidence ($\theta_i = 60°$ and $\phi_i = 0°$, i.e., in the xz-plane) and scattering angles ($\theta_s = 50°$). We note that the bistatic RCS of this shallow target is considerably larger than that of the deeper

target (Fig. 8), and the agreement between the MoM and FMM results is again excellent.

IV. Conclusion

The FMM has been extended to the case of 3-D perfectly conducting targets situated above or within a lossy half space. The "near" FMM terms are handled as in a MoM analysis, wherein the half-space dyadic Green's function is characterized rigorously via the method of complex images [26]–[29]. For the "far" FMM terms, the reflected terms in the dyadic half-space Green's function are approximated via real images with polarization-dependent amplitudes [30]; this technique having proven useful in previous MoM research [31]. The location of the real image is the same for each component of the dyadic. Thus, in addition to being accurate, this approximation yields significantly reduced computational and memory requirements.

Results were presented for targets above and below a half space, with good agreement *vis-a-vis* a reference MoM solution. Nevertheless, the subsurface problem is worthy of further examination. In particular, the real images are required for accurate treatment of the target-interface interaction. However, for deep targets and for relatively lossy soils, the images may be unnecessary, yielding a half-space FMM that replicates the free-space version [13], [14] (although the incident field and the scattered far-field calculations are slightly more complicated). This is likely to be a strong function of target type, depth, and orientation, as well as of the detailed soil characteristics. Moreover, recall that the "far" FMM interactions are handled via the spectral propagator in (4b). If the target is buried in a lossy half space (i.e., complex k_i), T_L is small after sufficiently large intercluster distance $X_{m'm}$. Consequently, for the buried-target case, a further algorithm acceleration and reduction in memory can be achieved by ignoring intercluster interactions for sufficiently large $X_{m'm}$.

References

[1] R. F. Harrington, *Field Computations by Moment Methods*. New York: Macmillan, 1968.

[2] A. F. Peterson, S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*. Piscataway, NJ: IEEE Press, 1998.

[3] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering from surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 30, pp. 409–418, May 1982.

[4] D. R. Wilton, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak, and C. M. Butler, "Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains," *IEEE Trans. Antennas Propagat.*, vol. 32, pp. 276–281, Mar. 1984.

[5] K. A. Michalski and D. Zheng, "Analysis of microstrip resonators of arbitrary shape," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 112–119, Jan. 1992.

[6] ——, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, Parts I and II," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–352, Mar. 1990.

[7] N. Geng, M. Ressler, and L. Carin, "Wideband VHF scattering from a trihedral reflector situated above a lossy dispersive half space," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2609–2617, Sept. 1999.

[8] T. K. Sarkar and E. Arvas, "On a class of finite step iterative methods (conjugate directions) for the solution of an operator equation arising in electromagnetics," *IEEE Trans. Antennas Propagat.*, vol. 33, pp. 1058–1066, Oct. 1985.

[9] F. X. Canning, "Improved impedance matrix localization method," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 659–667, May 1993.

[10] R. L. Wagner and W. C. Chew, "A study of wavelets for the solution of electromagnetic integral equations," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 802–810, Aug. 1995.

[11] A. Boag and R. Mittra, "Complex multipole beam approach to electromagnetic scattering problems," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 366–372, Mar. 1994.

[12] E. Bleszynski, M. Bleszynski, and T. Jaroszewicz, "AIM: Adaptive integral method compression algorithm for solving large scale electromagnetic scattering and radiation problems," *Radio Sci.*, vol. 31, pp. 1225–1251, Sept.–Oct. 1996.

[13] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas Propagat. Mag.*, vol. 35, pp. 7–12, June 1993.

[14] J. M. Song and W. C. Chew, "Fast multipole method solution using parametric geometry," *Micr. Opt. Techn. Lett.*, vol. 7, pp. 760–765, Nov. 1994.

[15] ——, "Multilevel fast multipole algorithm for solving combined field integral equations of electromagnetic scattering," *Micr. Opt. Techn. Lett.*, vol. 10, pp. 14–19, Sept. 1995.

[16] J. M. Song, C. C. Lu, W. C. Chew, and S. W. Lee, "Fast Illinois solver code (FISC)," *IEEE Antennas Propagat. Mag.*, vol. 40, pp. 27–33, June 1998.

[17] J. M. Song, C. C. Lu, and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 1488–1493, Oct. 1997.

[18] L. Gurel and M. I. Aksun, "Electromagnetic scattering solution of conducting strips in layered media using the fast multipole method," *IEEE Microwave Guided Wave Lett.*, vol. 6, pp. 277–279, Aug. 1996.

[19] J. S. Zhao, W. C. Chew, C. C. Lu, E. Michielssen, and J. Song, "Thin-stratified medium fast-multipole algorithm for solving microstrip structures," *IEEE Trans. Microwave Theory Tech.*, vol. 46, pp. 395–403, Apr. 1998.

[20] L. Carin, N. Geng, M. McClure, J. Sinchina, and L. Nguyen, "Ultra-wideband synthetic aperture radar for mine field detection," *IEEE Antennas Propagat. Mag.*, vol. 41, pp. 18–33, Feb. 1999.

[21] C. C. Chen and L. Peters, "Buried unexploded ordnance identification via complex natural resonances," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 1645–1654, Nov. 1997.

[22] S. Vitebskiy, L. Carin, M. A. Ressler, and F. H. Le, "Ultra-wideband, short-pulse ground-penetrating radar: Simulation and measurement," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 762–772, May 1997.

[23] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse plane-wave scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 143–151, Feb. 1996.

[24] N. Geng and L. Carin, "Wideband electromagnetic scattering from a dielectric BOR buried in a layered lossy, dispersive medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.

[25] Y. Rahmat-Samii, R. Mittra, and P. Parhami, "Evaluation of Sommerfeld integrals for lossy half-space problems," *Electromagn.*, vol. 1, no. 1, pp. 1–28, 1981.

[26] Y. L. Chow, J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 588–592, Mar. 1991.

[27] J. J. Yang, Y. L. Chow, and D. G. Fang, "Discrete complex images of a three-dimensional dipole above and within a lossy ground," *Proc. Inst. Elect. Eng.—H*, vol. 138, pp. 319–326, Aug. 1991.

[28] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.

[29] M. I. Aksun, "A robust approach for the derivation of closed-form Green's functions," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 651–658, May 1996.

[30] I. V. Lindell, *Methods for Electromagnetic Field Analysis*. Piscataway, NJ: IEEE Press, 1995.

[31] U. Jakobus, Erweiterte Momentenmethode zur Behandlung kompliziert aufgebauter und elektrisch groer elektromagnetischer Streuprobleme (in in German), VDI Forschungsberichte, 1994.

[32] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1970.

[33] A. D. McLaren, "Optimal numerical integration on a sphere," *Math. Computation*, vol. 17, pp. 361–383, 1963.

[34] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*. Englewood Cliffs, NJ: Prentice-Hall, 1973.

[35] K. Sarabandi and C. C. Chiu, "Optimum corner reflectors for calibration of imaging radars," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1348–1361, Oct. 1996.

[36] T. Griesser and C. A. Balanis, "Backscatter analysis of dihedral corner reflectors using physical optics and the physical theory of diffraction," *IEEE Trans. Antennas Propagat.*, vol. AP-35, pp. 1137–1147, Oct. 1987.

[37] N. Geng, A. Sullivan, and L. Carin, "Fast multipole method for scattering from 3D PEC targets situated in a half-space environment," *Micr. Opt. Techn. Lett.*, vol. 21, pp. 399–405, Feb. 1999.

[38] J. E. Hipp, "Soil electromagnetic parameters as functions of frequency, soil density, and soil moisture," *Proc. IEEE*, vol. 62, pp. 98–103, Jan. 1974.

Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Germany, in 1991 and 1996, respectively.

From 1991 to 1996 he was with the Institute for Microwaves and Electronics at the University of Karlsruhe, working on full-wave propagation modeling for radio communications systems. In January 1997 he joined the Department of Electrical and Computer Engineering at Duke University, Durham, NC, as a visiting post—doctoral Fellow for 18 months. While visiting Duke University, and for another 18 months back at the University of Karlsruhe, his research focused on numerical techniques in computational electromagnetics for planar-stratified media. In January 2000 he joined SIEMENS Corporate Technology in Munich, Germany, working as a Research Scientist on the design of broadband wireless communications systems.

Dr. Geng received the Mannesmann Innovation Award in 1997 for his Ph.D. dissertation.

Anders Sullivan (M'93) [redacted] in Staten Island, NY. He received the B.S. and M.S. degrees in aerospace engineering from the Georgia Institute of Technology, Atlanta, GA, in 1985 and 1987, respectively, and the Ph.D. degree in electromagnetics from Polytechnic University, Brooklyn, NY, in 1997.

From 1988 through May 1998, he was with the Air Force Research Laboratory at Eglin Air Force Base, FL. From June 1998 through September 1999 he was with the Electrical Engineering Department at Duke University as a Research Associate. Since September 1999, he has been with the Army Research Laboratory in Adelphi, MD. His current research interests include modeling complex targets and short-pulse scattering.

Dr. Sullivan is a member of the Tau Beta Pi and Sigma Gamma Tau.

Lawrence Carin (F'01) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989 he joined the Electrical Engineering Department at Polytechnic University (Brooklyn) as an Assistant Professor and became an Associate Professor there in 1994. In September 1995 he joined the Electrical Engineering Department at Duke University, where he is an Associate Professor. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION. He is a member of the Tau Beta Pi and Etea Kappa Nu.

Time-Domain Sensing of Targets Buried Under a Gaussian, Exponential, or Fractal Rough Interface

Traian Dogaru and Lawrence Carin, *Fellow, IEEE*

***Abstract*—We numerically examine subsurface sensing via an ultrawideband ground penetrating radar (GPR) system. The target is assumed to reside under a randomly rough air–ground interface and is illuminated by a pulsed plane wave. The underlying wave physics is addressed through application of the multiresolution time-domain (MRTD) algorithm. The scattered time-domain fields are parametrized as a random process and an optimal detection scheme is formulated, accounting for the clutter and target signature statistics. Detector performance is evaluated via receiver operating characteristics (ROCs) for variable sensor parameters (polarization and incident angle) and for several rough-surface statistical models.**

***Index Terms*—Detection, radar, sensing, time domain.**

I. Introduction

ULTRAWIDEBAND (UWB) radar techniques have been applied successfully to a broad range of problems in target detection, classification, and imaging. Of particular interest here, UWB radar has been employed in the context of short-pulse sensing of subsurface targets [1]. In this paper, we simulate such a system, by analyzing the time-domain fields scattered from a target buried under a rough air-ground interface. The time-domain scattered fields are processed via an optimal detector, from which we investigate the effects of surface roughness on the detection of subsurface targets.

The statistical analysis of time-domain electromagnetic scattering from a rough surface has been investigated previously [2], [3]. However, the vast majority of work on rough-surface scattering has been addressed in the frequency domain [4]–[7]. Moreover, previously, most such work has addressed the rough-surface scattering problem alone, without consideration of scattering from a target buried under such a surface. Motivated by new UWB systems that operate in the time domain [1], here we concentrate on time-domain scattering from a rough surface, and in particular on subsurface sensing in such an environment. Recently, there has been significant interest in the development of time-domain electromagnetic models, such as the finite-difference time-domain (FDTD) [8] and, more recently, the multiresolution time-domain (MRTD) [9], [10]. The latter has been employed in modeling rough surface scattering problems [10], wherein its advantages *vis-à-vis* the FDTD scheme have been discussed. Motivated by the attractiveness of the MRTD for rough-surface scattering, it is applied here for all electromagnetic simulations.

It is demonstrated that a Bayesian detector, explicitly accounting for the statistical character of the buried-target signature, often yields markedly improved detector performance *vis-à-vis* a matched filter. In practice, one would typically employ an approximate model to quantify the statistical characteristics of the target signature. In the work presented here, in which we investigate the fundamental phenomenological issues in such a detection scenario, the target-signature statistics are computed via our electromagnetic scattering model. In particular, we parametrize the statistics of the backscattered target signature numerically through consideration of multiple realizations of the rough surface, the latter dictated by a prescribed surface statistics. The scattered fields employed to test the performance of the detector are likewise computed via an electromagnetic model.

Before proceeding, we note that there has been previous work on the detection of targets under a rough surface. For example, Tsang and colleagues [12]–[15] have investigated the angular correlation function (ACF) to discriminate man-made targets from rough-surface clutter. In [12]–[15] the authors have addressed the three-dimensional (3-D) problem. The ACF exploits the statistical properties of a rough surface from multiple aspects (orientations) and across multiple frequency bands. The advantage of ACF is that it is applicable to a given target buried under a specific rough surface, and therefore it may be implemented in practice. In the work discussed here, we consider an optimal (Bayesian) detector for time-domain excitation. Such a detector requires an ensemble of scattered waveforms from a buried target, implying knowledge of the target signature under an ensemble of rough surfaces. This is therefore less practical than the ACF, requiring that the ensemble of scattered waveforms be measured or computed *a priori*. However, the Bayesian processor discussed here is an optimal detector and therefore, this study allows us to place bounds on the ultimate performance of detecting a target under a rough surface, against which more practical techniques such as ACF can be compared. In addition to addressing the detection problem *per se*, we also examine the degree to which the target signature is randomized by the rough surface. In particular, as noted above, even if the target identity is known, the randomization of the incident and transmitted fields by the rough surface cause the target signature to be stochastic. This effect is addressed for several rough surfaces, this issue being of importance for any detection scheme.

II. MRTD and Rough Surface Scattering

A. MRTD Formulation

We summarize the formulation of the MRTD algorithm for a two-dimensional (2-D) configuration using an expansion of the fields in terms of Haar scaling functions and a single level of

Manuscript received February 28, 2000; revised February 18, 2001.

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0196-2892(01)05477-8.

0196–2892/01$10.00 © 2001 IEEE

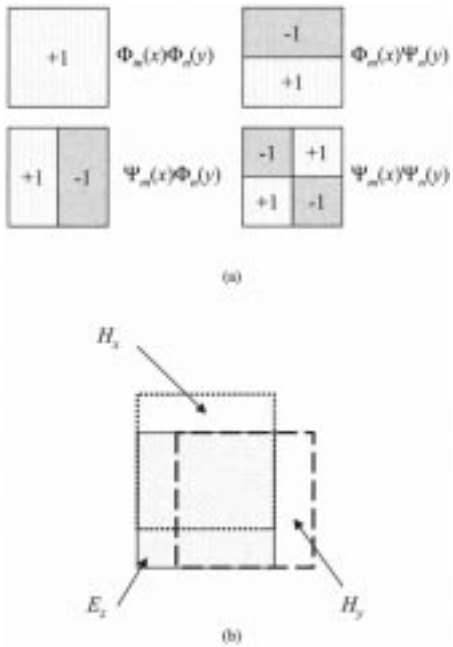


Fig. 1. (a) Prototype 2-D Haar scaling and wavelet functions and (b) support of the basis functions in the electric and magnetic field expansions using one level of Haar wavelets (two dimensions, TE polarization).

wavelets [16]. The model can be applied to an arbitrary number of Haar wavelet levels [17] at the cost of increased complexity. To simplify the presentation, we consider only a single wavelet level, this typically sufficient for scattering problems of interest. The MRTD is not limited to Haar wavelets, although algorithm complexity increases with the complexity of the underlying wavelet basis functions.

Fig. 1(a) shows the prototype Haar basis functions in the 2-D case. They are obtained from combinations of one-dimensional (1-D) Haar scaling and first-level wavelet functions Φ and Ψ, respectively. We consider a uniform grid with step Δx in both spatial directions. The prototype Haar basis functions are shifted by an integer number of Δx in both directions in order to cover the entire computational domain. For time discretization rectangular pulse functions are used, with a temporal step Δt (equivalent to the Yee temporal discretization scheme). Fig. 1(b) represents the relative position of the basis-function support for the three components of a TE polarized wave (the TM case is similar if one interchanges the E and H components).

The field components are expanded using the basis described previously. At each cell of the computational grid and each time step, there are four coefficients per field component (corresponding to the combinations $\Phi_x\Phi_y$, $\Phi_x\Psi_y$, $\Psi_x\Phi_y$, $\Psi_x\Psi_y$). Maxwell's equations are sampled in time and space using a Galerkin procedure [18]. The result is a set of equations relating the set of field coefficients, which are updated every time step. A detailed description of the sampling procedure and the resulting equations in two dimensions is given in [17].

It can be demonstrated that the Haar-MRTD scheme, using scaling functions and one wavelet level, leads to a formulation equivalent to the Yee FDTD algorithm applied on a grid with half the cell size [17]. However, the MRTD formulation has the advantage that it can employ Haar wavelets in regions where high resolution is required (equivalent to FDTD at a sampling rate of $\Delta x/2$), and scaling functions alone in regions where coarser spatial sampling is sufficient (equivalent to FDTD at a sampling rate of Δx). In other words, we obtain a self-consistent procedure for multiscale spatial gridding. Further wavelet levels can be added in certain regions, leading to an effective resolution increase in those regions of 2^n (where n is the maximum wavelet level employed by the scheme).

B. Modeling Rough Dielectric Interfaces with MRTD

In addition to the aforementioned multiresolution properties of MRTD, another advantage of this model *vis-a-vis* the classic FDTD Yee scheme becomes apparent when one models media interfaces that do not conform to the computational grid. It is well known that the classical FDTD Yee scheme typically relies on a staircase approximation of such an interface. This approximation may introduce errors, especially when essential features of the scatterer are only a fraction of a cell dimension. Several techniques have been developed in the context of the FDTD algorithm to alleviate this problem [8]. The MRTD formulation allows a self-consistent treatment of interfaces of arbitrary shape [9], [10]. This becomes particularly useful for modeling scattering by a rough interface dividing two dielectric media. In a previous paper [10], we demonstrated that the MRTD scheme increases the accuracy of the rough surface scattering model as compared to the staircase FDTD model, assuming analogous spatial sampling rates. Although the errors introduced by the staircase FDTD model may be acceptable for surfaces with relatively "mild" roughness, the more accurate MRTD model may be necessary when the roughness is increased.

In our formulation, the dielectrics can be lossy, with the material properties (ε and σ) independent of frequency. Dispersive materials can be handled in a manner analogous to that applied to the FDTD [8]. Because of the presence of arbitrary inhomogeneity $\varepsilon(x, y)$ and $\sigma(x, y)$, the usual MRTD discretization procedure yields cross-terms between the Φ and Ψ components of the electric field. Thus, the equations that update the various coefficients in the electric field expansion, at the cell (m, n), become coupled, and they can be expressed as a matrix-vector equation for the field coefficients at that cell [9], [10]. The matrix involves the average permittivity and conductivity within cell (m, n), weighted by the appropriate scaling/wavelet functions. If the medium is homogeneous inside one cell, the matrix is diagonal and the equations for that particular cell are decoupled. Consequently, the special matrix-vector equations need to be solved only for cells at the boundaries of two or more media.

The usual model for a 2-D rough surface involves a function $y = y(x)$, representing one realization of a random surface of particular statistics. The continuous function $y(x)$ is discretized (sampled) in the x direction and, between the discretization points, it is approximated by straight segments. Fig. 2 shows a small portion of an arbitrary interface and the two models: MRTD (dashed line) and staircase (interface between the shaded and blank areas). While in the staircase FDTD case the material properties are generally sampled at the same rate as the fields (i.e., one grid cell), in the MRTD scheme the surface can be sampled more finely than the rate at which the fields are sampled.

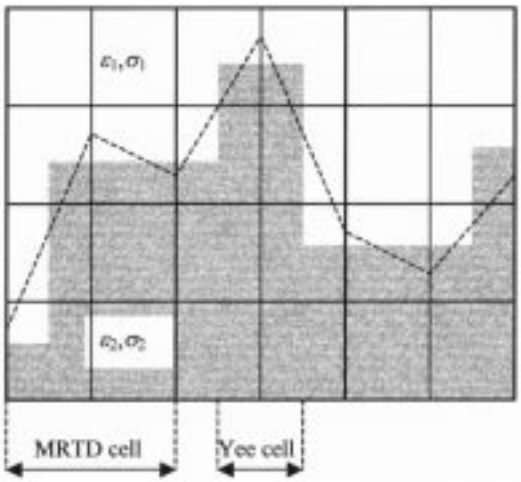


Fig. 2. Detail on the MRTD and staircase FDTD models of a rough surface (two dimensions, TE polarization).

The typical FDTD sampling rate Δx, normally dictated by the shortest wavelength in the electromagnetic spectrum, may not be small enough for representing $y(x)$ without aliasing. Therefore, for rough-surface scattering computations, the FDTD Δx must be made small enough to represent the rough surface well. The uniform gridding of the FDTD implies that this same spatial discretization must be used throughout, generally yielding over sampling away from the surface. With the MRTD algorithm, the rough interface can be sampled at a rate in excess of that used to represent the fields. The result is that, for the MRTD algorithm, the cell size is dictated by the shortest wavelength in the electromagnetic spectrum, and not by the finest feature of material inhomogeneity. Wavelets are added to the scaling functions to track fast field variations in the vicinity of the rough surface, with the fields away from the surface represented by scaling functions alone.

III. Rough Surface Models

A. Traditional Models

As mentioned in the previous section, the interface between the two-dielectric media is modeled as a random process with particular statistical parameters. We restrict our model to the 2-D case. For the "traditional" random processes considered in this section, the samples $y(x)$ are assumed to have a Gaussian probability density function. Scattering from a surface with an exponential density function (non-Gaussian) has been discussed in [13]. The most widely applied type of rough surface model is characterized by a Gaussian autocorrelation function [5]. For a surface with variance σ^2, the autocorrelation is expressed as

$$R(\Delta x) = E\left[y(x)y(x+\Delta x)\right] = \sigma^2 \exp\left(-\frac{\Delta x^2}{l^2}\right) \quad (1)$$

with the parameter l termed the correlation length. Another frequently employed autocorrelation is the exponential function. This correlation function has generally been avoided in analytic studies of scattering from rough surface, because it is not differentiable at the origin. Since we are concerned with the numerical analysis of rough-surface scattering, this does not present an obstacle in our study. The exponential correlation functions has the form

$$R(\Delta x) = E\left[y(x)y(x+\Delta x)\right] = \sigma^2 \exp\left(-\frac{|\Delta x|}{L}\right). \quad (2)$$

Analogous to the Gaussian case, the parameter L represents the correlation length. Authors have proposed other models, constituting hybrids of the Gaussian and exponential functions, characterized by two or more correlation-length-type parameters, taking into account the roughness at different scales. Although the methods presented in this paper are applicable to hybrid surface models, we restrict ourselves to Gaussian and exponential surfaces.

In practice, the random process (surface) is generated in the Fourier domain, by passing a Gaussian white-noise process through a filter with a spatial-frequency response corresponding to the desired rough-surface spectrum [5].

B. Bandlimited Fractal Models

Although simple correlation functions like the Gaussian or the exponential are commonly used in many theoretical studies of rough-surface scattering, they are not necessarily representative of many categories of natural surfaces. In particular, natural scenes such as sea, forest and other vegetated regions present a multiscale structure that cannot be described by only one parameter (the correlation length). The recently developed fractal framework [19] has proven a better approach for describing a multitude of natural phenomena, including natural rough interfaces. Recent papers have studied the scattering of electromagnetic waves from fractal surfaces, using analytical as well as numerical methods [20]–[22]. Theoretically, random fractal surface models take into account the roughness over an infinite range of scales. However, fractal surfaces cannot be handled rigorously in practice because of their peculiar mathematical properties, among which are nonstationarity and the consequent inappropriateness of defining a power spectrum. For practical purposes we therefore must restrict ourselves to the study of bandlimited fractals [20]. The finite length of the sequence is equivalent to considering a lower limit of the frequency spectrum, whereas sampling at finite intervals truncates the spectrum at high frequencies. In fact, any measurement process would present similar limitations, therefore making the bandlimited fractal a realistic model.

In the following, we consider the bandlimited Weierstrass process [20]–[23]:

$$w(x) = \sigma\sqrt{\frac{2(1-b^{-2H})}{b^{-2HN_1} - b^{-2H(N_2+1)}}} \cdot \sum_{n=-N_1}^{N_2} b^{-Hn}\cos(2\pi b^n x + \varphi_n) \quad (3)$$

where

σ	standard deviation;
H	Hurst exponent with values between 0 and 1;
$b > 1$	constant (we take $b = \sqrt{\pi}$);
φ_n	random phases, uniformly distributed between 0 and 2π.

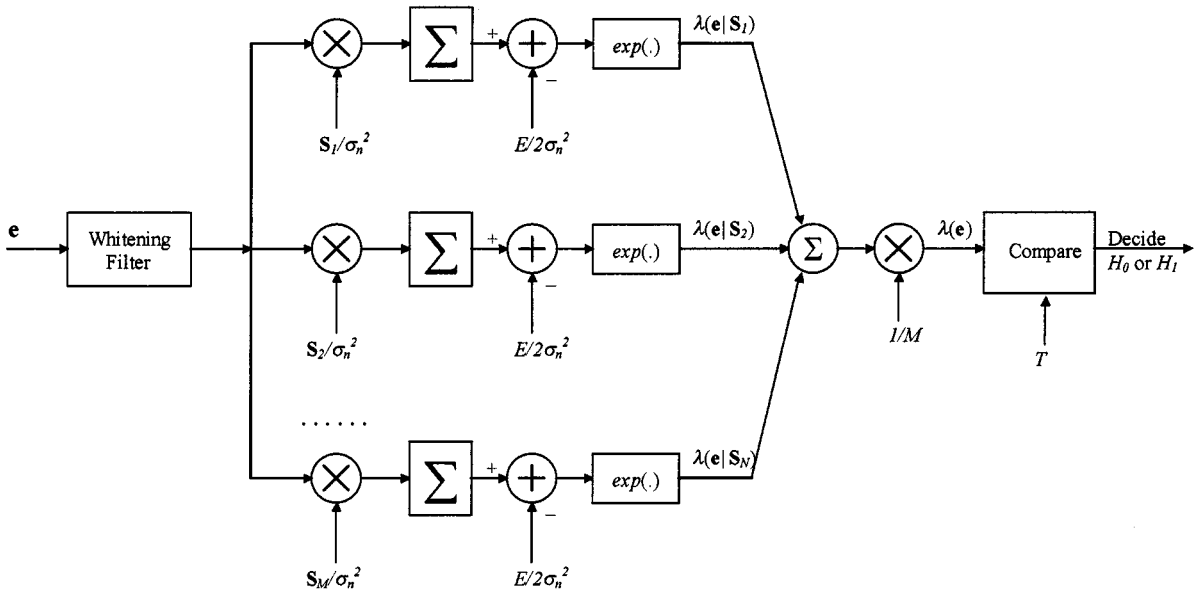


Fig. 3. Optimal detector for a random signal in additive clutter, using M realizations of the random signature. $s_1, s_2, \ldots, s_M$ are M independent realizations of the target signature, E is the average target signature energy, and σ_n is the clutter variance. Each channel computes the likelihood ratio for one given realization of the target signature.

The fractal dimension of this process is $D = 2 - H$. It can be shown that the spectrum of a Weierstrass process follows a power law [i.e., $S(k) \propto k^{-2H-1}$] [20]–[23].

The integers N_1 and N_2 in (3) represent the lowest and the highest harmonics in the spectrum. In a pure fractal process, we should have $N_1 = -\infty$ and $N_2 = \infty$, such that the frequency runs from 0 to ∞. However, in this case, the amplitude of the zero-frequency (dc) component would go to infinity, so we must keep N_1 finite. Also, the sampling process requires N_2 to be finite. One constraint in choosing these integers is to include enough discrete spectral modes such that the highest-frequency component has small amplitude compared to the lowest-frequency component. We notice that the amplitude decay is reduced as H decreases (higher fractal dimension), thus requiring more spectral modes to be taken into account. On the other hand, this increases the maximum frequency in the spectrum, putting another constraint on the discretization rate (one should notice that this tradeoff is independent of b). Summarizing, for higher fractal dimensions (close to 2), it is difficult to obtain a good representation of the Weierstrass process, unless we consider very fine discretization rates. Therefore, in this paper, we limit ourselves to the study of surfaces with fractal dimension less than or equal to 1.5.

It should be noted that the value of N_1 is extremely important in characterizing the statistics of both the rough surface and the associated scattered fields. In particular, changing N_1 for a set of surfaces leads to significant variation in the energy of the fields scattered from those surfaces. The parameter N_1 should be taken greater (or at least equal) to zero, otherwise the surface would include less than a period of the lowest frequency component (which has the largest amplitude). Since the phase of this component is random, it is possible to obtain a surface average other than zero, which is undesirable. It is not difficult to see that the larger N_1, the smaller the energy of the return in the backscatter direction (which is of interest in our problem), because the large-scale features in the rough surface (corresponding to the lower frequencies), are eliminated. It is also worth mentioning that, although a pure Weierstrass fractal surface is nonstationary [19], its bandlimited version is stationary and its PSD can be computed in closed-form [20]. We have found that the fields scattered by such a surface are stationary as well. This is important for the detector design (see Section IV), where one of the basic assumptions is the stationarity of the clutter.

C. Truncating the Surface

In any numerical model of rough surface scattering, the surface must be finite, making edge effects an important issue. One solution to this problem is to consider illumination by a beam [4]–[7]. This is often employed as the incident field in frequency-domain integral-equation solvers [4]–[6]. However, the beam approach is difficult to implement with time-domain methods, if one tries to model wideband electromagnetic scattering. This is because the beam width increases with the wavelength, which means that for the low frequencies of the spectrum we need an impractically large computational domain. In this paper, we are interested in the time-domain statistical characterization of the fields scattered by a rough surface. For this purpose, we consider scattering from a finite-extent surface, under pulsed plane-wave excitation. As mentioned, the plane-wave excitation gives rise to edge effects at the ends of the finite rough surface. The received time sequence is windowed temporally, with the scattered-field statistics computed by considering only a portion of the transient signature in the middle of the time-domain response, this weakly affected by edge diffraction (in the beam approach, the beam width windows the problem spatially). We also emphasize that, as indicated in Fig. 4, the MRTD model considers a finite 2-D rough surface in the vicinity of an infinite dielectric half space (the latter accounted for by the con-

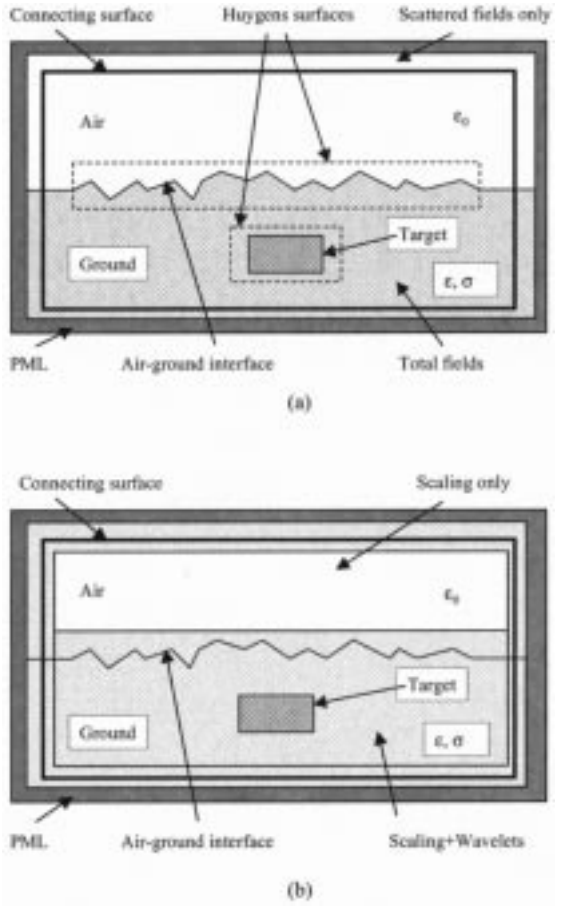


Fig. 4. MRTD computational domain. (a) Incident field implementation, PML, and near-to-far zone transformation and (b) repartition of scaling and wavelet basis functions.

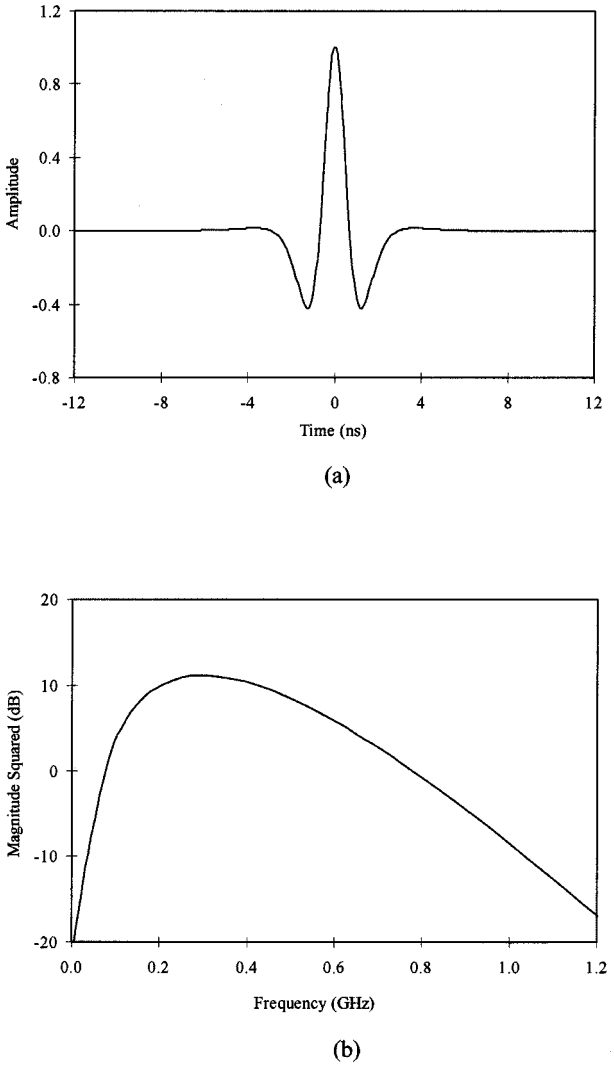


Fig. 5. Incident pulse (Rayleigh, fourth order) in (a) time and (b) frequency domain.

necting surface that gives rise to the half-space plane-wave excitation). The diffracted fields from the ends of the rough surface are therefore generated by the relatively smooth transition from a half space to a rough surface. Moreover, the rough surface is tapered at the ends, to avoid sharp peaks in the vicinity of the terminal points.

IV. Optimal Detection of Buried Targets

The problem of interest is detection of a buried target in the presence of a rough interface. There are two signals involved in the detection scheme: the target response (which is the signal we wish to detect) and the fields scattered by the interface (clutter). In the following, the various sampled signals are considered as vectors in an N-dimensional signal space, where $\mathbf{e} = [e_1 e_2 \cdots e_N]^T$ represents the received signal, $\mathbf{c} = [c_1 c_2 \cdots c_N]^T$ represents the random clutter response, and $\mathbf{s} = [s_1 s_2 \cdots s_N]^T$ represents the target response (assuming time-domain scattered waveforms with N temporal samples). At the detector output, we must decide between hypothesis H_0, in which $\mathbf{e} = \mathbf{c}$ (the received signal consists of clutter alone) and hypothesis H_1, in which $\mathbf{e} = \mathbf{s} + \mathbf{c}$ (the target is present).

We assume that the clutter is a Gaussian, wide sense stationary and uncorrelated random process. In numerical experiments, we have verified that the first two assumptions are valid for the rough surfaces modeled here. However, the fields scattered by a correlated random surface are correlated themselves. In order to satisfy the third assumption, we precede the detector by a whitening filter [11].

It is important to recognize that the target signature $\mathbf{s}$ is a random process, even if the target identity and depth are assumed known. This is manifested because the fields transmitted through the rough interface are stochastic, due to transmission through a randomly rough surface. Similar stochastic distortion occurs when the fields scattered by the target travel from the ground into the air through the rough interface. The stochastic character of $\mathbf{s}$ has important implications in the detector design, because a simple matched filter is suboptimal in such a situation [11]. In particular, if the target signature $\mathbf{s}$ is deterministic, the optimal detector is the likelihood-ratio test

$$\Lambda(\mathbf{e}) = \frac{\mathrm{p}_{\mathbf{c}}(\mathbf{e}-\mathbf{s})}{\mathrm{p}_{\mathbf{c}}(\mathbf{e})} \underset{H_0}{\overset{H_1}{\gtrless}} T \tag{4}$$

where $\mathrm{p}_c(\mathbf{c})$ represents the distribution of the clutter. After employing the aforementioned whitening filter $\mathbf{c}$ is an uncorrelated

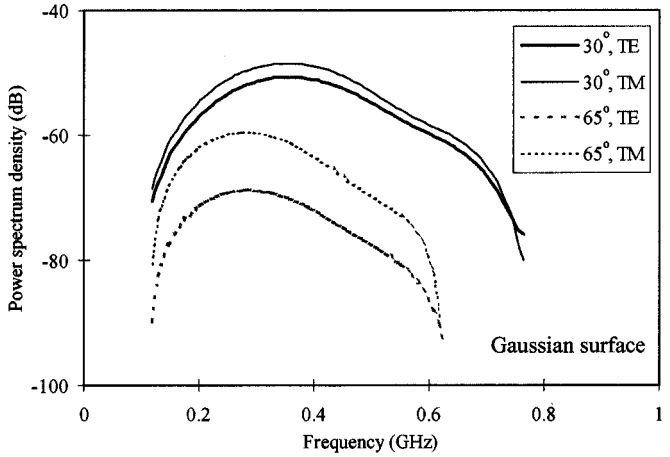


Fig. 6. Power spectrum density produced by a Gaussian surface with $\sigma = 3.95$ cm and $L = 25$ cm for the excitation described in Fig. 5.

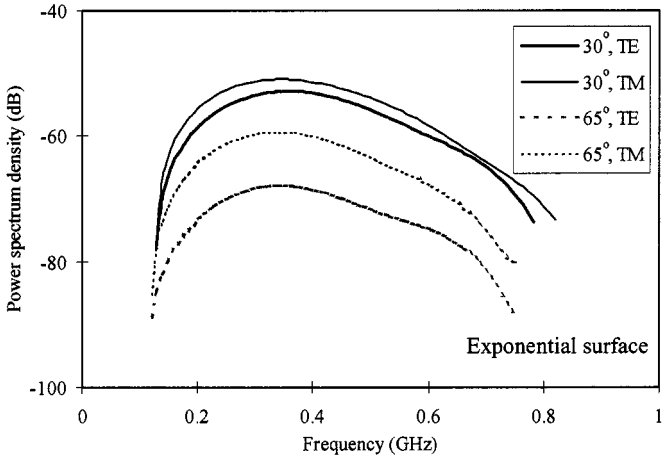
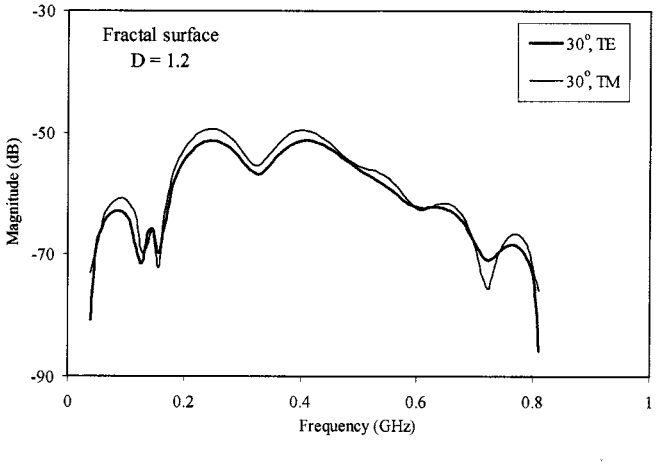
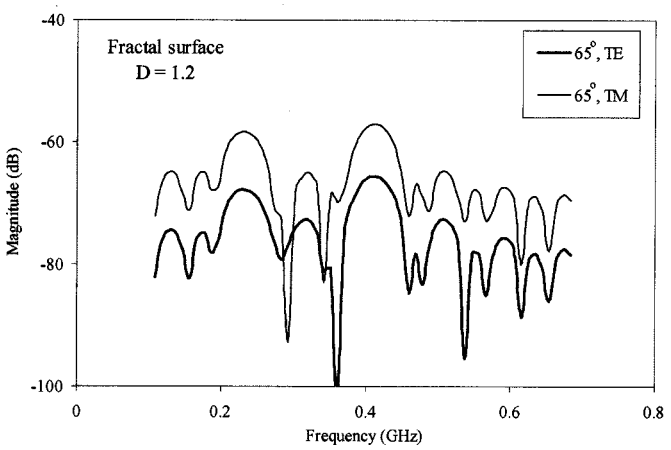


Fig. 7. Power spectrum density produced by an exponential surface with $\sigma = 3.95$ cm and $L = 25$ cm for the excitation described in Fig. 5.



(a)



(b)

Fig. 8. Power spectrum density produced by a fractal surface with $\sigma = 6.25$ cm and $D = 1.2$ for the excitation described in Fig. 5. (a) Incidence at 30° and (b) incidence at 65°.

Gaussian process, for which the likelihood ratio in (4) reduces to a form that is implemented as a matched filter. The threshold T is varied to generate the receiver operating characteristic (ROC) [11].

One can model the randomness in the target response by introducing a generalized stochastic parameter θ. Instead of assuming the target signature $\mathbf{s}$ is known exactly, we introduce an uncertainty, which we symbolize by the notation $\mathbf{s}(\theta)$, where θ represents a generalized vector of stochastic parameters responsible for the random nature of $\mathbf{s}$. Thus, under the hypothesis H_1, the received signal becomes $\mathbf{e} = \mathbf{s}(\theta) + \mathbf{c}$. The simple likelihood ratio in (4) is now generalized as

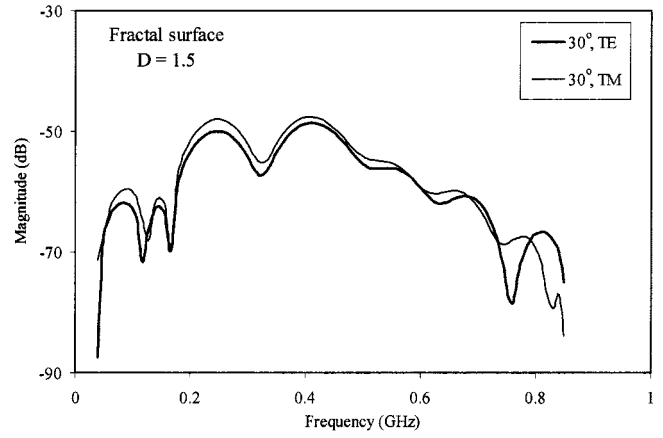
$$\Lambda(\mathbf{e}) = \frac{\int \mathrm{p}_{\mathbf{c}}\left(\mathbf{e} - \mathbf{s}(\theta)\right) d\theta}{\mathrm{p}_{\mathbf{c}}(\mathbf{e})} = \int \Lambda\left(\mathbf{e}\,|\theta\right) \mathrm{p}_{\hat{\mathbf{c}}}(\theta) d\theta \underset{H_0}{\overset{H_1}{\gtrless}} T \quad (5)$$

where $\Lambda(\mathbf{e}\,|\theta)$ represents the likelihood ratio for a particular $\mathbf{s}(\theta)$, and $\mathrm{p}_{\theta}(\theta)$ is the probability density function (pdf) of the vector θ.

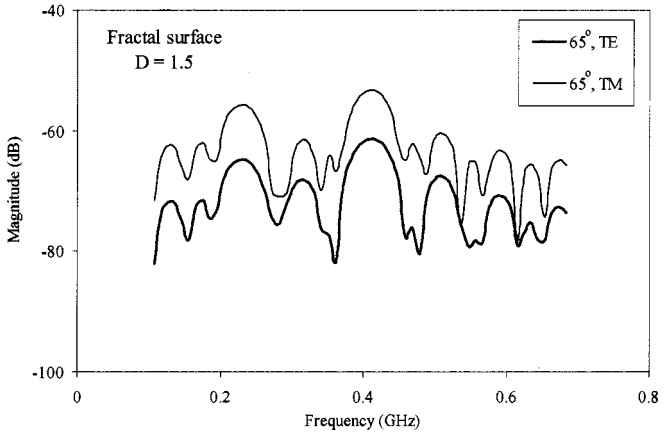
Note that we have not specified $\mathrm{p}_{\theta}(\theta)$, nor do we quantify its distribution. Although one could attempt to model θ as a physical quantity, we avoid this and consider an alternative manner of computing the likelihood ratio. Considering (5), note that the expression on the right side is simply the ensemble average of $\Lambda(\mathbf{e}\,|\theta)$, computed for all possible values of θ. We can approximate this quantity numerically by performing Monte Carlo integration. Thus, we consider M realizations of the random vector θ (physically corresponding to M realizations of the rough surface), the mth of which is represented by θ_m and compute the approximate likelihood ratio as

$$\Lambda(\mathbf{e}) = \int \Lambda\left(\mathbf{e}\,|\theta\right) \mathrm{p}_{\hat{\mathbf{c}}}(\theta) d\theta \approx \frac{1}{M} \sum_{m=1}^{M} \Lambda\left(\mathbf{e}\,|\theta_m\right) \underset{H_0}{\overset{H_1}{\gtrless}} T. \quad (6)$$

Consequently, we obtain the structure of the optimal receiver described in Fig. 3. The number of projections M is determined empirically by studying the convergence of the detector output as a function of M. In all the simulations presented in the next section, 40–50 projections were typically sufficient for convergence [3]. The ensemble of target scattered waveforms $\mathbf{s}(\theta_m)$ are computed via our MRTD scattering model, for the target of interest under M realizations of the rough surface, the latter described by a prescribed statistical model.



(a)



(b)

Fig. 9. Power spectrum density produced by a fractal surface with $\sigma = 6.25$ cm and $D = 1.5$, for the excitation described in Fig. 5. (a) incidence at 30° and (b) incidence at 65°.

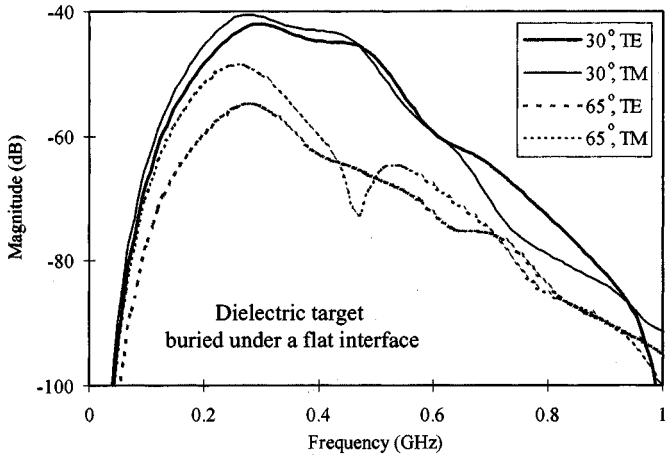


Fig. 10. Energy spectrum density for a 37.5 × 12.5 cm dielectric target with $\varepsilon_r = 2$, buried 25 cm beneath a flat air–ground interface. The soil has $\varepsilon_r = 6$ and $\sigma = 0.005$ S/m. The excitation is described in Fig. 5.

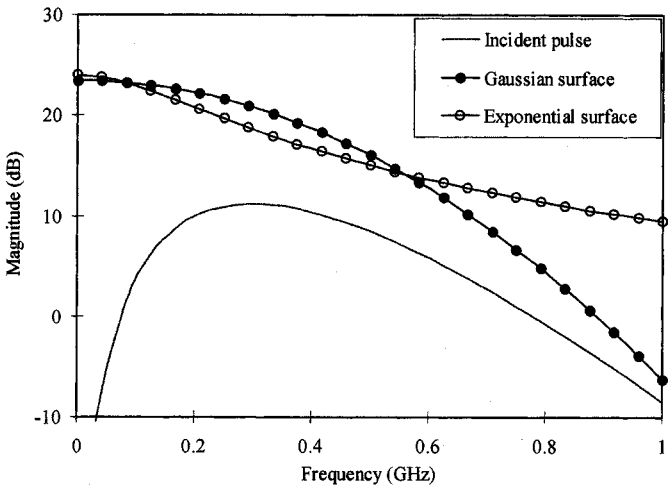


Fig. 11. Comparison between the incident pulse spectrum ($\lambda_c = 1$ m) and the rough-surface power spectrum density. The Gaussian and exponential surfaces have $L = 0.25$ m.

V. Results

A. Implementation

We consider short-pulse electromagnetic scattering from a dielectric target buried under a rough air-ground interface. The clutter statistics are computed by considering time-domain electromagnetic scattering from the rough surface, in the absence of a subsurface target. The clutter statistics are employed in the design of the whitening filter. Moreover, as discussed above, the model is used to compute the ensemble of stochastic target signatures $\mathbf{s}(\theta_m)$ characteristic of the target buried under the rough surface. Each realization of the target signature $\mathbf{s}(\theta_m)$ is computed by considering the target under a particular realization of the rough surface. We also compute the scattered fields from the same surface, in the absence of the target. By subtracting the latter from the former, we yield one realization of the target signature. To test the detector, we generate a set of stochastic scattered waveforms when the target is absent (hypothesis H_0) and when it is present (hypothesis H_1), and variation of the detector threshold T yields the detector performance in the form of ROC. In order to obtain meaningful statistical data, we consider many realizations of the rough surface for a given set of parameters. Distinct computations are used to compute the clutter statistics, the ensemble of $\mathbf{s}(\theta_m)$, as well as the waveforms employed to generate the ROCs (we do not test and train on the same data).

The large number of rough-surface calculations (typically 600 rough-surface calculations are used in training and testing the detector) requires an efficient implementation of the MRTD algorithm. A schematic drawing of the MRTD computational domain is shown in Fig. 4. The soil is modeled as a lossy dielectric, with $\varepsilon_r = 6$ and $\sigma = 0.005$ S/m, independent of frequency. These parameters are typical for dry soil. A more realistic soil model should account for a variation of these parameters with the frequency. However, for shallow targets, the propagation distance inside the soil (in the region of interest) is electrically very small, making the dispersion effects negligible. The 2-D target is modeled as a rectangular cylinder (infinite in the transverse direction), with dimensions 37.5 cm × 12.5 cm, composed of a lossless dielectric with $\varepsilon_r = 2$. The target is buried at a depth of 25 cm (measured from the average position of the rough interface to the top of the target). Alternative targets could be considered, but the

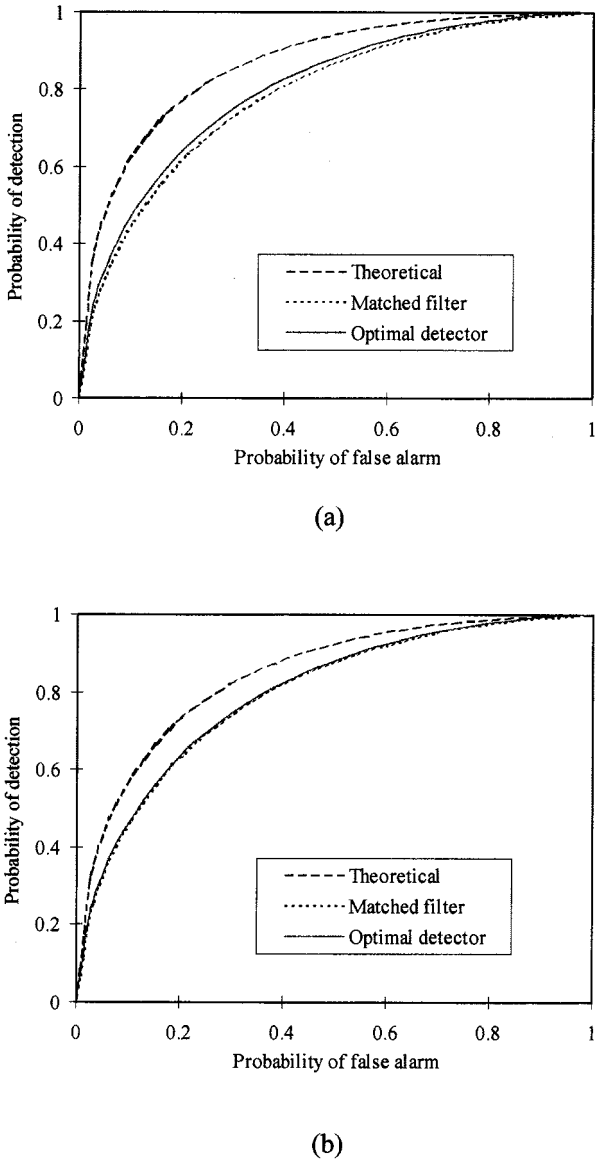


Fig. 12. Receiver operating characteristics for a Gaussian surface with $\sigma = 3.95$ cm and $L = 25$ cm, incidence at 30°. (a) TE polarization and (b) TM polarization.

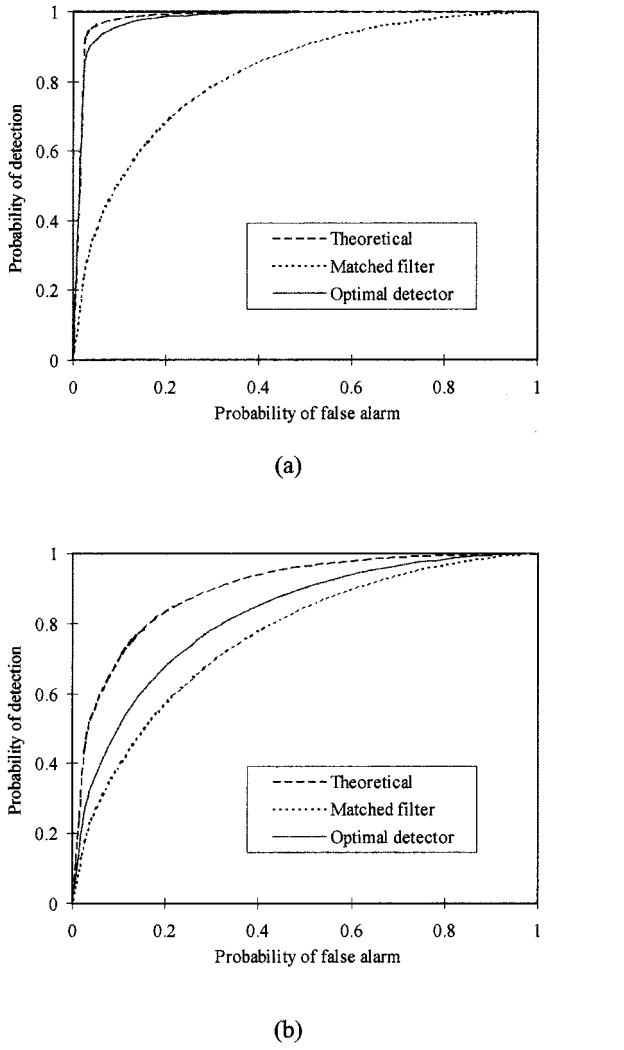


Fig. 13. Receiver operating characteristics for a Gaussian surface with $\sigma = 3.95$ cm and $L = 25$ cm, incidence at 65°. (a) TE polarization and (b) TM polarization.

focus here is on the impact of the rough-surface statistics on detection performance, and therefore the emphasis is placed on addressing several rough-surface statistical models.

The incident field is implemented as a pulsed plane wave, radiating in the presence of an infinite planar half space. The rough surface and target constitute perturbations to this incident wave, and yield the scattered fields. The incident time-domain waveform and associated frequency spectrum are shown in Fig. 5. This type of pulse shape (Rayleigh fourth order [24]) is typical of ultrawideband synthetic aperture radar (SAR) systems, and its central frequency (300 MHz) is also characteristic of applications involving subsurface sensing [1]. The discretization rate for the MRTD scheme is 40 Haar scaling functions per central wavelength in air, which implies about six Haar scaling functions per wavelength at the smallest wavelength in the spectrum, in the denser medium (when wavelets are considered as well, this corresponds to 12 cells per smallest wavelength in terms of equivalent Yee cells). The observations are always made in backscatter at a distance of $1000\lambda_c$ (far zone) from the target, where λ_c is the central wavelength of the pulse in air.

Of interest is the length of the rough surface employed in the computations. As indicated in Fig. 4, the rough surface is modeled as a finite-length perturbation to an infinite half space. The onset of the rough surface is smoothed such that diffraction at the edge of the rough surface is minimized. Moreover, recall that we are only interested in the clutter signature $\mathbf{c}$ in the vicinity of the (time-limited) target signature $\mathbf{s}$. Consequently, for calculation of the stochastic $\mathbf{e}$, $\mathbf{s}$ and $\mathbf{c}$, we only use the time-domain scattered fields for times, during which the incident pulsed plane wave is well separated temporally from diffraction at the rough-surface edges. To determine an appropriate rough-surface length, we compute the statistics of $\mathbf{s}$ and $\mathbf{c}$ for surface length L and separately for a length larger than L. When we no longer see a change in the scattered-field statistics (e.g., the first few coefficients of the autocorrelation sequence), the surface L is deemed long enough. For the cases studied here, we have found

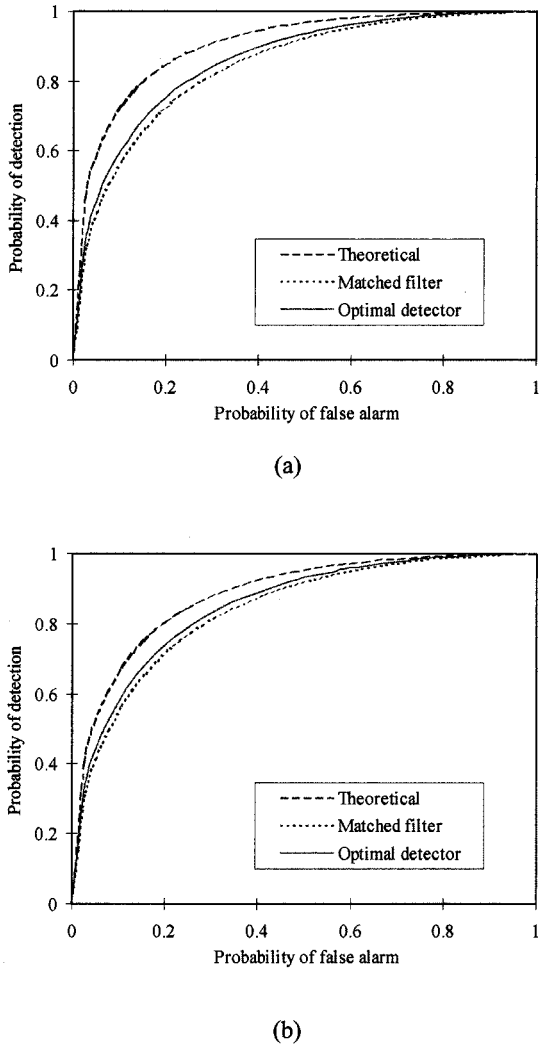


Fig. 14. Receiver operating characteristics for an exponential surface with $\sigma = 3.95$ cm and $L = 25$ cm, incidence at 30°. (a) TE polarization and (b) TM polarization.

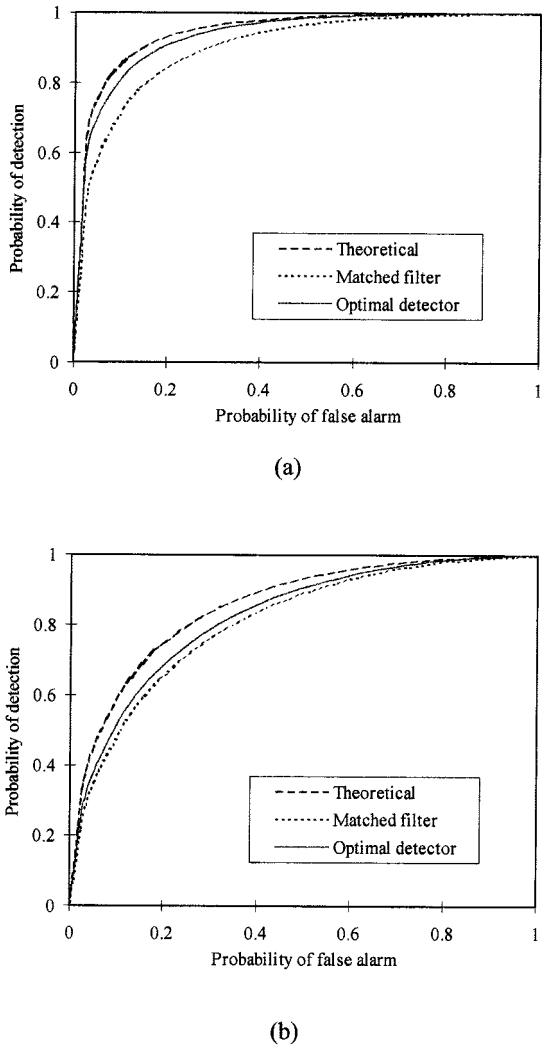


Fig. 15. Receiver operating characteristics for an exponential surface with $\sigma = 3.95$ cm and $L = 25$ cm, incidence at 65°. (a) TE polarization and (b) TM polarization.

that a surface length of about $12.5\lambda_c$ is normally sufficient, even at incidence angles of 65°. Numerical details on this issue can be found in [17].

B. Spectral Estimation

In Figs. 6–9, we plot the estimated power spectral density (PSD) of the clutter $\mathbf{c}$ for Gaussian, exponential, and Weierstrass fractal surfaces, at different angles of incidence, for both TE and TM polarization. The Gaussian (Fig. 6) and exponential (Fig. 7) surfaces under consideration have standard deviation $\sigma = 3.95$ cm and 25 cm correlation length, while the central wavelength of the incident pulse is $\lambda_c = 1$ m. The fractal surfaces are characterized by the standard deviation $\sigma = 6.25$ cm and fractal dimension $D = 1.2$ (Fig. 8) and $D = 1.5$ (Fig. 9). Other important parameters of the Weierstrass fractal surfaces are the order of the lowest and highest harmonics, N_1 and N_2. For fractal dimension $D = 1.2$, $N_1 = 2$ and $N_2 = 12$, while for fractal dimension $D = 1.5$, $N_1 = 2$ and $N_2 = 15$.

We note that the fundamental character of the PSD for the Gaussian and exponential surfaces (Figs. 6 and 7) is markedly different from that of the fractal surfaces (Figs. 8 and 9). This is interpreted as follows. Considering the Weierstrass fractal surfaces described in (3), we note that these surfaces are characterized by the superposition of harmonics with random phase. Each harmonic corresponds to a finite periodic surface, and therefore, diffraction from each is characterized by scattering in terms of a set of Floquet modes [20]. Each Floquet mode propagates at a frequency-dependent angle, with the angle of propagation representative of the backscattered direction at particular frequencies. This yields the fractal-surface induced clutter in Figs. 8 and 9, in which particular Floquet modes yield a strong backscatter response at particular frequencies. The number of Floquet modes excited is dependent on the excitation incidence angle, increasing as one gets closer to grazing.

Detector performance will be dictated largely by the ratio of the target-signature energy to the clutter energy. It is therefore of interest to consider the frequency-dependent target signature. In Fig. 10, we plot the frequency dependence of the target signature for the incidence angles and polarizations addressed in Figs. 6–9, for a flat air–soil interface. We notice that the target signature decreases in strength as the incidence angle increases.

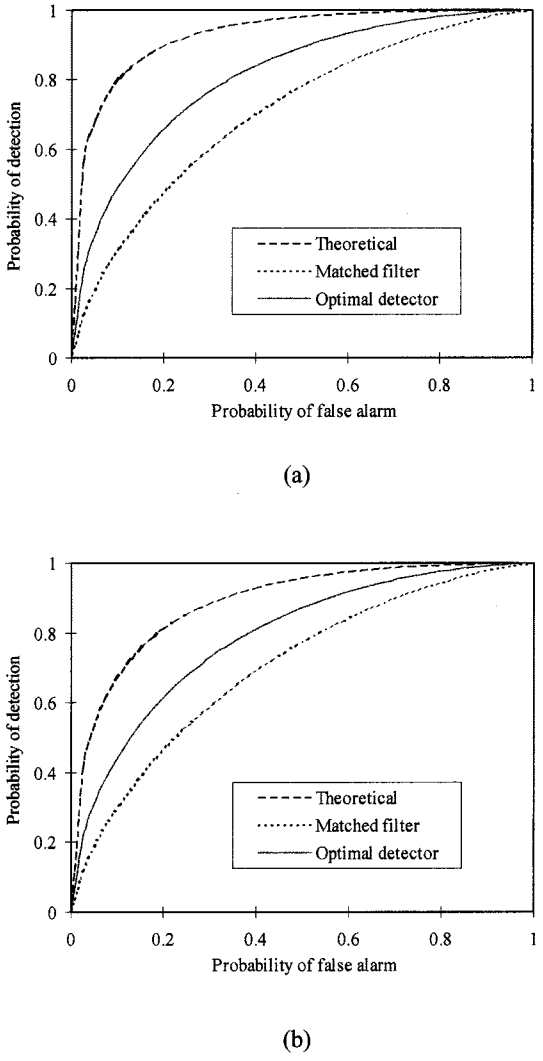


Fig. 16. Receiver operating characteristics for a fractal surface with $\sigma = 6.25$ cm and $D = 1.2$, incidence at 30°. (a) TE polarization and (b) TM polarization.

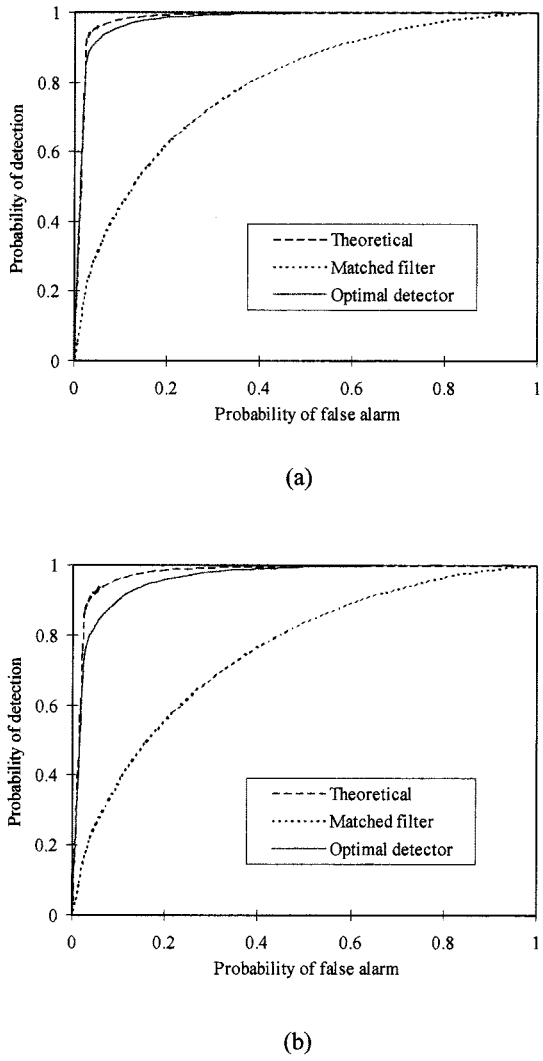


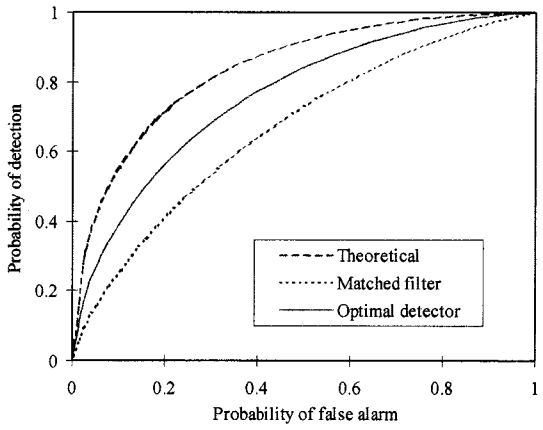
Fig. 17. Receiver operating characteristics for a fractal surface with $\sigma = 6.25$ cm and $D = 1.2$, incidence at 65°. (a) TE polarization and (b) TM polarization.

However, the clutter energy also decreases markedly with increasing incidence angle and as a result the target-to-clutter ratio is larger with increased incidence angle, thereby yielding improved detector performance with increasing incidence angle, as proved in the next section.

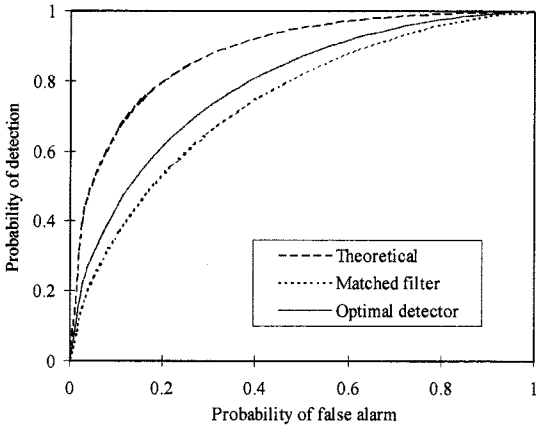
In Fig. 11, we plot the power spectral densities of the Gaussian and exponential surfaces considered here, as well as the incident-pulse spectrum. Note that the exponential surface has more high-frequency character, this manifested by more fine structure in the surface roughness. By contrast, at frequencies at which the incident pulse has maximum energy, the Gaussian surface is stronger. These curves will play an important role in analyzing the detector performance, in the next section.

C. Detector Performance

The detector performance is evaluated by plotting the ROC. This represents the variation of the probability of detection P_D as a function of the probability of false alarm P_F [by continuously varying the detector threshold T, see (4)–(6)]. All the relevant target and surface parameters have been described in the previous two sections. We plot three curves on each graph. One represents the idealized matched filter performance. In this case, the target signature is assumed to be deterministic and represented by its response under a flat interface. We obtain multiple data for hypothesis H_1 (target present) by superposing the flat-surface target signature with multiple clutter realization $\mathbf{c}$ obtained from multiple simulations of the rough interface. This idealized data is then processed with a matched filter, designed for the flat-surface target signature, yielding an upper bound on sensor performance (since in reality, as elucidated above, the target signature is random if the surface is rough). For this case, the ROC is based on the signal energy to clutter-variance ratio [11]. The second curve represents the actual matched filter performance, when the target signature is simulated rigorously under a rough interface, yielding $\mathbf{e} = \mathbf{c} + \mathbf{s}$ (with $\mathbf{s}$ random). In general, we see that the matched filter, designed for a flat-surface target response, performs quite poorly on the actual data. The third curve represents the performance of the optimal detector (Fig. 3), as applied to the rigorous data, with the perfor-

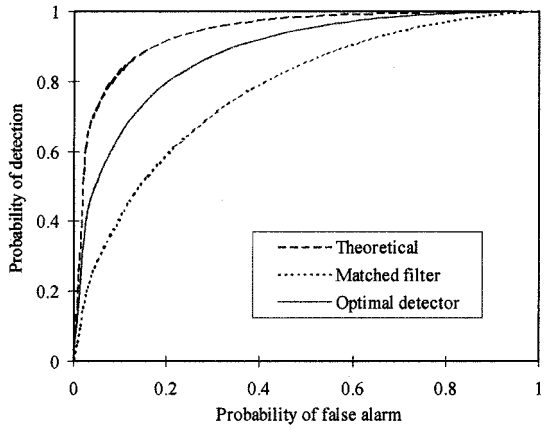


(a)

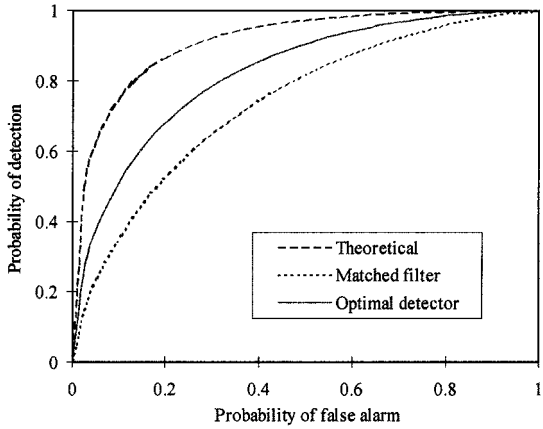


(b)

Fig. 18. Receiver operating characteristics for a fractal surface with $\sigma = 6.25$ cm and $D = 1.5$, incidence at 30°. (a) TE polarization and (b) TM polarization.



(a)



(b)

Fig. 19. Receiver operating characteristics for a fractal surface with $\sigma = 6.25$ cm and $D = 1.5$, incidence at 65°. (a) TE polarization and (b) TM polarization.

mance of this detector generally falling between the aforementioned two curves. In all cases, the data is prewhitened.

In the detector results, we consider the same physical parameters as addressed in Section V-B: Gaussian, exponential, and fractal surfaces, for both polarizations and incidence angles 30° and 65°. In Figs. 12 and 13, we consider Gaussian surfaces. Figs. 14 and 15 address exponential surfaces, Figs. 16 and 17 consider Weierstrass fractal surfaces with $D = 1.2$ and Figs. 18 and 19 investigate Weierstrass fractal surfaces with $D = 1.5$. The statistical parameters of the above surfaces are as in Section V-B. As discussed above, the fractal surfaces constitute a different class of physics than the Gaussian and exponential surfaces. Therefore, in the discussions below, we compare the relative characteristics of the detector for the Gaussian and exponential surfaces, and separately examine detector performance for the fractal surfaces.

The difference in performance between the optimal detector and the matched filter is largely dictated by the amount of randomization induced in the fields as they are transmitted through the rough interface. If there is little randomization of the transmitted incident fields, the target signature is nearly deterministic, for which the matched filter is optimal. To first order, the random surface manifests a *local* angle of incidence that varies randomly, as prescribed by the surface statistics. This is particularly true for relatively smooth rough surfaces, like the Gaussian surfaces. For such cases, the variation of the power transmitted into the ground, as a function of incidence angle, is therefore a good indication of this induced randomization. For the soil considered here, the variation of the transmitted power into the soil varies weakly for incident angles in the vicinity of 30°, for both polarizations. Consequently, for relatively smooth rough surfaces and an incident angle of 30°, there is minimal surface-induced randomization of the incident fields, and the matched filter (which assumes a deterministic targets signature) works relatively well. The TM-polarized plane wave has a Brewster angle of approximately 66° for the soil considered, and therefore variation of the TM transmitted fields is weak for incident angles in the vicinity of 65°. By contrast, the TE-polarized fields have no Brewster angle, and there is substantial variation of the transmitted fields for angles in the vicinity of 65°.

This is consistent with all the simulated results, which exhibit the largest optimal detector vs. matched filter performance improvement for 65° and TE polarization [this is particularly evident in Fig. 13(a)].

We also note that, for the cases considered, detection performance is generally enhanced for incident angles closer to grazing (i.e., performance is generally better for 65° incidence relative to 30°). As discussed previously, this is manifested because the ratio of target-signature energy to clutter energy is increased for larger incidence angles, for the examples examined. Concerning the relative detector performance for Gaussian and exponential surfaces, we note that the detector performance is similar for these two surfaces. However, we typically see more performance enhancement manifested by the optical detector for the Gaussian surface, relative to the exponential. This is because, for the same correlation length, the main difference between the Gaussian and exponential surfaces is more fine structure in the latter. The Gaussian surface is characterized by more large-scale variation, this having a greater impact on randomization of the transmitted fields over the bandwidth considered here. In particular, in Fig. 11 we notice that the Gaussian surface PSD has larger magnitude than the exponential surface in the region close to the central frequency of the incident pulse, where most of the electromagnetic energy is concentrated.

For the surfaces considered, the clutter from the Weierstrass fractal surfaces was considerably more intricate than that of the Gaussian and exponential surfaces considered (see Figs. 6–9). One would therefore expect that the fractal surfaces considered likewise randomize the transmitted fields more significantly than the Gaussian and exponential surfaces. This is manifested in Figs. 16–19, in which a noticeable increase in performance is realized via the optimal detector, *vis-à-vis* the matched filter. We note that the degree of randomization appears to be an intricate function of the incidence angle and the fractal dimension. In particular, we notice that the largest improvement in detector performance occurs at 65° (for both polarizations), and that this improvement is more marked for the surfaces with $D = 1.2$ than for $D = 1.5$.

VI. Conclusions

We have applied time-domain electromagnetic simulations to the problem of ultrawideband sensing of targets buried under a rough air-ground interface. Our main focus has been to examine time-domain electromagnetic scattering from three classes of surface statistics, and to examine the effects of such randomization on optimal subsurface target detection. In many cases, for which the target signature was significantly randomized, the optimal detection performance improvement, relative to a match filter, was substantial. For the surfaces considered, greater target-signature randomization was realized for the fractal surface, *vis-à-vis* the Gaussian and exponential surfaces. However, the details of these surfaces are very different, and therefore it is difficult to make direct comparisons.

There are many fertile areas of future research. In particular, the Haar-wavelet expansion represents the simplest realization of the MRTD model. One direction for future research consists of improving the numerical modeling scheme by considering other wavelet basis, with better smoothness properties. Another interesting direction involves studying scattering from Weierstrass fractal surfaces in greater detail, both analytically and numerically.

References

[1] M. A. Ressler and J. W. McCorkle, "Evolution of the Army Research Laboratory ultra-wideband test bed," in *Ultra-Wideband Short-Pulse Electromagnetics 2*, L. Carin and L. B. Felsen, Eds. New York: Plenum, 1995, pp. 109–123.

[2] T. Dogaru and L. Carin, "Time-domain sensing of targets buried under a rough air-ground interface," *IEEE Trans. Antennas Propagat.*, vol. AP-46, pp. 360–372, Mar. 1998.

[3] T. Dogaru, L. Collins, and L. Carin, "Optimal time-domain detection of a deterministic target buried under a randomly rough interface," *IEEE Trans. Antennas Propagat.*, vol. 49, pp. 313–326, Mar. 2001.

[4] M. Saillard and D. Maystre, "Scattering from metallic and dielectric rough surfaces," *J. Opt. Soc. Amer. A*, vol. 7, pp. 982–990, June 1990.

[5] E. I. Thorsos, "The validity of the Kirchoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Amer.*, vol. 83, pp. 78–92, 1988.

[6] E. I. Thorsos and D. R. Jackson, "The validity of the perturbation approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Amer.*, vol. 86, pp. 261–277, 1989.

[7] F. D. Hastings, J. B. Schneider, and S. L. Broschat, "A Monte–Carlo FDTD technique for rough surface scattering," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 1183–1191, Nov. 1995.

[8] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech, 1995.

[9] M. Krumpholz and L. Katehi, "MRTD: New time-domain schemes based on multiresolution analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 555–571, Apr. 1996.

[10] T. Dogaru and L. Carin, "Multiresolution time-domain analysis of scattering from a rough dielectric surface," *Radio Sci.*, vol. 35, pp. 1279–1292, Nov.–Dec. 2000.

[11] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.

[12] L. Tsang, G. F. Zhang, and K. Pak, "Detection of a buried object under a single random rough surface with angular correlation function in EM scattering," *Microw. Opt. Tech. Lett.*, vol. 12, p. 375, Aug. 20, 1996.

[13] G. F. Zhang and L. Tsang, "Angular correlation function of wave scattering by a random rough surface and discrete scatterers and its application in the detection of a buried object," *Waves Random Media*, vol. 7, pp. 467–478, July 1997.

[14] G. F. Zhang, L. Tsang, and K. Pak, "Angular correlation function and scattering coefficient of electromagnetic waves scattered by a buried object under a two-dimensional rough surface," *J. Opt. Soc. Amer. A.*, vol. 12, pp. 2995–3002, Dec. 1998.

[15] G. F. Zhang, L. Tsang, and Y. Kuga, "Studies of the correlation function of scattering by random rough surfaces with and without a buried object," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 444–453, Mar. 1997.

[16] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[17] T. Dogaru, "Modeling and signal processing for electromagnetic subsurface sensing,", Ph.D. dissertation, Duke University, Durham, NC, 1999.

[18] R. F. Harrington, *Field Computation by Moment Methods*. Malabar, FL: Krieger, 1968.

[19] B. B. Mandelbrot, *The Fractal Geometry of Nature*. New York: W. H. Freeman, 1983.

[20] D. L. Jaggard and Y. Kim, "Diffraction by band-limited fractal screens," *J. Opt. Soc. Amer. A*, vol. 4, pp. 1055–1062, June 1987.

[21] S. Rouvier, P. Borderies, and I. Chenerie, "Ultra-wideband electromagnetic scattering of a fractal profile," *Radio Sci.*, vol. 32, pp. 285–293, Mar.–Apr. 1997.

[22] E. Bachelier, C. Ruiz, P. Borderies, I. Chenerie, and M. Davidson, "Weierstrass functions determination for soil modeling," in *Proc. Int. Geoscience and Remote Sensing Symp.'98*, Seattle, WA, 1998.

[23] M. V. Berry and Z. V. Lewis, "On the Weierstrass–Mandelbrot fractal function," *Proc. R. Soc. London*, pp. 459–484, 1980.

[24] P. Hubral and M. Tygel, "Analysis of the Rayleigh pulse," *Geophysics*, vol. 54, pp. 654–658, 1989.



Traian Dogaru He received the engineering degree from the Polytechnic University of Bucharest in 1990, the M.S. degree in electrical engineering in 1997, and the Ph.D. degree in 1999, both from Duke University, Durham, NC.

From 1992 to 1995, he held different engineering positions in the magnetic recording industry. He is currently a Research Associate with Duke University, and his main research interests are in electromagnetic wave theory, computational electromagnetics, rough surface scattering, and radar-related signal processing.

Lawrence Carin (SM'96–F'01) and received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor, and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department at Duke University, Durham, NC, where he is now a Professor. He is the Principal Investigator on a Multidisciplinary University Research Initiative (MURI) on demining. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies and is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.

Rigorous Modeling of Ultrawideband VHF Scattering From Tree Trunks Over Flat and Sloped Terrain

Jiangqi He, Norbert Geng, *Member, IEEE*, Lam Nguyen, and Lawrence Carin, *Fellow, IEEE*

***Abstract*—Three electromagnetic models are employed for the investigation of ultrawideband VHF scattering from tree trunks situated over flat and sloped terrain. Two of the models are numerical, each employing a frequency-domain integral-equation formulation solved via the method of moments (MoM). A body-of-revolution (BoR) MoM formulation is applied for a tree trunk on a flat terrain, implying that the BoR axis is perpendicular to the layers of an arbitrary layered-earth model. For the case of sloped terrain, the BoR model is inapplicable, and therefore the MoM solution is performed via general triangular-patch basis functions. Both MoM models are very accurate but are computationally expensive. Consequently, we also consider a third model, employing approximations based on the closed-form solution for scattering from an infinite dielectric cylinder in free space. The third model is highly efficient computationally and, despite the significant approximations, often yields accurate results relative to data computed via the reference MoM solutions. Data from the three models are considered, and several examples of application to remote sensing are addressed.**

I. Introduction

SCATTERING from tree trunks plays an important role in many remote sensing applications. For example, in foliage-penetrating (FOPEN) radar, scattering from tree trunks constitutes the principal source of clutter [1]. FOPEN radar systems typically operate at VHF and low-UHF frequencies, for which foliage penetration is reasonably effective. Consequently the canopy-induced clutter [2] is typically much weaker than the clutter signatures induced by dihedral scattering at tree trunks. In particular, several proximate tree trunks can yield composite radar signatures that are very similar to those produced by man-made targets. Before addressing scattering from multiple tree trunks, we must first fully understand the scattered fields induced by a single tree trunk, as a function of terrain topography, this constituting the subject of the present paper.

The scattering of electromagnetic waves via finite-height tree trunks is investigated through development of a method-of-moment (MoM) model for dielectric targets in the presence of a layered medium. Two distinct MoM formulations are considered. In particular, it is often appropriate to model the tree trunk as a lossy, dielectric body of revolution (BoR), with BoR axis perpendicular to the planar surfaces in the layered medium. Such a MoM model is often termed a 2.5-dimensional (2.5-D) solution, since the basis functions exist along a two-dimensional (2-D) generating arc, with the third dimension accounted for efficiently via cylindrical Fourier harmonics [3]. Consequently, the BoR model is typically computationally much more efficient than a general MoM solution, for which the induced electric and magnetic surface currents on the dielectric surface are expanded in terms of general triangular-patch basis functions [4]. The BoR tree-trunk model is applicable to a general generating arc, and therefore it can be used to investigate the effects of tree curvature on the scattered fields, *vis-à-vis* a simpler model that assumes the tree trunk is a constant-radius cylinder [5].

Despite the attractiveness of the BoR model, it is limited to the case of a flat terrain, i.e., the tree-trunk axis must be perpendicular to the air–ground interface. For the case of a sloped terrain, the BoR model is inappropriate. Consequently, we have developed a separate MoM algorithm, applicable to general dielectric targets embedded in an arbitrary multilayered environment. This MoM model employs the general Rao, Wilton, and Glisson [4] triangular-patch basis functions, and allows the consideration of sloped terrain as well as general tree-trunk shapes.

The two MoM models discussed above are quite powerful for gaining insight into the phenomenology associated with VHF and low-UHF scattering from tree trunks over soil. However, as the target size grows relative to wavelength, they require prohibitive computer memory as well as computation time. Consequently, where possible, it is desirable to use the rigorous MoM models as reference solutions for alternative, approximate models that are computationally efficient. In this context we consider the following approximate model. If we assume that the tree trunk is composed of a constant radius, the electric and magnetic currents on the tree-trunk surface can be approximated by considering scattering from an infinite cylinder in free space, while assuming two incident fields: (1) the direct plane wave, as in free space, plus (2) the contribution due to plane-wave reflection induced by the soil (ground bounce). Moreover, when computing the scattered fields, we must also account for the ground-bounce-induced fields, in a manner similar to that employed for the incident fields. The above model assumes that the surface currents are weakly affected by diffraction induced at the ends of the finite-height tree trunk. Lin and Sarabandi [5] considered a similar model, in which they also accounted for the effects of the tree bark. However, the focus of [5] was on microwave scattering, for which the tree bark is often important. Here we are primarily interested in FOPEN radar, and at the VHF and low-UHF frequencies associated with such, the

Manuscript received December 9, 1999; revised October 25, 2000.

J. He and L. Carin are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

N. Geng is with the Institute for Microwaves and Electronics, University of Karlsruhe, Karlsruhe, Germany.

L. Nguyen is with the Army Research Laboratory, AMSRL-SE-RU, Adelphi, MD 20783 USA.

Publisher Item Identifier S 0196-2892(01)08838-6.

0196–2892/01$10.00 © 2001 IEEE

tree bark is expected to be of secondary consideration. We note that in [5], the authors did not have access to a numerically rigorous model such as the MoM, and therefore, the accuracy of the aforementioned approximate model could not be confirmed. Our reference MoM solution allows us to perform such a calibration.

The scattering of electromagnetic waves from trees has been investigated by several authors using numerous techniques [6]–[11]. In addition to the aforementioned infinite-cylinder model [5], authors have examined the use of volumetric electric-field integral equations, solved via an electric-dipole model [9]. This algorithm is applicable to tree models far more realistic [9] than a simple finite-length cylinder. In [9], the free-space Green's function was employed with appropriate reflection coefficients appended to account for soil interactions. Authors have also considered scattering from multiple trees [6], [8], employing, for example, an iterative procedure to account for intertree electromagnetic interaction [6]. In [6], the fields within a given tree (finite-length dielectric cylinder) are computed via an infinite-cylinder model, however the scattered fields from the cylinders are computed using the dyadic half-space Green's function. In the work employed here, in the context of our MoM solutions, we employ a surface-integral-equation formulation, and the currents on the target surface are computed rigorously through explicit use of the half-space Green's function (the scattered fields are also computed via the half-space Green's function).

The remainder of the paper is organized as follows. The details of the two MoM models employed here have been elucidated elsewhere [3], [12], and therefore in Section II, we provide a brief overview of these techniques. Their application to computing the tree-trunk radar cross section (RCS) is provided in some detail, such that it is understood how we account for diffractive effects induced at the top of the (modeled) tree trunk. Finally, we also briefly summarize the approximate model, the details of which are found in [5]. In Section III, a large set of results are presented, in which we examine tree-trunk scattering from several perspectives. In particular, we compare the relative accuracy of the approximate and two MoM solutions, for flat and sloped terrain. Several phenomenological examples are addressed, including the aspect-dependent synthetic aperture radar (SAR) signature of a tree trunk on a sloped terrain.

II. Theoretical Formulation

A. Method-of-Moments (MoM) Analyses

We have developed MoM models applicable to a dielectric body of revolution (BoR) embedded in an arbitrary layered medium, as well as for a general dielectric target also embedded in an arbitrary layered medium [3], [12]. For the former, along the target-generating arc, we employ the subsectional basis functions discussed in [3], while for the latter, we employ the general triangular-patch basis functions introduced in [4]. MoM scattering models for BoR and general dielectric targets have been available for over a decade for the case of the target residing in free space [13]. To consider scattering from such a target embedded in an arbitrary layered medium, the problem is complicated significantly by the need to compute the layered-medium dyadic Green's function [14]. As is well known, in general, each component of the dyadic Green's function cannot be represented in closed form, but must be expressed in terms of a generally highly oscillatory Sommerfeld integral [14], [15]. The efficient evaluation of such integrals has been the principal bottleneck to generalizing the free-space models in [13] to the case of general layered media.

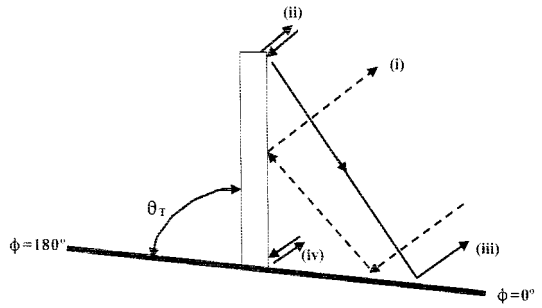


Fig. 1. Schematic of a lossy dielectric cylinder (model tree trunk) on a sloped terrain. Four scattering mechanisms, discussed in the text, are identified.

Recently the method of complex images has been developed, in which a fit is performed to the spectral Green's function, along a salutary path in the complex spectral space. The fit is performed using a sum of exponentials, each with a complex argument [3], [16], [17]. The attractiveness of this approach is that, once the sum of exponentials is fit to the spectral Green's function, a closed-form inversion exists for the Sommerfeld integral associated with each exponential, cumulatively yielding the space-domain Green's function. Here we use a Prony [18] fitting procedure, although any of a number of such techniques are available [19]. The complex-image technique is most effective when the target lies entirely within a single layer, since in this case all source and observation points are also within the same layer. Under this circumstance the spectral fit is performed once, for each dyadic component, for all source and observation points, significantly accelerating computational efficiency. For scattering from tree trunks, the dielectric scatterer is entirely above the soil (we ignore the root system) and therefore the complex-image technique offers an attractive means of implementing both of the aforementioned MoM models. The interested reader is referred to [3], [16], [17] for more details on the complex-image technique.

B. Scattering Mechanisms

As indicated, the details of the MoM formulations can be found elsewhere [3], [12] and therefore, here we focus on issues of particular interest to tree-trunk scattering. In particular, consider plane-wave scattering from a finite-length tree trunk over soil, as shown in Fig. 1. There are four principal scattering mechanisms: i) a dihedral response due to the tree-soil interaction; ii) direct diffraction from the top of the tree trunk; iii) diffraction from the top of the tree trunk followed by a ground bounce; and iv) diffraction from the bottom of the tree trunk (at the air-ground interface). In addition to these scattering mechanisms, there are multiple interactions between the four, as well as other higher order effects. Scattering mechanisms (ii)-(iv) are

always present in the far-zone *backscattered* fields, while i) is present when the tree trunk is normal to the air-ground interface (e.g., flat terrain) and for special incident angles [5] when the terrain is sloped.

It is important to emphasize that, for a real tree with a branch and leaf structure, the diffraction mechanisms from the "top" of the tree will be different than those in Fig. 1. However, if there is a backscattered dihedral effect observed at the sensor (flat terrain), it is typically much stronger than leaf and branch scattering at VHF frequencies, and therefore mechanism i) dominates. To explore this backscattered effect in detail, when it exists, we perform the following procedure. Considering Fig. 1, scattering mechanism ii) is observed earliest in time, for transient excitation of sufficient resolution (wide enough bandwidth). To investigate the radar cross section (RCS) of mechanism i), we window out in time the contribution from ii), leaving principally the time domain response of i), iii), and iv). Mechanisms iii) and iv) can be viewed as components of i), when the latter is observed in backscatter, with the sum of the two generally being much weaker than mechanism i), which typically occurs during an overlapping time window. After windowing out ii), the transient scattered waveform is converted back to the frequency domain, the incident-pulse spectrum removed (deconvolved), yielding an approximation for the RCS of i). As elucidated above, this model will well represent the backscattered dihedral response at VHF frequencies, when it exists, despite the fact that the leaf and branch structure are ignored.

For the case of sloped terrain, scattering mechanism i) is generally *not* observed in backscatter, leaving mechanisms ii)–iv). Like the dihedral effect i), scattering mechanism iv) is generally weakly affected by the leaf and branch structure. However, mechanisms ii) and iii) are driven primarily by diffraction at the "top" of the tree trunk, this ignoring effects of the leaves and branches. As mechanisms ii)–iv) will generally drive the backscattered signature for the case of sloped terrain, it is important to emphasize that the aforementioned approximations must be confirmed via VHF measurements, this constituting an important area of future research. We note as well that the MoM model with triangular-patch basis functions, discussed above, can in principle be used to model the tree branch structure, with the leaves generally being of secondary importance at VHF frequencies. This too is an area of future work.

C. Data Format

Below we present both frequency- and time-domain results, with the former presented as RCS. For examples in which the terrain is flat (tree trunk normal to the air-ground interface), the RCS is principally characterized by the dihedral scattering represented by i) in Fig. 1, where we have windowed out the effects of ii) as discussed above. For sloped terrain, mechanism i) is typically not present in backscatter, and therefore the RCS in this case is representative of mechanisms ii)-iv), plus associated higher-order effects. All scattering mechanisms are retained in the time-domain results, such that the associated phenomena are observed. The far-zone time-domain backscattered results are normalized as follows. The frequency-domain, far-zone polarimetric scattering matrix can be expressed as

$$\begin{bmatrix} E_V^{scat} \\ E_H^{scat} \end{bmatrix} = \frac{\exp(-jk_1 r)}{r} \begin{bmatrix} S_{VV} & S_{VH} \\ S_{HV} & S_{HH} \end{bmatrix} \begin{bmatrix} E_V^{inc} \\ E_H^{inc} \end{bmatrix} \quad (1)$$

where k_1 is the wavenumber of the top region (usually air), and r is the nominal far-zone distance from the target to the observer. The *normalized* time-domain results are defined as

$$\begin{bmatrix} w_{VV}(t) & w_{VH}(t) \\ w_{HV}(t) & w_{HH}(t) \end{bmatrix} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \begin{bmatrix} S_{VV}(\omega) & S_{VH}(\omega) \\ S_{HV}(\omega) & S_{HH}(\omega) \end{bmatrix} \begin{bmatrix} P_V(\omega) \\ P_H(\omega) \end{bmatrix} \exp(j\omega t)\, d\omega \quad (2)$$

where P_V and P_H represent the frequency dependence of the V (vertical) and H (horizontal) polarized incident-field components.

D. Series Approximation

The MoM yields a rigorous solution for the electric and magnetic surface currents induced on the target surface, from which the scattered fields can be computed. If the tree trunk is assumed to be a constant-radius dielectric target, with possibly multiple concentric layers, one can employ the exact solution for free-space scattering from an infinite dielectric cylinder [5]. The tangential electric and magnetic fields on this infinite target can be used to approximate the electric and magnetic surface currents on a finite-length dielectric cylinder. Moreover, the solution must be modified slightly to account for ground-bounce scattering [5]. For comparison, we consider such an approximate model, for a lossy dielectric cylinder over a flat and sloped half space. Here we assume the cylinder is homogeneous, although it is straightforward to extend the model to account for tree layering [5].

III. Example Results

A. Comparison Between Two MoM Models

As discussed in Section II, two MoM formulations have been implemented, one assuming that the dielectric target is a body of revolution (BoR) and the other employing general triangular-patch basis functions [4], the latter applicable to general dielectric targets. We begin by comparing the results computed by these two models, for the special case of a BoR tree trunk, and therefore we assume flat terrain. Moreover, although the MoM models are applicable to general layered media, here and throughout the paper we assume the soil is a lossy, dispersive half space. We perform the comparisons in the time domain, assuming time-domain plane-wave excitation with the incident pulse in Fig. 2. This waveform encompasses the VHFand low-UHF frequency bands. The upper frequency considered is dictated by the memory and computation-time (CPU) requirements of the MoM models. In particular, the size of the subsectional basis functions must be small relative to the wavelength in the tree trunk [3], [13], and therefore the number of basis functions N grows precipitously as the target size increases relative to wavelength i.e., as the frequency

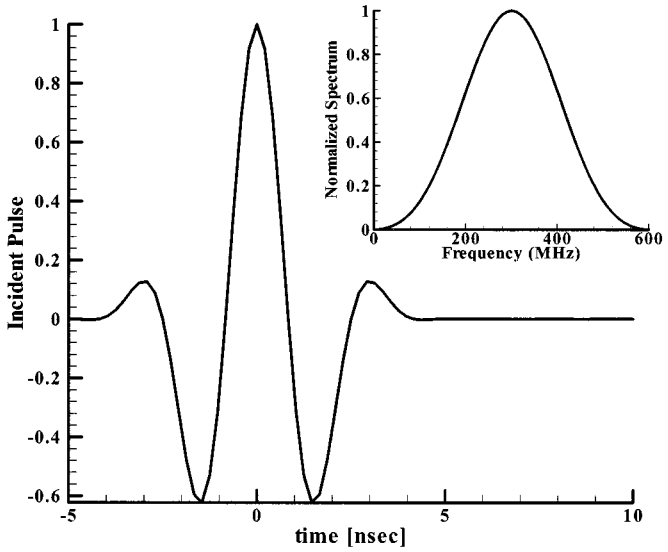


Fig. 2. Incident pulse and associated spectrum, used for the time-domain scattering examples.

increases). For a traditional MoM analysis, the memory requirements are order N^2 and the CPU requirements are order N^3, for a direct solution of the MoM matrix equation (the CPU is order PN^2 for a conjugate-gradient solver, where P is the number of conjugate-gradient iterations). We note that the BoR model is typically more efficient from both a memory and CPU perspective than the triangular-patch model, since the basis functions for the former only exist along a 2-D generating arc [3]. However, the number of azimuthal modes [3] required of the BoR model increases with target electrical size, i.e., with increasing frequency), and therefore the BoR model is as well limited in the target size that can be considered.

In Fig. 3, we compare the time-domain scattered fields computed via the two models, considering tree trunks of 30 cm diameter and heights of 3, 4, and 5 m, for an incidence angle of 50° relative to the normal (40° relative to grazing). The tree trunk is characterized by the complex permittivity $\epsilon = \epsilon_0(3 - \mathrm{j}0.5 - \mathrm{j}\sigma/\omega\epsilon_0)$ with conductivity $\sigma = 0.005$ S/m, where ϵ_0 is the free-space permittivity. The soil is representative of Yuma soil with 10% water content (see [20]). The VV and HH scattered fields are presented in Fig. 3(a) and 3(b), respectively. For this BoR target, the cross-polarized *backscattered* fields are zero [21]. We first note that the agreement between the two MoM solutions is generally good, although there are slight differences near the peaks of the waveforms, these associated with the relatively high-frequency components of the waveform. This is attributed as follows. For the BoR model, we employed a minimum of ten subsectional basis functions per wavelength inside the tree trunk, while for the triangular-patch model, at the highest frequencies (see Fig. 2) we applied approximately six basis functions per wavelength. The latter discretization was applied due to the required three-dimensional (3-D) current expansion on the target surface, this becoming computationally expensive as the frequency increases (see discussion in previous paragraph). Nevertheless, the two solutions are in good agreement.

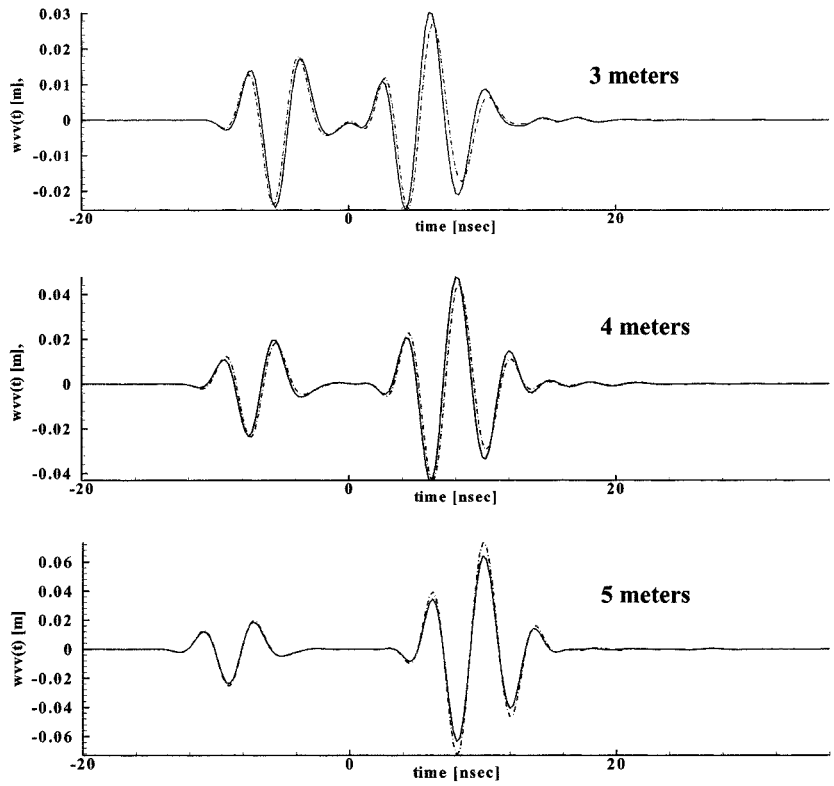
Focusing now on the phenomenology reflected in Fig. 3, we note that, as expected, the time-domain scattered fields are characterized initially by direct diffraction from the "top" of the tree trunk. The tree height and bandwidth are sufficient that this effect, mechanism ii) in Fig. 1, is resolvable in time from the other scattering effects. As expected, the initial scattered waveform is independent of the tree height. By contrast, the second return, this representing a composite of mechanisms i), iii), and iv) from Fig. 1, grows with increasing tree-trunk height. In fact, this portion of the scattered waveform is almost a linear function of the tree-trunk height. A notable distinction seen in Fig. 3 is that the dihedral response i) in Fig. 1 is significantly larger for the HH scattered fields relative to the VV case. This is attributed to generally enhanced penetration at a dielectric interface for vertically polarized incidence relative to the horizontal case, thereby diminishing the dihedral interaction between the soil and the tree trunk. For the soil considered here, the Brewster angle is approximately 65° (the soil is dispersive so the Brewster angle is frequency dependent, and moreover, since the soil is lossy, the concept of a Brewster angle is itself approximate). The 50° angle of incidence at the soil interface therefore yields considerable penetration for V polarization, with this significantly less for H polarization.

Before leaving Fig. 3, we note that for both the VV and HH cases, there are relatively little scattered fields after the principal dihedral response. This implies that, for the target considered, there is little multiple interaction between the various scattering mechanisms.

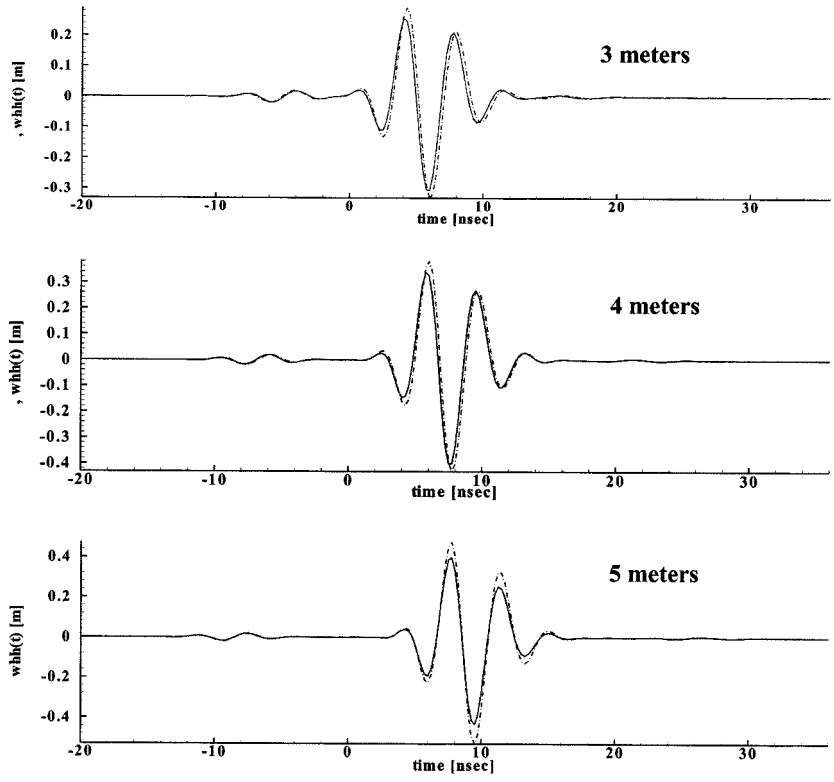
B. Diameter and Electrical Properties of Tree Trunk

The previous example served to validate the relative accuracy of the two MoM solutions and to provide an introduction to the tree-trunk scattering phenomenology. However, the results in Fig. 3 addressed a single tree diameter and tree trunks of fixed electrical parameters. Addressing the latter, the electrical parameters of tree trunks are expected to be a strong function of the tree type as well as of the season. In the next set of examples, we consider trees composed of the electrical parameters considered in Fig. 3$\epsilon = \epsilon_0(3 - \mathrm{j}0.5 - \mathrm{j}\sigma/\omega\epsilon_0)$ and $\sigma = 0.005$ S/m as well as tree trunks characterized by $\epsilon = \epsilon_0(10 - \mathrm{j}2 - \mathrm{j}\sigma/\omega\epsilon_0)$ and $\sigma = 0.02$ S/m. An investigation of these relatively disparate examples will help quantify the effect of the tree electrical parameters on the scattered signature. Moreover, the tree height is fixed at 5 m, the terrain is assumed flat, composed of Puerto Rico clay soil [22], and several tree-trunk diameters are considered. Although we consider a fixed tree height in these examples, the data in Fig. 3 indicate that these results will scale linearly with varying tree-trunk height. Finally, as discussed in Section II-B and Section II-C, the frequency-domain RCS calculations presented here do not include scattering mechanism ii) in Fig. 1 (it being windowed out in the time domain).

Having demonstrated the relative accuracy of the two MoM solutions, for the case in which the tree trunk is normal to the air-ground interface, here we only consider computations performed with the BoR MoM model. Further, for comparison, we also consider scattered fields computed via the approximate series-solution-based model [5] discussed in Section II-D. In Fig. 4, we consider the RCS computed by these two models for tree trunks characterized by $\epsilon = \epsilon_0(3 - \mathrm{j}0.5 - \mathrm{j}\sigma/\omega\epsilon_0)$ and



(a)



(b)

Fig. 3. Time-domain waveform backscattered from a lossy dielectric cylinder (complex permittivity $\epsilon = \epsilon_0(3 - \mathrm{j}0.5 - \mathrm{j}\sigma/\omega\epsilon_0)$ with conductivity $\sigma = 0.005$ S/m) situated atop a half space composed of Yuma soil with 10% water content [20]. The tree trunk has a 30 cm diameter, and is normal to the air-ground interface. The pulsed plane wave has the shape shown in Fig. 2, and is incident at 50° relative to the normal, and results are shown for trunk heights of 3, 4, and 5 m. The solid curves were computed via the body of revolution MoM model [3], and the dashed via the triangular-patch model [12]. (a) VV polarization and (b) HH polarization.

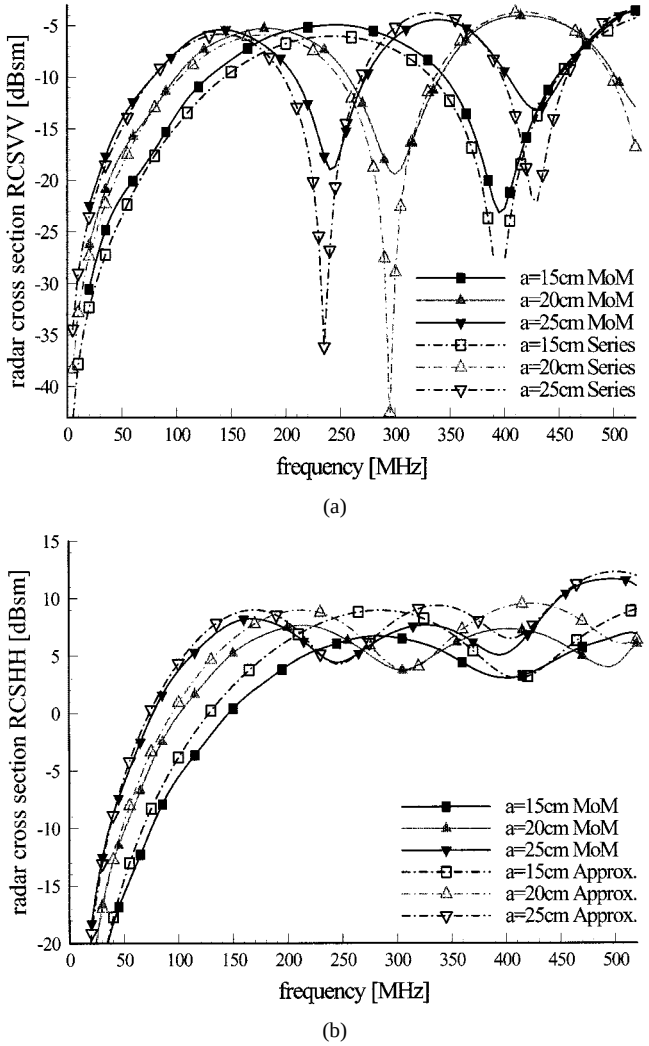


Fig. 4. Radar cross section (RCS) of a tree trunk normal to a half-space interface, with the half space electrical properties characteristic of Puerto Rico clay [22], and the plane wave is incident at 60° relative to the air-soil normal. The tree trunks are characterized by $\epsilon = \epsilon_0(3 - \mathrm{j}0.5 - \mathrm{j}\sigma/\omega\epsilon_0)$ with $\sigma = 0.005$ S/m, and tree diameters of 30, 40, and 50 cm are considered. The solid curves represent the BoR MoM solution [3], and the dashed curves the approximate series solution [5]. (a) VV polarization and (b) HH polarization.

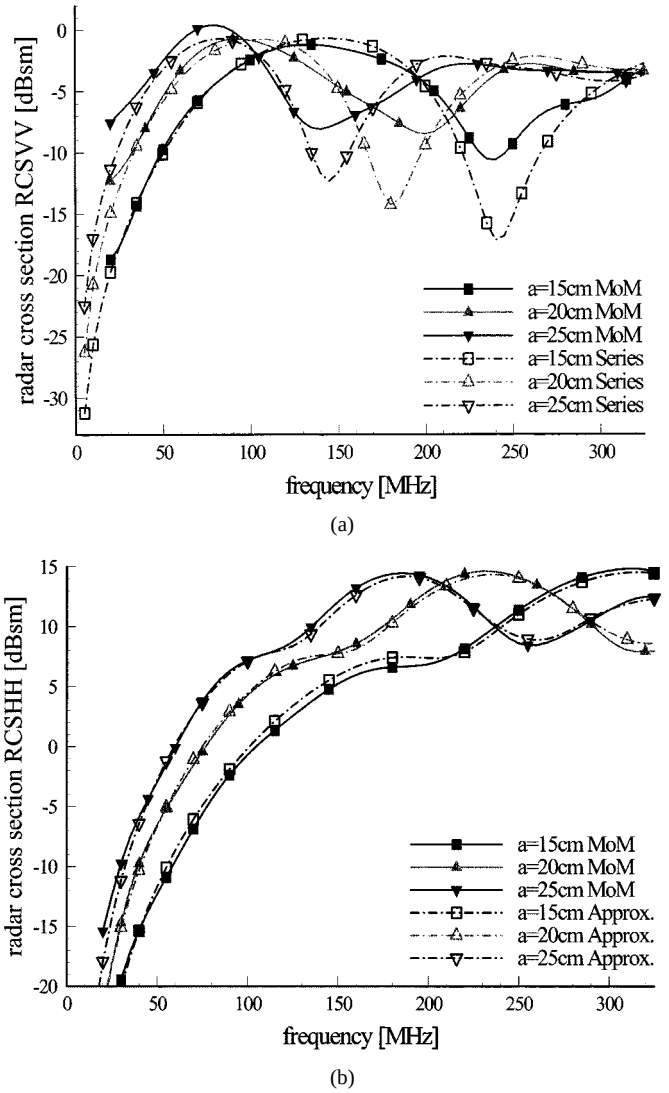
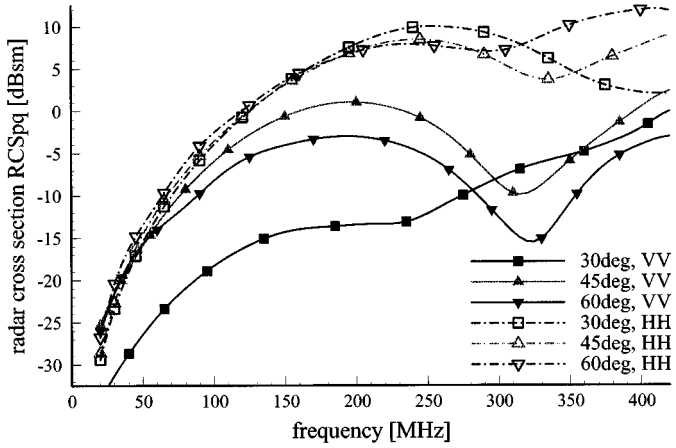


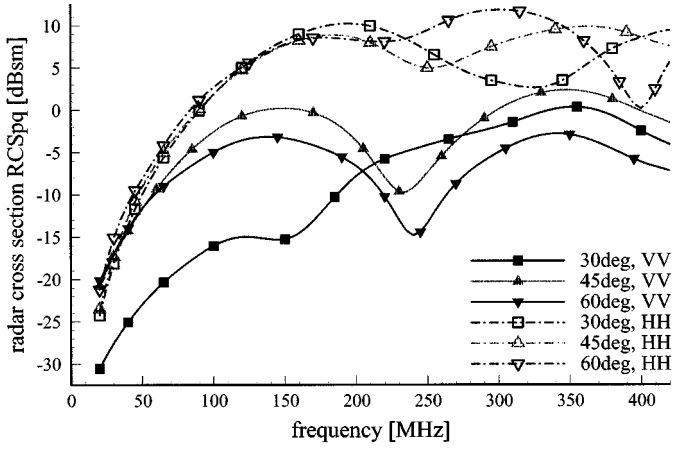
Fig. 5. As in Fig. 4, but now the tree trunks are characterized by $\epsilon = \epsilon_0(10 - \mathrm{j}2 - \mathrm{j}\sigma/\omega\epsilon_0))$ with $\sigma = 0.02$ S/m. (a) VV polarization and (b) HH polarization.

$\sigma = 0.005$ S/m for tree diameters of 30, 40, and 50 cm. The plane-wave incident angle is 60° relative to the air-soil normal. We again note that the VV response [Fig. 4(a)] is markedly smaller than that for the HH polarization [Fig. 4(b)]. We also note that, as expected, the RCS increases with frequency as the wavelength becomes smaller relative to the tree-trunk diameter (approaching the "optical" regime). Consequently, the frequency at which the RCS starts to reach its maximum increases as the tree diameter decreases. There are several other interesting scattering mechanisms playing a role in this example. In particular, note that, in the high-frequency region, the RCS is characterized by frequency-dependent dips in the RCS response, with these dips particularly strong for VV polarization. This phenomenon is also closely related to the tree-trunk diameter since the dips scale to lower frequencies as the diameter increases. We therefore attribute these dips to frequencies at which, in backscatter, the superposition of the fields scattered from the front and back of the tree trunk destructively interfere. This phenomenon is expected to be stronger for VV polarization, since the penetration into the tree trunk is expected to be larger in this case.

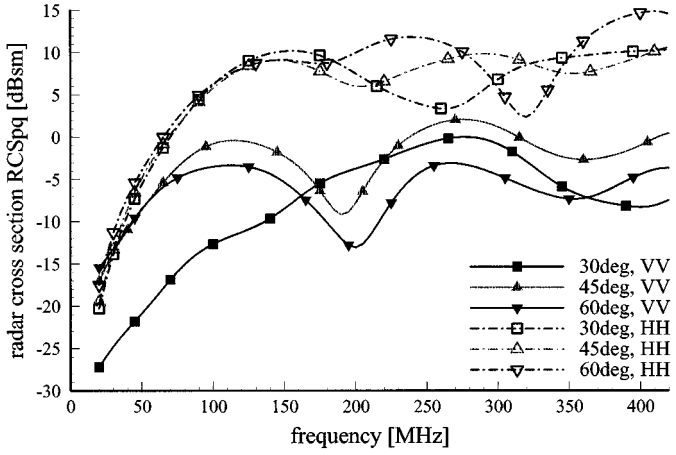
Another matter of note is the relative agreement between the MoM and the approximate series solution [5] for the target RCS. In Fig. 4(a), the agreement between these two models is good for the VV case, while for the HH polarization considered in Fig. 4(b) the agreement is excellent, especially for the largest diameter considered. Recall from Fig. 1 that, in addition to the dihedral response characterized by mechanisms i) and iv), there are scattered fields induced by top diffraction followed by a ground bounce, reflected by mechanism iii). For the HH case, the cumulative effects of i) and iv) are expected to be large, with this less so for VV polarization (due to enhanced penetration for the latter). Consequently, mechanism iii) is expected to be of stronger *relative* strength for VV than for HH polarization. The approximate RCS calculation, based on the series solution for an *infinite* cylinder [5], does not provide a good representation for the diffraction induced at the top of the tree trunk. As noted, this effect is of greater relative importance for the VV



(a)



(b)

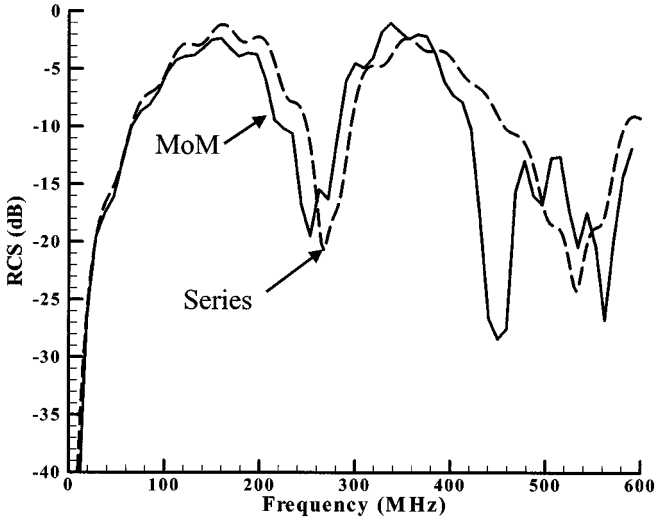


(c)

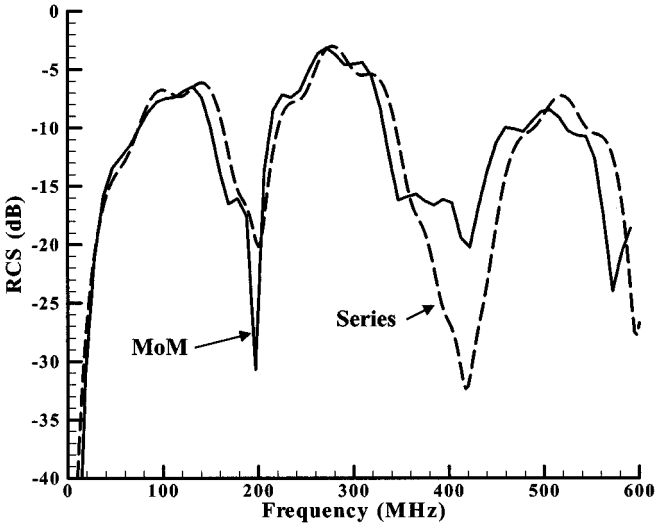
Fig. 6. Radar cross section of tree trunks composed of permittivity $\epsilon = \epsilon_0(5 - \mathrm{j}1 - \mathrm{j}\sigma/\omega\epsilon_0)$, with conductivity $\sigma = 0.01$ S/m, 5 m trunk height, and the half space is modeled as Puerto Rico clay [22]. The trunks are normal to the air-soil interface, and the plane wave is incident at 30°, 45°, and 60° relative to the normal to the air-soil interface. (a) 30-cm tree diameter, (b) 40-cm tree diameter, and (c) 50 cm tree diameter.

case, and therefore the approximate series solution is expected to be less accurate for this polarization.

In Fig. 5, we consider the same example as in Fig. 4, but now the permittivity and conductivity of the target are raised to $\epsilon = \epsilon_0(10 - \mathrm{j}2 - \mathrm{j}\sigma/\omega\epsilon_0)$ and $\sigma = 0.02$ S/m. We expect less



(a)



(b)

Fig. 7. Radar cross section for HH polarization, for a 10° sloped terrain with electrical properties characteristic of Yuma soil with 10% water content [20]. The tree-trunk diameter is 30 cm, its height is 5 m, and its electrical properties are characterized by $\epsilon = \epsilon_0(3 - \mathrm{j}0.5 - \mathrm{j}\sigma/\omega\epsilon_0)$ with conductivity $\sigma = 0.005$ S/m. The azimuthal angle $\phi = 0°$ corresponds to viewing up the slope, with $\phi = 180°$ viewing downward (see Fig. 1). The solid line represents the MoM solution [12] and the dashed curve the approximate series-based solution [5], assuming 50° incidence relative to the axis of the tree trunk. (a) $\phi = 0°$ and (b) $\phi = 180°$.

penetration into the tree trunk for this high-contrast case, and therefore, the dips seen in Fig. 4 as a function of frequency are less dramatic for this case. Note also that the frequency at which the RCS becomes relatively constant with frequency is smaller in Fig. 5 than in Fig. 4 due to the enhanced electrical size of the target with increased permittivity. In this context, note that in Fig. 5, the MoM computations were only performed up to 325 MHz, while those in Fig. 4 were performed up to 520 MHz. This again is reflective of the target electric size, and the associated enhanced computational requirements (see Section IIIA) as the target permittivity increases. It is interesting to also note that, for the higher permittivity in Fig. 5, the agreement between the approximate and MoM-computed RCS is very good for HH polarization, while being slightly less good for VV.

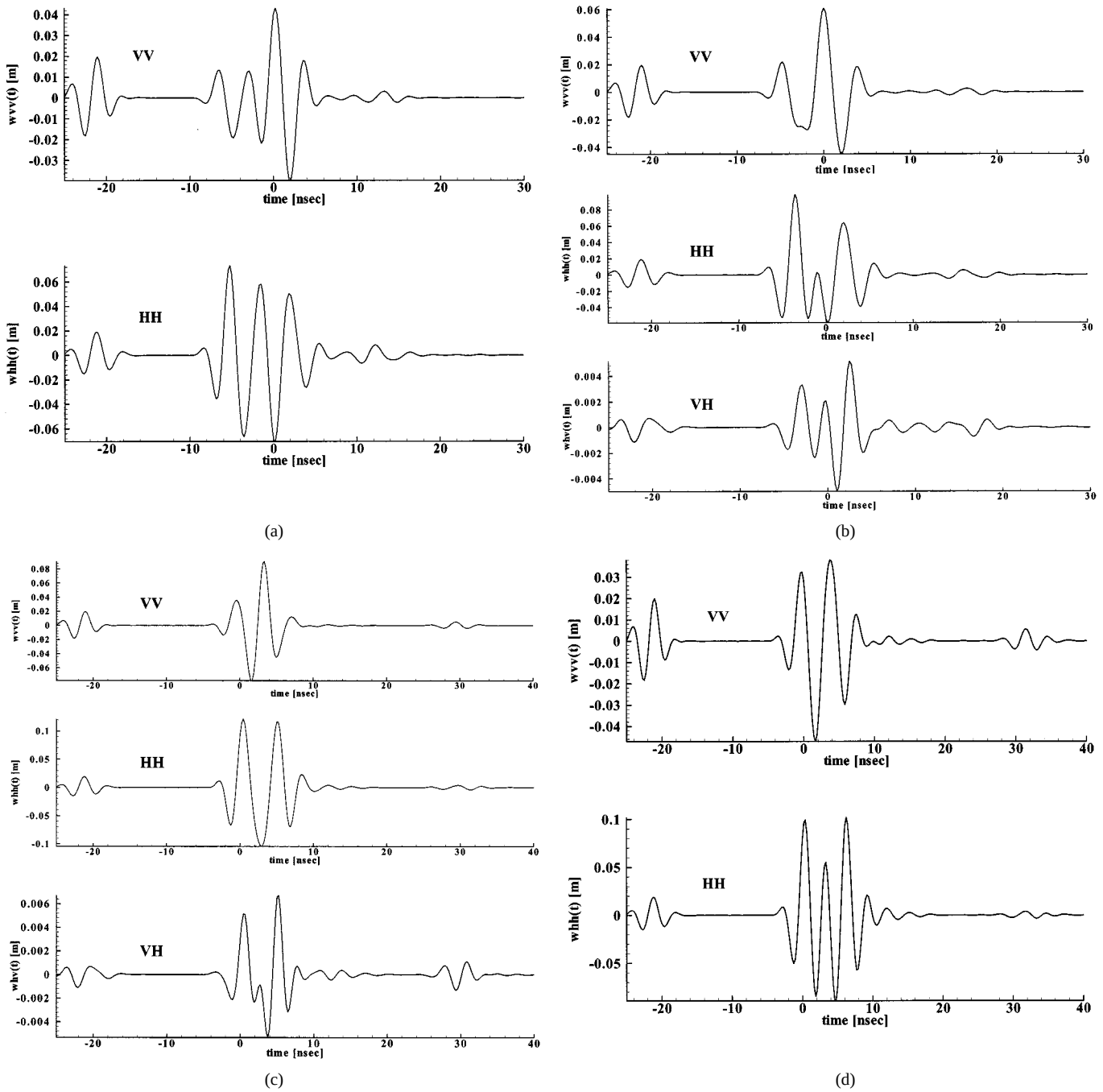


Fig. 8. Time-domain backscattered fields from the tree trunk considered in Fig. 7, assuming an incident pulse as shown in Fig. 2. The pulsed plane wave is incident 50° from the axis of the tree trunk. (a) $\phi = 0°$, (b) $\phi = 45°$, (c) $\phi = 135°$, and (d) $\phi = 180°$.

C. Angle of Incidence

In the previous set of examples, the angle of incidence was held constant while other parameters were varied. When performing synthetic aperture radar (SAR) measurements, the angle between the tree trunk and sensor is dictated, for a fixed sensor height, by the distance between the sensor and scatterer. Consequently, two identical tree trunks at different portions of the SAR image will be interrogated with distinct angles of incidence. We therefore examine the effect of the incidence angle on the RCS of a tree trunk on flat terrain, assuming the same soil type as in Figs. 3–5. We consider tree trunks composed of permittivity $\epsilon = \epsilon_0(5 - \mathrm{j}1 - \mathrm{j}\sigma/\omega\epsilon_0)$ and conductivity $\sigma = 0.01$ S/m, with 5 m height. In Fig. 6(a)–(c), we plot the computed RCS for tree diameters of 30, 40, and 50 cm. The frequency-dependent soil parameters are again characterized as Puerto Rico clay [22]. As expected, the increasing tree-trunk diameter (with all other parameters fixed) principally results in a shifting of the RCS characteristics to lower frequencies. As we have found throughout, at frequencies greater than approximately 30 MHz, the HH RCS is consistently stronger than its VV counterpart.

From Fig. 6, we note that the angle of incidence has an important effect on the tree-trunk RCS, with this particularly true for VV polarization. In fact, at low frequencies the HH RCS is almost independent of incidence angle, with a larger variation occurring with increased frequency. This is to be expected, since for VV polarization there is relatively strong penetration into the tree trunk, and the scattering induced by the transmitted compo-

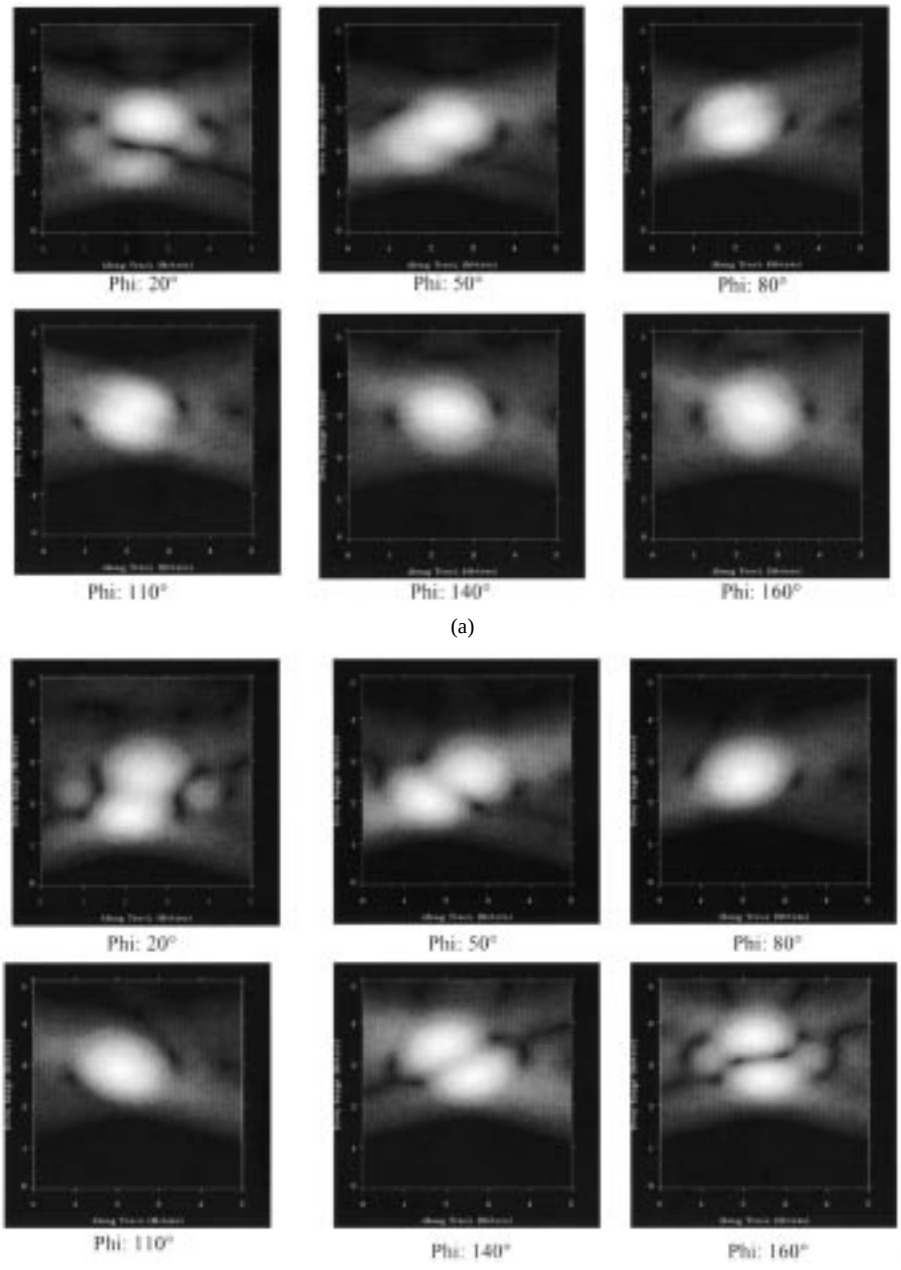


Fig. 9. Synthetic aperture radar (SAR) images of the tree trunk considered in Figs. 7–8. The SAR has the incident waveform in Fig. 2, and the angle between the sensor and tree-trunk axis is fixed at 50°, and a 35° SAR aperture is assumed, relative to the tree axis. Results are shown for apertures centered at the azimuthal positions (see Fig. 1) $\phi = 20°$, $\phi = 50°$, $\phi = 80°$, $\phi = 110°$, $\phi = 140°$, and $\phi = 160°$. In all images the dynamic range is fixed at 40 dB (white strongest, black weakest). The peak image strength is weakest for azimuthal positions near $\phi = 0°$ and $\phi = 180°$, and strongest near $\phi = 90°$ (see Fig. 8). (a) VV polarization and (b) HH polarization.

nent is expected to be a strong function of the incidence angle. By contrast, this is expected to be a much smaller effect for HH polarization, for which the penetration is relatively small.

D. Tilted Terrain

We next consider the scattered return from a tree trunk over a sloped terrain, with soil electrical parameters characteristic of Yuma soil with 10% water content [20]. A 10° slope is considered, with a tree trunk of 30 cm diameter, 5 m height, permittivity $\epsilon = \epsilon_0(3 - \mathrm{j}0.5 - \mathrm{j}\sigma/\omega\epsilon_0)$, and conductivity $\sigma = 0.005$ S/m. The tree trunk axis makes an angle $\theta_{\mathrm{T}} = 80°$ with respect to the terrain slope (see Fig. 1). The BoR MoM model is inappropriate for this example, and therefore, all associated MoM data have been computed with the triangular-patch model. We consider an incidence angle of 50° relative to the tree-trunk axis for all azimuthal positions considered. This was selected to be consistent with a SAR system flying at a fixed height, and at a constant radius relative to the tree axis. For reference, the azimuthal positions $\phi = 0°$ and $\phi = 180°$ are shown in Fig. 1.

In Fig. 7, we consider the RCS for HH polarization. Since for this incidence angle and terrain slope mechanism i) is not observed in backscatter, the RCS calculations in Fig. 7 include mechanisms ii)-iv). For the case $\phi = 0°$ and $\phi = 180°$ in

Fig. 7(a) and (b), respectively, we see the agreement between the MoM and approximate solutions is relatively good. It is interesting to note that in this case, unlike in the previous HH results (for flat terrain), there are sharp dips in the frequency-dependent RCS response. This is because here, unlike in Figs. 4–6, there is no dihedral response i) in Fig. 1 to dominate the RCS. Hence, the frequency-dependent interference between the other scattering mechanisms dominates the RCS, yielding the characteristics in Fig. 7. We also performed similar comparisons between the MoM and approximate solutions for the VV polarization, and in that case the agreement between the MoM and approximate solution was considerably worse.

When the tree trunk is normal to the air-ground interface, the scattered fields are independent of the azimuthal angle Φ, and the cross-polarized backscattered fields are zero. We now address the Φ-dependent signature for the sloped-terrain example considered in Fig. 7. The time-domain backscattered fields for $\phi = 0°, 45°, 135°$ and $180°$ are shown in Figs. 8(a)–8(d), respectively. Due to bisectional symmetry, the cross-polarized backscattered fields vanish at $\phi = 0°$ and $\phi = 180°$ (see Fig. 1). In all cases the incident plane wave has an angle $\theta_i = 50°$ relative to the tree axis, with an incident-pulse shape as in Fig. 2. A close inspection of the results in Fig. 8 for HH and VV polarizations reveals that, for a given polarization, the initial scattered waveform, representative of mechanism ii) in Fig. 1, is independent of the azimuthal orientation ϕ. Further, theoretically, there should be no cross-polarized component to mechanism ii), since the top of the target is theoretically rotationally symmetric, and mechanism ii) is independent of the ground slope. However, in Fig. 8(b) and (c) we note that there is a small cross polarized diffracted component representative of mechanism ii). This is due to the fact that the triangular-patch discretization yields a target mesh that is not exactly rotationally symmetric [4], [12]. Hence, the small initial diffracted component in Fig. 8(b) and (c) is reflective of this asymmetry.

After the arrival of mechanism ii), there is a stronger scattered waveform, which appears to be the composite of *two* pulses, slightly delayed with respect to each other. This representation of two delayed pulses is particularly clear, for example, in the HH signature of Fig. 8(b). We attribute these two time-delayed waveforms to mechanisms iii) and iv) in Fig. 1. For some of the examples, there is appreciable scattered energy that comes later in time, this reflective of multiple interactions. It is clear from Fig. 8 that the transient scattered fields from a tree trunk over sloped terrain are a strong function of orientation.

E. Synthetic Aperture Radar

The results in Fig. 8 indicate that the time-dependent scattered waveform is a strong function of orientation for a tree trunk over a sloped terrain. A synthetic aperture radar (SAR) system implicitly views a target from multiple orientations. To address the impact of the aspect-dependent signatures in Fig. 8 on a subsequent SAR image, we consider a SAR system that is flying in a circular path around the tree trunk considered in Fig. 8 (with the same 10% Yuma soil [20]). The SAR is positioned such that a constant angle of incidence exists between the sensor and the tree axis and, as in Fig. 7, this is assumed to be $\theta_i = 50°$. The SAR is assumed to span a 35° aperture with respect to the target center, and here each aperture position is weighted equally i.e., the antenna pattern is assumed uniform over approximately 35°). Simple backprojection of the time domain fields (as in Fig. 8) is used to form the SAR image. In Fig. 9, we present SAR images using data computed via the triangular-patch MoM model, for six different center positions of the 35° aperture. In particular, we consider apertures centered at 20°, 50°, 80°, 110°, 140°, and 160°, where Fig. 9(a) and 9(b) are for VV and HH polarization, respectively. As expected from Fig. 8, the SAR image is a strong function of the aperture position. At particular positions [e.g., $\phi = 160°$ in Fig. 9(b)], the image is characterized by two strong regions, these likely due to separated responses from mechanisms iii) and iv), while for other sensor positions, these two returns coalesce into a single strong return.

IV. Conclusions

Three models have been presented for electromagnetic scattering from a tree trunk over a half space. Two of the models were based on the MoM, with one MoM model restricted to the special case of a body of revolution (BoR) and the other employing a general triangular-patch basis function decomposition. In addition, we considered an approximate model based on the series-expansion solution for scattering from an infinite dielectric cylinder. For the VHF frequencies considered, we found that the approximate solution was in excellent agreement with the numerical solutions for HH polarization, with this agreement degrading slightly for VV polarization (for the sloped-terrain case, the approximate solution was particularly poor for VV polarization). The numerical models were employed to investigate the backscattered signature from a tree trunk in both the frequency and time domain. Results were presented as a function of tree-trunk height, diameter, and electrical parameters. Moreover, we considered the backscattered RCS for various incident angles. We also considered backscattering from a tree trunk on sloped terrain, both in the time domain and in the SAR image domain. These examples have demonstrate that, despite the relative simplicity of the tree-trunk target, the associated scattering phenomenology is quite complicated and must be considered carefully in the context of foliage penetrating (FOPEN) radar.

Despite the fact that, to our knowledge, this paper represents the most rigorous analysis of tree-trunk scattering to date, there are still many issues that need to be addressed further. In particular, we demonstrated in Fig. 1 that there are four principal scattering mechanisms of interest for tree-trunk scattering. Two of these, ii) and iii) in Fig. 1, are driven by diffraction from the top of the tree trunk. Unfortunately, this is the least realistic aspect of the scattering model. Dihedral scattering between the tree-trunk surface and the soil interface is accounted for accurately, but this phenomenon is not generally present for the case of sloped terrain. For such circumstances, the scattered fields computed by the simple model in Fig. 1 are driven in large part by diffraction at the "top" of the tree trunk, this being a simple approximation for the scattering mechanisms of an actual tree. Thus, for the flat-terrain case the principal (dihedral) scattering mechanism is well represented, but for the tilted-terrain case there are significantly more approximations. Consequently, the scattering physics predicted by the model must be validated by measurements. Moreover, the triangular-patch model is appli-

cable to general dielectric targets in the presence of a layered medium (soil). Future work will involve the use of this model to more realistically handle the large branches in the tree, these potentially being important at the VHF and low-UHF frequencies of interest. We have emphasized throughout, however, that the computational burden of such electrically large problems is prohibitive for the MoM solver. The limitations in the MoM have been ameliorated of late through development of the fast multipole method (FMM) and the multilevel fast multipole algorithm (MLFMA) [23]–[25]. These methods have been used primarily for free-space scattering [17], [18] and for perfectly conducting targets over a half space [25]. For the targets of interest here, the FMM/MLFMA must be extended to general dielectric targets over a half space.

We close by acknowledging that, despite the large number of examples presented here, such a presentation can never be entirely complete within the constraints of brevity. We note, therefore, that in addition to being applicable to constant-radius tree trunks, like those considered here, the BoR and triangular-patch MoM models are also applicable to more general tree-trunk shapes. Orientation-independent curvature in the tree-trunk shape can be handled via the BoR model, while the triangular-patch model is applicable to even more general tree trunk shapes (e.g., elliptical cross section). We found in the course of this work, for example, that small changes in the height-dependent tree radius (appropriate for the BoR model) produce relatively modest changes in the backscatter RCS. A further investigation of such higher-order effects may also be warranted, the focus here being on the principal scattering mechanisms.

REFERENCES

[1] D. R. Sheen, N. P. Malinas, D. W. Kletzli, T. B. Lewis, and J. F. Roman, "Foliage transmission measurements using a ground-based ultra-wideband (300–1300 MHz) SAR system," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 118–130, Jan. 1994.

[2] S. G. Leblanc, P. Bicheron, J. M. Chen, M. Leroy, and J. Cihlar, "Investigation of directional reflectance in boreal forests with an improved four-scale model and airborne POLDER data," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 1396–1414, May 1999.

[3] N. Geng and L. Carin, "Wide-band electromagnetic scattering from a dielectric BOR buried in a layered lossy dispersive medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.

[4] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 30, pp. 409–418, May 1982.

[5] Y. C. Lin and K. Sarabandi, "Electromagnetic scattering model for a tree trunk above a tilted ground plane," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 1063–1070, July 1995.

[6] L. Tsang, K. H. Ding, G. Zhang, C. Hsu, and J. A. Kong, "Backscattering enhancement and clustering effects of randomly distributed dielectric cylinders overlying a dielectric half space based on Monte-Carlo simulations," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 488–499, May 1995.

[7] L. Tsang, C. H. Chan, J. A. Kong, and J. Joseph, "Polarimetric signatures of a canopy of dielectric cylinders based on first and second order vector radiative transfer theory," *J. Electromagn. Waves Appl.*, vol. 6, pp. 19–51, 1992.

[8] K. Sarabandi, P. F. Polatin, and F. T. Ulaby, "Monte carlo simulation of scattering from a layer of vertical cylinders," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 465–475, 1993.

[9] G. Zhang, L. Tsang, and Z. Chen, "Collective scattering effects of trees generated by stochastic Lindenmayer systems," *Microwave Opt. Tech. Lett.*, vol. 11, pp. 107–111, Feb. 1996.

[10] S. H. Yueh, J. A. Kong, J. K. Jao, R. T. Shin, Shin, and T. LeToan, "Branching model for vegetation," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 390–402, 1992.

[11] Z. Chen, "Microwave remote sensing of vegetation: stochastic lindenmayer systems, collective scattering effects, and neural network inversions," Ph.D. dissertation, Dept. Elect. Eng., Univ. Washington, Seattle, 1994.

[12] J. He, T. Yu, N. Geng, and L. Carin, "Method-of-moments analysis of electromagnetic scattering from a general three-dimensional dielectric target embedded in a multi-layered medium," *Radio Sci.*, Feb. 2000.

[13] A. W. Glisson, "Integral equation techniques ," in *Dielectric Resonators*, D. Kaifez, Ed. Dedham, MA: Artech House, 1986, pp. 259–325.

[14] K. A. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, parts I and II," *IEEE Trans. Antennas Propagat.*, vol. 38, pp. 335–352, Mar. 1990.

[15] Y. Rahmat-Samii, R. Mittra, and P. Parhami, "Evaluation of sommerfeld integrals for lossy half-space problems," *Electromagnetics*, vol. 1, no. 1, pp. 1–28, 1981.

[16] R. M. Shubair and Y. L. Chow, "A simple and accurate complex image interpretation of vertical antennas present in contiguous dielectric half-spaces," *IEEE Trans. Antennas Propagat.*, vol. 41, pp. 806–812, June 1993.

[17] Y. L. Chow, J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, vol. 39, pp. 588–562, Mar. 1991.

[18] M. L. V. Blaricum and R. Mittra, "A technique for extracting the poles and residues of a system directly from its transient response," *IEEE Trans. Antennas Propagat.*, vol. 23, pp. 777–781, Nov. 1975.

[19] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acousti., Speech, Signal Process.*, vol. 38, pp. 814–824, May 1990.

[20] L. Carin, N. Geng, M. McClure, J. Sichina, and L. Nguyen, "Ultra-wide-band synthetic-aperture radar for mine-field detection," *IEEE Antennas Propagat. Mag.*, vol. 41, pp. 18–33, Feb. 1999.

[21] L. Carin, R. Kapoor, and C. E. Baum, "Polarimetric SAR imaging of buried landmines," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1985–1988, Nov. 1998.

[22] J. E. Hipp, "Soil electromagnetic parameters as functions of frequency, soil density, and soil moisture," *Proc. IEEE*, vol. 62, pp. 98–103, Jan. 1974.

[23] J. M. Song and W. C. Chew, "Multilevel fast multipole algorithm for solving combined field integral equations of electromagnetic scattering," *Microwave Opt. Tech. Lett.*, vol. 10, pp. 14–19, Sept. 1995.

[24] X. Q. Sheng, J. M. Jin, J. Song, W. C. Chew, and C. C. Lu, "Solution of combined-field integral equations using multilevel fast multipole algorithm for scattering by homogeneous bodies," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 1718–1726, Nov. 1998.

[25] N. Geng, A. Sullivan, and L. Carin, "Multi-level fast-multipole algorithm for scattering from conducting targets above or embedded in a lossy half space," *IEEE Trans. Geosci. Remote Sensing*, to be published.

Jiangqi He was born in China. He received the M.S.E.E. degree from New Mexico State University, Las Cruces, in 1998 and the Ph.D. degree in electrical engineering from Duke University, Durham, NC, in 2001.

He joined Intel Corporation, Phoenix, AZ, in 2001, where he is developing techniques for the analysis of high-frequency integrated circuits. His research interests include numerical methods for wideband electromagnetic scattering and propagation as well as high-speed integrated circuits.

Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Germany, in 1991 and 1996, respectively.

From 1991 to 1996, he was with the Institute for Microwaves and Electronics, University of Karlsruhe, Karlsruhe, Germany, working on full-wave propagation modeling for radio communications systems. In January 1997, he joined the Department of Electrical and Computer Engineering, Duke University, Durham, NC, as a visiting Postdoctoral Researcher for 18 months. While visiting Duke University, and for another 18 months at the University of Karlsruhe, his research focused on numerical techniques in computational electromagnetics for planar-stratified media. In January 2000, he joined SIEMENS Corporate Technology, Munich, Germany, working as a Research Scientist on the design of broadband wireless communications systems.

Dr. Geng received the Mannesmann Innovation Award in 1997 for his dissertation on full-wave propagation modeling for radio communications systems.

Lam Nguyen [redacted] in Vietnam. He received the B.S.E.E. degree from Virginia Polytechnic Institute, Blacksburg, the M.S.E.E. degree from George Washington University, Washington, DC, and the M.S.C.S. degree from Johns Hopkins University, Baltimore, MD, respectively.

He was with General Electric from 1984 to 1985. He has been with the Army Research Laboratory and its predecessor organization, Harry Diamond Laboratory since 1986. He has been primarily engaged in the research and development of foliage penetratuib and pround penetration via ultrawideband (UWB) radars. His primary research interests include digital signal processing and algorithm development.

Lawrence Carin (SM'96–F'01) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor, and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department at Duke University, Durham, NC, where he is now a Professor. He is the Principal Investigator on a Multidisciplinary University Research Initiative (MURI) on demining. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies and is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.

Multiresolution Time-Domain Analysis of Plane-Wave Scattering From General Three-Dimensional Surface and Subsurface Dielectric Targets

Xianyang Zhu, *Member, IEEE,* and Lawrence Carin, *Fellow, IEEE*

***Abstract*—The multiresolution time domain (MRTD) is used to analyze wide-band plane-wave scattering from general dielectric targets embedded in a lossy half-space, with free-space scattering as a special case. A Haar wavelet expansion is used for simplicity, this constituting a generalization of the widely used finite-difference time-domain (FDTD) method. In addition to developing the mathematical formulation, example results are presented for several targets, with the MRTD results validated through comparison with an independent frequency-domain method-of-moments solution and an FDTD model.**

***Index Terms*—Buried object detection, electromagnetic scattering, multiresolution techniques, time-domain analysis.**

I. Introduction

SUBSURFACE sensing is of interest for many applications, including detection of buried pipes, unexploded ordnance (UXO), and land mines [1]–[7]. There has consequently been significant interest in the development of modeling tools for the analysis of electromagnetic-based subsurface sensing, with application to radar [1]–[4] and electromagnetic induction [5]. In particular, authors have considered both frequency- and time-domain models, the former including the method of moments (MoM) [3], [4], the extended-Born method [5], and fast-multipole methods [6]. With regard to time-domain methods, the finite-difference time-domain (FDTD) method has been applied widely [2], [7]. The aforementioned frequency-domain methods typically employ a Green's function (e.g., for layered media or a half-space [3]–[6]), and therefore they assume some regularity in the background medium. The advantage of such an approach is that one need only discretize the target surface, rather than the entire region of interest, yielding computational efficiency. By contrast, the FDTD discretizes the entire computational domain, in both space and time, and therefore it is applicable to very general background media [2], [7].

The FDTD can be viewed in terms of a pulse expansion of the electromagnetic fields, in both space and time [8]. With this understanding, one can generalize the FDTD by considering an alternative expansion. For example, there has recently been interest in expanding the fields in terms of a wavelet basis, thus realizing what has been termed a multiresolution time-domain (MRTD) analysis [8], [9]. The multiresolution property is a consequence of the wavelet expansion, with the lowest resolution manifested in terms of scaling functions, with higher resolution (refinement) added by wavelets. Wavelets of successively higher resolution can be added to the scaling-function expansion to realize a field representation of a desired resolution. There are an infinite number of possible orthogonal or biorthogonal wavelet classes [10], but one typically imposes requirements (e.g., differentiability and/or vanishing low-order moments [10]) that restrict the wavelet form. As illustrated further in Section II, one of the important requirements of the scaling and wavelet functions, in the context of MRTD, is that they have compact support (zero outside a prescribed region).

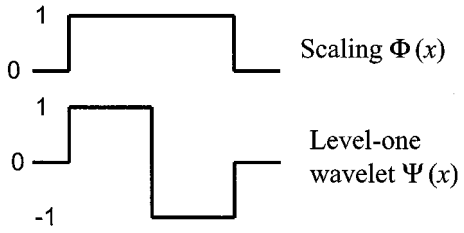


Fig. 1. Haar scaling function and first-level wavelet on one dimension.

With regard to the choice of scaling and wavelet functions, for implementation of the MRTD, one can consider smooth functions such as the Battle–Lemarie [8], [9] orthogonal wavelets or the Cohen–Daubechies–Fouveau [10], [11] biorthogonal wavelets. The Battle–Lemarie wavelets are not of strictly compact support, although these infinite-support functions can be truncated approximately in the context of MRTD [8], [9]. In this paper, we focus on the simpler Haar wavelets, the scaling function and mother wavelet for which are shown in Fig. 1 (in one dimension). One notices that the Haar scaling function is simply the pulse expansion used in the traditional FDTD, with the wavelets providing refinement to this representation. The MRTD with Haar wavelets is therefore a natural extension of the traditional FDTD, thereby constituting a good starting point for MRTD consideration. One can show, for example, that Haar-based MRTD with n wavelet levels is equivalent to FDTD with an expansion in terms of pulses $1/2^n$ the width of the Haar scaling function. One might consequently argue that Haar-based MRTD and FDTD are equivalent. The difference

Manuscript received September 11, 2000; revised February 9, 2001. This work was supported by the U.S. Army Research Office under Grant DAAD19-99-1-0192.

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291 USA.

Publisher Item Identifier S 0018-926X(01)09192-X.

0018–926X/01$10.00 © 2001 IEEE

is manifested when one realizes that the MRTD offers significantly more flexibility than FDTD. In particular, in the MRTD, one need not use the same number of wavelet levels throughout the computational domain. In regions of relatively slow spatial field variation, one can use scaling functions alone. Where fields vary more quickly, successively higher wavelet levels can be added. Consequently, in the context of a Haar-based MRTD, one effectively realizes a natural multiscale gridding of the traditional FDTD. There has been previous work on the development of multiscale FDTD gridding schemes, although in that work there have been issues regarding algorithmic stability [12]–[14], which is not a problem for the MRTD.

As discussed further in Section II, other advantages of the MRTD are found in the natural way in which general material interfaces and heterogeneity are handled, without the need for a staircase interface approximation, or the introduction of special conformal interface cells. As with the multiscale gridding discussed above, authors have developed techniques for improved accommodation of general material interfaces within the FDTD [15]. The point is not that general interfaces cannot be handled via FDTD, but rather that such issues are treated naturally and easily within the MRTD construct [8], [9]. We also reiterate that the basic MRTD is applicable to a general wavelet expansion, not only the simple Haar wavelets [8], [9], [11] considered here, and further benefits are manifested in the reduced MRTD numerical dispersion with smooth wavelets, *vis-à-vis* the traditional FDTD [9].

Previous published work with MRTD has focused primarily on its application to high-speed integrated circuits [8], [9]. There has recently been investigation of two-dimensional electromagnetic scattering via MRTD, with primary focus on interaction with rough interfaces [16]. It is believed that this paper constitutes the first fully three-dimensional analysis of electromagnetic scattering with MRTD, with the principal focus here on subsurface sensing, although free-space scattering examples are also presented as a special case.

The remainder of this paper is organized as follows. In Section II, we present a detailed discussion of the MRTD scattering formalism, with an implementation in terms of Haar wavelets. We consider plane-wave excitation and therefore discuss extension of the traditional FDTD connecting surface to the MRTD problem. We also discuss MRTD implementation of the perfectly matched layer (PML) absorbing boundary condition. Several example results are presented in Section III, with the MRTD data compared to results computed via an independent MoM solution and to FDTD-generated data. Conclusions and future directions are discussed in Section IV.

II. MRTD Scattering Formulation

A. Basic Formulation

We begin the MRTD analysis by expanding the electromagnetic fields in terms of a wavelet basis, composed of scaling and wavelet functions, in a manner analogous to the well-known method of moments [3], [4]. Scaling functions are employed in regions characterized by smoothly varying fields (coarse field representation), while additional sampling points (refinement) are introduced by incorporating wavelets in regions of fast field variation. This yields a natural multiresolution field representation.

In the results presented subsequently, comparisons are made between MRTD and FDTD results [2], [7], [15]. The FDTD can be viewed as a field expansion in terms of pulse basis functions (cubes in three dimensions), these pulses being analogous to the Haar scaling function. Haar wavelets therefore represent a good starting point for examination of MRTD. Let the Haar scaling function be represented by ϕ, with the mother wavelet denoted ψ. These functions are depicted in Fig. 1. Although a wavelet decomposition can be performed in time as well, here the time variation is expanded in a pulse basis h. Assuming an expansion in terms of scaling functions and a single level of wavelets, we have

$$\begin{aligned} F(x,y,z,t) = \sum_{i,j,k,m=-\infty}^{\infty} & h_{m+\delta t}(t) \\ & \times \Big[F_{i,j,k}^{\phi,\phi,\phi} \phi_{i+\delta x}(x)\,\phi_{j+\delta y}(y)\,\phi_{k+\delta z}(z) \\ & + F_{i,j,k}^{\phi,\phi,\psi} \phi_{i+\delta x}(x)\,\phi_{j+\delta y}(y)\,\psi_{k+\delta z}(z) \\ & + F_{i,j,k}^{\phi,\psi,\phi} \phi_{i+\delta x}(x)\,\psi_{j+\delta y}(y)\,\phi_{k+\delta z}(z) \\ & + F_{i,j,k}^{\phi,\psi,\psi} \phi_{i+\delta x}(x)\,\psi_{j+\delta y}(y)\,\psi_{k+\delta z}(z) \\ & + F_{i,j,k}^{\psi,\phi,\phi} \psi_{i+\delta x}(x)\,\phi_{j+\delta y}(y)\,\phi_{k+\delta z}(z) \\ & + F_{i,j,k}^{\psi,\phi,\psi} \psi_{i+\delta x}(x)\,\phi_{j+\delta y}(y)\,\psi_{k+\delta z}(z) \\ & + F_{i,j,k}^{\psi,\psi,\phi} \psi_{i+\delta x}(x)\,\psi_{j+\delta y}(y)\,\phi_{k+\delta z}(z) \\ & + F_{i,j,k}^{\psi,\psi,\psi} \psi_{i+\delta x}(x)\,\psi_{j+\delta y}(y)\,\psi_{k+\delta z}(z) \Big] \end{aligned} \tag{1}$$

where F denotes any electric or magnetic field component. The indexes i, j, k, and m are the discrete space and time indexes related to the space and time coordinates, analogous to the traditional FDTD [2],[7],[15], with δx, δy, δz, and δt representing displacements along the x, y, z, and t directions. Table I summarizes the relative shifts between the different field components, constituting the MRTD unit cell, assuming Haar expansion with scaling functions and a single level of wavelets.

Considering (1) more closely, we note that each field component is represented by eight terms (e.g., $\phi(x)\phi(y)\phi(z)$, $\phi(x)\phi(y)\psi(z)$, etc.), each of which is a three-dimensional basis function realized here in terms of a Haar scaling function and a single wavelet level. The eight basis functions are analogous to taking a single FDTD cubic basis and dividing it in half in three dimensions, yielding eight smaller cubes where there was one. In general, a single scaling function plus n wavelet levels will produce 8^n basis functions per unit cell, for a given field component, which is equivalent to FDTD with 8^n cubes within the volume of the scaling functions alone. As discussed further below, by selectively pruning the number of wavelets utilized in a given region, one can realize a multigrid FDTD scheme without the stability concerns of previous multigrid FDTD algorithms [12]–[14]. We also note from (1) and Table I that the 8^n basis functions used to represent *different* field components are in general shifted spatially by a quarter-cell with respect to one another (a cell corresponding to the size of the scaling function), analogous to FDTD in which the different field

TABLE I
DISPLACEMENTS FOR DIFFERENT FIELD COMPONENTS

	E_x	E_y	E_z	H_x	H_y	H_z
$\delta x\,(cells)$	0.25	0	0	0	0.25	0.25
$\delta y\,(cells)$	0	0.25	0	0.25	0	0.25
$\delta z\,(cells)$	0	0	0.25	0.25	0.25	0
$\delta t\,(\Delta t)$	0	0	0	0.5	0.5	0.5

components are in general also shifted in space with respect to one another [2], [7], [15]. However, in FDTD, the shift is by half a cell.

Consider a homogeneous medium with electric permittivity ϵ, magnetic permeability μ, and conductivity σ. An example equation from Ampere's law can be expressed in Cartesian coordinates as

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = \varepsilon \frac{\partial E_x}{\partial t} + \sigma E_x. \tag{2}$$

Inserting (1) in (2), we obtain the MRTD equations by employing a Galerkin's testing procedure [4], with the subsequent difference-equation representative of (2) given as

$$[E_x]_{i,j,k}^{(m+1)} = \alpha\,[E_x]_{i,j,k}^{(m)} + \beta \left\{ \begin{array}{l} \Big([C_1]\,[H_z]_{i,j,k}^{(m)} - [C_2]\,[H_z]_{i,j-1,k}^{(m)} \Big) \\ -\Big([D_1]\,[H_y]_{i,j,k}^{(m)} - [D_2]\,[H_y]_{i,j,k-1}^{(m)} \Big) \end{array} \right\} \tag{3}$$

where the space discretization (size of the scaling function) in the x-, y-, and z-directions is assumed to be the same. The constants $\alpha = (1-(\sigma\Delta t/2\varepsilon))/(1+(\sigma\Delta t/2\varepsilon))$ and $\beta = \Delta t/(\varepsilon\Delta x(1+\sigma\Delta t/2\varepsilon))$ corresponded to the material characteristics of the cell. The vector $[E_x]_{i,j,k}$ represents the eight basis functions discussed above, for the field component E_x at the discrete spatial coordinates (i,j,k), with the superscripts in (3) representing the discrete time index m, as in the FDTD. The eight components of the fields are arranged as shown in (4) at the bottom of the page for E_x, and the matrices are

$$[C_1] = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 3 \end{bmatrix} \tag{5a}$$

$$[C_2] = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix} \tag{5b}$$

$$[D_1] = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 3 \end{bmatrix} \tag{5c}$$

$$[D_2] = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \tag{5d}$$

with these found through the testing procedure and via the relations

$$\int \psi_{i+1/4}(x)\,\phi_i(x)\,dx = 0 \tag{6a}$$

$$\int \frac{\partial \phi_{i+1/4}(x)}{\partial x}\phi_i(x)\,dx = \delta_{i,i} - \delta_{i-1,i} \tag{6b}$$

$$\int \frac{\partial \psi_{i+1/4}(x)}{\partial x}\phi_i(x)\,dx = \delta_{i-1,i} - \delta_{i,i} \tag{6c}$$

$$\int \frac{\partial \phi_{i+1/4}(x)}{\partial x}\psi_i(x)\,dx = \delta_{i,i} - \delta_{i-1,i} \tag{6d}$$

$$\int \frac{\partial \psi_{i+1/4}(x)}{\partial x}\psi_i(x)\,dx = \delta_{i-1,i} + 3\delta_{i,i} \tag{6e}$$

where $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ if $i \neq j$. The MRTD equations for the other field components can be deduced similarly.

$$[E_x]_{i,j,k} = [\,E_x^{\phi,\phi,\phi} \quad E_x^{\phi,\phi,\psi} \quad E_x^{\phi,\psi,\phi} \quad E_x^{\phi,\psi,\psi} \quad E_x^{\psi,\phi,\phi} \quad E_x^{\psi,\phi,\psi} \quad E_x^{\psi,\psi,\phi} \quad E_x^{\psi,\psi,\psi}\,]_{i,j,k}^{T} \tag{4}$$

The matrix-based representation of the update (3)–(5) constitutes a compact MRTD implementation and is applicable to a general wavelet basis and an arbitrary number of wavelet levels. If higher order wavelet levels are considered, the matrices increase in size. Through consideration of the integrals in (6), it also clear that wavelets and scaling functions of compact support are desirable, such that the matrices in (3)–(5) are as small as possible.

In (3)–(5), scaling and wavelet functions are used throughout the computational domain. As discussed in Section III, in practice scaling functions are used everywhere and wavelets are only utilized in regions of relatively fast field variation. In the results presented below, we consider regions that are represented in terms of scaling functions alone, for each of the three dimensions, and other regions in which spatial variation in each Cartesian direction is expanded in terms of scaling functions and a single wavelet level. This implies that in some regions, one basis function is used within the volume of the three-dimensional scaling function (scaling functions alone), and in other regions, eight basis functions are used (scaling function plus single wavelet level). This is analogous to a multigrid FDTD scheme [12]–[14], in which a certain cubic basis is used in some regions and, in other regions, eight smaller cubes are used in this same volume. However, we note that within the context of MRTD, we need not restrict ourselves to one or eight basis functions within a given cubic volume. We can utilize a subset of the components in (4), ranging from one to eight. For example, if we apply wavelets in the x- and y-directions, but only scaling functions in z, we realize bases that are analogous to FDTD with *noncubic* basis functions (the basis length in z is twice that in x and y). This is appropriate for geometries in which the fields may vary quickly in x and y, but not in z. We do not consider this higher level pruning scheme here, but this discussion points out the flexibility of MRTD, even with simple Haar basis functions.

With regard to the stability condition, as indicated in the Introduction, a Haar-based MRTD with n wavelet levels, assuming a scaling function of width w, is identical to an FDTD scheme with spatial sampling of $w/2^n$, and therefore the Haar-based MRTD stability condition can be easily borrowed from that of the traditional FDTD [2], [7], [15]. We note that for a more general wavelet expansion, the stability condition is different than that of the FDTD [9]. Finally, as in the FDTD analysis, one need only store the time-dependent basis-function coefficients on a Huygen's surface enclosing the target. The scattered fields at any position outside the Huygen surface are calculated by convolving the Huygen equivalent currents with the Green's function characteristic of the outside region.

B. Dielectric Interfaces

The surfaces of arbitrary-shaped dielectric objects are generally not conformal to the Haar-MRTD rectangular grid. Similar issues occur in the FDTD, in which one can consider a staircase approximation to the interface [17], averaging [15], or conformal FDTD cells [18]. While conformal FDTD cells can be tailored to general surfaces, the mesh for each target must be considered anew. Within the context of MRTD, there are no such issues. The same wavelet-based mesh is applied to all inhomogeneities, without requiring a staircase approximation. Again considering (2), for example, the wavelet expansion and Galerkin testing yield a generalized form of (3), expressed as

$$\begin{aligned}&\left(\frac{1}{\Delta t}[\varepsilon]_{i,j,k}+\frac{1}{2}[\sigma]_{i,j,k}\right)[E_x]_{i,j,k}^{(m+1)}\\&\quad-\left(\frac{1}{\Delta t}[\varepsilon]_{i,j,k}-\frac{1}{2}[\sigma]_{i,j,k}\right)[E_x]_{i,j,k}^{(m)}\\&=\frac{1}{\Delta x}\left\{\left([C_1][H_z]_{i,j,k}^{(m)}-[C_2][H_z]_{i,j-1,k}^{(m)}\right)\right.\\&\quad\left.-\left([D_1][H_y]_{i,j,k}^{(m)}-[D_2][H_y]_{i,j,k-1}^{(m)}\right)\right\}\end{aligned}\tag{7}$$

where we have (8), shown at the bottom of the page. We observe that there are only eight elements required in (8), expressed as

$$\begin{aligned}\varepsilon_{i,j,k}^{\chi,\xi,\zeta}=\frac{1}{\Delta x^3}\iiint\limits_{cell\,i,j,k}&\varepsilon(x,y,z)\phi_{i+\delta x}(x)\,\phi_{j+\delta y}(y)\,\phi_{k+\delta z}(z)\\&\chi_{i+\delta x}(x)\,\xi_{j+\delta y}(y)\,\zeta_{k+\delta z}(z)\,dxdydz\end{aligned}\tag{9}$$

where $\varepsilon(x,y,z)$ is the distribution function of the permittivity and χ, ξ, and ζ can be function ϕ or ψ. We also obtain the analogous matrix $[\sigma]_{i,j,k}$. From (7), we see that the inverse of $((1/\Delta t)[\varepsilon]+(1/2)[\sigma])$ is required for each cell filled with inhomogeneous media to obtain an explicit updating equation. This is not a big burden since all matrices are computed once and subsequently utilized in the time-stepping algorithm. Moreover,

$$[\varepsilon]_{i,j,k}=\begin{bmatrix}\varepsilon_{i,j,k}^{\phi,\phi,\phi} & \varepsilon_{i,j,k}^{\phi,\phi,\psi} & \varepsilon_{i,j,k}^{\phi,\psi,\phi} & \varepsilon_{i,j,k}^{\phi,\psi,\psi} & \varepsilon_{i,j,k}^{\psi,\phi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\psi}\\ \varepsilon_{i,j,k}^{\phi,\phi,\psi} & \varepsilon_{i,j,k}^{\phi,\phi,\phi} & \varepsilon_{i,j,k}^{\phi,\psi,\psi} & \varepsilon_{i,j,k}^{\phi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\phi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\psi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi}\\ \varepsilon_{i,j,k}^{\phi,\psi,\phi} & \varepsilon_{i,j,k}^{\phi,\psi,\psi} & \varepsilon_{i,j,k}^{\phi,\phi,\phi} & \varepsilon_{i,j,k}^{\phi,\phi,\psi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\psi} & \varepsilon_{i,j,k}^{\psi,\phi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi}\\ \varepsilon_{i,j,k}^{\phi,\psi,\psi} & \varepsilon_{i,j,k}^{\phi,\psi,\phi} & \varepsilon_{i,j,k}^{\phi,\phi,\psi} & \varepsilon_{i,j,k}^{\phi,\phi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\psi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\phi,\phi}\\ \varepsilon_{i,j,k}^{\psi,\phi,\phi} & \varepsilon_{i,j,k}^{\psi,\phi,\psi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\psi} & \varepsilon_{i,j,k}^{\phi,\phi,\phi} & \varepsilon_{i,j,k}^{\phi,\phi,\psi} & \varepsilon_{i,j,k}^{\phi,\psi,\phi} & \varepsilon_{i,j,k}^{\phi,\psi,\psi}\\ \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\phi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\psi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\phi,\phi,\phi} & \varepsilon_{i,j,k}^{\phi,\phi,\psi} & \varepsilon_{i,j,k}^{\phi,\psi,\psi} & \varepsilon_{i,j,k}^{\phi,\psi,\phi}\\ \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\psi} & \varepsilon_{i,j,k}^{\psi,\phi,\phi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\phi,\psi,\phi} & \varepsilon_{i,j,k}^{\phi,\psi,\psi} & \varepsilon_{i,j,k}^{\phi,\phi,\phi} & \varepsilon_{i,j,k}^{\phi,\phi,\psi}\\ \varepsilon_{i,j,k}^{\psi,\psi,\psi} & \varepsilon_{i,j,k}^{\psi,\psi,\phi} & \varepsilon_{i,j,k}^{\psi,\phi,\phi} & \varepsilon_{i,j,k}^{\psi,\phi,\psi} & \varepsilon_{i,j,k}^{\phi,\psi,\phi} & \varepsilon_{i,j,k}^{\phi,\psi,\psi} & \varepsilon_{i,j,k}^{\phi,\phi,\psi} & \varepsilon_{i,j,k}^{\phi,\phi,\phi}\end{bmatrix}.\tag{8}$$

the matrixes $[\varepsilon]_{i,j,k}$ and $[\sigma]_{i,j,k}$ are diagonal if the cell (i,j,k) is homogeneous. For this case, (7) reduces to (3). Therefore, we only use the form in (7) for cells in which the material is inhomogeneous.

C. Connecting Surface

When the source is distant from the target of interest, the incident fields can typically be well approximated in terms of a plane wave. We here employ a connecting surface [15], which divides the fields into total fields (inside surface) and scattered fields (outside surface). Connecting surfaces have been discussed previously in the context of FDTD, and here they are discussed in terms of MRTD. To implement a connecting surface, one requires the time-domain incident plane-wave fields on a closed surface enclosing the target. For free-space scattering, these fields can be expressed in closed form, while for more general media, more care is required. In particular, in the context of subsurface sensing, one is interested in a pulsed plane wave incident in the presence of a lossy half-space [3], [4], [6]. For this problem, the incident fields are frequency dependent, and therefore they are calculated in closed form in the frequency domain, with the time-domain incident fields on the connecting surface synthesized via Fourier transform. Note that in the half-space in which the incident fields first arrive, the incident fields due to an infinite half-space include the initial incident fields plus the Snell's-law reflected component, while in the subsurface the incident fields are represented by the transmitted component.

For the cells in the vicinity of the connecting surface, the update equations, e.g., (3) and (7), contain both total- and scattered-field terms. Correction terms are introduced to balance this disparity, with the terms represented by the incident fields [15]. We again consider (3), for example. For MRTD cells around the connecting surface, the fields on one side of the surface are scattered fields alone and on the other side total fields. Consider the bottom of a connecting surface, with z-position k in the total field region and $k-1$ in the scattered-field region. The coefficients $H^{\phi,\phi,\phi}_{y_{i,j,k-1}}$ and $H^{\phi,\phi,\psi}_{y_{i,j,k-1}}$ are related to the scattered-field region, while the other coefficients are related to the total-field region. Therefore, to keep the equation consistent, a correction term $E^{\phi,\phi,\phi\ (m+1)}_{x_{i,j,k}}|_{\text{correction}}$ is added on the right side of (3). In this case, the correction $E^{\phi,\phi,\phi\ (m+1)}_{x_{i,j,k}}|_{\text{correction}}$ is expressed as

$$E^{\phi,\phi,\phi\ (m+1)}_{x_{i,j,k}}|_{\text{correction}} = \beta\left[H^{\phi,\phi,\phi\ (m)}_{y_{i,j,k-1}}|_{\text{incident}} - H^{\phi,\phi,\psi\ (m)}_{y_{i,j,k-1}}|_{\text{incident}}\right] \quad (10)$$

We see therefore that the correction terms in the MRTD are handled analogously to how they are treated in the FDTD.

D. Perfectly Matched Layer (PML)

A PML [19], [20] is employed as the absorbing boundary conditions. To save computer memory, we only split the fields [19] in the PML region. We again consider E_x as an example. This field component can be split into two parts, E_{xy} and E_{xz} [19], which propagate along the y- and z-direction, respectively. The corresponding governing equations for lossy medium are therefore expressed as

$$\varepsilon\frac{\partial E_{xy}}{s_0\partial t} + (\sigma+\sigma_y)E_{xy} + \frac{\sigma\sigma_y}{\varepsilon}\varsigma_{xy} = \frac{\partial H_z}{\partial y} \quad (11a)$$

$$\varepsilon\frac{\partial E_{xz}}{s_0\partial t} + (\sigma+\sigma_z)E_{xz} + \frac{\sigma\sigma_z}{\varepsilon}\varsigma_{xz} = -\frac{\partial H_y}{\partial z} \quad (11b)$$

$$\frac{\partial \varsigma_{xy}}{\partial t} = E_{xy} \quad (11c)$$

$$\frac{\partial \varsigma_{xz}}{\partial t} = E_{xz} \quad (11d)$$

where ς_{xy} and ς_{xz} are two auxiliary variables, and the appropriate selection of the parameters s_0, σ_y, and σ_z can be found in [20]. The corresponding MRTD updating equations are now rewritten as

$$[E_{xy}]^{(m+1)}_{i,j,k} = \alpha_{xy}[E_{xy}]^{(m)}_{i,j,k} + \beta_{xy}\left([C_1][H_z]^{(m)}_{i,j,k} - [C_2][H_z]^{(m)}_{i,j-1,k}\right) - s_0\frac{\sigma\sigma_y}{\varepsilon}E^{(m)}_{xy} \quad (12a)$$

$$[E_{xz}]^{(m+1)}_{i,j,k} = \alpha_{xz}[E_{xz}]^{(m)}_{i,j,k} - \beta_{xz}\left([D_1][H_y]^{(m)}_{i,j,k} - [D_2][H_y]^{(m)}_{i,j,k-1}\right) - s_0\frac{\sigma\sigma_y}{\varepsilon}E^{(m)}_{xz} \quad (12b)$$

where the coefficients α_υ and β_υ (index υ can be xy or xz) are given by

$$\alpha_\upsilon = \frac{1}{s_0}\exp\left(-\frac{(\sigma+\sigma_{y,\,or\,z})\Delta t}{\varepsilon}\right) \quad (13a)$$

$$\beta_\upsilon = \frac{1}{s_0}\frac{1-\exp\left(-\frac{(\sigma+\sigma_{y,\,or\,z})\Delta t}{\varepsilon}\right)}{\sigma\Delta x} \quad (13b)$$

where the exponential time-stepping [19] has been used. Consequently, like the connecting surface, the PML is handled in a manner analogous to FDTD [19].

III. Results

Several scattering scenarios are presented for dielectric targets in vacuum as well as targets in the vicinity of a lossy half-space, the latter motivated by subsurface-sensing applications. The examples are selected to underscore the strengths and limitations of the MRTD, as implemented via Haar wavelets. In all examples, plane-wave excitation is considered, with connecting-surface implementation as discussed in Section II.C.

A. Dielectric Sphere in Free Space

We first consider a lossless dielectric sphere of relative permittivity $\epsilon_r = 4$ and 16 cm radius, in free space, with results compared with data calculated via an MoM analysis tailored for a body of revolution (BoR) [21]. The comparison is performed in terms of radar cross section (RCS) as a function of frequency. Concerning the MRTD results, a Fourier transform is taken of the time-domain fields on the surface of a Huygen's surface that

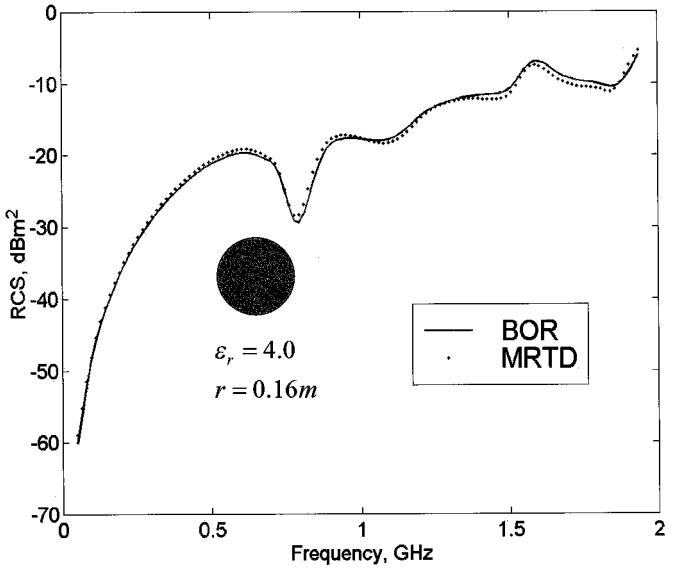


Fig. 2. Monostatic RCS of a 16-cm-radius dielectric sphere of dielectric constant; $\epsilon_r = 4$, as computed via MRTD and a BoR-MoM model [21].

encloses the target, with these frequency-domain fields spatially convolved with the appropriate Green's function, to compute the far-zone scattered fields. These fields are then divided by the frequency response of the incident waveform to deconvolve the effects of the incident pulse. In this example, we use the free-space Green's function; in examples below, the half-space Green's function is employed. We see in Fig. 2 that the agreement between the frequency-dependent MRTD and BoR-MoM results is generally good. In this example, the PML is placed 4 cm from the surface of the sphere, and therefore almost the entire computational domain is occupied by the target. Consequently, for this example, we do not exploit the multiresolution properties of MRTD, since the air region outside the target is a very small fraction of the target volume. In particular, there is little computational savings in using scaling functions alone in the air region (outside the target) and scaling functions and wavelets inside the target. Therefore, the Haar scaling function and a single level of wavelets are used throughout.

In this example, the principal advantage of the MRTD is manifested in the analysis of the curved spherical surface, which is not matched to a Cartesian lattice. In particular, we have used both scaling functions and wavelets throughout the computational domain, with the smallest dimension cubes in this analysis having a volume of $1 \times 1 \times 1$ cm^3, which constitutes a relatively coarse representation for the sphere considered (an FDTD staircase approximation with this discretization yields errors on the order of 2.0 dB). Within the context of MRTD, the inhomogeneity is handled as discussed in Section II-B.

B. Two Dielectric Cylinders in Free Space

The example in Section III-A underscores that not all targets necessarily require a multiresolution algorithm. However, there are many complex targets for which such an analysis is particularly useful. As such an example, we consider the bistatic RCS from two proximate dielectric cylinders in free space. The scattering geometry is shown in Fig. 3, with dimensions $D_1 = 20$ cm, $D_2 = 12$ cm, $h_1 = 102$ cm, $h_2 = 42$ cm, and $l = 80$ cm. The relative permittivity of each target is $\epsilon_r = 3$. In Fig. 4, we compare the bistatic RCS as computed via MRTD and FDTD, for a frequency of 700 MHz, with incident angles $\theta_{\text{inc}} = 45°$ and $\varphi_{\text{inc}} = 0°$, with observation at the angles $\theta_s = 90°$ and $\varphi_s = 0 - 180°$. While one is unlikely to perform such a bistatic measurement in practice, we have found bistatic comparisons to be a particularly good test of model accuracy. The frequency-domain data are computed from the time-domain MRTD and FDTD fields, as discussed in Section III-A.

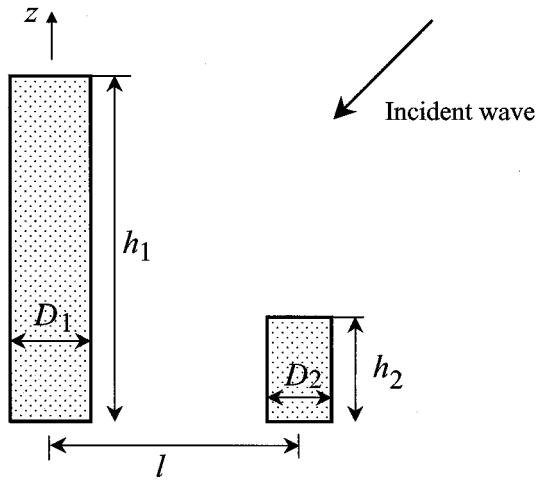


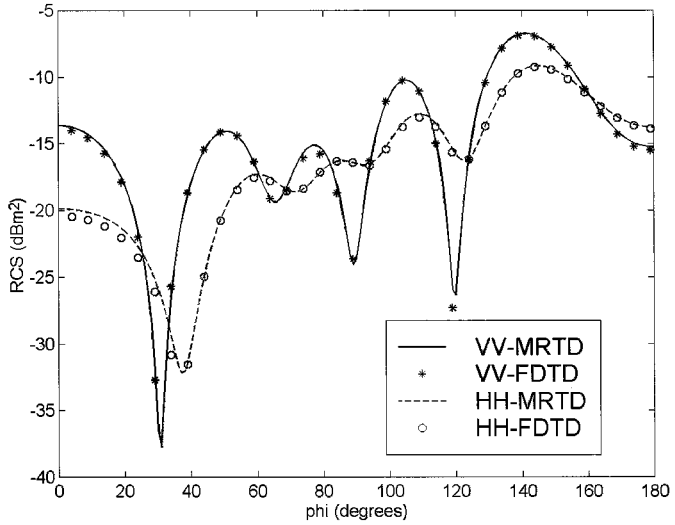
Fig. 3. Two dielectric cylinders in free space with plane-wave excitation.

In the FDTD computations, a cell size of $1 \times 1 \times 1$ cm^3 is used throughout. In the MRTD computations, wavelets are used in the vicinity of the targets, with scaling functions alone used outside. Where wavelets are employed, the resolution is as in the FDTD, and in the regions of scaling functions alone the sampling is half as fine (one MRTD cube, where FDTD employs eight). In the vicinity of the connecting surface, for the incident fields, both scaling functions and wavelets are used. For the FDTD computations the material interface at the target surface is treated via averaging the respective permittivities [15], while in the MRTD the material inhomogeneity is handled as discussed in Section II-B. We consider the copolarized fields in Fig. 4(a) and the cross-polarized case in Fig. 4(b). We see that the agreement between the FDTD and MRTD computations is generally good.

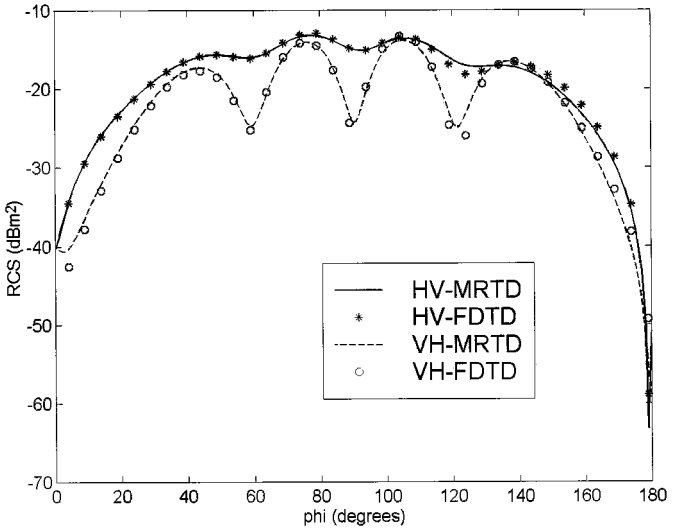
For the computations considered in Fig. 4, 28% of the MRTD volume is modeled via scaling functions alone, with the remainder modeled via a single level of wavelets as well. In the time-domain computations, a fourth-order Rayleigh pulse [22] (central frequency is 600 MHz) is chosen as the source, and the MRTD calculations required 23% less computation time (CPU) *vis-à-vis* the FDTD. The savings in computation time is largely dictated by the spatial separation between the two cylinders in Fig. 3, with further advantages of MRTD realized as this separation is increased.

C. Canonical Buried Targets

Many land mines are cylindrical in shape, and with this problem in mind we consider a dielectric cylinder of 4 cm height and 8 cm diameter, with 4 cm between the top of the target and the air–ground interface. The relative permittivity



(a)



(b)

Fig. 4. Bistatic scattering for the target in Fig. 3, at frequency 700 MHz, with $D_1 = 20$ cm, $D_2 = 12$ cm, $h_1 = 102$ cm, $h_2 = 42$ cm, and $l = 80$ cm. The relative permittivity of each target is $\epsilon_r = 3$. The plane-wave incident angles are $\theta_{\rm inc} = 45^\circ$ and $\varphi_{\rm inc} = 0^\circ$, with observation at the angles $\theta_s = 90^\circ$ and $\varphi_s = 0 - 180^\circ$. The data are computed via MRTD and FDTD. (a) Copolarized and (b) cross-polarized fields.

of the target is $\varepsilon_r = 2$, and the soil is modeled with a relative permittivity $\varepsilon_r = 5$ and conductivity $\sigma = 0.01$ S/m. Note that even a real permittivity with frequency-independent conductivity yields a dispersive soil medium. However, often one wishes to employ a more-sophisticated model for the dispersive properties of soil [2]–[4], [6], [7]. Such dispersion has been incorporated into the FDTD formalism [2], [7], and similar modifications can be made to MRTD, although we have not considered such here.

The plane-wave excitation has an incidence angle $\theta_{\rm inc} = 45^\circ$, and the bistatic scattered fields are viewed as a function of θ_s, in the same azimuthal plane as the incident wave. Results are shown for frequencies of 300, 500, and 700 MHz, with MRTD results compared with data computed via a BoR-MoM model [21]. As indicated in Fig. 5, the bistatic results from these two disparate models are in good relative agreement.

In Fig. 6, we consider the same scattering scenario as in Fig. 5, but now we consider monostatic scattering as a function of frequency. The excitation is incident as in Fig. 4. The agreement between the two models is consistent with our previous examples. In these computations, the PML in the air region is placed close to the air–soil interface, and therefore the air region in these computations constitutes a small percentage of the total computational volume. Moreover, since the target-soil contrast is relatively small, there is less need for a multiresolution analysis. In these MRTD computations, wavelets and scaling functions are employed throughout, with the principal utility of the MRTD again found in the manner by which the curved target surface is handled. The smallest cube in these computations had dimensions $0.5 \times 0.5 \times 0.5$ cm^3, and the central frequency of the incident pulse is 900 MHz. Note that in the context of a Haar expansion, when wavelets are used throughout, one could simply utilize scaling functions alone, with cell size split in half (in each dimension).

D. Buried Target Near Clutter

As indicated in Section III-D, for many subsurface-sensing problems, the utility of the MRTD multiresolution property is less clear, since the PML allows the air region to be very small (the PML is close to the air–ground interface). In this case, the air region, which permits coarser spatial sampling than the dielectric target and soil, only constitutes a small fraction of the computational domain. However, there are problem classes for which MRTD yields clear advantages. For example, rough-surface scattering is a problem well suited to MRTD, this having been demonstrated previously for two-dimensional problems [16]. In particular, a general rough surface often manifests fine roughness structure that is difficult to model via special conformal FDTD cells [18], due to the generality of the roughness. On the other hand, the natural manner with which MRTD handles dielectric interfaces is well suited to the analysis of such scattering scenarios. Rough-surface scattering is a discipline unto itself and requires a statistical parametrization of the surface and scattered fields [16]. In lieu of addressing this problem directly, we consider a simpler but related scattering scenario. In Fig. 7, we again consider a buried dielectric cylinder, but now there is a protruding bump ("hill") in the surface, characterized by radius r and height s. This bump constitutes one portion of a general rough surface, allowing consideration of relevant numerical issues. In particular, for a "hill" of height s, assume that k FDTD cells are employed along the height dimension s (Fig. 7). As the height s diminishes, the size of the FDTD cells must diminish (k must increase). If the FDTD cells are sampled uniformly across the computational domain, as in a traditional FDTD analysis, the accurate representation of this fine structure will significantly increase the numerical size of the problem. Similar issues hold in the context of representing the detail in a rough surface [16].

We consider scattered fields from the situation in Fig. 7, as computed via FDTD and MRTD. In MRTD scaling, functions are used throughout, and a single level of wavelets is used in the subsurface and about the "hill." Assume that the scaling functions have sides of length w. In the FDTD analysis, we initially use cells of length $w/2$, corresponding to an MRTD

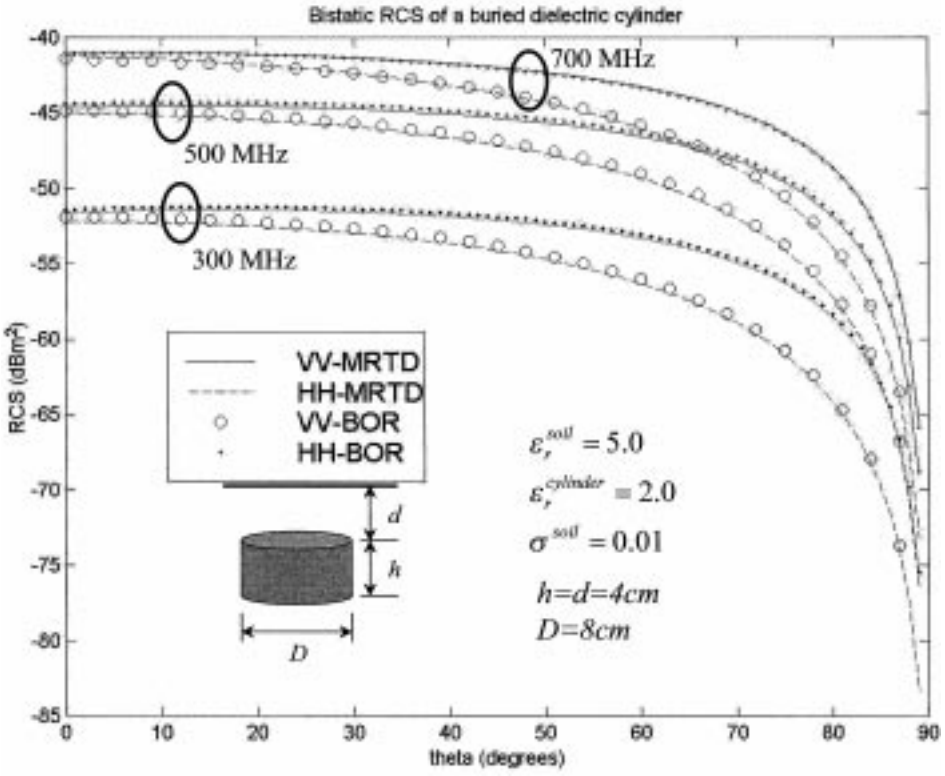


Fig. 5. Copolarized bistatic scattering from a lossless dielectric buried in a lossy half-space (see insert for geometric and electrical parameters). The incident plane wave is incident at $\theta_{\rm inc} = 45^\circ$, and the bistatic fields are computed as a function of θ, in the same azimuthal plane as the incident wave. Data are computed via the MRTD and BoR-MoM [21] at 300, 500, and 700 MHz.

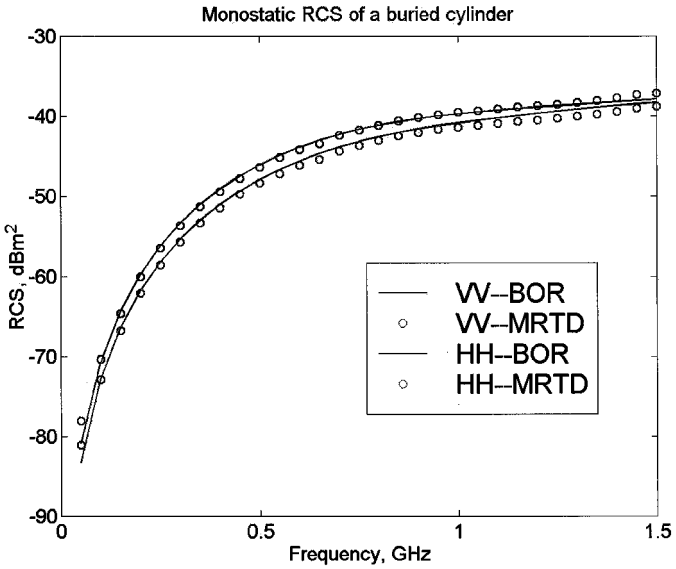


Fig. 6. Monostatic RCS as a function of frequency for the target geometry and incidence angle considered in Fig. 5.

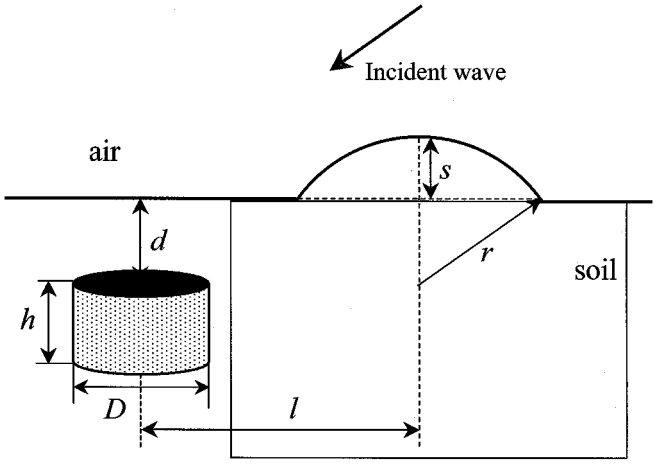


Fig. 7. Scattering geometry for results in Fig. 8Fig. 9. The plane-wave excitation is incident in the azimuthal plane bisecting the centers of the cylindrical target and spherically shaped surface clutter.

scheme with Haar wavelets throughout. In the MRTD, the "hill" inhomogeneity is modeled as discussed in Section II-B, and the FDTD performs averaging across discrete inhomogeneities [15]. One could also employ conformal FDTD cells [18], but this is not considered here, as this would be difficult to implement for a general rough surface.

The results in Fig. 8 correspond to $h = 8$ cm, $d = 4$ cm, $D = 16$ cm, $r = 16$ cm, and $l = 30$ cm and the relative permittivity of the lossless target is $\epsilon_r = 2.0$. The soil is modeled as in Figs. 5 and 6, and we consider $s = 3$ cm, $s = 5$ cm, and $s = 8$ cm. In the initial examples, the spatial sampling is $w = 2$ cm, corresponding to $k = 3$, $k = 5$, and $k = 8$ for $s = 3$ cm, $s = 5$ cm and $s = 8$ cm, respectively.The bistatic RCS is plotted as a function of θ_s in a plane bisecting the center of the two scatterers (Fig. 7), with the excitation in this plane at $\theta_{inc} = 45°$. As expected, as the size of the hill diminishes, so do the scattered fields. It is also clear that the differences between the MRTD and FDTD increase as the hill size diminishes, which is manifested

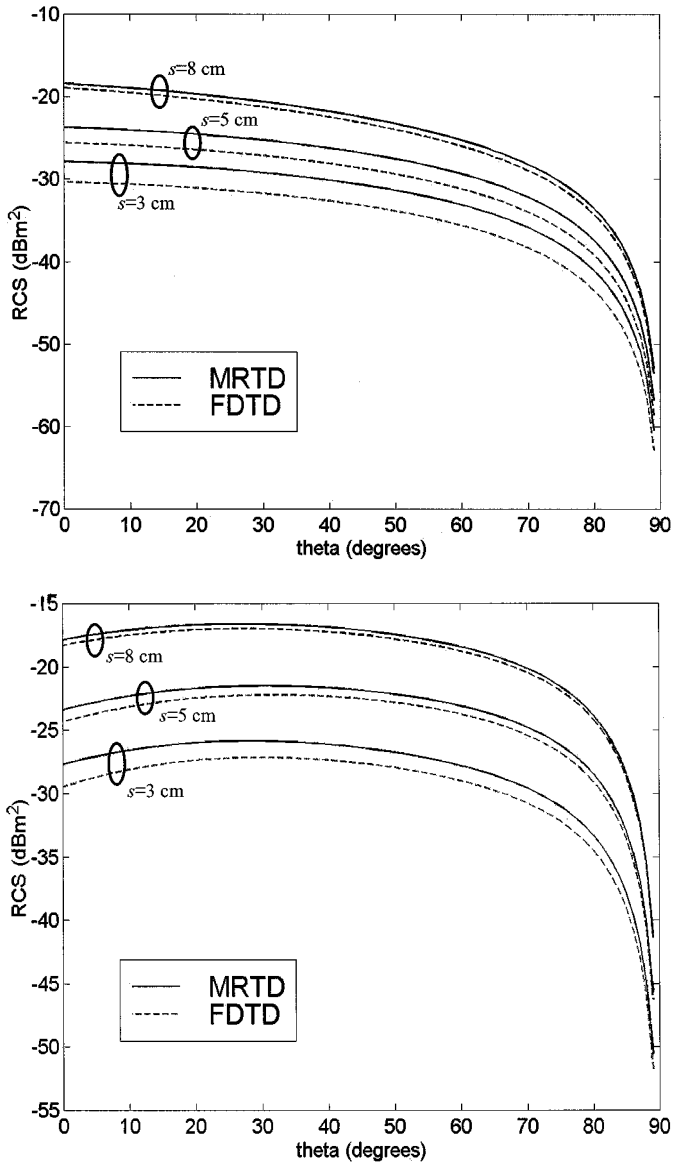


Fig. 8. Bistatic scattering from the target in Fig. 7, with parameters $h = 8$ cm, $d = 4$ cm, $D = 16$ cm, $r = 16$ cm, and $l = 30$ cm, and the relative permittivity of the lossless target is $\epsilon_r = 2.0$. The soil is modeled as in Figs. 5 and 6, and results are presented for $s = 3$ cm, $s = 5$ cm, and $s = 8$ cm. The plane-wave excitation is at $\theta_{\rm inc} = 45°$. The scaling function has a width w and the FDTD cells have width $w/2$, where $w = 2$ cm. (a) HH polarization and (b) VV polarization.

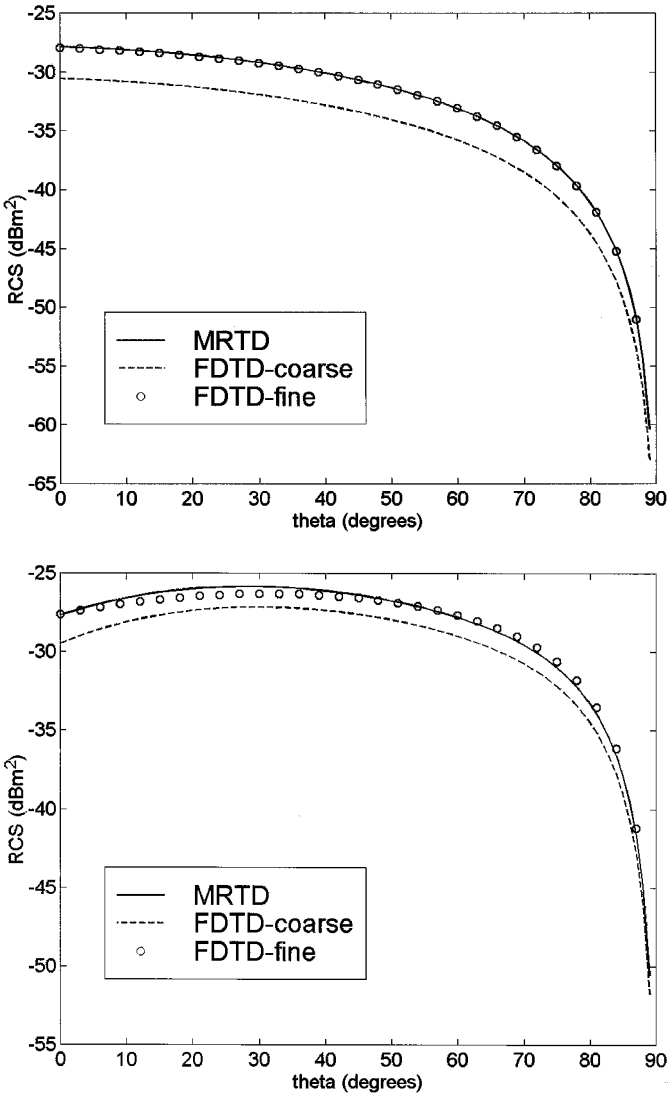


Fig. 9. Bistatic scattering as in Fig. 8, for the case $s = 3$ cm. The MRTD results are as in Fig. 8, but the FDTD results employ spatial sampling twice as dense as Fig. 8. In particular, the scaling function has a width w and the FDTD cells have width $w/4$, where $w = 2$ cm. The "coarse" results correspond to data from Fig. 8. (a) HH polarization and (b) VV polarization.

by the fact that, with the discretization level considered, the hill is not well represented via the FDTD.

In Fig. 9, we again consider the $s = 3$ cm case but increase the FDTD sample rate. In particular, the FDTD cell width is now $w/4$, where w is again the length of the Haar scaling functions. We see that now the FDTD and MRTD results are in nearly exact agreement for HH polarization and in closer agreement for VV polarization. For the case $s = 3$ cm, $k = 3$ in Fig. 8 and $k = 6$ in Fig. 9. As applied here, a uniform spatial sampling has been used throughout the FDTD model, and therefore the FDTD computations in Fig. 9 required 83% more CPU time than the corresponding MRTD results with scaling-function width w. A more detailed analysis of the benefits of MRTD for rough-surface scattering is discussed in [16].

IV. Conclusion

In this paper, we have examined the analysis of three-dimensional scattering via the MRTD method, with a Haar-wavelet expansion. Comparisons have been made to data generated via an MoM model [21] and to FDTD-generated data. Several observations can be accrued from these initial studies. First, an advantage of the MRTD is manifested in its natural multigrid capabilities. In particular, within the context of a Haar expansion, regions in which wavelets are employed correspond to cubic sampling at a higher density relative to where scaling functions are used alone. A good example of where such a multiresolution scheme might appear salutary is for subsurface sensing. For such a problem, one can often spatially sample the air region more coarsely than the soil, since the wavelength in air is typically much larger than that in the soil. However, new absorbing-boundary conditions (ABCs), such as the perfectly matched layer [19], [20], are so good that the ABC can be placed

very close to the air–soil interface. Therefore, the air region constitutes a small fraction of the computational domain, realizing less computational benefit to a multiresolution analysis. We found the same to be true of free-space scattering from a dielectric sphere, for example, in which the PML ABC was again placed very close to the target, minimizing the volume occupied in the computational domain by the surrounding air. Nevertheless, there are problems for which such a multiscale analysis is appropriate, e.g.,multiple dielectric targets with relatively large intertarget spacing. Moreover, if there is a large contrast between a buried target and the surrounding soil, a multiresolution analysis may again be desirable.

In the examples considered here, in which the targets had a curved surface that did not conform to the Cartesian FDTD or MRTD grid, the principal utility of MRTD was manifested in the handling of general inhomogeneities. As discussed in Section II-B, for regions in the vicinity of discrete dielectric contrasts, the MRTD difference equations are augmented by matrices that account for the material medium. These matrices are easily computed (and stored) within the context of MRTD, and it was demonstrated that in some examples this can lead to significant numerical savings vis-à-vis a traditional (simple) FDTD analysis. In particular, in FDTD one typically samples the computational domain uniformly in space, and therefore the computational complexity is dictated by the smallest scatter to be discretized. It should be noted that conformal cells have been introduced within the context of FDTD, which could also be employed to handle local curvature [18]. Multigrid FDTD schemes have also been addressed, although these typically require significant care to assure numerical stability [12]–[14]. It is important to emphasize that the examples presented in this paper, comparing FDTD and MRTD, handicapped the FDTD scheme in that the simplest FDTD implementation has been considered. In this context, we note that numerous improvements have been implemented in the FDTD, allowing it to consider multigridding (as mentioned above) as well as conformal meshes [23]–[25].

We have undertaken an initial investigation of the utility of MRTD for three-dimensional scattering problems, employing Haar wavelets. There are many issues that deserve further study. In particular, one of the principal advantages of MRTD involves reduced numerical dispersion compared to FDTD, assuming that higher order wavelets are employed. For example, the Battle–Lemarie wavelets have been demonstrated to yield markedly reduced numerical dispersion relative to FDTD [9], although the integrals required for material inhomogeneity are more complicated than with a Haar expansion. One can also consider a biorthogonal wavelet expansion [10], which offers advantages relative to orthogonal wavelets. Therefore, while a principal utility of MRTD, as investigated here, has centered around the ability to handle general dielectric inhomogeneity, to this advantage can be added reduced numerical dispersion through use of more sophisticated wavelets.

REFERENCES

[1] L. Peters Jr, J. J. Daniels, and J. D. Young, "Ground penetrating radar as an environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1822, Dec. 1994.

[2] J. M. Bourgeois and G. S. Smith, "A fully three-dimensional simulation of ground penetrating radar: FDTD theory compared with experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 36–28, Jan. 1996.

[3] S. Vitebskiy, K. Sturgess, and L. Carin, "Short-pulse plane-wave scattering from buried perfectly conducting bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 143–151, Feb. 1996.

[4] S. Vitebskiy and L. Carin, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy dispersive half-space," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.

[5] T. J. Yu and L. Carin, "Analysis of the electromagnetic inductive response of a void in a conducting-soil background," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1320–1327, May 2000.

[6] N. Geng, A. Sullivan, and L. Carin, "Multilevel fast-multipole algorithm for scattering from conducting targets above or embedded in a lossy half space," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1561–1573, July 2000.

[7] T. P. Montoya and G. S. Smith, "Land mine detection using a ground-penetrating radar based on resistively loaded vee dipoles," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 1795–1806, Dec. 1999.

[8] M. Krumpholz and L. Katehi, "MRTD: New time-domain schemes based on multiresolution analysis," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 555–571, 1996.

[9] E. T. Tentzeris, R. L. Robertson, J. Harvey, and L. Katehi, "Stability and dispersion nalysis of Battle-Lemarie based MRTD schemes," *IEEE Trans. Microwave Theory Tech.*, vol. 47, pp. 1004–1013, 1999.

[10] T. Dogaru and L. Carin, "Multiresolution time-domain using CDF biorthogonal wavelets," *IEEE Trans. Microwave Theory Tech.*, vol. 49, pp. 902–912, May 2001.

[11] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[12] M. W. Chevalier, R. J. Luebbers, and V. P. Cable, "FDTD local grid with material traverse," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 411–421, 1997.

[13] M. J. White, M. F. Iskander, and Z. Huang, "Development of a multigrid FDTD code for three-dimensional applications," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 1512–1517, 1997.

[14] M. Okoniewski, E. Okoniewska, and M. A. Stuchly, "Three-dimensional subgridding algorithm for FDTD," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 422–429, 1997.

[15] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 1995.

[16] T. Dogaru and L. Carin, "Multiresolution time-domain analysis of scattering from a rough dielectric surface," *Radio Sci.*, vol. 35, pp. 1279–1292, Nov.-Dec. 2000.

[17] F. D. Hastings, J. B. Schneider, and S. L. Broschat, "A Monte-Carlo FDTD technique for rough surface scattering," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 1183–1191, 1995.

[18] T. G. Jurgens, A. Taflove, K. R. Umashankar, and T. G. Moore, "Finite-difference time-domain modeling of curved surfaces," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 357–366, 1992.

[19] J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, vol. 114, pp. 185–200, 1994.

[20] J. Fang and Z. Wu, "Generalized perfectly matched layer—An extension of Berenger's perfectly matched layer boundary condition," *IEEE Microwave Guided Wave Lett.*, vol. 5, pp. 451–453, Dec. 1995.

[21] N. Geng and L. Carin, "Wide-band electromagnetic scattering from a dielectric BOR buried in a layered lossy dispersive medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.

[22] P. Hubral and M. Tygel, "Analysis of the Rayleigh pulse," *Geophysics*, vol. 54, pp. 654–658, 1989.

[23] W. H. Yu and R. Mittra, "A conformal FDTD algorithm for modeling perfectly conducting objects with curve-shaped surfaces and edges," *Microwave Opt. Tech. Lett.*, vol. 27, pp. 136–138, Oct. 20, 2000.

[24] M. W. Yang, Y. C. Chen, and R. Mittra, "Hybrid finite-difference/finite-volume time-domain analysis for microwave integrated circuits with curved PEC surfaces using a nonuniform rectangular grid," *IEEE Trans. Microwave Theory Tech.*, vol. 48, pp. 969–975, June 2000.

[25] M. Celuch-Marcysiak and W. K. Gwarek, "Mproved and simpler FDTD formulation for axisymmetrical problems," in *IEEE Antennas Propagation Soc. Int. Symp.*, vol. 1, 2000, pp. 31–31.



Xianyang Zhu (M'01) [redacted] He received the B.S. and Ph.D. degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1988 and 1994, respectively.

Since March 1999, he has been a Research Associate in the Center for Applied Remote Sensing, Department of Electrical and Computer Engineering, Duke University, Durham, NC. From November 1997 to February 1999, he was a Postdoctoral Research Fellow with the Center for Computational Electromagnetics, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign. From 1994 to 1997, he was with the Electromagnetics Institute, Southwest Jiaotong University, where he became an Associate Professor in 1995. His main research interests include computational electromagnetics, radar cross-section analysis, electromagnetic compatibility, and antennas analysis.

Lawrence Carin (SM'96–F'01) [redacted] in Washington, DC. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, Brooklyn, NY, as an Assistant Professor. He became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, Durham, NC, where he is now a Professor. He is the Principal Investigator on a Multidisciplinary University Research Initiative (MURI) on demining. His current research interests include short-pulse scattering, subsurface sensing, and wave-based signal processing.

Prof. Carin is a member of Tau Beta Pi and Eta Kappa Nu. He is an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.

Cross-Reactive Chemical Sensor Arrays

Keith J. Albert,[†] Nathan S. Lewis,*[,‡] Caroline L. Schauer,[†] Gregory A. Sotzing,[‡] Shannon E. Stitzel,[†] Thomas P. Vaid,[‡] and David R. Walt*[,†]

The Max Tishler Laboratory for Organic Chemistry, Department of Chemistry, Tufts University, Medford, Massachusetts 02155, and Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received August 17, 1999

Contents

I. Introduction 2595
II. Tin Oxide Arrays 2598
  A. Theory 2598
  B. Tin Oxide Characteristics 2599
  C. Array Design 2600
  D. Applications 2600
  E. Analysis 2601
III. MOSFET Arrays 2602
  A. Principles of Operation 2602
  B. Fabrication and Operation 2602
  C. Applications 2603
  D. Alternative Implementations of MOSFETs 2604
  E. Hybrid Metal Oxide Containing Arrays 2605
IV. Intrinsically Conductive Polymer Chemiresistor Arrays 2605
  A. Principles of Operation 2605
  B. Array Fabrication 2606
  C. Applications 2607
  D. Future Directions 2608
V. Conductive Polymer Composite Chemiresistor Arrays 2609
  A. Theory 2609
  B. Fabrication 2609
  C. Applications 2611
VI. Optical Vapor Sensing Arrays 2611
  A. Introduction 2611
  B. Polymer-Deposited Optical Sensor Arrays 2611
    1. Array Fabrication 2611
    2. Sensor Diversity 2612
  C. Self-Encoded Bead Sensors 2613
  D. Sensor Sensitivity 2613
  E. Data Processing for Optical Sensor Arrays 2614
VII. Electrochemical Sensor Arrays 2614
  A. Introduction 2614
  B. Potentiometric (Equilibrium) Measurements 2615
    1. Principles of Operation 2615
    2. Examples of Potentiometric Arrays 2616
  C. Voltammetric (Nonequilibrium) Measurements 2618
    1. Principles of Operation 2618
    2. Examples of Voltammetric Arrays 2618
VIII. Acoustic Wave Devices 2620
  A. Introduction 2620
  B. TSM Arrays 2620
  C. SAW Arrays 2621
  D. Response Prediction 2622
IX. Conclusions and Future Prospects 2622
X. Acknowledgments 2623
XI. References 2623

I. Introduction

Conventional approaches to chemical sensors have traditionally made use of a "lock-and-key" design, wherein a specific receptor is synthesized in order to strongly and highly selectively bind the analyte of interest.[1−6] A related approach involves exploiting a general physicochemical effect selectively toward a single analyte, such as the use of the ionic effect in the construction of a pH electrode. In the first approach, selectivity is achieved through recognition of the analyte at the receptor site, and in the second, selectivity is achieved through the transduction process in which the method of detection dictates which species are sensed. Such approaches are appropriate when a specific target compound is to be identified in the presence of controlled backgrounds and interferences. However, this type of approach requires the synthesis of a separate, highly selective sensor for each analyte to be detected. In addition, this type of approach is not particularly useful for analyzing, classifying, or assigning human value judgments to the composition of complex vapor mixtures such as perfumes, beers, foods, mixtures of solvents, etc.

An emerging strategy that is complementary to the conventional chemical sensing approach involves the use of sensor arrays. The utilization of sensor arrays is inspired by the superb performance of biological olfactory systems in odor detection, identification, tracking, and location tasks. Recent work has shown that the mammalian olfactory system contains approximately 1000 different olfactory receptor genes and that, upon odor stimulation, responses from many receptors are sent to the olfactory bulb and then on to the olfactory cortex for processing.[7−10] Furthermore, recent experiments have shown that the olfactory receptors are not highly selective toward specific analytes; in fact, one receptor responds to many analytes and many receptors respond to any given analyte.[8,10−12] Pattern recognition methods are thus thought to be a dominant mode of olfactory

[†] Tufts University.
[‡] California Institute of Technology.

10.1021/cr980102w CCC: $35.00 © 2000 American Chemical Society
Published on Web 06/24/2000



Keith Albert was [redacted] He attended Colby College where he did a one year exchange program at UCC in Cork, Ireland. He received his bachelor's degree in 1996 and is presently a graduate student at Tufts University with Prof. David Walt. His research interests include optical sensors, explosives-like vapor detection with microsensor array platforms, and instrumental designs for field use.



Dr. Nathan Lewis has been on the faculty at the California Institute of Technology since 1988 and has served as Professor since 1991. He has also served as the Principal Investigator of the Beckman Institute Molecular Materials Resource Center at Caltech since 1992. From 1981 to 1986, he was on the faculty at Stanford, as an Assistant Professor from 1981 to 1985 and a tenured Associate Professor from 1986 to 1988. Dr. Lewis received his Ph.D in Chemistry from the Massachusetts Institute of Technology. Dr. Lewis has been an Alfred P. Sloan Fellow, a Camille and Henry Dreyfus Teacher–Scholar, and a Presidential Young Investigator. He received the Frensenius Award in 1990 and the ACS Award in Pure Chemistry in 1991. He has published over 170 papers and has supervised approximately 50 graduate students and postdoctoral associates. His research interests include semiconductor electrochemistry and photochemistry, scanning tunneling microscopy of organic monolayers, and artificial olfactory systems using arrays of chemical sensors.

signal processing in the broadly responsive portion of the olfactory system of higher mammals.

In the array approach, the strict "lock-and-key" design criterion of traditional sensing devices is abandoned. Instead, in this alternative sensor architecture, an array of different sensors is used, with every element in the sensor array chosen to respond to a number of different chemicals or classes of chemicals. The elements of such an array need not be individually highly selective toward any given analyte, so this stressing constraint on sensor design is relaxed. Instead, the collection of sensors should contain as much chemical diversity as possible, so that the array responds to the largest possible cross-section of analytes. In practice, most chemical sensors



Caroline Schauer was born in Washington, DC, and graduated in 1991 from Beloit College in Beloit, WI with a B.S. degree in Chemistry. She completed her M.S. and Ph.D. from State University of New York at Stony Brook under the direction of Professors Frank W. Fowler and Joseph W. Lauher working on organic synthesis and small-molecule crystallography. She spent 9 months in the laboratory of David N. Reinhoudt as a postdoctoral fellow at the University of Twente, The Netherlands, working on hydrogen-bonded polymers. Currently she is a postdoctoral fellow in the laboratory of David R. Walt at Tufts University working on enzymatic biosensors. Her research interests include structural chemistry and biology, biosensors for proteins, development of an optical array for the diagnoses of disease states, and crystallography.

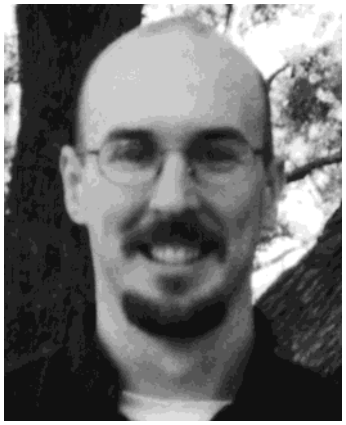


Gregory Sotzing was born in Fredericksburg, VA and graduated in 1993 from Mary Washington College with a B.S. degree in Chemistry with highest honors. He received the ACS Polyed Undergraduate Polymer Chemistry Award for his work on the synthesis of fluorinated polyimides. After completing his Ph.D. degree in 1997 under the advisement of John Reynolds in the George and Josephine Bulter Polymer Research Laboratory at the University of Florida, he accepted a postdoctoral position at the California Institute of Technology working jointly with Nathan Lewis and Robert Grubbs. Presently, Greg is an Assistant Professor of Chemistry in the Institute of Materials Science, Polymer Program, at the University of Connecticut, Storrs, CT. His research interests are primarily directed toward the application fields of sensors and actuators with an emphasis on materials design and polymer synthesis (www.mail.ims.uconn.edu/~sotzing).

suffer from some interference by responding to chemical species that are structurally or chemically similar to the desired analyte. This interference is an inevitable consequence of the "lock" being able to fit a number of imperfect "keys". Differentially responsive arrays take advantage of this interference or "cross reactivity" by deliberately attempting to use the nonspecific response patterns for analyte recognition. In this design, identification of an analyte cannot be accomplished from the response of a single sensor element; a distinct pattern of responses pro-



Shannon Stitzel [redacted] She received her B.S. in Chemistry from Davidson College in 1997. Presently she is a graduate student in Chemistry at Tufts University with Prof. David Walt. Her research interests include designing fiber-optic chemical sensors for vapor and liquid-phase analytes and alternative fuel sources.



Thomas Vaid [redacted]. He received his B.S. in Chemistry from the University of Illinois at Urbana-Champaign in 1992. His graduate work was with Prof. Peter T. Wolczanski at Cornell University, where he received his Ph.D. in 1997. Since then he has been a postdoctoral scholar at the California Institute of Technology with Prof. Nathan Lewis. In the fall of 2000, he will join the faculty of Washington University in St. Louis as an Assistant Professor in Chemistry. His research interests include the synthesis and study of new inorganic and organic materials with unusual magnetic ordering or electrical transport properties.



Dr. David R. Walt is Robinson Professor of Chemistry at Tufts University. He received a B.S. in Chemistry from the University of Michigan and a Ph.D. in Organic Chemistry and Pharmacology from SUNY at Stony Brook. After postdoctoral studies at MIT, he joined the chemistry faculty at Tufts. Dr. Walt served as Chemistry Department Chairman from 1989 to 1996. Dr. Walt serves on many government advisory panels and editorial advisory boards. He has received numerous national and international awards and honors recognizing his work. Dr. Walt has published over 120 papers, holds over 30 patents, and has given hundreds of invited scientific presentations. His research interests include micro and nano sensors, genomics and organized assembly of nano materials.

duced over the collection of sensors in the array can provide a fingerprint that allows classification and identification of the analyte (Figure 1). The pattern can be obtained from equilibrium or kinetic responses with the latter often providing additional discriminating power. The response mechanism for such systems is highly varied, as described in the sections below. Both binding and colligative properties can be interrogated with such arrays. For example, broadly responsive receptors can be employed to allow a range of structurally similar molecules to bind, membranes may be used that are size selective, and polymers may be employed that select on the basis of polarity. All these recognition mechanisms, as well as others described in this review, are often employed simultaneously in these arrays. These types of systems, which are the topic of this review, are thus commonly designated as artificial or electronic noses.

The advantage of this approach is that it can yield responses to a variety of different analytes, including those for which the array was not necessarily originally designed to detect. An array of sensors naturally performs an integration to yield a unique signal for complex but distinctive odors (e.g., cheeses, beers, etc.) without requiring that the mixture be broken down into its individual components prior to, or during, the analysis. This is a disadvantage when the precise chemical composition of a complex mixture is required but is advantageous when the only required information is the composite composition of the odor of concern. Some additional information can also be obtained by identifying unique spatial and/or temporal characteristics of certain analytes, so that the composition of even modestly complex mixtures can sometimes be obtained from sensor array signals using such methods. Another potential disadvantage to an array system is the possibility that other unknowns may give the same "unique" signal as a specific analyte of interest. However, these arrays are no different from other sensor types, in that there is always the potential for species other than the analyte of interest to provide a response that may be misconstrued as the target analyte. As will be described later in this review, cross-reactive arrays can be trained to evaluate more complex aspects of a sample, such as "freshness", and the fidelity of such an analysis may pose additional stringencies on the quality of the information produced by the array in order for it to not be "fooled". The sensors themselves are typically low power and simple in concept and operate at ambient or near ambient temperature and pressure. Their simplicity eliminates the need to solve the power and complexity challenges involved in miniaturizing traditional laboratory analytical chemical systems that involve high power and high vacuum (mass spectrometers, for example), high pressures and/or gas flows (e.g., gas chromatography), or other operational constraints that present severe mismatches between the optimal instrumental operating conditions and those likely to be encountered in an out-of-lab setting.

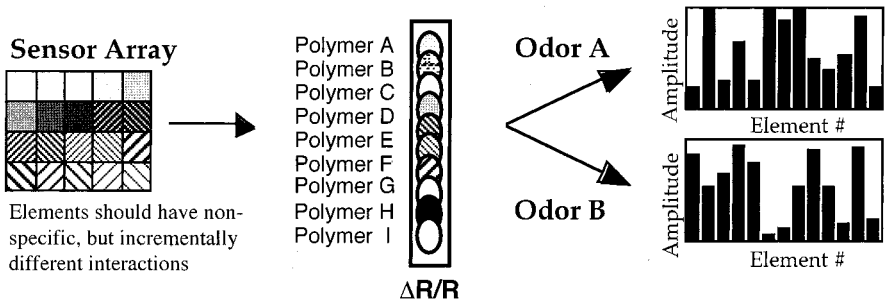


Figure 1. Response of a collection of incrementally different but nonspecific sensors, used to generate a complex pattern, or fingerprint, characteristic of a given analyte. Pattern recognition processing, including neural networks, can then be used to identify analytes on the basis of these patterns.

Signal processing methods and algorithms are intimately associated with the performance of broadly cross-reactive sensor arrays as a vapor detection system. Signal processing algorithms can comprise statistically based chemometric methods, pattern recognition algorithms, neural networks, or some combination thereof. One of the design tradeoffs in deploying any specific sensor array system for a given vapor detection task will thus clearly involve assessment of the computational needs to achieve a robust classification of the desired target analyte in the presence of background, environmental variability, and interfering signals.

There is some controversy in the literature regarding whether it is advantageous to use large numbers of sensors in an array device. This arises because approximately five fundamental molecular descriptors have been sufficient to describe much of the variance in the gas−solid partition coefficients for sorption of vapors into various polymers. Although it is possible to describe the gross features of such partitioning with a small number of descriptors, we note that there are over 1000 olfactory genes in humans and over 100 million olfactory cells in a canine's nose. Even if the dimensionality of odor space is fairly small, say on the order of 10^1, it is not likely that ideal sensors that produce optimal resolution along the fundamental directions of odor space could be identified. In practice, correlations between the elements of a sensor array will likely necessitate a much larger number of sensors to distinguish successfully between any two molecules in a complex environment. Also, it is beneficial to measure the same property in many different ways due to noise limitations in a practical system. For example, if sufficient precision could be obtained, it might be possible to identify uniquely any molecule merely from a 38 bit precision measurement of two parameters, perhaps its dipole moment and its polarizability. But of course, it is not practical to make such measurements with this precision; hence at lower precision, useful information on the nature of the analyte is gained by making measurements of the molecular parameters through many independent determinations on different sensor elements. The dimensionality of odor space will depend on the sensor array and the recognition mechanism employed. The dimensionality of odor space is thus inherently coupled to the precision of the sensor array that is being used to make the determination as well as to the diversity in the analytes that is being used to define the space as a whole. Furthermore, the above arguments on needing a limited number of sensors only hold if one is tasked to distinguish between a series of pure substances that are maintained at one fixed, known concentration. In contrast, if the background is unknown, if mixtures are present, or if the background gases are changing in concentration, many more sensors are needed simply to avoid ambiguity in interpreting the output signal pattern, and even more are needed if optimal discrimination is to be accomplished between a given target signature and a wide possible range of background clutter and false alarm signatures. Having large numbers of sensors also allows redundancy, which improves the signal-to-noise ratio and provides the ability to veto the output of poorly performing sensors. Redundancy is clearly important in biological systems where each odorant receptor type is expressed clonally on thousands of individual cells. Because of all of these issues, the number of sensors required to successfully span odor space in a practical device will rapidly multiply from the minimum value defined by the rank of smell space, and this is a requirement that is readily met using versatile sensor array architectures that can incorporate high levels of chemical diversity and redundancy into the system.

Sensor types for electronic noses can be quite diverse. Sensor array elements that will be discussed in this review include metal oxide devices, intrinsically conducting organic polymers, conducting polymer composites, dye-impregnated polymers coated onto optical fibers, electrochemical devices, and polymer-coated surface or bulk acoustic wave oscillators. While the focus of this review is primarily on the chemistry of these arrays, there are a number of application areas for the technology that will be of interest to the reader. These cross-reactive arrays have a wide variety of application areas including food and beverages, fragrances, environmental monitoring, chemical and biochemical processing, medical diagnostics, transportation, and a host of others. Some aspects of these areas as well as specific examples from each will be illustrated in the appropriate sections of the text that follows.

II. Tin Oxide Arrays

A. Theory

Tin oxide (SnO_2) gas sensors were first demonstrated in the early 1960s. Since that time, SnO_2

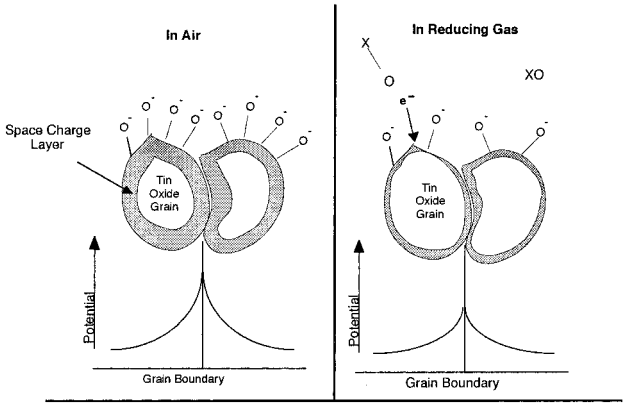


Figure 2. Potential barrier at the grain boundary (a) in air and (b) in a reducing gas.

sensors have become commercially available for detecting fuel gas, carbon monoxide, general purpose combustible gases, ammonia, water vapor, etc.[13,14] As early as 1954, the concept of a cross-reactive array of sensors for odor detection was discussed in the literature.[15] However, it was not until 1982 that a cross-reactive array of SnO_2 sensors was demonstrated by Persaud and Dodd to mimic olfaction.[16] This publication started the "modern era" of cross-reactive arrays and electronic noses that combine broad sensitivity with convergent signal processing to perform a wide range of analyte identification. Although many different arrays have been fabricated, the sensing mechanism is similar for all SnO_2 elements.

The general mechanism for vapor sensing with tin oxide is a change in the resistance (or conductance) of the sensor when it is exposed to an analyte relative to the sensor resistance in background air. This change in resistance is due to irreversible reactions between the analyte and oxygen-derived adsorbates such as O^-, O_2^-, and O^{2-} on the semiconductor surface. The most reactive of these species is O^-, which tends to dominate the resistance of the semiconductor through chemical control of vacancy sites in the SnO_2. Oxygen adsorbs to the semiconductor surface according to the following equation:

$$^1/_2O_2 + e^- \rightarrow O^- \text{ (ad.)} \qquad (1)$$

With n-type semiconductors, adsorption of O^- creates a space-charge layer on the surface of each SnO_2 grain, which creates a potential barrier to conduction at each grain boundary. The depth of the space charge layer changes in relation to the concentration of oxygen adsorbates on the surface. Therefore, when the surface is exposed to air, the oxygen concentration is high and the material displays a high electrical resistivity. In contrast, when the sensor is exposed to a reducing gas X, the gas reacts with some adsorbed oxygen species O_m^- as follows:

$$X + O_m^- \rightarrow XO_m + e^- \qquad (2)$$

This leads to oxygen consumption, the return of electrons to the oxide grains, and a decrease in the semiconductor resistance (Figure 2). When the sensor is exposed to an oxidizing gas such as NO_2, the resistance increases as the gas chemisorbs as a negatively charged species to the semiconductor surface. The change in resistance is therefore due to the chemisorption of the oxidizing gas, assuming that the concentration of oxygen adsorbates remains constant.[13,14,17] Another hypothesis for the mechanism of conductivity change in SnO_2 is that reducing gases react with the material and thereby reduce the number of oxygen vacancies in the material, lowering its conductance, while oxidizing environments regenerate the oxygen-derived vacancies and thus restore the electrical conductivity of the film.

B. Tin Oxide Characteristics

Various properties of tin oxide arrays have been manipulated in attempts to enhance and/or broaden the sensitivities of such devices. The sensitivities of SnO_2 sensors to various gases can be enhanced by doping with metals. Metals promote the catalytic activity of the semiconductor surface with gases, leading to chemical sensitization. The addition of a metal to SnO_2 also causes electronic sensitization due to the effects on the space charge layer. Metals that have work functions greater than the electron affinity of the semiconductor take electrons from the semiconductor, leading to even greater resistances in air. Because oxygen adsorbs to both the metal and semiconductor surfaces, when it desorbs from both surfaces due to gas interactions, an enhanced sensitivity is produced in the resistance change of the semiconductor.[14] Typical doping metals include Pt and Pd, but others have also been used, such as Al[18] and Au[19] (although neither Al nor Au has provided much improvement in sensitivity). Doping with Pt and Pd has been shown to increase the sensitivity of tin oxide sensors to gases such as benzene and toluene, with lower doping concentrations leading to greater sensitivity. Doped sensors have demonstrated greater sensitivity to oxygenated volatile organics than to aliphatic, aromatic, or chlorinated compounds.[18]

Temperature is another factor that affects the sensitivity of tin oxide sensor arrays. Typically, these arrays are operated at temperatures greater than 300 °C to increase the reactivity of the semiconductor surfaces. Tin oxide sensors are also operated at elevated temperatures to desorb water produced from the reactions on the catalyst surface. One study looked at the effect of various temperatures between 300 and 450 °C on an array of three elements.[20] It was determined that if a sensor's response changed linearly with change in temperature, it did not affect the array's overall performance. An approach to utilize the change in sensitivity with temperature is to use a thermal cycling technique to extract more information from each sensor. This approach was first demonstrated by Heilig et al.[21] and then used by Corcoran et al., who cycled an eight sensor array between 250 and 500 °C. These workers used the data in conjunction with feature extraction techniques to classify teas. The thermally cycled array had a 90% classification rate, whereas the same array at a fixed temperature only had a 69% classification rate.[22]

Another study investigated the response and recovery times of tin oxide sensors.[19] Some general conclusions could be drawn, for example, that as analyte concentration increased, response time decreased while sensor recovery times increased. For concentrations between 0 and 400 ppm, the response times varied from 5 to 35 s, while recovery times ranged from 15 to 70 s, depending on the analyte. A comparison between tin oxide, conducting organic polymers and carbon-black polymers found that tin oxide sensors responded more quickly (~7 s) than the other sensor types (20−200 s). Tin oxide sensors were also found to have larger responses in general but were not found to be as good at resolving volatile organic compounds as carbon black−polymer composites.[23]

C. Array Design

The first array developed with tin oxide sensors used commercially available sensor elements (Figure 3).[16] A typical commercially available element has a

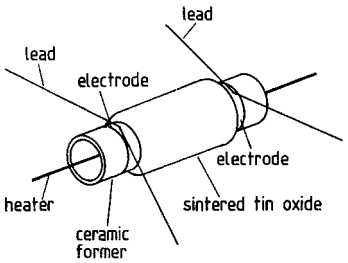


Figure 3. Diagram of the typical commercially available tin oxide gas sensor.

thick film of SnO_2 on a ceramic tube. A heater runs through the center of the tube, and electrical contacts are provided at both ends. Each sensor is individually packaged and is ~1 cm in diameter.[13,17] Commercially available elements are the most widely used SnO_2 array format described in the literature because such sensors are easy to obtain and have a broad range of sensing properties. An example of the Warwick nose, with 12 SnO_2 sensors in the array, is shown in Figure 4.

Tin oxide arrays can also be fabricated on microchips trough silicon-based microfabrication technology to make many identical and miniaturized sensors (Figure 5). Silicon technology allows control over chip array size, power consumption, and sensor reproducibility.[24−26] It also allows for several types of metal oxide layers to be combined within an array.[27,28] In light of the microfabrication technology available, a new design for an array has been developed with built-in circuitry in proximity to each array element. This circuitry performs preprocessing for pattern recognition by using analogue VLSI to translate the sensor inputs to a binary digital output.[29−31] The most recent version has an output generated by using the mean and median values of all sensors, comparing each individual sensor response, and assigning a 1 or 0 corresponding to a high or low value relative to the mean/median. This system was able to discriminate between acetone, butanol, ethanol, methanol, and xylene.[31]

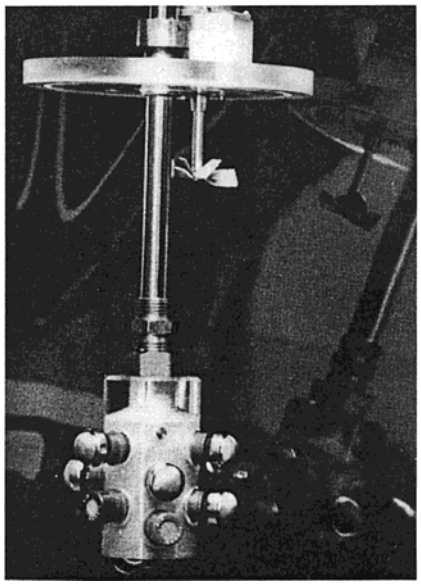


Figure 4. The Warwick Nose, an array of 12 commercially available SnO_2 sensors. Reprinted from *Sens. Actuators B*, Vol. 1, Shurmer, H. V., et al., Intelligent Vapour Discrimination Using a Composite 12-Element Sensor Array, pp 256−260, Copyright 1990, with permission from Elsevier Science.

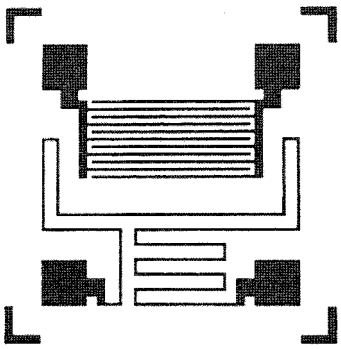


Figure 5. Pattern for interdigitated electrodes (top) and heater (bottom). Reprinted from *Sens. Actuators B*, Vol. 4, Gardner, J. W. et. al., Integrated Tin Oxide Odour Sensors, pp 117−121, Copyright 1991, with permission from Elsevier Science.

One drawback to tin oxide array fabrication is the necessity of incorporating a heating element to operate the array at high temperatures. To avoid this, an array that can operate at room temperature has been demonstrated by treating the sensor elements with an oxygen plasma. When the semiconductor surface is exposed to an oxygen plasma, which has many oxygen species present including both positively and negatively multiple charge species, the SnO_2 surface is covered with a high concentration of oxygen that then reacts with gases and allows sensitivity at room temperature. An untreated array has mostly O^-, O_2^-, and O^{2-}, adsorbed which are mainly unreactive at room temperature. It was found that the plasma-treated array was sensitive to carbon monoxide, carbon tetrachloride, methanol, and light petroleum gas at room temperature, whereas the untreated array was not.[32,33]

D. Applications

Tin oxide sensor arrays have been used to detect a variety of different analytes under varying conditions.

Three major groups of analytes for these arrays are toxic gases,[19,20,32,34,35] volatile organic compounds (VOCs),[16,18,23,30,31,36−42] and food-related species.[22,29,43−53] Toxic gases and VOCs are important components of industrial emissions, and monitoring such analytes can be challenging with a single sensor approach due to drift and interferences. Food analysis benefits from cross-reactive arrays because the technique is noninvasive and does not require breaking down a complex mixture into its individual components merely to establish whether the sample has changed or not from a prior batch.

Work with tin oxide arrays began with VOC's when Persaud and Dodd used compounds such as methanol, ethanol, and diethyl ether to demonstrate that an array of cross-reactive sensors could discriminate between odors without being analyte-specific.[16] Abe showed that with an array of eight sensors, 30 different VOCs could be classified into four categories, ethereal, ethereal-minty, pungent, and ethereal-pungent, on the basis of the odor.[36] Work with volatile organics has continued and is directed toward the more difficult task of analyzing mixtures of organics such as benzene, toluene, octane, and propanol.[37] An array of SnO_2 sensors with a humidity sensor was used to detect and quantify mixtures of methane and ethanol in a variable humidity environment for domestic cooking applications.[54] Arrays for the detection and quantification of sulfides, sulfur oxides, and nitrogen oxides have also been investigated.[34,35]

The use of tin oxide arrays for food and beverage classification is well established. These arrays have been applied to evaluate fish freshness on the basis of various alcohols and biogenic amines given off as fish spoils. A SnO_2 array was shown to have similar sensitivities to fish spoilage as chemical and electrical methods.[49] Coffee beans have been classified on the basis of their aromas, and one study was able to discriminate between different blends of coffee beans as well as to differentiate between various roasting times.[47] SnO_2 arrays have also been used to determine the ripeness of bananas. Llobet et al. were able to classify bananas into seven categories of ripeness on the basis of the headspace sampling of stored bananas.[48]

Various beverages have also been investigated with tin oxide arrays. Di Natale et al. were able to discriminate 1990, 1991, and 1992 vintages of one type of wine.[44] This task was accomplished using a metal oxide-sensing array, with three of the five elements being SnO_2 sensors, and the data were processed with principal component analysis. These workers then used a five-element SnO_2 array to determine which vineyard a particular type of wine had come from, and the method was found to be superior to current standard methods of analysis.[45] In another study, a thermally cycled, eight-sensor array in combination with an artificial neural network was able to discriminate between three different teas.[22]

E. Analysis

Many different parameters affect the outcome of a data analysis process. Some of the key features for tin oxide arrays are the type of signal used for analysis and the normalization process used on these signals before processing and classification is performed. Gardner found that using the fractional conductance gave improved results compared to using the relative or log absolute conductance changes.[55] Other studies have found that using transient signals combined with steady-state signals gives improved discrimination ability in the final processing.[30,39] Preprocessing algorithms can also have important effects on the analysis outcome. For example, by development of a microscopic model of conduction for the array, the appropriate sensor response parameters were selected to improve analyte classification. The normalization procedure developed is shown in eq 3, where S_{ij} is the fractional conductance of the

$$\tilde{S}_{ij} = S_{ij}/(\sum_{i=1}^{n} S_{ij}^{2})^{1/2} = \alpha_{ij}/(\sum_{i=1}^{n} \alpha_{ij}^{2})^{1/2} \qquad (3)$$

ith sensor element to the j_{th} gas and α_{ij} is the ratio of the reaction rates between oxygen desorption and oxygen reaction with gas j. This normalization procedure removes the effect of a change in sign of S_{ij}. A 12-element array was used to distinguish between types of alcohol, including two lagers, two beers, and two spirits. Cluster analysis (CA) of the original data could not separate the lagers from the beers. However, by choice of the appropriate normalization technique, the CA and principal component analysis (PCA) results were enhanced compared to the original results, with lagers and beers being clearly separated.[46]

Chemometric techniques such as principal component analysis, cluster analysis, and least squares methods have been used as techniques to process responses from multianalyte arrays. Using a 12-sensor tin oxide array, discrimination between methanol, ethanol, propan-2-ol, and butan-1-ol was achieved through use of a weighted fault-tolerant least-squares analysis method.[40] A study to evaluate PCA, CA, and partial least squares (PLS) showed that each method was a useful chemometric technique for classifying analytes.[41] Another version of cluster analysis, called transformed cluster analysis (TCA), uses transformation equations based on the mean and variance values for the training data set. This analysis was tested with a four-element metal oxide-sensor array using four individual VOCs, and was found to correctly cluster these analytes.[56] With use of PCA and CA and creation of star symbol plots of sensor array responses to VOCs, it was found that these techniques could help in the development of sensor devices by identifying which sensors provided useful or redundant information.[38,42] The ability to select sensors is useful for creating arrays that can perform specific tasks and also for creating more diverse sensor arrays within the practical constraints of a limited sensor count being available in a fielded device.

A different approach to multianalyte array analysis is to use artificial neural networks (ANN) to perform the data analysis. These systems can handle highly nonlinear data and can deal with noise or sensor drift

with less of a negative impact than classical approaches such as PCA and PLS.[55] As with most data analysis approaches, there are many variations on the ANN theme. Back-propagation networks have been useful to discriminate between similar analytes[39,55] and to analyze gas mixtures containing hydrogen, methane, and carbon monoxide.[57] Self-organizing artificial neural networks (SOM) create unsupervised, statistical descriptions of the environment with no supplementary information. These systems have been shown to perform gas classification well and have been investigated for their ability to learn while operating and thereby counteract sensor drift.[58,59] Fuzzy logic has been incorporated into some neural networks to provide them with the ability to give more than just a yes or no response. This type of neural net in combination with a six element, tin oxide array has been shown to discriminate between carbon monoxide, ethanol, and methane.[60] Fuzzy ARTMAPs seem to be superior to back-propagation networks and learning vector quantization, due to their ability to learn incrementally without forgetting previous information. This feature combined with the ability to classify as well or better than these other techniques could make fuzzy neural nets the best choice for cross-reactive SnO_2 array analysis.[48]

III. MOSFET Arrays

A. Principles of Operation

Metal oxide semiconductor field effect transistors (MOSFET) have been studied as sensors since 1975 when Lundstrom et al. reported a hydrogen-sensitive MOSFET with palladium as the gate metal.[61] These gas sensitive MOSFETs are sometimes referred to as GASFETs. In GASFETs, the structure is the same as a MOSFET, but the traditional aluminum gate is replaced with a catalytically active metal. Since 1975, a significant amount of research has been directed toward finding the best materials for the gate to give better sensitivity and selectivity for a variety of analytes. Several other variations on MOSFETs have been developed, such as open gate field effect transistors (OGFET), ion sensitive field effect transistors (ISFET) (which have an electrolyte solution as the conducting layer), and CHEMFETs (which are ISFETs coated with an organic membrane).[62,63] Although many variations on MOSFETs have been developed, the GASFET seems to be the only type that has been fabricated into cross-reactive arrays.

The MOSFET structure consists of a metal gate on top of an oxide layer, typically SiO_2, and a p-type silicon base with n-doped channels on either side of the gate (Figure 6). The surface potential, φ_s, of the semiconductor layer is dependent on the applied gate voltage V_G and the work function of both the metal, W_m, and semiconductor, W_s, materials.

$$\varphi_s = f_1(V_G - W_{ms}/q) \qquad W_{ms} = W_m - W_s \quad (4)$$

According to eq 4, a change in W_{ms} leads to a change in the surface potential. Therefore, to maintain a constant surface potential, the applied gate voltage must be adjusted in relation to the change in the work function.[62,64]

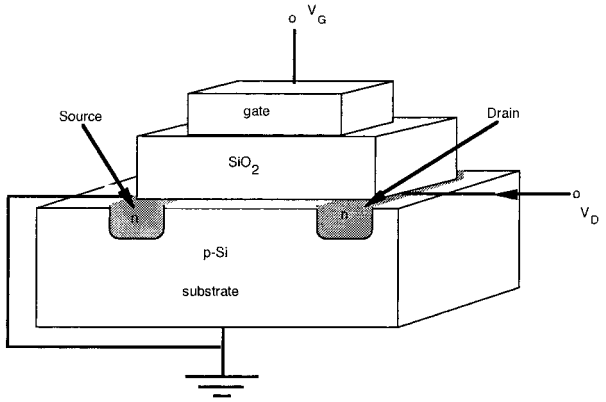


Figure 6. Typical MOSFET structure.

GASFETs operate on the principle that the work functions of the metal and semiconductor are affected by gases adsorbed to the surfaces. The manner in which the W_{ms} is affected depends on whether the metal gate is continuous (thick film) or discontinuous (thin film). On a continuous surface, the metal surface catalyzes dehydrogenation reactions of hydrogen-containing gases using O_2 as the electron acceptor. The hydrogen adsorbs to the metal surface and can diffuse through the metal to the SiO_2/metal interface where a dipole layer is created. The hydrogen dipole layer leads to a change in W_{ms}. On a discontinuous surface, changes are due not only to a hydrogen dipole layer but, in addition, the work function is affected by adsorbates on the metal surface and on exposed portions of the SiO_2 layer.[62,64]

B. Fabrication and Operation

Although many variations on the actual fabrication details for MOSFET sensor arrays are possible, some guidelines are generally followed. Dry oxidation at 1100−1200 °C is first used to create an oxide layer ∼100 nm thick.[65,66] The gate metal can be thermally evaporated onto the oxide surface through a mask, to produce either a layer ∼100−400 nm thick or an ultrathin metal 3−30 nm in thickness.[65,67] The distance between the *n*-regions of the gate is usually 1−10 μm, while the gate depth is between 10 and 100 μm. The actual design of the array can vary significantly, depending on the application. Arrays have been designed with 2−12 MOSFET sensors, and some hybrid arrays have been made with MOSFETs and other sensors (such as tin oxide sensors) combined into the same array.

It is well-known that the type of gate metal as well as the temperature at which the MOSFET is operated will affect the catalytic properties of the sensor. Platinum, palladium, and iridium are the three most common catalytic metals used as gate materials. Pd is a good hydrogen sensor, while Pt and Ir have sensitivities for analytes such as ammonia and ethanol. Modifying the device operating temperature can sometimes enhance the catalytic activity of a sensor toward particular analytes. For example, Pt has a higher affinity for ethylene at 200 °C than at 100 °C (Figure 7).[64]

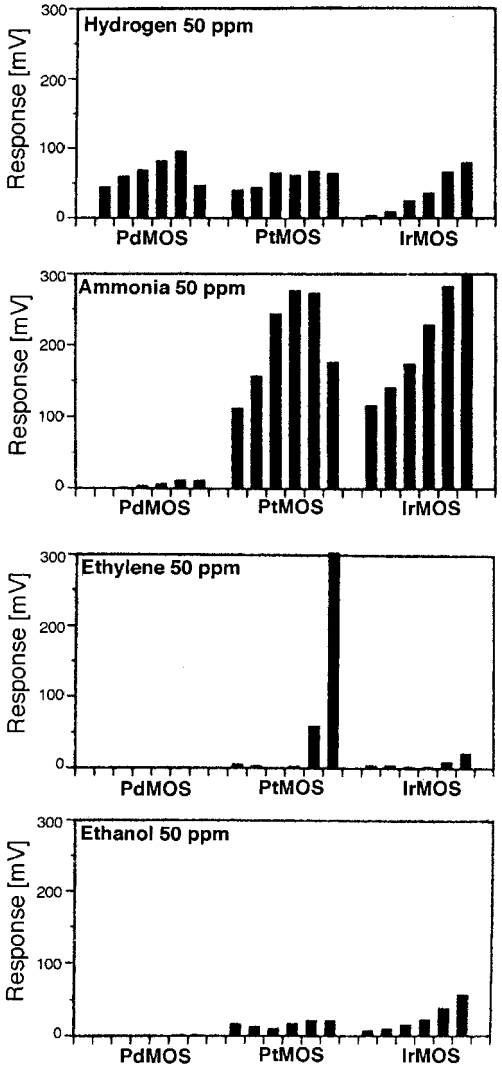


Figure 7. Responses of GASFETs with different gate metals ~100 nm Pd, ~5 nm Pt, and ~5 nm Ir. Each device type was tested with four analytes at 50 ppm in synthetic air and at six temperatures: 75, 100, 125, 150, 175 and 200 °C running from left to right in the diagrams. The responses are the change in gate voltage during a 5 min vapor pulse. Reprinted from ref 64 with kind permission from Kluwer Academic Publishers.

C. Applications

One of the early MOSFET arrays was developed by Muller and Lange and consisted of four palladium gate MOSFET elements. Three of the four elements had zeolite layers with different pore sizes deposited on top of the Pd gate. This array was used to identify hydrogen, methane, and acetylene on the basis of the characteristic changes in capacitance with time. By using correlation coefficients, the three different gases could be classified, regardless of concentration. This approach was only used with single-component gas samples.[66] Muller modified the array to identify binary mixtures of hydrogen and methanol vapors. The cross-sensitivity of a single sensor was compared to the sensitivity of a two element, Pd and Pt MOSFET array. Using the transformed least-squares method, the cross-reactivity patterns of the array could be deconvoluted to determine the concentration of hydrogen within the binary mixture.[68]

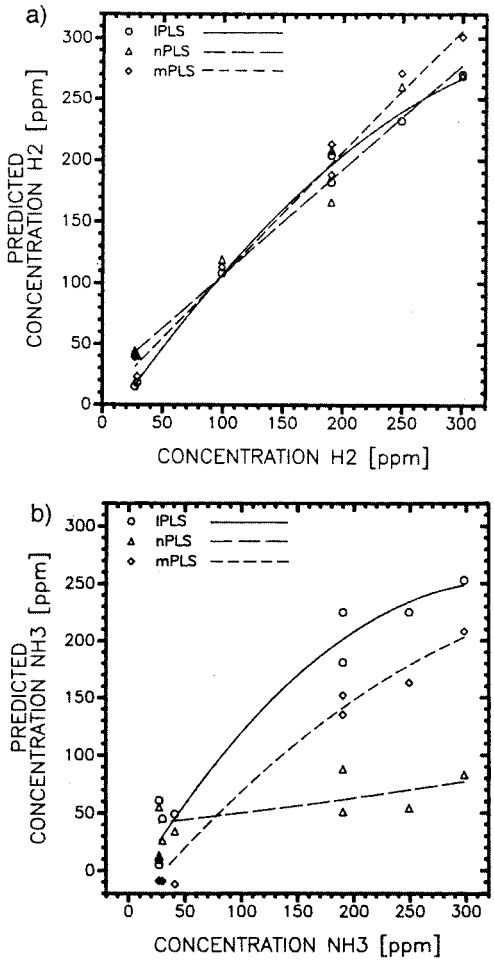


Figure 8. Predicted vs actual concentration of (a) H_2 and (b) NH_3 for the 1PLS, nPLS, and mPLS models. Reprinted from *Sens. Actuators B*, Vol. 2, Sundgren, H., et al., Evaluation of a Multiple Gas Mixture with a Simple MOSFET Gas Sensor Array and Pattern Recognition, pp 115−123, Copyright 1990, with permission from Elsevier Science.

Much of the work with MOSFET cross-reactive arrays has been performed at the Linköping Institute of Technology in Sweden. In 1990, Sundgren et al. showed that an array of commercially available MOSFETs could be used to identify components in a quaternary mixture of gases. The six-element array used in that work consisted of three pairs of Pd and Pt MOSFETs, with each pair operated at different temperatures, 100, 150, and 200 °C. The array was tested with a mixture of hydrogen, ammonia, ethylene, and ethanol in a carrier gas of air and argon. The change in the gate voltage due to exposure of the array to the gas mixtures was monitored over time. The three characteristics of this curve used for analysis were the difference between the pre- and post-exposure values of the gate voltage, the maximum positive derivative, and the maximum negative derivative of the signal with time. Data were analyzed by using three different versions of partial least squares. Using these methods, the concentration of hydrogen and ammonia in the gas mixtures could be predicted, while ethanol and ethylene could not be separated.[67] Figure 8 shows the predicted vs actual concentrations of hydrogen and ammonia with

the three versions of PLS. A 45° line through the origin represents a perfect prediction.

An essential aspect to any type of cross-reactive array is the pattern recognition method employed to determine the composition of the analyte. One study by Sundgren et al. compared the analysis of two types of MOSFET arrays with PLS and an artificial neural network (ANN). The first array was identical to the previously described array[67] and was tested with a mixture of hydrogen, ammonia, ethanol, and ethylene. PLS was able to predict the concentrations of hydrogen and ammonia. ANN could not achieve accurate predictions for ethanol or ethylene either but made closer predictions to the true concentrations of hydrogen and ammonia than PLS. The second test was to use a six-element array consisting of three types of gate materials, Pt, Pd, and Ir. This array was tested with a binary mixture of hydrogen and acetone. Both PLS and ANN were able to make reasonable predictions of the hydrogen and acetone concentrations in these mixtures, with ANN again giving the more accurate prediction. Although ANN was more accurate than PLS in both tests, the authors warned that this might not be the general case.[69]

One other data analysis technique was used with the previously described two type array data.[69] An adaptive learning network using the abductory induction mechanism (AIM) was used to predict the hydrogen and ammonia concentrations. This method was found to have similar prediction capabilities as PLS and ANN, but it was unable to predict the concentration of ethanol and ethylene. This software is relatively fast because it develops the networks in a feed-forward fashion that allows for faster learning than feed-backward ANNs. AIM was found easier to use than PLS or ANN to ascertain which sensor information is being used for each analyte. Since each analyte has its own network, this method of analysis is well suited for analyzing single components of a mixture.[70] A later comparison of ANN and AIM points out that while AIM is faster, ANN can provide a more accurate prediction over a wider concentration range.[71]

D. Alternative Implementations of MOSFETs

Lundstrom et al. have also taken a completely different approach to MOSFET arrays by employing a method called the scanning light-pulse technique.[64,65,72,73] The premise of this approach is that when light is shined on the surface of a MOSFET coated with a thin metal film, the light will penetrate the metal and induce a photocapacitive current (i_{ph}) in the semiconductor. The depletion layer determines this current similar to the surface potential. To maintain a constant i_{ph}, the applied gate voltage must be varied in response to changes in the W_{ms}, just as it would to maintain a constant surface potential. Therefore, when the change in the gate voltage is monitored a map of $\Delta V(x, y)$ over the sensing surface can be obtained by taking the difference between the gate voltage in air vs in gas at the same i_{ph}. In this approach, the MOSFET array employed three continuous strips of Pt, Pd, and Ir along a 4 mm × 6 mm rectangle, instead of separately defined sensor elements. The sensor surface was divided into an 18 × 18 grid, and a temperature gradient (110−180 °C) was established down the length of the sensor surface. This temperature variation allowed for different sensitivity and selectivity at each point of the sensor grid. Hydrogen, ammonia, and ethanol were tested individually as mixtures with air. The change in voltage (ΔV) was determined for each point of the sensor grid, and the plot of ΔV for each gas was then processed by taking the average of the ΔV for each point and its eight nearest neighbors. An average higher than 65 mV was assigned white, and lower averages were assigned black. In this manner, three distinct image maps of the gases were created (Figure 9). The possible uses for such image mapping include gas mixture identification, investigation of new sensing materials simultaneously with the same gas sample, and mapping spatially inhomogeneous reactions.[64,65]

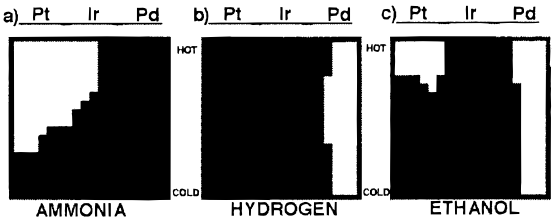


Figure 9. Image maps of (a) ammonia, (b) hydrogen, and (c) ethanol. Reprinted from *Sens. Actuators B*, Vol. 6, Winquist, F., et al., Visual Images of Gas Mixtures Produced with Field-Effect Structures, pp 157−161, Copyright 1992, with permission from Elsevier Science.

Light-pulsed sensing combines many types of information, including the catalytic activity of the gate metals, gas flow turbulence, edge effects, etc. Distributed chemical sensing was inspired by the light-pulsed technique but sought to separate only the effects of catalysis. A cell was designed with a continuous catalytic surface of Pd, with seven MOSFET sensors spaced 2.8 mm apart along the length of the cell. The top and/or bottom of the cell were coated with thick Pd layers and served as continuous catalytic surfaces along the length of the cell. Hydrogen and ethanol were mixed with O_2 and N_2 and were passed through the cell at a constant flow of 10 mL min^{-1}. The monitored response was the change in gate voltage prior to gas exposure and at the end of a gas exposure. Ethanol and hydrogen could be discriminated on the basis of their different catalytic profiles by using information only from the first and last sensors in the cell.[74] A recent paper on distributed chemical sensing utilized two types of MOSFETs and two catalyst surfaces for a total of four different sensor/catalyst combinations. Each array had five identical MOSFET sensors evenly distributed along the length of the cell. The four different array combinations were tested with a quaternary mixture of hydrogen, ammonia, ethyne, and ethanol. Information from all four array types were combined and analyzed with PCA and ANN, and used to estimate all four components of the gas mixture. It was also found that using information from various positions in the cell afforded improved results compared to using only the first and last sensors of the arrays.[75]

E. Hybrid Metal Oxide Containing Arrays

Since the early 1990s, interest has been increasing in creating hybrid arrays, or arrays that combine more than one type of sensing element. Some of these hybrid arrays are already commercially available as electronic noses. Alpha MOS sells The Fox system which is a hybrid array of tin oxide, conducting polymer, and thickness shear mode (TSM) sensors.[76] Another nose available is the eNOSE 5000 from EEV Limited, which consists of conducting polymer, metal oxide, and TSM sensors.[77] In the semiconductor-based sensing area, combining MOSFETs and tin oxide sensors has been especially popular. A common hybrid array setup is using ten MOSFETs, four tin oxide, and one CO_2 sensor. At Linköping University these types of arrays have been used to investigate various food products. For example, these arrays were used to both estimate the storage time of ground pork and beef and to tell these two meats apart.[50,52] Ninety percent of the wheat, barley, and oat samples tested were correctly classified as good or bad based on the mold/musty smells of the grains. This classification rate exceeded the agreement of the two grain inspectors who originally classified the grains.[53] They have also used this array to classify five different types of cardboard, although the best discrimination was found using only 4 of the 15 sensor elements.[78] Recently, this 15-element hybrid array format was used to monitor the different stages of recombinant bioprocesses. The array followed the cultivation of mammalian cells secreting the blood coagulant factor VIII. In combination with PCA, four different stages of the process were determined, being cell growth, draining the reactor, growth medium exchange, and production of blood coagulant.[79]

A slightly smaller hybrid array of eight FETs and four tin oxide sensors combined with a neural network was able to identify 9 out of 10 different types of cheese correctly.[52] A modular approach to hybrid arrays was presented by Ulmer et al. and utilized a variety of sensors such as MOSFET, tin oxide, TSM, and electrochemical sensors. This system was demonstrated to discriminate between different olive oils, whiskeys, tobaccos, coffees, and plastics.[51] A recent paper from the Linköping University has utilized a hybrid array in parallel with a conductive polymer array. A 32-element conductive polymer array was used in parallel with a hybrid array comprised of 10 MOSFETs and 6 tin oxide sensors. This parallel array was used to predict the fermentation of wood hydrolysates and to estimate quantities of components of the hydrolysates. Although some success was demonstrated, the problem was nontrivial.[80]

IV. Intrinsically Conductive Polymer Chemiresistor Arrays

A. Principles of Operation

The experimental and theoretical behaviors of intrinsically conducting polymers (ICP's) have been discussed in several reviews and books,[81−84] so only the basic properties of ICP's that are required to understand the operation of ICP-based sensor arrays are discussed in this section. Typically, the fundamental structural unit of an ICP is a linear backbone comprised of repeating conjugated organic monomers such as acetylene, pyrrole, thiophene, or aniline (Figure 10).

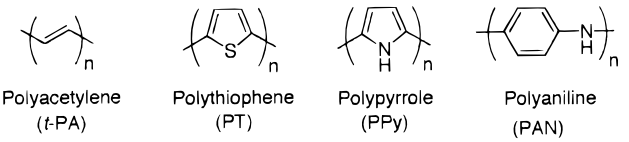


Figure 10. Structure of four intrinsically conductive polymers in their insulating form.

These materials are insulating in their neutral state. However, chemical or electrochemical reduction (n-doping) or oxidation (p-doping) of these materials renders the polymers electrically conductive. The conductivity is produced through band structure transformation and/or the generation of charge carriers. A vast majority of ICPs act as one-dimensional conductors because the electrons in an n-type ICP (or holes in a p-type ICP) mainly travel through the linear conjugated chains. Formally, the materials are usually electrical semiconductors as opposed to metallic conductors because the materials have an electronic energy state band gap at room temperature between their valence and conduction bands so that their intrinsic electrical conductivity decreases as the temperature is lowered.

ICPs are useful as chemical sensors because the electrical properties, typically the dc electrical resistance of these systems, are responsive to the presence of a diverse set of analytes in the vapor phase. Sorption of a vapor into an ICP will induce physical swelling of the material and will affect the electron density on the polymeric chains. The change in dc conductivity, $\Delta\sigma$, that results from sorption of an analyte can be divided conceptually into three components:

$$\Delta\sigma = (\Delta\sigma_c^{-1} + \Delta\sigma_h^{-1} + \Delta\sigma_i^{-1})^{-1} \qquad (5)$$

Here, $\Delta\sigma_c$ is the overall change in intrachain conductivity of the ICP, $\Delta\sigma_h$ is the change in intermolecular conduction due to electron hopping across polymer chains modulated by the presence of the analyte in the film, and $\Delta\sigma_i$ is the change in ionic conductivity between chains upon analyte sorption. The value of $\Delta\sigma_i$ is not only a function of the ion migration upon condensation of analyte but also is a function of proton tunneling rates if hydrogen bonding to the ICP backbone is significant.[85] These conductivity changes may, or may not, be linearly dependent on the concentration of analyte presented to the sensor, depending on the particular transduction mechanism involved in the ICP of concern.

The pathway for conduction through the conductive polymer backbone is more favorable energetically than across polymer backbones. Therefore, large changes in conductivity can be attributed to changes in the extrinsic conductivity of the material. For

example, highly oriented polyaniline is more conductive than amorphous polyaniline. Miller et al. have reported that a large increase in conductivity was observed upon exposure of a dinonylnaphthalenesulfonic acid doped polyaniline film to ethanol vapor.[86] Using both electron microscopy and wide-angle X-ray scattering, these workers were able to attribute this response to the crystallization of polyaniline. Hence, the hydrogen-bonding interactions between ethanol and polyaniline were sufficient to bring the polyaniline chains closer together, thereby decreasing the energy barrier for electron hopping and increasing the mobility of the charge carrier.

Intrinsically conducting polymers possess several potential advantageous features for use in sensor arrays for vapor detection. A diverse range of polymers is available via the electrochemical or chemical polymerization of a range of monomer types. The sensitivity of a sensor element can be readily altered by changing the polymerization conditions and the charge compensating counterion. The oxidation state of the polymer can be altered after deposition in order to tune the electronics of the polymer to be compatible with the analyte of interest such that an optimum charge transfer interaction may take place. Fabrication is easily performed through either electrochemical deposition and/or film casting from solution (for soluble analogues), which readily allows for miniaturization and mass production of sensors. Biomaterials such as enzymes, antibodies, and cells may be readily incorporated into such structures, and finally (but not exhaustively) reversible responses are typically obtained in relatively short time periods at ambient temperatures. Drawbacks are the relatively poorly understood signal transduction mechanism, high sensitivity to changes in humidity that necessitate operation in a controlled, conditioned, environment, and a relatively low diversity in affinity of the conducting polymers toward a very diverse set of analytes, thereby producing less than optimal separation of many analytes in the sensor array data space.

The response commonly reported for ICP sensors is the normalized change in resistance obtained from a two-probe resistance measurement, $(R_2 - R_1)/R_1$, where R_1 is the resistance before exposure to the analyte of interest and R_2 is the resistance of the sensor at steady-state conditions upon exposure to the analyte. The response has also been reported as either the change in resistance (2 probe) or the change in conductivity (4 probe).

ICPs have reached the commercial market as detectors in "electronic" noses (sensor arrays) sold by companies such as AromaScan,[87] Bloodhound,[88] AlphaMOS,[76] and Zellweger Analytics.[89] The AromaScan A32/50S Multisampler consists of an array of 32 conducting polymer detectors and is based on research initiated at the University of Manchester Institute of Science & Technology (UMIST) by Krishna Persaud. The Bloodhound BH114 consists of an array of conducting polymer and discotic liquid crystal detectors and is a product of research efforts at the University of Leeds. AlphaMOS has a commercial nose on the market that combines sensor technologies using arrays of conductive polymer resistive detectors, metal oxide semiconductor detectors, and TSM resonator detectors, each fulfilling a specific duty. The role of the conducting polymer array is to detect polar molecules, whereas the metal oxide semiconductor detectors are sensitive to nonpolar molecules and are insensitive to water. Finally, Zellweger Analytics, which purchased Neotronics in 1996, sells a portable version of its electronic nose for environmental monitoring applications.

B. Array Fabrication

Conductive polymer sensors are typically prepared electrochemically.[90−97] The electrodes are typically interdigitated electrodes or are a pair of metal leads separated by approximately 10−50 μm. The metal is typically Au and the substrate is an insulator, such as glass.[97]

The electrochemical deposition of ICPs is reasonably controllable because films can be electrodeposited onto metallized areas and the film thickness can be varied by monitoring the total charge passed during the deposition process. Polymers can be deposited using a number of different electrochemical waveforms, for example, potentiostatically, galvanostatically, cycled potential, or by pulsed potentiometry. The polymerization is usually carried out in a three-electrode configuration in which the working electrode is the sensor substrate. A positive oxidation potential is often applied to the working electrode that is sufficient to oxidize the monomer at which time the monomer undergoes oligomerization to a point at which the oligomers electroprecipitate onto the electrode surface. As the material grows it bridges the gap between interdigitated leads, thus making a simple resistor. Measurement of the total charge passed gives an approximation of the film thickness, and the final potential applied to the material will control the extent of doping. In addition to this technique, many conducting polymers substituted with pendant solubilizing groups are soluble in solvents that are normally employed for spin coating.[98−100]

Several strategies have been employed to create an array of ICP detectors with each detector having a different response toward a variety of analytes. First, many different single-ring heterocycles and multiring fused/unfused heterocycles readily undergo either electrochemical or chemical polymerization and can be used as ICP elements. Some of these include pyrrole, thiophene, aniline, indole, and carbazole. Second, each heterocycle can be substituted with a number of different side groups. Third, different counterions can be used to compensate the positive charge (for a p-doped polymer) generated upon oxidation of the ICP to produce the electrically conductive state of the polymer. Fourth, different oxidation states are attainable in each ICP and different polymerization conditions (i.e., by changing the oxidation potential, oxidant, temperature, solvent, electrolyte concentration, monomer concentration, etc.) can be used to generate a number of different sensors from the same monomer, because properties such as

Table 1. Twelve Conductive Polymers That Constitute an ICP Sensor Array That Were Prepared by Electrochemical Polymerization Using a Conventional 3-Electrode Electrochemical Cell[a]

sensor no.	polymer system	monomer concn (mol/dm^{-3})	electrolyte concn (mol dm^{-3})	solvent	growth potential (V)	final potential (V)	resistance (Ω)
1	PPy-BSA	Py 0.1	0.1	H_2O	0.85	0.00	1650
2	PPy-PSA	Py 0.1	0.1	H_2O	0.85	0.00	193
3	PPy-HxSA	Py 0.1	0.1	H_2O	0.85	0.00	27
4	PPy-HpSA	Py 0.1	0.1	H_2O	0.85	0.00	16
5	PPy-OSA	Py 0.1	0.1	H_2O	0.85	0.00	35
6	PPy-DSA	Py 0.1	0.1	H_2O	0.85	0.00	37
7	PPy-TSA (Na)	Py 0.1	0.1	H_2O	0.80	0.80	19
8	PPy-TSA (m)	Py 0.1	0.1	EtOH	1.20	0.00	70
9	PPy-TEATS	Py 0.1	0.1	H_2O	0.75	0.00	34
10	PPy-TEATS	Py 0.1	0.1	PC	1.10	0.00	37
11	PAN−$NaHSO_4$	AN 0.44	0.1	H_2O	0.90	0.90	44
12	P3MT-TEATFB	3MT 0.1	0.1	CH_3CN	1.65	1.65	13

[a] Reproduced from ref 101. The distance between gold electrodes was 15 μm. PPy is polypyrrole, BSA is butanesulfonic acid, HXSA is hexanesulfonic acid, HPSA is heptane sulfonic acid, OSA is octane sulfonic acid, DSA is decanesulfonic acid, TSA (Na) is *p*-toluenesulfonic acid sodium salt, TSA (m) is *p*-toluenesulfonic acid monohydrate, TEATs is tetraethylammonium toluenesulfonate, PAN is polyaniline, $NaHSO_4$ is sodium hydrogen sulfate, P3MT is poly(3-methylthiophene), TEATFB is tetraethylammonium tetrafluoroborate, Py is pyrrole, AN is aniline, 3MT is 3-methylthiophene, EtOH is ethanol, PC is propylene carbonate, and CH_3CN is acetonitrile.

morphology, molecular weight (conjugation length), connectivity of monomers, conductivity, band gap, etc., are dependent upon the polymerization conditions.

Polypyrrole is the predominant polymer in the Neotronics NOSE. Chemical variation is achieved by either changing the dopant ion or changing the polymerization conditions. The use of 12 different polypyrrole-based sensors produces classification between chemically similar samples over a wide range of analytes.[94] In a separate paper, Pearce et al. constructed a 12-element electronic nose.[101] Table 1 lists the 12 different sensors and the electrochemical conditions used to prepare them.

Ten of the elements consisted of polypyrrole prepared with different counterions and polymerization conditions while the eleventh and twelfth elements were a polyaniline and a poly(3-hexylthiophene) film, respectively. Hatfield et al. have reported the construction of a 20-element sensor array composed of derivatized polypyrroles, thiophenes and other heterocyclic polymers.[102] Serra et al. have reported the differences in sensitivities of various analytes to three different poly(3,3′-dipentoxy-2,2′-bithiophene) sensors.[103] Chemical variation in these films was achieved by performing the vapor phase polymerization of the monomer in the presence of different oxidizing salts. Serra et al. conclude that, under different oxidizing conditions, different sensors are prepared with sensitivities (percentage variation of the resistance) ranging from −6.22 to 1.77.

Jahnke et al. utilized an array of poly(2,5-furylene vinylene) sensors to construct an electronic nose[85] using a soluble precursor route involving an aldol addition reaction of 5-methylfuran-2-carbaldehyde with a basic catalyst to obtain the precursor poly(2,5-furylenehydroxyethylene). The polymers were then spin-coated onto the sensor substrate and thermally dehydrated to varying degrees, thus making different sensors by the amount of elimination that occurred. In a similar approach, De Wit et al. prepared an array of poly(2,5-thenylene vinylene) copolymer conductometric sensors.[98] Preparation involved making a sulfonium precursor polymer, dissolving this polymer in acetone, and spin casting it onto the substrate. Thermal elimination and chemical doping yielded the conductive PTV. The authors prepared both the unsubstituted PTV and a methoxy-substituted derivative. Thermal elimination to varying degrees created chemically different sensors in that the conjugation lengths of the conductive segments were different. Sensors were reported to exhibit an extreme baseline drift that rendered them inoperable after ca. 1 month, although they could be regenerated by redoping the polymers with iodine vapor. The responses were reported to be linear with concentration between 5% and 100% of the saturated vapor pressure for a series of nine test vapors (toluene, water, propanol, acetone, acetic acid, diethyl ether, ethyl acetate, methanol, and ethanol). In using both the relative response and the recovery time response, a single detector was able to distinguish between six of the nine analytes.

C. Applications

One of the more prevalent fields of application presented in the literature for the ICP electronic noses is monitoring the quality of foods and beverages.[104] A nose consisting of four different polythiophene sensor elements was used in conjunction with a taste sensor consisting of eight different polymer/lipid membranes. Using principal component analysis, four different wines were correctly identified and discrimination was successfully performed among five aged samples of the same wine. The authors concluded that sensor fusion of both the smell and taste sensor arrays led to enhanced discrimination.[99]

Numerous papers have been published on using the electronic nose to monitor the quality of beer.[101,105,106] In a detailed analysis performed by Pearce and Gardner, an array of 21 conducting polymers was used to predict organoleptic scores. In this study, a control lager beer was spiked with different reference compounds (diacetyl, dimethyl sulfide, and hop essence). For high concentrations of the reference compound in beer (>40 ppb), the

electronic nose performed worse than both gas chromatography and flavor profile analysis; whereas, at low concentrations, the accuracies were reported to be similar.[107]

The Aroma Scanner has been reported to be an efficient instrument which complements gas chromatography/mass spectrometry in the analysis of volatiles from both earth-almond and carob, which were both analyzed raw and roasted for different periods of time.[108] The correlation between the sensors and the GC/MS was reported to be 0.98 for earth-almond; whereas a lower correlation of 0.74 was obtained for carob. Schaller et al. reported that an electronic nose consisting of ICPs exhibited poor sensitivity to the volatile components of cheese with the main problem being attributed to sensor drift.[109] The Neotronics e-nose has been successfully applied to discriminate between different levels of boar taint.[110]

Many microorganisms are known to expel volatile chemicals. Work at the University of Leeds has led to both the detection and discrimination of 13 different types of microorganisms utilizing a 16-component sensor array.[111] Through use of a three layer perceptron network on data arising from 244 sample responses, a majority of the microorganisms were reported to be classified with a quoted success rate of 100%. Arnold et al. have investigated the use of an artificial nose to assess bacteria isolated from processed poultry.[112]

Conducting polymer-based electronic nose technology also has been reported to offer a rapid, reproducible, and objective method for sewage odor assessment.[113−116] In averaging all of the data from 10 different sewage treatment plants, the authors reported that at low concentration ranges (below 4000 odor units (ou)/m^3) the Neotronics eNOSE output correlated well with olfactometry data.[113] However, this was not the case at higher concentrations, due to saturation of the polypyrrole sensors. The data from a single sewage treatment plant exhibited a relationship between the electronic nose responses and the threshold odor number (TON) over a range of concentrations from 125 to 781 066 ou/m^3. Further work was carried out by Stuetz et al. using the eNOSE[114] in which canonical discriminant analysis was used to distinguish between different sewage samples from different treatment works. In this study, linear correlations were established between the electronic nose responses and the 5 day biochemical oxygen demand (BOD5) over time intervals of about 1 month. This array could be used for measuring chemical activity, making it a useful process control since it has been estimated that about 40% of the energy could be saved by using an on-line aeration control system. The e-NOSE has also been used to discriminate between untainted water samples and water tainted with geosmin, methylisoborneol, 2-chlorophenol, phenol, diesel, and 2-chloro-6-methylphenol.[115]

Researchers have utilized the electronic nose to monitor air quality in relation to the assessment of malodor in agriculture. Fresh liquid pig slurry was analyzed using GC/MS to identify the major components in the vapor phase. A 20-element ICP sensor array exhibited a linear response to methanol vapor, whereas nonlinear behavior was reported for acetic acid. Using Sammon mapping, each of the individual components and both basic and acidic pig slurry were discriminated.[117] Byun et al. have reported that the best way to visualize the data obtained from pig slurry is to combine both principal component analysis and Sammon mapping.[118] Furthermore, the AromaScanner has also been used to assess malodor concentration after the application of cattle slurry to grassland.[119] Masila et al. have demonstrated the usefulness of the Aroma Scanner to detect environmentally unfriendly halogenated organic compounds such as 2,4,5-trichlorophenol, pentachlorophenol, 1,1-bis(4-chlorophenyl)-2,2,2-trichloroethane (DDT), and 2,2-bis(4-chlorophenyl)-1,1-dichloroethylene (DDE), just to name a few. Using cluster and Euclidean distance analysis, both the identification and the quantification of these organic compounds were reported.[120]

The electronic nose has been successfully applied to several medical and veterinary science applications. To increase the likelihood of successful artificial insemination in cows it is beneficial to detect the occurrence of estrus. Lane et al. have been able to utilize the Neotronics eNose to associate odors with the estrus period.[121] For five Holstein-Friesian cows, both the luteal phase and the estrus phase were discriminated using principal components analysis.[121] The AromaScanner was used to identify the source of an off-odor from a pharmaceutical inhalant.[122] Furthermore, the AromaScanner was used to screen for bacterial vaginosis.[123] Following training of the nose to four positive and four negative cases, 16 of the 17 cases were recognized as being positive. Of the 43 negative cases, 33 were correctly identified as being negative. Thus, the positive predictive value of the test was calculated to be 61.5%.[123]

Recently, a conducting polymer sensor array has been used for monitoring the environment of a confined system. An artificial nose consisting of an array of 20 conducting polymer (UMIST) and 6 TSM resonators (HKR Sensor Systems) was utilized in the MIR-95 mission and was reactivated in the DARA-MIR 97 mission.[124] In this work, it was concluded that the artificial nose was useful in detecting a fluid leak of ethylene glycol in the cooling system, which ultimately led to contamination of the water supply. The sensor array was also was useful in monitoring the atmospheric quality after a small fire. Furthermore, it was reported that after 1.5 years in the MIR space station, the system showed little drift or degradation.[124]

D. Future Directions

Amrani et al. have carried out numerous studies investigating the use of impedance techniques to increase both the sensitivity and selectivity of ICP sensor arrays.[125−129] In this work, the authors report both the relative changes in resistance and the relative changes in reactance upon sweeping the measurement frequency for each sensor in a 20-element ICP array. Both the resistance and reactance

were reported to give a linear regression when the relative values versus the concentration of the analyte were plotted. The specificity of the array was enhanced when the dissipation factor as a function of frequency were plotted.[129]

The dissipation factor can be expressed as the ratio of the energy dissipated (purely resistive) per cycle divided by the energy stored (purely reactive) per cycle. Therefore, it was calculated as the ratio of the resistive part of the impedance divided by the reactive part of the impedance of the sensor. When the dissipation factor vs the applied frequency was plotted characteristic resonances were observed for a single sensor exposed to air, acetone, methanol, and ethyl acetate. Not only was excellent discrimination achieved but these resonances were claimed to be useful to identify components of mixtures of analytes.[129] Furthermore, in using alternating current (ac) techniques, sensitivity to a particular analyte was increased 18-fold. The authors noted that the impedance analyzer is not practical for real applications due to the time that it would take to make a high-resolution scan across a range of frequencies.[129] They suggested that a way to overcome this drawback would be to employ periodic signals containing many different frequencies such that all of the frequencies are presented to the sensor at once. Furthermore, it was suggested that Fourier transform techniques could be used to isolate the response of a sensor at any number of specific frequencies.[129]

V. Conductive Polymer Composite Chemiresistor Arrays

A. Theory

Electronic noses based on an array of polymer composite sensors have been developed using either carbon black[130] or polypyrrole[131] as a conductive filler and nonconductive organic polymers as the insulating matrix. The transduction mechanism for the polypyrrole composites is potentially quite complex in that the analyte can interact with both the insulating matrix and the conductive polymer; whereas the transduction mechanism of the carbon black composite sensors has been described on the basis of percolation theory. Upon exposure to an odorant, the composites will swell to varying degrees depending on the polymer−odorant interactions, and this swelling results in a change in the conductivity of the composite film. This response can be easily monitored using a conventional ohmmeter. Therefore, each sensor element in the array, consisting of a chemically unique insulating matrix, responds differently to a given odorant, resulting in a distinctive pattern. For example, carbon black composite sensor arrays consisting of 17 sensor elements can easily distinguish between a chemically diverse set of analytes and mixtures consisting of two chemically similar analytes.

Percolation theory[132,133] predicts that the resistivity, ρ, of a conducting composite will be given by

$$\rho = \frac{(z-2)\rho_c\rho_m}{A + B + [(A+B)^2 + 2(z-2)\rho_c\rho_m]^{1/2}} \quad (6)$$

where

$$A = \rho_c[-1 + (z/2)[1 - (\nu_c/f)]] \quad (7)$$

$$B = \rho_m[(z\nu_c/2f) - 1] \quad (8)$$

and where ρ_c is the resistivity of the conductive filler, ρ_m is the resistivity of the insulating matrix, ν_c is the volume fraction of the conductive filler in the composite, z is the coordination number of the conductive filler particles, and f is their total packing fraction ($\nu_c < f$). Since these sensors are believed to operate by this percolation conduction, it is possible to fabricate high gain sensors by working close to the percolation threshold, ν_p, which is given by $2f/z$, such that a small volume change would induce a large resistance change.

B. Fabrication

Sensor fabrication entails casting a thin film of the composite over two electrical leads. For polypyrrole-based sensors this was achieved by chemically polymerizing pyrrole using phosphomolybdic acid in a solution containing the insulating polymer. The solution was then used to dip coat the substrate, a cut 22 μF capacitor (interdigitated electrode pair with a lead separation of 15 μm).[131] Carbon black composite sensors were prepared simply by either dip coating an interdigitated electrode capacitor or by either dip coating or spin casting onto a glass substrate containing two gold leads separated by a 5 mm spacing.[130]

Lonergan et al. prepared a diverse carbon black-based conducting polymer composite sensor array consisting of 17 chemically different insulating polymers as listed in Table 2.[130] Responses were typically measured as the maximum relative differential resistance $(R_2 - R_1)/R_1$, where R_1 is the baseline resistance of the film prior to exposure and R_2 is the maximum resistance upon exposure to the analyte. Each of the analyte exposures produced an increase in the resistance of the sensor, which is consistent with percolation theory. However, this was not always the case with polypyrrole composite sensors, indicating a more complex transduction mechanism that involves both the swelling of the insulating matrix and/or the polypyrrole or specific electronic interactions of the analyte with the conducting polymer, which would change its intrinsic conductivity.[96]

The normalized, relative differential resistance data for a set of carbon black sensors produced a different pattern upon exposure to nine different analytes as would be expected by the chemical nature of both the insulating phase of the composite and the analyte. For instance, sensor 16 (poly(ethylene-*co*-vinyl acetate)/carbon black) is hydrophobic and thereby responds best to benzene, whereas sensor 6 (poly(N-vinylpyrrolidone)/carbon black) is hydrophilic and responds very well to methanol. Principal component

Table 2. Materials Employed in the Carbon Black-Based Conducting Polymer Composite Sensor Array As Seen in Ref 130

sensor no.	polymer	T_g[a]	T_m[b]	δ[c]
1	poly(4-vinyl phenol)			
2	poly(styrene-*co*-allyl alcohol), 5.7% hydroxyl			
3	poly(α-methylstyrene)	49		
4	poly(vinyl chloride-*co*-vinyl acetate), 10% vinyl acetate			
5	poly(vinyl acetate)	30		9.35
6	poly(*N*-vinylpyrrolidone)	175		
7	poly(carbonate bisphenol A)	150	267	
8	poly(styrene)	100	237.5	9.1
9	poly(styrene-*co*-maleic anhydride), 50% styrene			
10	poly(sulfone)	190		
11	poly(methyl methacrylate)	105		9.3
12	poly(methyl vinyl ether-*co*-maleic anhydride), 50% maleic anhydride			
13	poly(vinyl butyral)	51		
14	poly(vinylidene chloride-*co*-acrylonitrile), 80% vinylidene chloride			
15	poly(caprolactone)	−60	60	
16	poly(ethylene-*co*-vinyl acetate), 82% ethylene			
17	poly(ethylene oxide)	−67	65	9.9

[a] Glass transition temperature (°C). [b] Melting temperature (°C). [c] Solubility parameter $(cal/cm^3)^{1/2}$.

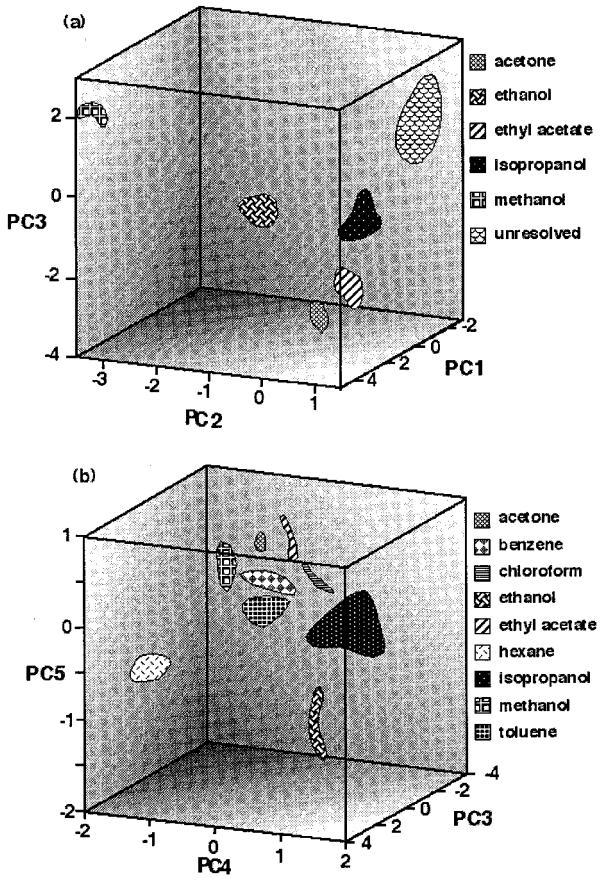


Figure 11. Results from the exposure of the 17-element array to nine solvents as represented in (a) the first three dimensions of principal component space and (b) the third, fourth, and fifth dimensions of principal component space. These five principal components contain over 98% of the total variance in the data. Reprinted from ref 130. Copyright 1996 American Chemical Society.

analysis was used to map the odorant patterns. The authors reported that the first five principal components contained greater than 98% of the total variance in the data. PCs 1−3 and 3−5 for the array are shown in Figure 11. This array distinguishes between a number of different classes of analytes and was able to distinguish between a homologous series of alcohols as shown in Figure 11 by the separation that was achieved for methanol, ethanol, and propanol vapors.

The carbon black-insulating polymer composite sensor array was exposed to a mixture of methanol and ethanol. The sensors in the array were found to give a linear plot of the response vs concentration for both of the pure components. Separate exposures to different compositions of the ethanol/methanol mixture yielded distinct pseudolinear paths between the responses of the pure component in principal component space. Therefore, the sensor array has demonstrated the potential to determine the absolute concentrations of a binary mixture.

One method used to increase the diversity of an array was to use compatible polymer blends as the insulating matrix.[134] A series of different sensors consisting of a poly(methyl methacrylate)/poly(vinyl acetate) blend in a number of different ratios and of both the pure polymers were studied. A plot of the detector response vs the mole fraction of poly(vinyl acetate) yielded nonlinear behavior for five different analytes. The compatible blend detectors clearly demonstrated an enhancement in maximizing the discrimination between analytes. Severin et al. have demonstrated that the incorporation of the chiral polymer poly((*R*)-3-hydroxybutyrate-*co*-(*R*)-3-hydroxyvalerate) into a carbon black composite detector allowed for the differentiation between enantiomeric odorants.[135] Doleman et al. have shown that the response intensity of a carbon black composite sensor is, to first order, determined by the thermodynamic activity effects that dictate the concentration of the analyte in the polymer matrix.[136] When the concentration of the analyte in the gas phase was maintained at a constant fraction of its vapor pressure, the mean response intensity taken as an average across a 13-element detector array remained constant for both a homologous series of *n*-alkanes and a homologous series of *n*-alcohols. This trend has been observed in human detection threshold judgments for the same homologous series of alkanes and alcohols.

C. Applications

Doleman et al. evaluated the performance tasks for three separate arrays of different conductometric sensors.[23] In this work they compared the resolving power of an array of intrinsically conducting polymers, an array of tin oxide detectors, and an array of carbon black conductive composite detectors by calculating the resolution factor according to the Fisher linear discriminant method for pairwise combinations of 19 different solvents. In general, it was shown that for all three sensor arrays that the array performance increased as the number of detectors increased. Furthermore, it was reported in their work that the carbon black−polymer composite detectors significantly outperformed the tin oxide and intrinsically conducting polymers arrays in resolving both the full set of 19 analytes and subsets thereof, as determined by the criteria of best resolving the analytes on average or best resolving the worst-resolved analyte pairs using the raw data responses.

One advantage to the use of detectors whose conductivities follow percolation theory is that their sensing properties can be controlled by changing the composition or loading of conductor in the insulating matrix. Having detectors that operate closer to the percolation threshold could, theoretically, lead to significant signal enhancement. The detection thresholds for carbon black detectors were initially estimated to be at 0.25 parts per billion.[130] Since it is more appropriate to report detection thresholds in terms of activity instead of concentration,[136] the detection thresholds have been measured to be 4×10^{-5} of an analyte's vapor pressure.[137] The concentration detection threshold data obtained both for a homologous series of alkanes with vapor pressures ranging from 5 to 100 Torr and a homologous series of alcohols with vapor pressures ranging from 2 to 100 Torr ranged from 0.1 to 100 ppm and from 0.1 to 1 ppm, respectively.

VI. Optical Vapor Sensing Arrays

A. Introduction

Optical sensors employ analyte-sensitive indicators for detecting chemical species. A multitude of transduction mechanisms may be utilized for detecting the analyte, including fluorescence intensity and lifetime, polarization, spectral shape, absorbance, wavelength, and reflectance. A typical optical sensor is composed of an optical fiber with an indicator immobilized on the fiber tip.[138] By immobilization of different polymer layers onto the fiber tip, the sensor can be tuned to distinguish between a particular analyte or a range of analytes. By combination of multiple fibers into an array format, such sensors can be designed to detect many analytes simultaneously. An optical cross-reactive array also employs multiple fiber sensors, but these sensors exhibit a broad, rather than a specific, response to different analytes. Several recent reviews describe these optical cross-reactive array systems.[139,140] While several other cross-reactive optical vapor sensing systems exist, they are not in an array format.[72,141−146] This section focuses on two formats of cross-reactive optical array: polymer-deposited fiber arrays and self-encoded bead arrays.

B. Polymer-Deposited Optical Sensor Arrays

Solvatochromic fluorescent dyes have been immobilized in various polymer layers to produce a sensor array for detecting organic vapors. The interaction between the immobilized sensing chemistry and the analyte generates local fluorescence signals that can be monitored over time.[147] One of the major advantages to using an optical approach for designing a cross reactive array sensor is the ability to collect many kinds of complex information simultaneously, including (but not limited to) changes in intensity, fluorescence lifetime, wavelength, and spectral shape. Increasing the dimensionality of a system, by observing many different parameters at one time, makes it possible to build a more sensitive, multianalyte detecting device with fewer sensors. A single solvatochromic dye (Nile Red) is immobilized within different polymer matrixes, each having its own baseline polarity.[148] Nile Red exhibits large shifts in its emission wavelength maximum with changes in local polarity. Typically, when exposed to solvents with increasing polarity, solvatochromic indicators will exhibit progressively more red-shifted absorption and/or emission spectra. The sorption of organic vapors into the polymer induces a change in the microenvironmental polarity, which is reported by the solvatochromic dye, as seen in Figure 12. By

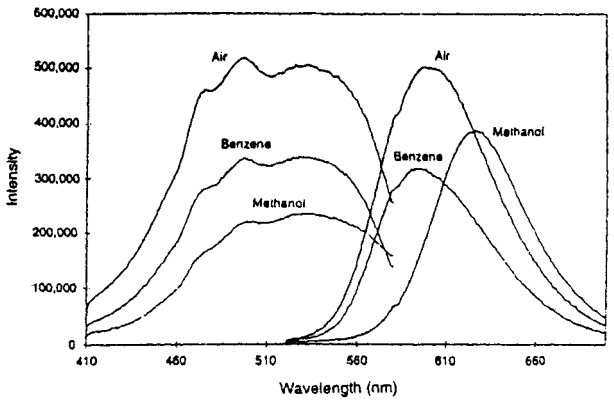


Figure 12. Fluorescence spectral changes of polymer-immobilized Nile Red with polarity changes imposed by different vapors (excitation on left, emission on right). Reprinted with permission from ref 150.

immobilization of the dye in polymer matrixes of varying polarity, hydrophobicity, pore size, flexibility, and swelling tendency, unique sensing regions are created that elicit different fluorescence responses when exposed to organic vapor pulses. The temporal response signals during these vapor pulses form the basis for unique patterns, which can then be used to train a computational network for subsequent analyte identification.

1. Array Fabrication

The distal ends of each fiber in the bundle are individually coated with different polymer/Nile Red combinations using either photopolymerization or dip-coating techniques. Photopolymerization involves

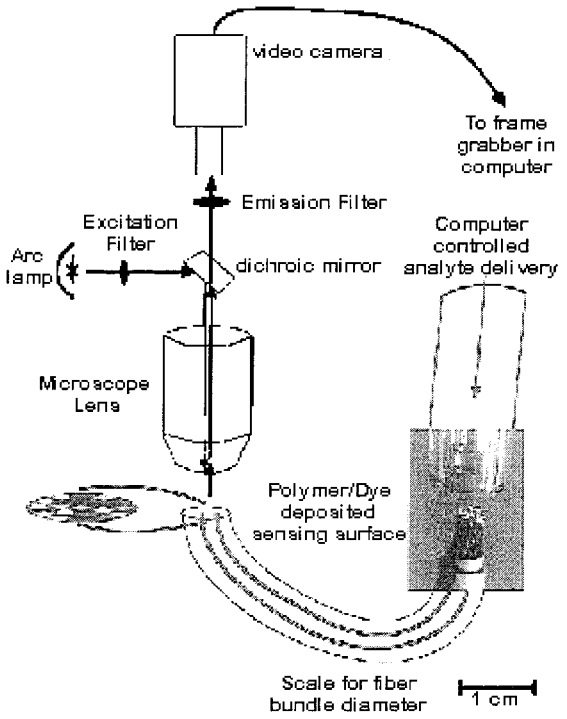


Figure 13. Schematic diagram of the video imaging and analyte delivery setup for the recording of responses of the optical fiber sensor array to applications of analytes. Reprinted from ref 151. Copyright 1996 American Chemical Society.

propagating ultraviolet light through the fiber to polymerize dye/monomer/initiator solutions onto the fiber tip. In the dip coating method, silanized fiber tips are repeatedly dipped in prepared Nile Red/polymer solutions and the solvent is allowed to evaporate between each dip. Both approaches result in Nile Red polymers attached to the fiber tips. Changing the amount of exposure to UV light or the number of solvent dips can alter the thickness of the polymer layers. Arrays can be assembled by bundling together multiple individual sensors with each sensor comprised of a different polymer. Alternatively, the entire array can be prepared on the ends of 350-μm-diameter fibers called image guides. Image guides consist of thousands of 3-μm optical fibers bundled together in a coherent fashion, such that each pixel's position is maintained from one end of the bundle to the other, allowing images to be transmitted down the length of the guide.[149] With these bundles one can photodeposit discrete polymer regions at selected positions on the face of the fiber using UV light, focused through a pinhole.

For detection, the proximal end of the fiber array is attached to a detection system comprised of a fluorescence microscope, light source, CCD camera, and computer. Filter wheels, shutters, and dichroic mirrors select and direct the appropriate wavelength of light used to analyze the optical arrays as shown in Figure 13. Vapor pulses are delivered through a vapor delivery system. Such pulses are analogous to a "sniff". Excitation light is introduced into the proximal end of the fiber and fluorescence from each sensor in the array returns through the fiber and is sent through an emission filter system and is then focused onto a CCD camera. The pattern of spectral

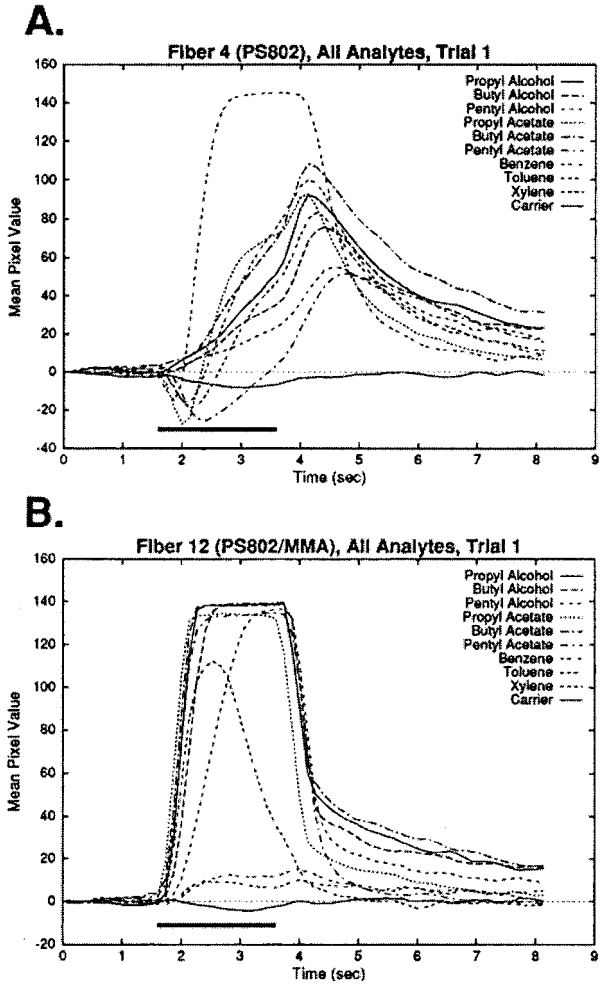


Figure 14. Examples of the responses of two different fibers (A, B) to all analytes presented as saturated vapor. The horizontal line at the bottom of each graph indicated the duration of analyte application. Reprinted from ref 151. Copyright 1996 American Chemical Society.

shifts exhibited by such an array upon exposure to different vapors is unique and characteristic for each vapor. These complex, time-dependent signals provide a "signature" for each vapor.[150−152]

The responses are monitored as a function of time versus mean pixel value (fluorescence). The fluorescence changes in these sensors are detected within milliseconds, and data are gathered by monitoring the fluorescence of each sensor in the array as pulses of different vapors are applied. As seen in Figure 14, two distinct responses are obtained when different vapors are pulsed onto two different polymer sensors. The time plots for fiber 4 demonstrate that the responses are distinct in their amplitudes and/or time courses for each analyte tested. On the other hand for fiber 12, the responses rapidly rise to a saturated plateau for five of the nine analytes tested.

2. Sensor Diversity

With the shift toward array formats comprised of large sets of cross-reactive sensors, a demand for new ways of creating diverse sensors has arisen throughout the sensor field. Polymerization reactions between different combinations of two starting materials lead to many new, different, nonlinearly related

sensing materials. This approach has been demonstrated in two ways: (a) the use of discrete polymer sensing cones each comprised of a specific monomer combination, (b) the fabrication of polymer gradient sensors. Gradient sensors were formed by scanning a pinhole-focused UV light beam across the proximal face of the fiber while steadily changing the concentrations of the two monomers in the solution being photopolymerized. The distal tip of the fiber resides, in the solutions, so that a polymer "stripe" forms as the photopolymerization occurs. The resulting polymer stripe contains all combinations of the starting and ending monomer concentrations. Responses from various regions of the gradient sensor were found to possess a far higher degree of diversity than did the corresponding region of a single component control stripe. In addition, responses from adjacent regions progressing down the length of the gradient stripe are not related to each other in a logical, linear fashion. One might expect different combinations of two monomers to yield sensors with responses that are simply related to the proportion of the two monomers used to construct the sensor. In actuality, such a progression is not observed: mixture responses do not appear to be linear combinations of the individual polymer responses. This individuality is most likely due to variation in polymerization kinetics and microstructure formation in the final polymer layer caused by the monomer ratio changes. In this way, it is possible to generate large collections of different sensing matrixes simply and efficiently from a limited set of starting materials.

C. Self-Encoded Bead Sensors

Decreasing the size of the sensor has enormous advantages in terms of both response time as well as the number of potential uses for such a device. Thousands of tiny beads (3.2 μm) can be immobilized in individual, acid-etched wells at the tip of an imaging fiber.[153] By introducing vapor-sensitive dyes into beads constructed from a wide range of polymeric and/or ceramic materials, it is possible to build an "ensemble" comprised of as many different types of sensor beads as necessary. Nile Red can be attached to different beads with the responses of all the beads from a particular preparation displaying nearly identical response behavior. Each bead class has a unique, well-defined response allowing individual beads to be easily identified after they have been deposited into wells. Since each bead type has its own unique response to a given vapor, the beads can easily be recognized *after* they have been placed into the image guide wells by simply exposing the array to a known test vapor and matching the resulting temporal response plots to those obtained beforehand for each bead class (Figure 15). In this way, the bead sensors are "self-encoding". The ability to prepare literally billions of identical beads in a single preparation should enable the computational network training to be transferred from one image guide ensemble to another.[154]

By using microspheres as sensor matrixes, the size of the overall array can be drastically reduced. This decrease in size in turn significantly shortens sensor response times while simultaneously enhancing sensor sensitivity. Silica bead sensors, for instance, were found to have response times on the order of ~150 ms, with signal changes of up to 1700%. On average, a 5-s experiment is ample time for most beads to fully respond to and recover from a pulse of vapor at any given concentration.

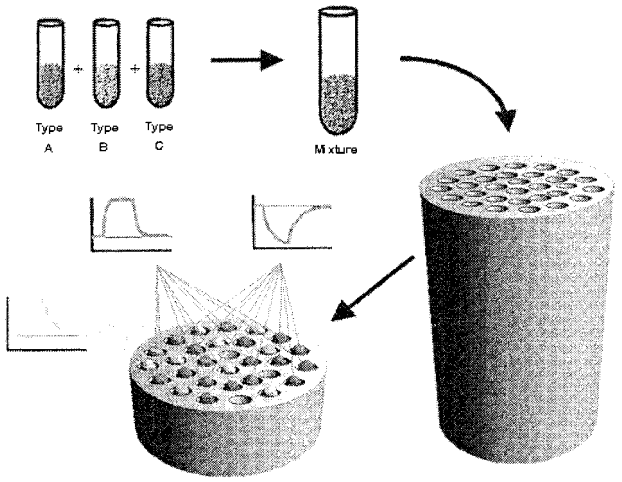


Figure 15. Schematic depiction of the self-encoded bead arrays concept. A mixture of sensor beads is prepared by combining aliquots from three stock solutions, each containing a different type of polymer/dye sensor suspended in a Nanopure water/0.01% Tween solution. A drop of the final mixture is then placed onto the distal tip of an etched imaging fiber. After they have settled in random locations throughout the well array, the beads are identified and categorized by their characteristic responses to a test vapor pulse, without the need for any additional encoding schemes. Reprinted from ref 154. Copyright 1999 American Chemical Society.

The recent development of an "electronic tongue", which is selective for the individual solutions, but not specific in their recognition properties, is an important step in the progress of optical cross-reactive sensor arrays.[155] Four taste categories, sweet, sour, salty, and bitter, were created from four different bead types, which measure combinations of pH, Ca^{+}, Ce^{+}, and simple sugars. Red, green, and blue light intensities were acquired for the each of the individual beads and used for the analyte quantification. The polymer response times were under 1 min, and the array was integrated allowing for continuous measurements of the bead−analyte response. The use of the red, green, and blue light intensities is a novel approach to sensor fabrication and provides a basis for combining an "artificial nose" with an "electric tongue" to gather both vapor and solution information.

D. Sensor Sensitivity

An important feature of any artificial nose is its ability to detect and quantify low concentrations of odors. In an effort to improve sensor sensitivity, common adsorbents have been incorporated directly into the sensing layers.[156] The most effective approach has been to directly adsorb dye molecules to the surface of alumina and silica particles. These particles are then adhered to fiber distal tips using various polymers and epoxies. The resulting sensors

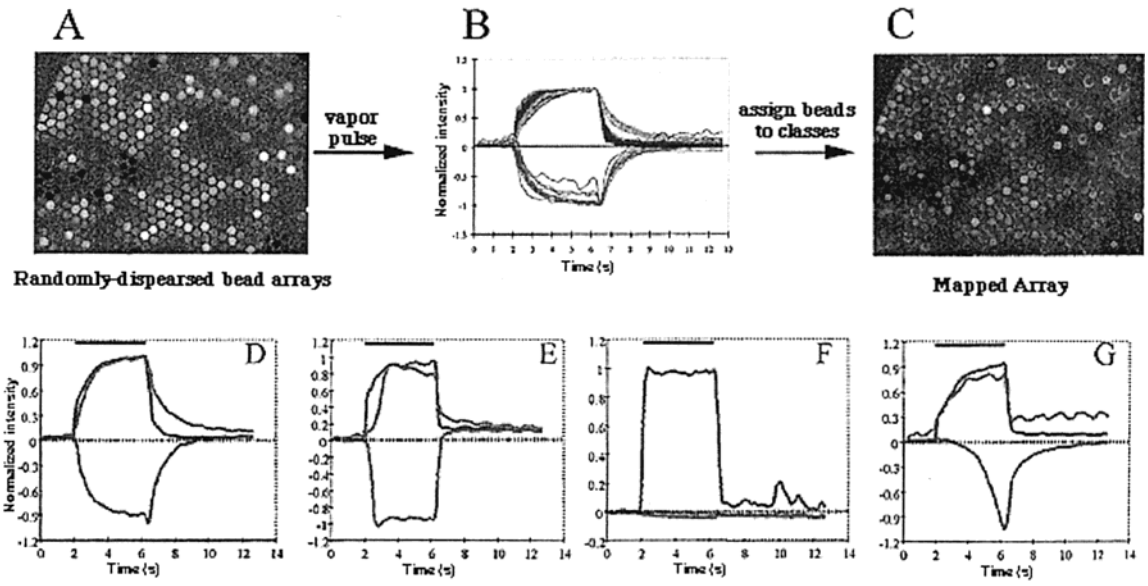


Figure 16. Data from a three-component self-encoded bead arrays (A). An initial pulse of methanol serves both as a vapor-sensing experiment and the array-decoding step. The three responses of the three bead types used are identified (B) to provide a map of the array (C). In this particular region are shown (D) methanol, (E) dichloromethane, (F) toluene, and (G) acetone. In D−G, traces represent the average response from each bead class. Reprinted from ref 154. Copyright 1999 American Chemical Society.

are capable of detecting vapors at much lower concentrations.

Another way to improve sensitivity is to sum the low-level responses of a large number of individual sensing elements (polymer beads), we have been able to produce large enhancements in the overall sensitivity of the array.[152,154] A similar approach has been used in the field of microelectrode arrays, where the advantages of microelectrodes (small size, rapid response times, small diffusion layers, and small *iR* drops) are combined with an array format in order to amplify the relatively small currents generated by individual microelectrodes. The primary difference, however, is that in the optical format, each element is individually addressable, making it possible to join the signals from large numbers of like elements randomly dispersed throughout the array. In addition, summing the responses from several beads at low vapor concentrations results in a substantial improvement in signal-to-noise ratios.

E. Data Processing for Optical Sensor Arrays

Pattern-recognition computational networks are used to analyze the data from the cross-reactive optical array sensors. The fluorescence changes are plotted versus time to produce a temporal response pattern for each sensing site for a particular vapor. The fluorescence response produces a series of characteristic patterns that are introduced as inputs to an neural network. Changes in fluorescence intensity data as a function of time contains sufficient information to accurately identify and quantify a broad range of organic vapors. In the first demonstration of the optical array's capabilities, a feed-forward neural network was able to correctly identify individual compounds, as well as components in binary mixtures, and to give a rough estimate of concentration.[150] In addition, the network was able to categorize compounds as aromatic, alcohol, or ester and to rank the relative molecular weights of compounds within a class. This latter capability seems to be due to the different response times of different analytes resulting from diffusion within the polymer layers. The sensor also has the ability to discriminate between vapors comprised of very similar functionality.[157]

In Figure 16, the data from an initial three-component system are described. Plotting the fluorescence intensity of the bead-containing wells with respect to time produces the characteristic response shapes. Methanol is used to decode the positions of the beads, as it provides a distinct response for the three bead types. Once the beads have been assigned, they are tested against three other vapors: dichloromethane, toluene, and acetone. Each vapor produces a distinct response to each sensor, and the combined use of the three sensors allows for the differentiation between the four saturated vapors. All of the resulting data are normalized to clarify the display and assist in the response−shape comparisons.

The self-encoded bead sensor array is an attractive approach because it can quickly accommodate new sensor bead types and is an improvement from the previous sensors in size, response and recovery times, ease of preparation, sensitivity, cost, and sensor reproducibility.

VII. Electrochemical Sensor Arrays

A. Introduction

In this section, we review multianalyte detection via cross-reactive electrochemical sensor arrays. By convention, most electrochemical array detection schemes incorporate identical and/or selective sensors that are not necessarily cross-reactive. These and

other array systems will only be discussed here if the sensor array was specifically designed to be cross-reactive in nature or if the cross-reactive responses from the sensor array are considered in the analysis.[158−174] Electrochemical arrays designed for the determination of one analyte,[175−178] multianalyte detection systems such as stripping voltammetry,[179−187] detection systems in which electrochemical sensor arrays are employed for chromatographic or electrophoresis separations,[188−200] or biosensor-related systems[201−203] will not be discussed. Individual sensors within a cross-reactive electrochemical array should show cross-sensitivity and/or cross-selectivity to multiple targets. As is the case for all the different cross-reactive arrays, it is also advantageous if individual sensors are sensitive or selective to varying degrees to target analytes so that different response patterns can be generated.

Over the last two decades, there has been a relaxed demand for highly selective sensors. More emphasis has been placed on sensor reproducibility, sensor stability, and the computational analysis software to retain the level of accuracy and/or detection limits for cross-reactive electrochemical sensor arrays.[158−163,166,169−174] Glass et al. has calculated that by incorporating a cross-reactive electrochemical sensor array in combination with a computational analysis program one can increase the amount of information content by at least 25% as compared to a single sensing element.[172] Thus, a cross-reactive array may supply the information necessary to predict target concentrations of multiple compounds in a known or an unknown matrix. As always, the information patterns generated by the sensor array are directly linked to the calibration techniques and experimental setup of the cross-reactive electrochemical array. There are also many variables which govern these response patterns, such as sensor material and size, solution or vapor composition, whether the reaction reaches equilibrium, and variation in sensor response, diffusion, and fluid dynamics (mixing, stirring, flow-injection, heat, etc.). There have been recent reviews in dynamic electrochemistry[204] and on chemical sensors[205] so many of these variables will not be discussed further, except where necessary.

B. Potentiometric (Equilibrium) Measurements

1. Principles of Operation

By definition, potentiometry refers to the difference in potential between two electrodes of a galvanic cell under the condition of zero current.[206] The indicator electrode is chosen so that it is responsive to the solution's target analyte and the reference electrode is invariant. This is an equilibrium measurement process because no current passes through the cell while the potential is measured. The potential of an electrochemical cell (E_{cell}) is

$$E_{cell} = E_{ind} - E_{ref} + E_{lj} \quad (9)$$

where E_{ind} and E_{ref} are the half-cell potentials for the indicator and reference electrodes, respectively, and E_{lj} is the potential at the electrode's liquid−junction interface. By employment of the Nernst equation, potentiometric measurements can provide accurate results for the concentration and activity coefficients of a target analyte. For example, in a general chemical reaction comprised of two half-cell redox reactions, $aA + bB \leftrightarrow cC + dD$, the Nernst equation would be

$$E_{cell} = E_{cell}^{\circ} - (RT/nF) \ln[(a_C^{\,c}\, a_D^{\,d})/(a_A^{\,a}\, a_B^{\,b})] \quad (10)$$

where E_{cell}° is the standard cell potential, R is the gas constant, T is the temperature, n is number of electrons involved in the reaction, F is the Faraday constant, and a is the activity of each chemical species. Many parameters affect the cell overpotential including temperature fluctuations, change of solution, electrode selectivity and sensitivity, and response to nontarget and interfering species in solution. Complexity in an electrochemical response is also related to adsorption of specific or nonspecific analytes on the electrodes and the nature of the electron transfer. Therefore employment of a "lock-and-key" sensor approach might only be feasible if the system is highly controlled. One may choose to control the temperature, employ solvents that minimize adsorption, or employ ion-selective electrodes (ISEs) to eliminate some of these variables. Theoretically, ISEs are "lock-and-key" sensors where each ISE is designed for one target ion. The ultimate goal, and biggest challenge in this field, is to build highly selective sensors that are not affected by interferents. However, nearly all ISEs developed to date have some cross-sensitivity (referred to as interferences) and respond in some fashion to other chemical species; herein lies the reason to employ these sensors in a cross-reactive array.

To discuss the use of ISEs in cross-reactive arrays, it is important to know how they generally operate and their original purpose. More background for this electroanalysis field can be found in a recent review.[207] ISEs are potentiometric sensors containing an ion-selective membrane that coats the electrode and sets up a potential due to transport of a single ion. The magnitude of this potential change is directly related to the specificity of the membrane for a target ion. A general schematic of an ISE is shown in Figure 17.

For an ISE, the potential of the membrane is considered and this changes eq 9 to

$$E_{cell} = E_{ref,ext} - E_{ref,int} + E_{mem} + E_{lj} \quad (11)$$

where $E_{ref,ext} - E_{ref,int}$ is the potential difference between the external and internal reference electrodes, respectively, and E_{mem} is the potential generated at the membrane. Usually the activity of the internal reference electrolyte is fixed and a plot of E_{cell} vs log(activity of analyte)[206] should be linear over the working range of the electrode according to eqs 12 and 14,

$$E_{mem} = (RT/zF) \ln[(a_i)_{samp}/(a_i)_{internal}] \quad (12)$$

where z is the charge on the ion and $(a_i)_{samp}$ and $(a_i)_{internal}$ are the potentials that develop on either side

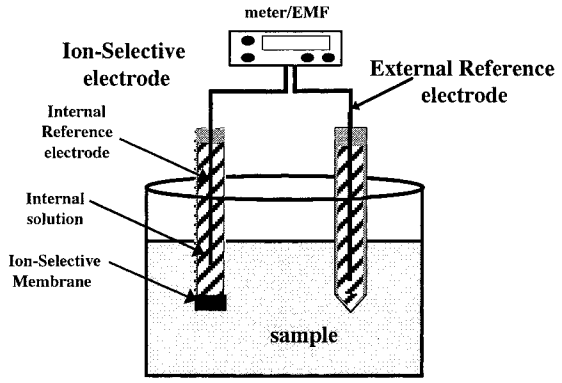


Figure 17. Basic schematic of an ion-selective electrode (ISE) used for potentiometric measurements. Upon interaction between the ion-selective membrane and the analyte, the potential across the membrane is altered. This potential change is dependent on the membrane specificity for a target ion in solution according to eq 11.

of the membrane for the target ion (i). For simplicity, the slope term (S) in eq 12 can be abbreviated

$$S = (RT/zF) \tag{13}$$

Assuming a fixed internal potential (a_i) and plugging E_{mem} into eq 11 produces the potential of the electrochemical cell as

$$E_{cell} = E_{ref,ext} - E_{ref,int} + (S)\ln[(a_i)_{samp}] + E_{lj} \tag{14}$$

The terms $E_{ref,\,ext} - E_{ref,\,int}$ and E_{lj} are constant so the only variables in the system are the target (i) ion activity and/or slope variation. These systems can still suffer from nontarget selectivities, interferents, and species buildup on the membrane. Therefore, ISEs find common application in solutions where there are low levels of interfering ions or where the concentrations are fixed to match those of calibration samples *unless* an ISE's cross-selectivity is used as an advantage. Many ISEs may be employed in the same sensor array. If each ISE within the array has varying degrees of selectivity/sensitivity for target analytes, this cross-reactivity can be an advantage and response patterns for each sensor can then be used to train a system for target determination. An array of ISEs may be used for multianalyte detection schemes as long as some type of computational analysis is employed. After modeling of an array of ISEs, the observed potential in the solution of unknown composition can be decoupled into those contributions arising from the target analytes and those arising from the interferents. This can improve precision when linked to an analysis scheme. The push to develop more sophisticated and intelligent systems involves acquiring the total analytical signal for accurate determination of the individual components. However, the acquisition of redundant (and nonuseful) information for such arrays does not benefit a system's prediction ability. Therefore, employing a variety of differentially responsive electrochemical sensors into an array and monitoring each sensor's response to a given analyte or analyte mixture will generate more useful information for target determination.

2. Examples of Potentiometric Arrays

As briefly stated above, ISEs are often unintentionally cross-reactive in nature because binding interactions between the membrane and nontarget chemical species. In this review, ISE array-types are not considered cross-reactive unless each electrode's nontarget responses and/or cross-selectivity are examined and/or incorporated into the analysis. For example, ISE arrays employing more than one type of sensor-element may be suited for multianalyte detection[208,209] but if sensor cross-selectivity is not considered or factored into the analysis these will not be discussed.

The membrane of the ISE will directly determine that sensor element's degree of cross-selectivity. Systems can be designed to be cross-reactive according to known membrane−analyte interactions.[207] For an ISE array to be considered cross-reactive, it must incorporate two or more ISE-types, be employed to determine more than one analyte, and use the electrochemical responses from many or all species in solution. Some groups have employed arrays of semiselective or sparingly selective electrodes in their cross-reactive arrays.[158−160] Sparingly selective electrodes are nonspecific ISEs and have been used for the specific purpose of generating more cross-reactivity within a sensor array because they respond to many analytes in varying degrees of selectivity. Otto and Thomas[158] may have been the first to perform these types of experiments in which nonspecific ISEs were employed in combination with a chemometrics program to identify concentrations of multiple targets within a mixture. They report on the simultaneous determination of free metal ions (Ca^{2+}, Mg^{2+}, K^+, Na^+) at physiological concentration levels. Such response data cause eq 14 to be transformed into a much more complicated system for each electrode−analyte response and can be thought of as an extended Nernst (Nicholski) equation

$$E_{ij} = E_j^\circ + S_j \ln(a_{ik} + \sum_l K_{jk} a_{il}) \tag{15}$$

These notations are employed in Beebe et al.'s[159] work where E_{ij} is the potential of the jth electrode, measured in the ith sample with respect to the reference electrode. E_j° is the intercept potential when the analyte activity a_{ik} is 1 M and the interferent levels, a_{il}, for all l interfering ions, are zero; K_{jk} is the selectivity coefficient of the jth electrode for kth ion, and S_j is the measured slope of the electrode in the absence of interferents. In the unique case where an ISE is not cross-reactive or there is no cross-selectivity toward nontarget ions, a selectivity coefficient for all lth ions would be zero, essentially eliminating the summation factor in eq 15. A more complicated version of this equation would be needed if ions of mixed valences affected the membrane potential, as suggested in eq 12. However, Otto and Thomas[158] found that they could use the valence (z) of the ion (i) regardless of the valences of the lth interfering ions. The most difficult problem that they encountered was the direct identification of Mg^{2+} in the presence of Ca^{2+} because the Mg^{2+} ISE

Table 3. Prediction of Remaining Nine Samples[a]

	predicted concns (% error)	
sample *i*	$[Na^+]$, M	$[K^+]$, mM
2	0.1198 (0.2)	3.73 (2.4)
4	0.1198 (0.1)	6.86 (2.0)
7	0.1351 (0.1)	3.89 (1.8)
9	0.1349 (0.1)	7.15 (2.1)
11	0.1498 (0.1)	2.03 (1.5)
14	0.1497 (0.1)	7.00 (0.0)
15	0.1486 (0.9)	8.45 (0.6)
17	0.1665 (0.9)	3.69 (3.4)
19	0.1641 (0.5)	7.03 (0.4)

[a] Reproduced from ref 159. Copyright 1993 American Chemical Society.

lacked selectivity for magnesium over that of calcium. They overcame this difficulty by employing more ISEs than they had targets, i.e., an overdetermined system, in combination with multiple regression analysis based on partial least squares (PLS). In a four component solution which simulated concentration levels in bodily fluids (blood serum or urine), they were still able to determine Mg^{2+} in the presence of Ca^{2+} using the responses from the cross-reactive array. Incorporating more sensors than they had target analytes, i.e., eight sensors, and incorporating PLS analysis, the prediction errors were approximately 4.5% (Ca^{2+}), 6.8% (Mg^{2+}), 2.3% (Na^+), and 1.4% (K^+).

Otto and Thomas[158] based their measurements on E_j° and K_{jk} and calibrated their sensors with knowledge of their slopes, as opposed to the work by Beebe et al.[159] who extended this work to see if they could calibrate an array of sensors without prior knowledge of sensor slope. Otto and Thomas[158] noted that multivariate calibration for the ISE array is hindered by the need to employ a constant overall ionic strength of the solution (fixed boundary conditions) while simultaneously varying the concentrations for each ion-type. However, they also used calibration solutions that did not contain interferents that may have caused some problems when moving to test solutions. Any errors associated with the calibration would also be incorporated into the test set. Beebe et al.[159] avoided these methods by varying the Nernstian slopes which led to an advantage in the calibration process where less error was predicted in the test set. They employed five to eight nonspecific ISEs in their cross-reactive array, and without having any potassium-selective sensors, they detected K^+ (2−10 mM) in the presence of much higher Na^+ levels (135−155 mM) in the matrix. According to their calibration methods, their use of only sparingly selective electrodes, and their analysis scheme, they attained lower prediction errors than Otto and Thomas;[158] however, unlike Otto and Thomas, they did not employ ions of mixed valences. Table 3 shows the results of Beebe et al.'s study.

This experiment demonstrated the feasibility of using a nonspecific sensor array to determine ion concentrations with reasonable accuracy. Beebe et al. went on to analyze their data to see if *each* of the five employed electrodes could be used alone to predict sodium concentration levels when potassium ions were ignored. For the single-sensor system(s), the prediction errors for nine samples varied between 0 and 38%, further emphasizing the benefit of cross-reactive electrochemical arrays and the increase in information content for analyte detection. Beebe and Kowalski[160] then attempted to expand this work by calibrating their cross-reactive array without a priori information about the functional relations between the responses and the ion concentrations. Unlike previous work[158,159] they did not base their calibration models on the assumption that the electrodes obey a certain equation. They merely used their array of nonspecific ISEs to prove their algorithms had the ability to determine binary mixtures of Na^+ (0.120−0.165 M) and K^+ (2.0−8.4 mM) in aqueous solution with average prediction errors lower than 5.3%.

Forster et al.[162] employed three highly specific ISEs in combination with one sparingly selective electrode in their cross-reactive array for determining sodium, potassium, and calcium ions in tertiary mixtures. They used this overdetermined system to poll results which enhanced the amount of information gained through cross-talk between sensors without introducing significant error from unmodeled interferents. It should be noted that although they employed three highly specific ISEs, the determined selectivity coefficients for the sodium, potassium, and sparingly selective electrodes suggested that these electrodes were not as selective as the calcium electrode (note different valence) in the presence of interferents. By incorporation of the sparingly selective electrode with the 3 selective sensor elements into an array, their prediction errors for 12 samples were improved from 4.5% to 2.8%, relative to the prediction errors for the 3 selective-element array alone. The array was able to determine each ion in a high interference background when the solutions did not contain unmodeled interferents. Forster and Diamond[163] used the same cross-reactive array model and analyte targets in combination with flow injection to improve prediction performance for mineral water and human plasma samples. They attributed the results to the kinetic factors (not equilibrium) associated with the flow injection sample delivery. Their array showed enhanced sensitivity for the target ions and they were able to determine these ions, without worrying about responses from unmodeled ion interferents such as Mg^{2+}.

In a related model, Diamond and Forster[164] employed three specific ISEs (for Na^+, K^+, Ca^{2+}) with a multiple ionophore electrode. They tailored the fourth electrode's selectivity by incorporating an ionophore for each of the target ions into the membrane. This array system, though more specifically designed, meets the cross-reactive array criteria. The array was more responsive to the three target ions relative to the multiple ionophore electrode as seen in previous work[163] with a sparingly selective electrode. However, even though this array was cross-reactive, the calibration and tests were specific for particular defined reaction conditions and could not be implemented into unknown systems. Finally, Legin et al.[165] used an array of electrodes to detect zinc ions in a background with three other metals (Cu^{2+}, Pb^{2+}, Cd^{2+}) even though there was no zinc-selective elec-

trode employed. The sensors were incorporated into an array, and the system was trained on complex solutions of heavy metals (Cu^{2+}, Pb^{2+}, Cd^{2+}, Zn^{2+}, Cl^-, F^-, and SO_4^{2-}) in concentration ranges typical of industrial waste. All the species were identified even without having specific sensors for each of the target analytes, i.e., zinc and sulfate.

C. Voltammetric (Nonequilibrium) Measurements

1. Principles of Operation

In contrast to potentiometry, where there is no applied current, voltammetry is the measurement of the current−potential relationship in an electrochemical cell where equilibrium is not established. There are many voltammetric methods: square wave, staircase, pulse and differential pulse, cyclic, anodic stripping, and amperometric titrations. Again, we will only deal with those areas that employ cross-reactive sensor arrays. In voltammetric measurements, information about a target analyte is generated from the measurement of the current as a function of the applied potential under conditions, which enhance polarization of the working electrode.[210] The reference electrode's potential is constant, and the working electrode assumes the value of the applied potential. Therefore, the working electrode is the site where electrolysis occurs and this generates the measured current. The current is quantitatively related to the speed of the electrolytic process, which in its simplest form, is a redox half-cell reaction:

$$M^{n+} + e^- \leftrightarrow M^{(n-1)+} \qquad (16)$$

Here M^{n+} is the electroactive species which is reduced to $M^{(n-1)+}$ (or the reverse reaction where $M^{(n-1)+}$ is the electroactive species oxidized to M^{n+}). Sensors based on detecting the current flow caused by oxidizing or reducing an analyte are highly successful because of their selectivity and high sensitivity coupled to the wide range of organic and inorganic analytes that can be detected using this technique.[211] If one were to employ an array of working electrodes, each element *could* generate an independent measurement signal. An important feature of many voltammetric-type arrays is the movement toward miniaturization. Arrays of microelectrodes used in voltammetry can have either *all* the electrodes interconnected to produce one overall measured current, or *each* of the electrodes can have different applied voltages and allow individual electrodes to register a current. These arrays can also offer significant improvements in sensitivity, *S*/*N* ratios, and detection limits so combining these advantages into a cross-reactive array with an incorporated computational analysis program can lead to highly effective and intelligent detection systems. One can use the time response or even use different potentials for different sensors to generate more information that can be used for target discrimination. Considerable effort has been applied to controlling the reactivity of amperometric sensors by modifying the surfaces with thin films, metal layers, polymers, or biological materials. Therefore, cross-reactive voltammetric arrays can be generated by employing different potentials to the individual elements, employing an array of different electrode types, modifying the electrodes, or using a combination of these approaches.

2. Examples of Voltammetric Arrays

An array of cross-reactive amperometric sensors was employed by Stetter et al.[166] to detect 22 organic vapors (20−300 ppm) in a portable system. This system consisted of four uncoated electrodes, each operating at a different potential and preceded by different filament catalysts to pretreat the incoming vapor flow before presentation to the array. The four sensors with their applied potentials were the following: (a) Au, −200 mV; (b) Au, +300 mV; (c) Pt, +200 mV; (d) Pt black, 0 mV. As noted, the sensors were set at different oxidation and reduction potentials and the four modes used to acquire data were the following: (1) no filament used; (2) platinum filament at fixed temperature; (3) rhodium filament at fixed temperature; (4) rhodium filament at a second fixed temperature. The heated filaments can partially oxidize some of the vapors as they flow to the amperometric sensors. In all, 16 separate channels of information were generated for each of the chemical species in the test set. Many of these data channels contained unique information that was analyzed with pattern recognition software.

One goal of Stetter et al.'s work was to detect and discriminate the vapors in the test set and minimize the number of sensors employed. The amperometric sensors used were known to fluctuate by as much as ±25% over the course of 1 month so they incorporated ±25% random noise into the response patterns to see how their prediction errors varied. When they removed redundant data and incorporated nonlinear mapping to check the original data set in combination with two sets containing random errors, only two analyte vapor pairs overlapped in the cluster analysis (nitrobenzene/acetone; cyclohexane/acetic acid).

Though more sensors need to be incorporated for better discrimination, these results show that by employing a four-element cross-reactive amperometric array, nearly 22 vapors could be discriminated and detected at ranges between 20 and 300 ppm using pattern recognition analysis even when ±25% random error was incorporated into the array. In a related study with the same array model and pattern recognition, Stetter et al.[169] classified grain quality according to patterns for "good", "sour", or "insect" wheat classes with good accuracy. In another similar approach, Schweizer-Berberich et al.[170] characterized fish freshness vs time with an eight-element cross-reactive array employing a filament catalyst varied over five temperatures. Forty channels of data were generated, and principal component analysis (PCA) and principal component regression (PCR) were used for data analysis.

Using the same four-element array model, Stetter et al.[167] aimed to use only *one* sensor's response from the array to perform analysis. In the defined system, it was determined that sufficient information was generated from the one-sensor/one-filament in order

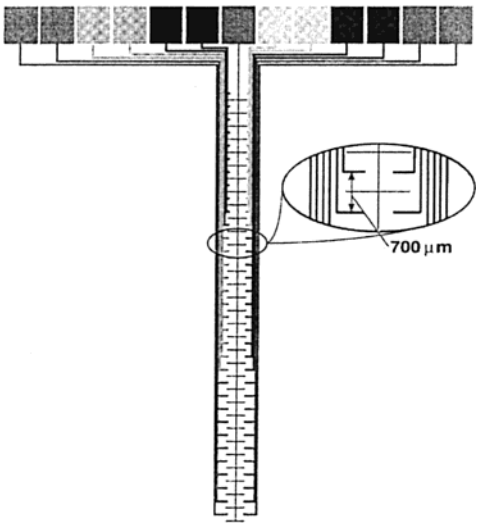


Figure 18. Diagram of a microelectrode array. The insert shows spacing between working electrodes. Electrical contacts are made at the top of the chips, and there are two pads for each shown electrode. From left to right (pad numbers): Pt (1,2), carbon (3,4), V (5,6), Pt auxiliary (7), Au (8,9), Ir (10,11), and Pt (12,13). For each material, one contact pad leads to a single electrode and the other pad leads to an array of 10 electrically connected together. There are 66 working electrodes, and the platinum auxiliary electrode runs down the center. Reprinted from ref 172. Copyright 1990 American Chemical Society.

to quantify and identify certain compounds and mixtures. In a related study, Vaihinger et al.[168] reasoned that one electrochemical sensor in combination with one catalytic filament could be employed to identify and quantify pure gases and mixtures by varying filament temperatures. Although this single-element system is not an array, the multiple information obtained from the sensor serves to increase the dimensionality of the data and thereby enables the sensor to improve its ability to solve analytical tasks. In a relatively similar effort as Stetter et al.[167] and Vaihinger et al.,[168] Glass et al.[172] tried to compare the amount of useful information one sensor element could generate as compared to a five-element cross-reactive array. The microelectrode arrays were produced by photolithography and employed two sensors for each of the five different sensor types (Figure 18).

For one electrode (Pt), they compared data at five different potentials (−0.2, −0.5, −0.8, −1.0, and −1.2 V) vs the data at one potential (−1.2 V) across the five-element electrode pairs. These data did not provide identification of analytes; however they show that the average information content gained for the cross-reactive array relative to one sensor element was 25%.

Wang et al.[171] took a different approach to creating sensor diversity for developing a cross-reactive sensor array. The four amperometric sensors were each coated with a different semiselective film of varying pore size, charge, and polarity, and unlike Stetter et al.,[166] all the sensors within the array were maintained at the same potential. They used this array as a thin film detector to quantify neurologically significant catechol compounds such as dopamine, (3,4-dihydroxyphenyl)acetic acid (DOPAC), epinephrine, norepinephrine, and catechol using flow injection to deliver the sample to the array. See Figure 19 for the array schematic.

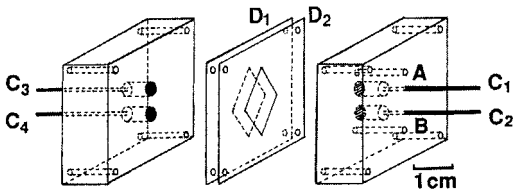


Figure 19. Enclosed view of the thin-layer flow cell: (A, B) solution inlet and outlet; (C_1−C_4) working electrodes; (D_1, D_2) spacers. Reprinted from ref 171. Copyright 1990 American Chemical Society.

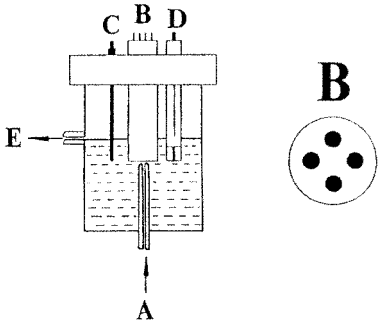


Figure 20. Schematic view of the large-volume wall-jet detector: (A) inlet; (B) electrode array; (C) counter electrode; (D) reference electrode; (E) outlet. To the right is the bottom view of the electrode array (B). Reprinted from ref 173. Copyright 1993 American Chemical Society.

Wang et al.'s array exhibits unique responses to the analytes because of partial selectivity and employs multiple linear regression analysis. Sensor diversity is directly related to the film employed and produced different current responses over time when all the sensors were held at the same potential. For instance, the smaller catechol compounds transport readily through the size-exclusion cellulose acetate layer to produce larger sensor responses than dopamine. In other data not shown, Wang et al. also used this cross-reactive array and applied different potentials to the four elements and showed that a third dimension can be used to generate more information for better analyte determination.

Chen, Wang, and co-workers[173] modified the electrode elements to create even more sensor diversity. They modified four carbon paste electrodes with metal oxide catalysts (Cu_2O, RuO_2, NiO, and CoO) and kept all the elements at the same potential to determine individual carbohydrates and amino acids in different sample mixtures by amperometric flow injection. Figure 20 is a schematic of the array system. Each modifier shows a different electrocatalytic behavior toward each analyte. Figure 21A (carbohydrates) and Figure 21B (amino acids) show the generated response patterns for the sensor array, i.e., the analyte fingerprint.

The distinct electrocatalytic properties for each sensor results in unique responses for the analytes when tested within the dynamic range of the sensor array. Using statistical regression analysis with two and three component mixtures, the prediction errors for analyte identification ranged from 0 to 11%, whereas the average value was 2.3% for the sensor array. Like Wang et al.[171] they show that more

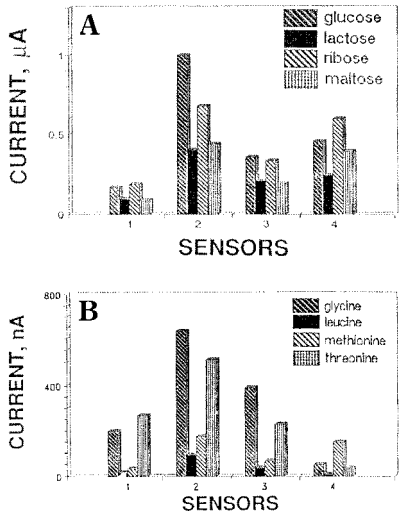


Figure 21. Sensor response patterns for the listed carbohydrates (A) and amino acids (B). The carbon paste electrodes are modified with the following metal oxides: (1) CoO; (2); Cu_2O (3) NiO; (4) RuO_2. Reprinted from ref 173. Copyright 1993 American Chemical Society.

information can be extracted when multiple potentials are applied to the individual sensor elements. In a recent publication, Amatore et al.[212] employed paired-band microelectrode assemblies to mimic physiological neuronal processing to observe current vs time relationships for two species using square pulse voltammetry. Their system offers many advantages for use in cross-reactive arrays because of on-line logic processing and the degree to which sensor variations could produce diversity in the system's response profiles. In a final related voltammetric approach, Wehrens and van der Linden[174] examined calibration data from an array of individually modified electrodes to determine the best analysis protocol for nonlinear voltammogram data from four chemically similar compounds. Different voltammograms were expected for this array because they used modified electrodes. Sixteen Ir electrodes were employed, six of which were modified with Au (2), Rh (2), and Pt (2), and they used PCR analysis for the calibration data and artificial neural networks to determine the best tool for analysis. They state that it is possible to use such an array to quantify multiple components in a sample although their results were not that convincing.

VIII. Acoustic Wave Devices

A. Introduction

Piezoelectric materials produce a voltage when mechanical stress is applied, and conversely will deform if a voltage is applied across them. When an oscillating potential is applied at a frequency near the resonant frequency of a piezoelectric crystal, a stable oscillating circuit is formed. Depending upon the geometry of the crystal and electrodes, a variety of wave modes can be established.[213] Two device types have been used to construct electronic noses: thickness-shear mode (TSM, also called quartz crystal microbalance (QCM) and bulk acoustic wave (BAW)) and surface acoustic wave (SAW) resonators. These

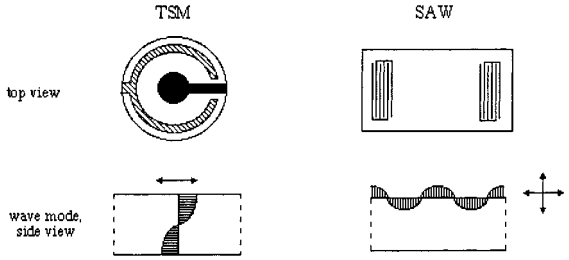


Figure 22. Thickness-shear mode (TSM) and surface acoustic wave (SAW) resonators. On top are shown the quartz crystals and their gold contacts (on two sides of the crystal for the TSM). On the bottom are schematics of the wave that propagates in each resonator.

devices are commonly constructed from quartz crystals, with the configurations shown in Figure 22. In a TSM resonator, the acoustic wave propagates through the bulk of the crystal in a direction perpendicular to the surface, with motion at the surface parallel to the surface. In a SAW device, motion occurs only at the surface, penetrating to a depth of approximately one acoustic wavelength into the crystal. The direction of propagation is parallel to the surface, while motion at the surface is both parallel and perpendicular to the surface. SAW devices can be constructed in two different configurations, delay-line and resonator, as pictured in Figure 22. TSM devices typically operate at frequencies from 5 to 30 MHz, while SAW devices generally operate between 100 and 400 MHz.

Adding mass to the surface of acoustic resonators changes their resonant frequency, and Sauerbrey derived eq 17 to describe the frequency shift of a

$$\mathrm{d}F = 2.3 \times 10^6 \times F^2(dM/A) \qquad (17)$$

quartz TSM resonator resulting from a change in mass on its surface.[214] The change in frequency (dF) is proportional to the original frequency squared and the change in mass per area of the crystal surface. A resonator coated with a material, such as a polymer, that absorbs organic molecules from the gas phase would be expected to change its resonant frequency upon exposure to organic vapors. The Sauerbrey equation is of course an approximation that is valid under certain conditions; for relatively thick films of hard polymers, sorption-induced changes in viscoelastic effects can also contribute to the observed frequency shift. King first utilized quartz TSM devices coated with common GC stationary phases as a detector of organic molecules exiting a GC column.[215] Wohltjen and Dessey were the first to use coated SAW resonators to detect organic vapors.[216]

B. TSM Arrays

In 1986, Kowalski and co-workers first examined arrays of polymer-coated TSM resonators.[217,218] Their data[219] consisted of the responses of 27 coatings on TSM devices exposed to 14 different analytes. Principal components analysis (PCA) was used to analyze the matrix of responses and indicated that 95% of the variance in the data was present in the first 7 principal components. The authors concluded that an

array of seven sensors comprised of those with the greatest contribution to each of the first seven PCs would represent a near-optimal array.[217] The same data set was analyzed to determine what subset of sensors could provide the best sensitivity, selectivity, signal-to-noise ratio, and limit of detection. Since the original data consisted of only one response for each sensor−analyte pair, an estimated random noise was added to the data. One seven-sensor array consisted of the most sensitive individual sensors, while a second array consisted of seven sensors chosen using PCA as outlined above. A set of equations developed by Lorber[220] that describe the sensitivity and selectivity of an array of partially selective sensors to particular analytes was used to compare the two sensor arrays. It was found that while the two arrays had similar overall selectivities, the PCA-selected array was substantially more sensitive to all the analytes tested. For one sensor, a detection limit can be defined, commonly at the analyte concentration at which the signal-to-noise ratio is three. For an array of sensors, an analogous *limit of determination* (LOD) can be defined for an analyte in the presence of other possible interfering components. The LOD for a variety of analytes were generally lower for the PCA-selected array of sensors. A comparison of the determined case (seven sensors, seven analytes) and the overdetermined case (seven sensors, three analytes) shows that in the overdetermined case the LOD decreased by nearly 2 orders of magnitude.

Kowalski and co-workers constructed a nine-sensor array and determined the concentration of individual analytes in two- and three-component mixtures.[221] The methods of multiple linear regression (MLR)[222] and partial least squares (PLS)[223] were used. For the two-component mixtures, PLS predicted concentrations approximately 5 times better than MLR did. The relative prediction error of PLS for two-component mixtures was 4.6%, while individual sensors have a 3−6% relative response error. For the three-component case, the average PLS prediction error was 10.1%, while that of MLR was 18.6%. Carey and Kowalski used a similar six-sensor array to monitor the solvent vapors exiting an industrial process dryer simulator.[224] Principal component regression (PCR) was used to calculate the concentrations of vapors exiting the chamber, and the values agree with those that would be expected.

Göpel and co-workers first investigated arrays of TSM resonators coated with both polymer and nonpolymer materials in 1991.[225] A variety of functionalized polysiloxanes on TSM resonators were exposed to a series of organic molecules, and the strengths of their various interactions were discussed.[226,227] A correlation was noted between log K (K is the polymer−gas partition coefficient) and T_b/T' (T_b is the solvent's boiling point, and T' is the temperature at which the measurements are made). After the boiling point correlation was corrected for, the expected trends were observed: relatively nonpolar organic molecules were absorbed strongly by poly(dimethylsiloxane), polar organic molecules were most strongly absorbed by poly(cyanopropylmethylsiloxane), and polarizable molecules were relatively well absorbed by polyphenylmethylsiloxane. Cellulose derivatives were also used as coating materials.[228] Both PLS and artificial neural networks (ANN) were tested for their ability to predict concentrations of analytes in mixtures, and ANN were found to slightly outperform PLS.[229,230] Neural networks were also used to dynamically monitor an analyte stream.[231] An array of TSM devices coated with siloxane polymer, metal complex-modified siloxane polymer, or pure metal complex was used to sense organic vapors and was found to be particularly sensitive to oxygen- and nitrogen-containing molecules.[232] A hybrid array containing polymer-coated TSM resonators, tin oxide gas sensors, electrochemical sensors, and metal oxide semiconductor field effect transistors (MOSFETs) was used to discriminate coffees, olive oils, tobaccos, and whiskeys.[51] Chiral GC stationary phases that consist of polysiloxanes with chiral side chains were coated on TSM resonators, and small arrays of these devices were used to discriminate enantiomeric odorants.[233,234] In another study, the coating materials consisted of γ-cyclodextrin derivatives dissolved in a polysiloxane matrix.[235] Polymer emulsions were used to create porous coatings on TSM resonators, and the porous polymers were subsequently coated with a variety of lipids.[236]

In 1989 Nakamoto, Moriizumi, and co-workers used an array of six TSM resonators coated with both polymeric and nonpolymeric materials to analyze 11 kinds of liquors.[237] Neural networks were 73% successful in categorizing the liquors. In another study, 8 coating materials were selected from a library of 18, in an attempt to find an optimized array. This array was used to discriminate 10 whiskys.[238,239] Flavors and fragrances were also studied.[240,241]

One group has used polymer and nonpolymer films that were applied via radio frequency sputtering of the materials.[242,243] Plasma-deposited organic films have properties significantly different from the original polymers because of the loss of atoms and fragments and the presence of dangling bonds in the deposited film. Conducting polymers have also been used as TSM resonator coating materials.[244−246]

There has been a second study on the discrimination of enantiomers.[247] Another compared the responses of a TSM resonator-based nose to those of human subjects.[248] Other examples of TSM resonator-based noses have also been studied.[249−252]

C. SAW Arrays

The first application of pattern recognition methods to data from an array of SAW devices was by Grate and co-workers in 1986.[253] A total of 12 different sensor coatings were exposed to 11 vapors. Principal components analysis was used to display in two dimensions the separation of various analytes. Hierarchical clustering categorized both the vapors and the sensor coatings. Clusterings were rationalized through their solubility parameters, including hydrogen bonding donor and acceptor ability and dipolarity/polarizability. On the basis of the classification methods used, four coatings were selected that could completely separate two classes of vapors. Mixtures of analytes were also studied.[254] An array

of four temperature-controlled SAW devices successfully detected and classified organophosphorus and organosulfur vapors.[255]

A total of 20 polymers on SAW devices were exposed to 5 organic vapors.[256] Three of the sensors were exposed to ternary mixtures of toluene, acetone, and dichloromethane. Partial least squares regression predicted the vapor concentrations, with errors of a few percent. In another study, four sensors were used to analyze ternary mixtures of water, methanol, and automotive fuel.[257] The experimental and known methanol concentrations agreed quite well. A simple two-sensor array was able to monitor H_2O and CO_2 levels in air.[258] An array of four polymer-coated SAW devices was employed to measure small amounts of tetrachloroethylene, trichloroethylene, and methoxyfurane in humid air.[259] The samples were meant to imitate exhaled breath containing trace amounts of organic contaminants; actual breath samples spiked with one of the analytes were also examined. Limits of detection were estimated to be 0.7, 0.6, and 4 ppm for tetrachloroethylene, trichloroethylene, and methoxyflurane, respectively. In a follow-up to this work, several adsorbents were tested as desorbable preconcentrators to increase the response to organic vapors in simulated exhaled breath.[260]

A combination of plasma-polymerized films, plasma-grafted films, and self-assembled monolayers was applied to SAW devices, and they were exposed to a variety of organic vapors.[261] The newly developed visually empirical region-of-influence (VERI) pattern recognition algorithm was used to analyze and categorize the data. The authors found that the best three-film array comprised one coating from each category, while the optimal arrays consisted of five to seven sensors. The same group has investigated the use of dendrimers as sensor coatings.[262]

The effects of temperature and humidity on arrays of polymer-coated SAW sensors have been studied.[263,264] Ultraviolet light-cross-linked polysiloxanes were found to have better long-term stability than noncross-linked polysiloxanes.[265]

Arrays of SAW devices coated with nonpolymeric materials have also been investigated, including liquid crystals,[266] Langmuir−Blodgett-deposited phospholipids and fatty acids,[267] self-assembled monolayers,[268] covalently bound organic molecules,[269] crown ethers,[270] and other nonvolatile organic molecules.[271]

D. Response Prediction

The frequency change of a TSM or SAW device depends on how much vapor is absorbed by the polymer film, which in turn is proportional to K, the polymer−vapor partition coefficient of the analyte. The frequency response of polymer-coated SAW resonators was predicted on the basis of the analyte's boiling point, solubility parameter, and linear solvation energy relationships.[272] The expected linear correlation was observed between log K and the vapor's boiling point, although the correlation was not as good for polar coatings. A modification of the predictions on the basis of the solvation parameters in linear solvation energy relationships (LSERs) gave the best results. Grate, Abraham, and co-workers have also extensively utilized LSERs in the prediction of sensor responses.[273−275] Hydrogen bond acidic polymers have been used specifically to improve detection and discrimination of basic vapors.[276] Other molecular orbital and molecular mechanics parameters have been used to predict SAW sensor responses.[277]

It has been found that, under certain conditions, SAW devices do not act solely as gravimetric sensors but also respond to the modulus change in polymers when a vapor is sorbed.[278] For thin films with low modulus, both SAW and TSM devices act only as gravimetric sensors, while, for polymers with initially high modulus, the SAW devices are also sensitive to the modulus change, but TSM resonators still act primarily as gravimetric sensors.[279,280] Under certain conditions, TSM resonators are also sensitive to modulus changes.[280]

Once a method for estimating responses is established, it is possible to choose the best array of sensors for detecting specific analytes or mixtures of analytes. Monte Carlo simulations of sensor responses allowed such a selection.[281]

Only one paper has been published that compares the actual analytical performance of polymer-coated TSM and SAW resonators.[282] TSM devices with frequencies of 10 and 30 MHz and SAW resonators with frequencies of 80 and 433 MHz were investigated. The influences of temperature and film thickness were studied. For each device, the thickest coating that could be applied without quenching the oscillation was applied, since thicker coatings will provide larger responses. The higher the frequency of the resonator, the less polymer could be applied. The 30 MHz TSM device had the best signal-to-noise ratio (S/N) and limits of detection (LOD) for the two polymers and two analytes that were tested.

The LOD (where $S/N = 3$) for n-octane for the 30 MHz TSM device and the 80 MHz SAW device coated with poly(dimethylsiloxane) was determined to be 2 ppm.[282] The other coating, poly(ether urethane), had LODs of 3−15 ppm. A comparable LOD of 0.6 ppm was found for n-octane on the carbon black−polymer composite nose (for the best detector in the array).[137]

IX. Conclusions and Future Prospects

Cross-reactive chemical sensor arrays are a promising alternative to conventional chemical sensors. By relying on patterns of response over a multisensor array, such systems have reduced the need to obtain the exquisite specificity of conventional sensors. The "electronic nose" or "artificial nose" moniker that has been given to such systems stems from the use of pattern recognition applied to the complex signals derived from the arrays. While not entirely misnomers, these terms do not do justice to the remarkable process of biological olfaction. Mammalian olfactory systems are continuously replacing dead cells while maintaining fidelity of the intricate neural wiring necessary to conserve responses. Signal amplification is a signature of such systems with multiples at each stage in the signal transduction process. The geometric complexity of turbinates in the nasal epithelium, the complex sniffing patterns when

animals are exposed to odors, and the convergence of similar receptor cell axons onto the olfactory bulb as well as the complex "wiring" patterns of the processing anatomy all contribute to the high sensitivity and discriminating power of the mammalian olfactory system. The systems described in this review are only rudimentary mimics of the biological system. By continuing to incorporate the operating principles of biological systems into the artificial ones, chemists may eventually be able to approach some of the capabilities of the olfactory system with these cross-reactive arrays.

As discussed in the review, there are many important and potentially exciting areas where existing arrays are beginning to have an impact. In particular, the food processing industry is relying increasingly on electronic nose systems to make determinations about quality and freshness. In addition, these industries are evaluating packaging material before adding products to make sure that such materials do not impart an off taste or odor to the food. One can imagine that as such systems get smaller, they will begin to show up as consumer items for evaluating meat, fish and dairy freshness in the kitchen as well as for monitoring potential home and workplace hazards. Such systems, by nature of their broad-bandedness, could be used for chemical and/or biological weapons detection; they could be anticipatory in the sense that they would not necessarily be keyed to particular agents but could look for responses likely to be correlated with toxicity or virulence. One day, such systems may be used to perform medical diagnostics using a simple breath test. More far-fetched, but within the realm of possibility, perhaps they may be used to recognize individuals by their signature odors. After all, dogs are able to recognize their owners by odor cues.

For purposes of discussion, this review deliberately separated the different sensor transduction mechanisms. Each of these different mechanisms offers potentially unique information about the samples being tested. The basis for the discriminating power of these arrays is the information content contained in the responses. In the ultimate manifestation of the technology, it is likely that hybrid systems will be of value because they offer the most information-rich signals.

Finally, before such systems can have a major impact, they must be manufacturable in large quantities. Because such systems rely on training computational networks on sensor response patterns, there is a need to maintain robust responses over extended time periods, otherwise retraining will be required as sensors change. Similarly, sensors must be able to be manufactured reproducibly and in large quantities so that training is transferable from one array to the next. The oldest of all chemical sensors is the pH electrode, yet today's pH electrodes are a long way from meeting such a requirement-they must be calibrated regularly (ideally before each use). Thus, there are practical impediments that must be overcome before these devices are used as routine analytical systems.

X. Acknowledgments

We acknowledge the Army Research Office, NASA, and DARPA for support of sensor array work at Caltech and acknowledge DARPA, the Office of Naval Research, the National Institutes of Health, and the Department of Energy for support at Tufts that made preparation of this review, along with some of the examples discussed herein, possible.

XI. References

(1) Elghanian, R.; Storhoff, J. J.; Mucic, R. C.; Letsinger, R. L.; Mirkin, C. A. *Science* **1997**, *277*, 1078.
(2) Göpel, W. *Mikrochim. Acta* **1997**, *125*, 179.
(3) Göpel, W. *Microelectron. Eng.* **1996**, *32*, 75.
(4) Zhu, S. S.; Carroll, P. J.; Swager, T. M. *J. Am. Chem. Soc.* **1996**, *118*, 8713.
(5) Rickert, J.; Weiss, T.; Göpel, W. *Sens. Actuators B* **1996**, *31*, 45.
(6) Schierbaum, K. D. *Sens. Actuators B* **1994**, *18*, 71.
(7) Buck, L. B.; Axel, R. *Cell* **1991**, *65*, 175.
(8) Malnic, B.; Hirono, J.; Sato, T.; Buck, L. B. *Cell* **1999**, *96*, 713.
(9) Axel, R. *Sci. Am.* **1995**, *273*, 154.
(10) Breer, H.; Wanner, I.; Strotmann, J. *Behav. Genet.* **1996**, *26*, 209.
(11) Lancet, D.; Ben-Arie, N. *Curr. Biol.* **1993**, *3*, 668.
(12) Ostrowski, J. C.; Kjelsberg, M. A.; Caron, M. G.; Lefkowitz, R. J. *Annu. Rev. Pharmacol. Toxicol.* **1992**, *32*, 167.
(13) Bartlett, P. N.; Gardner, J. W. *NATO Advanced Research Workshop on Sensors and Sensory Systems for an Electronic Nose*, Reykjavik, Iceland, 1991; Plenum Press: New York, 1991; p 31.
(14) Shimizu, Y.; Egashira, M. *MRS Bull.* **1999**, *24*(*6*), 18.
(15) Hartman, J. *Proc. Am. Soc. Hortic. Sci.* **1954**, *64*, 335.
(16) Persaud, K. C.; Dodd, G. H. *Nature* **1982**, *299*, 352.
(17) Gardner, J. W.; Bartlett, P. N. *Electronic Noses Principles and Applications*; Oxford University Press: Oxford, U.K., 1999.
(18) Gutierrez, J.; Getino, J.; Horrillo, M. C.; Ares, L.; Robla, J. I.; Garcia, C.; Sayago, I. *Thin Solid Films* **1998**, *317*, 429.
(19) Mishra, V. N.; Agarwal, R. P. *Microelectron. J.* **1998**, *29*, 861.
(20) Di Natale, C.; Davide, F. A. M.; Faglia, G.; Nelli, P. *Sens. Actuators B* **1995**, *B23*, 187.
(21) Heilig, A.; Barsan, N.; Weimar, U.; Schweizer-Berberich, M.; Gardner, J. W.; Göpel, W. *Sens. Actuators B* **1997**, *43*, 45.
(22) Corcoran, P.; Lowery, P.; Anglesea, J. *Sens. Actuators B* **1998**, *48*, 448.
(23) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177.
(24) Gardner, J. W.; Pike, A.; Rooij, N. F.; Koudelka-Hep, M.; Clerc, P. A.; Hierlemann, A.; Göpel, W. *Sens. Actuators B* **1995**, *26*−*27*, 135.
(25) Gotz, A.; Gracia, I.; Cane, C.; Lora-Tamayo, E.; Horillo, M. C.; Getino, J.; Garcia, C.; Gutierrez, J. *Sens. Actuators B* **1997**, *44*, 483.
(26) Gardner, J. W.; Shurmer, H. V.; Corcoran, P. *Sens. Actuators B* **1991**, *B4*, 117.
(27) Hong, H.-K.; Shin, H. W.; Park, H. S.; Yun, D. H.; Kwon, C. H.; Lee, K.; Kim, S.-T.; Moriizumi, T. *Sens. Actuators B* **1996**, *33*, 68.
(28) Wang, X. D.; Yee, S.; Carey, P. *Sens. Actuators B* **1993**, *13*, 458.
(29) Bednarczyk, D.; DeWeerth, S. P. *Sens. Actuators B* **1995**, *27*, 271.
(30) Wilson, D. M.; DeWeerth, S. P. *Sens. Actuators B* **1995**, *28*, 123.
(31) Wilson, D. M.; Deweerth, S. P. *Sens. Mater.* **1998**, *10*, 169.
(32) Chaturvedi, A.; Mishra, V. N.; Dwivedi, R.; Srivastava, S. K. *Microelectron. J.* **1999**, *30*, 259.
(33) Srivastava, R.; Dwivedi, R.; Srivastava, S. K. *Sens. Actuators B* **1998**, *50*, 175.
(34) Di Natale, C.; D'Amico, A.; Davide, F. A. M.; Faglia, G.; Nelli, P.; Sberveglieri, G. *Sens. Actuators B* **1994**, *20*, 217.
(35) Getino, J.; Gutierrez, J.; Ares, L.; Robla, J. I.; Horrillo, M. C.; Sayago, I.; Agapito, J. A. *Sens. Actuators B* **1996**, *33*, 128.
(36) Abe, H.; Yoshimura, T.; Kanaya, S.; Takahashi, Y.; Miyashita, Y.; Sasaki, S. *Anal. Chim. Acta* **1987**, *194*, 1.
(37) Getino, J.; Horrillo, M. C.; Gutierrez, J.; Ares, L.; Robla, J. I.; Garcia, C.; Sayago, I. *Sens. Actuators B* **1997**, *43*, 200.
(38) Haswell, S. J.; Walmsley, A. D. *Anal. Proc.* **1991**, *28*, 115.
(39) Llobet, E.; Brezmes, J.; Vilanova, X.; Sueiras, J. E.; Correig, X. *Sens. Actuators B* **1997**, *41*, 13.
(40) Shurmer, H. V.; Gardner, J. W.; Corcoran, P. *Sens. Actuators B* **1990**, *1*, 256.
(41) Walmsley, A. D.; Haswell, S. J.; Metcalfe, E. *Anal. Chim. Acta* **1991**, *250*, 257.
(42) Walmsley, A. D.; Haswell, S. J.; Metcalfe, E. *Anal. Chim. Acta* **1991**, *242*, 31.

(43) Alexander, P. W.; Di Benedetto, L. T.; Hibbert, D. B. *Field Anal. Chem. Technol.* **1998**, *2*, 145.
(44) Di Natale, C.; Davide, F. A. M.; D′Amico, A.; Sberveglieri, G.; Nelli, P.; Faglia, G.; Perego, C. *Sens. Actuators B* **1995**, *25*, 801.
(45) Di Natale, C.; Davide, F. A. M.; D′Amico, A.; Nelli, P.; Groppelli, S.; Sberveglieri, G. *Sens. Actuators B* **1996**, *33*, 83.
(46) Gardner, J. W. *Sens. Actuators B* **1991**, *4*, 109.
(47) Gardner, J. W.; Shurmer, H. V.; Tan, T. T. *Sens. Actuators B* **1992**, *6*, 71.
(48) Llobet, E.; Hines, E. L.; Gardner, J. W.; Franco, S. *Meas. Sci. Technol.* **1999**, *10*, 538.
(49) Olafsson, R.; Martinsdottir, E.; Olafsdottir, G.; Sigfusson, P. I.; Gardner, J. W. *NATO Advanced Research Workshop on Sensors and Sensory Systems for an Electronic Nose*, Reykjavik, Iceland, 1991; Plenum Press: New York, 1991; p 257.
(50) Winquist, F.; Hornsten, E.; Sundgren, H.; Lundström, I. *Meas. Sci. Technol.* **1993**, *4*, 1493.
(51) Ulmer, H.; Mitrovics, J.; Noetzel, G.; Weimar, U.; Göpel, W. *Sens. Actuators B* **1997**, *43*, 24.
(52) Winquist, F.; Sundgren, H.; Lundström, I. *Spec. Publ.—R. Soc. Chem.* **1998**, *167*, 170.
(53) Borjesson, T.; Eklov, T.; Jonsson, A.; Sundgren, H.; Schurer, J. *Cereal Chem.* **1996**, *73*, 457.
(54) Faglia, G.; Bicelli, F.; Sberveglieri, G.; Maffezzoni, P.; Gubian, P. *Sens. Actuators B* **1997**, *44*, 517.
(55) Gardner, J. W.; Hines, E. L.; Tang, H. C. *Sens. Actuators B* **1992**, *9*, 9.
(56) Nayak, M. S.; Dwivedi, R.; Srivastava, S. K. *Sens. Actuators B* **1993**, *12*, 103.
(57) Moore, S. W.; Gardner, J. W.; Hines, E. L.; Göpel, W.; Weimar, U. *Sens. Actuators B* **1993**, *16*, 344.
(58) Marco, S.; Ortega, A.; Pardo, A.; Samitier, J. *IEEE Trans. Instrum. Meas.* **1998**, *47*, 316.
(59) Davide, F. A. M.; Di Natale, C.; D'Amico, A. *Sens. Actuators B* **1994**, *18—19*, 244.
(60) Vlachos, D. S.; Avaritsiotis, J. N. *Sens. Actuators B* **1996**, *33*, 77.
(61) Lundström, I.; Shivaraman, M. S.; Svenson, C. S.; Lundkvist, L. *Appl. Phys. Lett.* **1975**, *26*, 55.
(62) Bergveld, P. *Sens. Actuators* **1985**, *8*, 109.
(63) Vlasov, Y. G. *Mikrochim. Acta* **1991**, *2*, 363.
(64) Lundström, I.; Hedborg, E.; Spetz, A.; Sundren, H.; Winquist, F. NATO *Advanced Research Workshop on Sensors and Sensory Systems for an Electronic Nose*, Reykjavik, Iceland, 1991; Plenum Press: New York, 1991; p 303.
(65) Winquist, F.; Sundgren, H.; Hedborg, E.; Spetz, A.; Lundström, I. *Sens. Actuators B* **1992**, *B6*, 157.
(66) Muller, R.; Lange, E. *Sens. Actuators* **1986**, *9*, 39.
(67) Sundgren, H.; Lundström, I.; Winquist, F.; Lukkari, I.; Carlsson, R.; Wold, S. *Sens. Actuators B* **1990**, *2*, 115.
(68) Gall, M.; Muller, R. *Sens. Actuators* **1989**, *17*, 583.
(69) Sundgren, H.; Winquist, F.; Lukkari, I.; Lundström, I. *Meas. Sci. Technol.* **1991**, *2*, 464.
(70) Sundgren, H.; Lundström, I.; Vollmer, H. *Sens. Actuators B* **1992**, *9*, 127.
(71) Sommer, V.; Tobias, P.; Kohl, D.; Sundgren, H.; Lundström, I. *Sens. Actuators B* **1995**, *28*, 217.
(72) Lundström, I.; Erlandsson, R.; Frykman, U.; Hedborg, E.; Spetz, A.; Sundgren, H.; Welin, S.; Winquist, F. *Nature* **1991**, *352*, 47.
(73) Lundström, I.; Sundgren, H.; Winquist, F. *J. Appl. Phys.* **1993**, *74*, 6953.
(74) Eklov, T.; Sundgren, H.; Lundström, I. *Sens. Actuators B* **1997**, *45*, 71.
(75) Eklov, T.; Lundström, I. *Anal. Chem.* **1999**, *71*, 3544.
(76) www.alpha-mos.com.
(77) www.eev.com.
(78) Holmberg, M.; Winquist, F.; Lundström, I.; Gardner, J. W.; Hines, E. L. *Sens. Actuators B* **1995**, *27*, 246.
(79) Mandenius, C.-F.; Hagman, A.; Dunas, F.; Sundgren, H.; Lundström, I. *Biosens. Bioelectron.* **1998**, *13*, 193.
(80) Mandenius, C.-F.; Liden, H.; Ekloev, T.; Taherzadeh, M. J.; Liden, G. *Biotechnol. Prog.* **1999**, *15*, 617.
(81) Reddinger, J. L.; Reynolds, J. R. *Adv. Polym. Sci.* **1999**, *145*, 57.
(82) *Handbook of Conducting Polymers*; 2nd ed.; Skotheim, T. A., Elsenbaumer, R. L., Reynolds, J. R., Eds.; Marcel Dekker: New York, 1998; Vol. 1.
(83) Roncali, J. *Chem. Rev.* **1992**, *92*, 711.
(84) Epstein, A. J.; Macdiarmid, A. G. *Makromol. Chem. Macromol. Symp.* **1991**, *51*, 217.
(85) Jahnke, S. A.; Niemann, J.; Kautek, W.; Bischoff, R.; Pfeiffer, C.; Kossmehl, G. *Int. J. Environ. Anal. Chem.* **1997**, *67*, 223.
(86) Svetlicic, V.; Schmidt, A. J.; Miller, L. L. *Chem. Mater.* **1998**, *10*, 3305.
(87) www.aromascan.com.
(88) www.bloodhound.co.uk.
(89) www.zelana.com.
(90) Swann, M. J.; Glidle, A.; Cui, L.; Barker, J. R.; Cooper, J. M. *J. Chem. Soc., Chem. Commun.* **1998**, 2753.
(91) Partridge, A. C.; Harris, P. D.; Andrews, M. K. *Analyst* **1996**, *121*, 1349.
(92) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators A* **1995**, *51*, 57.
(93) Gardner, J. W.; Pike, A.; de Rooij, N. F.; Koudelka-Hep, M.; Clerc, P. A.; Hierlemann, A.; Göpel, W. *Sens. Actuators B* **1995**, *26*, 135.
(94) Hodgins, D. M. *Sens. Actuators B* **1995**, *27*, 255.
(95) Slater, J. M.; Paynter, J.; Watt, E. J. *Analyst* **1993**, *118*, 379.
(96) Slater, J. M.; Watt, E. J.; Freeman, N. J.; May, I. P.; Weir, D. J. *Analyst* **1992**, *117*, 1265.
(97) Bartlett, P. N.; Gardner, J. W.; Whitaker, R. G. *Sens. Actuators A* **1990**, *23*, 911.
(98) De Wit, M.; Vanneste, E.; Geise, H. J.; Nagels, L. J. *Sens. Actuators B* **1998**, *50*, 164.
(99) Baldacci, S.; Matsuno, T.; Toko, K.; Stella, R.; De Rossi, D. *Sens. Mater.* **1998**, *10*, 185.
(100) Kawai, T.; Kojima, S.; Tanaka, F.; Yoshino, K. *Jpn. J. Appl. Phys., Part 1* **1998**, *37*, 6237.
(101) Pearce, T. C.; Gardner, J. W.; Friel, S.; Bartlett, P. N.; Blair, N. *Analyst* **1993**, *118*, 371.
(102) Hatfield, J. V.; Neaves, P.; Hicks, P. J.; Persaud, K. C.; Travers, P. J. *Sens. Actuators B* **1994**, *18*, 221.
(103) Serra, G.; Stella, R.; De Rossi, D. *Mater. Sci. Eng., C* **1998**, *5*, 259.
(104) Hodgins, D. M.; Simmonds, D. J. *J. Autom. Chem.* **1995**, *17*, 179.
(105) Bailey, T. P.; Hammond, R. V.; Persaud, K. C. *J. Am. Soc. Brew. Chem.* **1995**, *53*, 39.
(106) Gardner, J. W.; Pearce, T. C.; Friel, S.; Bartlett, P. N.; Blair, N. *Sens. Actuators B* **1994**, *18*, 240.
(107) Pearce, T. C.; Gardner, J. W. *Analyst* **1998**, *123*, 2047.
(108) Cantalejo, M. J. *Z. Lebensmittel Untersuchung Forschung A* **1999**, *208*, 373.
(109) Schaller, E.; Bosset, J. O.; Escher, F. *Chimia* **1999**, *53*, 98.
(110) Annor-Frempong, I. E.; Nute, G. R.; Wood, J. D.; Whittington, F. W.; West, A. *Meat Sci.* **1998**, *50*, 139.
(111) Gibson, T. D.; Prosser, O. C.; Hulbert, J. N.; Marshall, R. W.; Corcoran, P.; Lowery, P.; Ruck-Keene, E. A.; Heron, S. *Sens. Actuators B* **1997**, *44*, 413.
(112) Arnold, J. W.; Senter, S. D. *J. Sci. Food Agric.* **1998**, *78*, 343—348.
(113) Stuetz, R. M.; Fenner, R. A.; Engin, G. *Water Res.* **1999**, *33*, 453.
(114) Stuetz, R. M.; Fenner, R. A.; Engin, G. *Water Res.* **1999**, *33*, 442.
(115) Stuetz, R. M.; White, M.; Fenner, R. A. *J. Water Serv. Res. Technol.-Aqua* **1998**, *47*, 223.
(116) Stuetz, R. M.; Engin, G.; Fenner, R. A. *Water Sci. Technol.* **1998**, *38*, 331.
(117) Persaud, K. C.; Khaffaf, S. M.; Hobbs, P. J.; Sneath, R. W. *Chem. Senses* **1996**, *21*, 495.
(118) Byun, H. G.; Persaud, K. C.; Khaffaf, S. M.; Hobbs, P. J.; Misselbrook, T. H. *Comput. Electron. Agric.* **1997**, *17*, 233.
(119) Misselbrook, T. H.; Hobbs, P. J.; Persaud, K. C. *J. Agr. Eng. Res.* **1997**, *66*, 213.
(120) Masila, M.; Sargent, A.; Sadik, O. A. *Electroanalysis* **1998**, *10*, 312.
(121) Lane, A. J. P.; Wathes, D. C. *J. Dairy Sci.* **1998**, *81*, 2145.
(122) Schiffman, S. S.; Kermani, B. G.; Nagle, H. T. *Chem. Senses* **1997**, *22*, 119.
(123) Chandiok, S.; Crawley, B. A.; Oppenheim, B. A.; Chadwick, P. R.; Higgins, S.; Persaud, K. C. *J. Clin. Path.* **1997**, *50*, 790.
(124) Persaud, K. C.; Pisanelli, A. M.; Szyszko, S.; Reichl, M.; Horner, G.; Rakow, W.; Keding, H. J.; Wessels, H. *Sens. Actuators B* **1999**, *55*, 118.
(125) Amrani, M. E. H.; Ibrahim, M. S.; Persaud, K. C. *Mater. Sci. Eng.* **1993**, *C1*, 17.
(126) Amrani, M. E. H.; Persaud, K. C.; Payne, P. A. *Meas. Sci. Technol.* **1995**, *6*, 1500.
(127) Amrani, M. E. H.; Payne, P. A.; Persaud, K. C. *Sens. Actuators B* **1996**, *33*, 137.
(128) Amrani, M. E. H.; Dowdeswell, R. M.; Payne, P. A.; Persaud, K. C. *Sens. Actuators B* **1997**, *44*, 512.
(129) Amrani, M. E. H.; Payne, P. A. *IEE Proc. Sci. Meas. Technol.* **1999**, *146*, 95.
(130) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubb, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298.
(131) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2652.
(132) Brosseau, C.; Boulic, F.; Queffelec, P.; Bourbigot, C.; LeMest, Y.; Loaec, J.; Beroual, A. *J. Appl. Phys.* **1997**, *81*, 882.
(133) Kirkpatrick, S. R. *Rev. Mod. Phys.* **1973**, *45*, 574.
(134) Doleman, B. J.; Sanner, R. D.; Severin, E. J.; Grubbs, R. H.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 2560.
(135) Severin, E. J.; Sanner, R. D.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 1440.
(136) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5442.
(137) Doleman, B. J. Ph.D., California Institute of Technology, 1999.
(138) Seitz, W. R. *Crit. Rev. Anal. Chem.* **1988**, *19*, 135.
(139) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Trends Biotechnol.* **1998**, *16*, 250.

(140) Steemers, F. J.; Walt, D. R. *Mikrochim. Acta* **1999**, *131*, 99.
(141) Barnard, S. M.; Walt, D. R. *Nature* **1991**, *353*, 338.
(142) Orellana, G.; Gomez-Carneros, A.; de Dios, C.; Garcia-Martinez, A.; Moreno-Bondi, M. *Anal. Chem.* **1995**, *67*, 2231.
(143) Posch, H. E.; Wolfbeis, O. S.; Pusterhoffer, J. *Talanta* **1988**, *35*, 89.
(144) Ronot, C.; Archenault, M.; Gagnaire, H.; Goure, J. P.; Jaffrezic-Renault, N.; Pichery, T. *Sens. Actuators B* **1993**, *11*, 375.
(145) Ronot, C.; Gagnaire, H.; Goure, J. P.; Jaffrezic-Renault, N.; Pichery, T. *Sens. Actuators A* **1994**, *42*, 529.
(146) Gauglitz, G.; Kraus, G. *Fresenius' J. Anal. Chem.* **1993**, *346*, 572.
(147) Dickert, F. L.; Keppler, M. *Adv. Mater.* **1995**, *7*, 1020.
(148) Walt, D. R.; Dickinson, T.; White, J.; Kauer, J.; Johnson, S.; Engelhardt, H.; Sutter, J. M.; Jurs, P. C. *Biosens. Bioelectron.* **1998**, *13*, 695.
(149) Pantano, P.; Walt, D. R. *Anal. Chem.* **1995**, *67*, 481A.
(150) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Nature* **1996**, *382*, 697.
(151) White, J.; Kauer, J. S.; Dickinson, T. A.; Walt, D. R. *Anal. Chem.* **1996**, *68*, 2191.
(152) Albert, K. J.; Walt, D. R. *Anal. Chem.* **2000**, *72*, 1947.
(153) Micheal, K. L.; Taylor, L. C.; Schultz, S. L.; Walt, D. R. *Anal. Chem.* **1998**, *70*, 1242.
(154) Dickinson, T. A.; Michael, K. L.; Kauer, J. S.; Walt, D. R. *Anal. Chem.* **1999**, *71*, 2192.
(155) Lavigne, J. J.; Savoy, S.; Clevenger, M. B.; Ritchie, J. E.; McDoniel, B.; Yoo, S.-J.; Anslyn, E. V.; McDevitt, J. T.; Shear, J. B.; Neikirk, D. *J. Am. Chem. Soc.* **1998**, *120*, 6429.
(156) Healey, B. G.; Walt, D. R. *Anal. Chem.* **1997**, *69*, 2213.
(157) Sutter, J. M.; Jurs, P. C. *Anal. Chem.* **1997**, *69*, 856.
(158) Otto, M.; Thomas, J. D. R. *Anal. Chem.* **1985**, *57*, 2647.
(159) Beebe, K. R.; Uerz, D.; Sandifer, J.; Kowalski, B. R. *Anal. Chem.* **1988**, *60*, 66.
(160) Beebe, K. R.; Kowalski, B. R. *Anal. Chem.* **1988**, *60*, 2272.
(161) Bos, M.; Bos, A.; van der Linden, W. E. *Anal. Chim. Acta* **1990**, *233*, 31.
(162) Forster, R. J.; Regan, F.; Diamond, D. *Anal. Chem.* **1991**, *63*, 876.
(163) Forster, R. J.; Diamond, D. *Anal. Chem.* **1992**, *64*, 1721.
(164) Diamond, D.; Forster, R. J. *Anal. Chim. Acta* **1993**, *276*, 75.
(165) Legin, A. V.; Vlasov, Y. G.; Rudnitskaya, A. M.; Bychkov, E. A. *Sens. Actuators B* **1996**, *34*, 456.
(166) Stetter, J. R.; Jurs, P. C.; Rose, S. L. *Anal. Chem.* **1986**, *58*, 860.
(167) Stetter, J. R.; Findlay, M. W.; Maclay, G. J.; Zhang, J.; Vaihinger, S.; Göpel, W. *Sens. Actuators B* **1990**, *1*, 43.
(168) Vaihinger, S.; Göpel, W.; Stetter, J. R. *Sens. Actuators B* **1991**, *4*, 337.
(169) Stetter, J. R.; Findlay, M. W. J.; Schroeder, K. M.; Yue, C.; Penrose, W. R. *Anal. Chim. Acta* **1993**, *284*, 1.
(170) Schweizer-Berberich, P.-M.; Vaihinger, S.; Göpel, W. *Sens. Actuators B* **1994**, *18*, 282−290.
(171) Wang, J.; Rayson, G. D.; Lu, Z.; Wu, H. *Anal. Chem.* **1990**, *62*, 1924.
(172) Glass, R. S.; Perone, S. P.; Ciarlo, D. R. *Anal. Chem.* **1990**, *62*, 1914.
(173) Chen, Q.; Wang, J.; Rayson, G.; Tian, B.; Lin, Y. *Anal. Chem.* **1993**, *65*, 251.
(174) Wehrens, R.; van der Linden, W. E. *Anal. Chim. Acta* **1996**, *334*, 93.
(175) Yang, Z.; Sasaki, S.; Karube, I.; Suzuki, H. *Anal. Chim. Acta* **1997**, *357*, 41.
(176) Ross, B.; Cammann, K.; Mokwa, W.; Rospert, M. *Sens. Actuators B* **1992**, *7*, 758.
(177) Atkinson, J. K.; Shahi, S. S.; Varney, M.; Hill, N. *Sens. Actuators B* **1991**, *4*, 175.
(178) Augelli, M.; Nascimento, V. B.; Pedrotti, J. J.; Gutz, I. G. R.; Angnes, L. *Analyst* **1997**, *122*, 843.
(179) Kounaves, S. P.; Deng, W.; Hallock, P. R.; Kovacs, G. T. A.; Storment, C. W. *Anal. Chem.* **1994**, *66*, 418.
(180) Kovacs, G. T. A.; Storment, C. W.; Kounaves, S. P. *Sens. Actuators B* **1995**, *23*, 41.
(181) Reay, R.; Flannery, A. F.; Storment, C. W.; Kounaves, S. P.; Kovacs, G. T. A. *Sens. Actuators B* **1996**, *34*, 450.
(182) Uhlig, A.; Schnakenberg, U.; Hintsche, R. *Electroanalysis* **1997**, *9*, 125.
(183) Paeschke, M.; Dietrich, F.; Uhlig, A.; Hintsche, R. *Electroanalysis* **1996**, *8*, 891.
(184) Silva, P. R. M.; El Khakani, M. A.; Chaker, M.; Champagne, G. Y.; Chevalet, J.; Gastonguay, L.; Lacasse, R.; Ladouceur, M. *Anal. Chim. Acta* **1999**, *385*, 249.
(185) Feeney, R.; Herdan, J.; Nolan, M. A.; Tan, S. H.; Tarasov, V.; Kounaves, S. P. *Electroanalysis* **1998**, *10*, 1.
(186) Belmont, C.; Tercier, M.-L.; Buffle, J.; Fiaccabrino, G. C.; Koudelka-Hep, M. *Anal. Chim. Acta* **1996**, *329*, 203.
(187) Herdan, J.; Feeney, R.; Kounaves, S. P.; Flannery, A. F.; Storment, C. W.; Kovacs, G. T. A.; Darling, R. B. *Environ. Sci. Technol.* **1998**, *32*, 131.
(188) Matson, W. R.; Langlais, P.; Volicer, L.; Gamache, P. H.; Bird, E.; Mark, K. A. *Clin. Chem.* **1984**, *30*, 1477.
(189) Matson, W. R.; Gamache, P. G.; Beal, M. F.; Bird, E. D. *Life Sci.* **1987**, *41*, 905.
(190) Niwa, O.; Horiuchi, T.; Morita, M.; Huang, T.; Kissinger, P. T. *Anal. Chim. Acta* **1996**, *318*, 167.
(191) Liu, J.; Zhou, W.; You, T.; Li, F.; Wang, E.; Dong, S. *Anal. Chem.* **1996**, *68*, 3350.
(192) Niwa, O.; Morita, M. *Anal. Chem.* **1996**, *68*, 355.
(193) Jin, W. R.; Weng, Q.; Wu, J. *Anal. Chim. Acta* **1997**, *342*, 67.
(194) Iwasaki, Y.; Morita, M. *Curr. Sep.* **1995**, *14*, 2.
(195) Gavin, P. F.; Ewing, A. G. *J. Am. Chem. Soc.* **1996**, *118*, 8932.
(196) Gavin, P. F.; Ewing, A. G. *Anal. Chem.* **1997**, *69*, 3838.
(197) Chao, M. H.; Huang, H. J. *Anal. Chem.* **1997**, *69*, 463.
(198) Svendsen, C. N. *Analyst* **1993**, *118*, 123.
(199) You, T.; Wu, M.; Wang, E. *Anal. Lett.* **1997**, *30*, 1025.
(200) Slater, J. M.; Watt, E. J. *Analyst* **1994**, *119*, 2303.
(201) Skladal, P.; Kalab, T. *Anal. Chim. Acta* **1995**, *316*, 73.
(202) Wollenberger, U.; Paeschke, M.; Hintsche, R. *Analyst* **1994**, *119*, 1245.
(203) Wang, J.; Chen, Q. *Anal. Chem.* **1994**, *66*, 1007.
(204) Anderson, J. L.; Coury, L. A., Jr.; Leddy, J. *Anal. Chem.* **1999**, *70*, 519R.
(205) Janata, J.; Josowicz, M.; Vanysek, P.; DeVaney, D. M. *Anal. Chem.* **1998**, *70*, 179R.
(206) Strobel, H. A.; Heineman, W. R. *Chemical Instrumentation: A Systematic Approach*; John Wiley & Sons (Wiley-Interscience): New York, 1989.
(207) Bakker, E.; Buhlmann, P.; Pretsch, E. *Chem. Rev.* **1997**, *97*, 3083.
(208) Cosofret, V. V.; Erdosy, M.; Johnson, T. A.; Buck, R. P.; Ash, R. B.; Neuman, M. R. *Anal. Chem.* **1995**, *67*, 1647.
(209) Lauks, I. R. *Acc. Chem. Res.* **1998**, *31*, 317.
(210) Skoog, D. A.; Leary, J. J. *Principles of Instrumental Analysis*; 4th ed.; ed.; Saunders College Publishers: New York, 1992.
(211) Forster, R. J.; Diamond, D. *Anal. Commun.* **1996**, *33*, 1H.
(212) Amatore, C.; Thouin, L.; Warkocz, J.-S. *Chem. Eur. J.* **1999**, *5*, 456.
(213) Grate, J. W.; Martin, S. J.; White, R. M. *Anal. Chem.* **1993**, *65*, 940A.
(214) Sauerbrey, G. *Z. Physik.* **1959**, *155*, 206.
(215) King, W. H. *Anal. Chem.* **1964**, *39*, 1735.
(216) Wohltjen, H.; Dessey, R. *Anal. Chem.* **1979**, *51*, 1458.
(217) Carey, W. P.; Beebe, K. R.; Kowalski, B. R.; Illman, D. L.; Hirschfeld, T. *Anal. Chem.* **1986**, *58*, 149.
(218) Carey, W. P.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 3077−3084.
(219) Hirschfeld, T.; Olness, D. *Sorption Detector System for Chemical Agents Detection and Recoginition*; U.S. Army Chemical Research and Development Center:, 1984.
(220) Lorber, A. *Anal. Chem.* **1986**, *58*, 1167.
(221) Carey, W. P.; Beebe, K. R.; Kowalski, B. R. *Anal. Chem.* **1987**, *59*, 1529.
(222) Draper, N. R.; Smith, H. *Applied Regression Analysis*, 3rd ed.; Wiley: New York, 1998.
(223) Geladi, P.; Kowalski, B. R. *Anal. Chim. Acta* **1986**, *185*, 1.
(224) Carey, W. P.; Kowalski, B. R. *Anal. Chem.* **1988**, *60*, 541.
(225) Lucklum, R.; Henning, B.; Hauptmann, P.; Schierbaum, K. D.; Vaihinger, S.; Göpel, W. *Sens. Actuators A* **1991**, *25−27*, 705.
(226) Schierbaum, K. D.; Hierlmann, A.; Göpel, W. *Sens. Actuators B* **1994**, *18−19*, 448.
(227) Zhou, R.; Weimar, U.; Schierbaum, K. D.; Geckeler, K. E.; Göpel, W. *Sens. Actuators B* **1995**, *26−27*, 121.
(228) Zhou, R.; Hierlemann, A.; Schierbaum, K. D.; Geckeler, K. E.; Göpel, W. *Sens. Actuators B* **1995**, *24−25*, 443.
(229) Hierlemann, A.; Weimar, U.; Kraus, G.; Gauglitz, G.; Göpel, W. *Sensors Mater.* **1995**, *7*, 179.
(230) Hierlemann, A.; Weimar, U.; Kraus, G.; Schweizer-Berberich, M.; Göpel, W. *Sens. Actuators B* **1995**, *26−27*, 126.
(231) Schweizer-Berberich, M.; Göppert, J.; Hierlemann, A.; Mitrovics, J.; Weimar, U.; Rosenstiel, W.; Göpel, W. *Sens. Actuators B* **1995**, *26−27*, 232.
(232) Hierlemann, A.; Bodenhöfer, K.; Fluck, M.; Schurig, V.; Göpel, W. *Anal. Chim. Acta* **1997**, *346*, 327.
(233) Bodenhöfer, K.; Hierlemann, A.; Seeman, J.; Gauglitz, G.; Koppenhoefer, B.; Göpel, W. *Nature* **1997**, *387*, 577.
(234) Bodenhöfer, K.; Hierlemann, A.; Seeman, J.; Gauglitz, G.; Christian, B.; Koppenhoefer, B.; Göpel, W. *Anal. Chem.* **1997**, *69*, 3058.
(235) Bodenhöfer, K.; Hierlemann, A.; Juza, M.; Schurig, V.; Göpel, W. *Anal. Chem.* **1997**, *69*, 4017.
(236) Yano, K.; Bornscheuer, U. T.; Schmid, R. D.; Yoshitake, H.; Ji, H.-S.; Ikebukuro, K.; Masuda, Y.; Karube, I. *Biosens. Bioelectron.* **1998**, *13*, 397.
(237) Ema, K.; Yokoyama, M.; Nakamoto, T.; Moriizumi, T. *Sens. Actuators* **1989**, *18*, 291.
(238) Nakamoto, T.; Fukunishi, K.; Moriizumi, T. *Sens. Actuators B* **1990**, *1*, 473.
(239) Nakamoto, T.; Fukuda, A.; Moriizumi, T. *Sens. Actuators B* **1991**, *3*, 221.
(240) Nakamoto, T.; Fukuda, A.; Moriizumi, T. *Sens. Actuators B* **1993**, *10*, 85.

(241) Ide, J.; Nakamoto, T.; Moriizumi, T. *Sens. Actuators B* **1993**, *13*−*14*, 351.
(242) Nakamura, M.; Sugimoto, I.; Kuwano, H. *J. Intell. Mater. Syst. Struct.* **1994**, *5*, 315.
(243) Sugimoto, I.; Nakamura, M.; Kuwano, H. *Sens. Actuators B* **1996**, *37*, 163.
(244) Slater, J. M.; Watt, E. J. *Analyst* **1991**, *116*, 1125.
(245) Slater, J. M.; Watt, E. J.; Freeman, N. J.; May, I. P.; Weir, D. J. *Analyst* **1992**, *117*, 1265.
(246) Deng, Z.; Stone, D. C.; Thompson, M. *Analyst* **1996**, *121*, 671.
(247) May, I. P.; Byfield, M. P.; Lindström, M.; Wünsche, L. F. *Chirality* **1997**, *9*, 225.
(248) Yokoyama, K.; Ebisawa, F. *Anal. Chem.* **1993**, *65*, 673.
(249) Cao, Z.; Xu, D.; Jiang, J.-H.; Wang, J.-H.; Lin, H.-G.; Xu, C.-J.; Xhang, X.-B.; Yu, R.-Q. *Anal. Chim. Acta* **1996**, *335*, 117.
(250) Barkó, G.; Hlavay, J. *Talanta* **1997**, *44*, 2237.
(251) Zhang, S.; Li, S. F. Y. *Analyst* **1996**, *121*, 1721.
(252) Schmautz, A. *Sens. Actuators B* **1992**, *6*, 38.
(253) Ballantine, D. S.; Rose, S. L.; Grate, J. W.; Wohltjen, H. *Anal. Chem.* **1986**, *58*, 3058.
(254) Rose-Pehrsson, S.; Grate, J.; Ballantine, D. S.; Jurs, P. C. *Anal. Chem.* **1988**, *60*, 2801.
(255) Grate, J. W.; Rose-Pehrsson, S. L.; Venezky, D. L.; Klusty, M.; Wohltjen, H. *Anal. Chem.* **1993**, *65*, 1868.
(256) Amati, D.; Arn, D.; Blom, N.; Ehrat, M.; Saunois, J.; Widmer, H. M. *Sens. Actuators B* **1992**, *7*, 587.
(257) Reichert, J.; Coerdt, W.; Ache, H. J. *Sens. Actuators B* **1993**, *13*−*14*, 293.
(258) Hoyt, A. E.; Ricco, A. J.; Bartholomew, J. W.; Osbourn, G. C. *Anal. Chem.* **1998**, *70*, 2137.
(259) Groves, W. A.; Zellers, E. T. *Am. Ind. Hyg. Assoc. J.* **1996**, *57*, 1103.
(260) Groves, W. A.; Zellers, E. T.; Frye, G. C. *Anal. Chim. Acta* **1998**, *371*, 131.
(261) Ricco, A. J.; Crooks, R. M.; Osbourn, G. C. *Acc. Chem. Res.* **1998**, *31*, 289.
(262) Crooks, R. M.; Ricco, A. J. *Acc. Chem. Res.* **1998**, *31*, 219.
(263) Liron, Z.; Greenblatt, J.; Frishman, G.; Gratziani, N.; Biran, A. *Sens. Actuators B* **1993**, *12*, 115.
(264) Zellers, E. T.; Han, M. *Anal. Chem.* **1996**, *68*, 2409.
(265) Barié, N.; Rapp, M.; Ache, H. J. *Sens. Actuators B* **1998**, *46*, 97.
(266) Patrash, S. J.; Zellers, E. T. *Anal. Chim. Acta* **1994**, *288*, 167.
(267) Chang, S.-M.; Tamiya, E.; Karube, I.; Sato, M.; Masuda, Y. *Sens. Actuators B* **1991**, *5*, 53.
(268) Thomas, R. C.; Yang, H. C.; DiRubio, C. R.; Ricco, A. J.; Crooks, R. M. *Langmuir* **1996**, *12*, 2239.
(269) Moore, L. W.; Springer, K. N.; Shi, J.-X.; Yang, X.; Swanson, B. I.; Li, D. *Adv. Mater.* **1995**, *7*, 729.
(270) Lu, C.-J.; Shih, J.-S. *Anal. Chim. Acta* **1995**, *306*, 129.
(271) Katritzky, A. R.; Savage, G. P.; Pilarska, M. *Talanta* **1991**, *38*, 201.
(272) Patrash, S. J.; Zellers, E. T. *Anal. Chem.* **1993**, *65*, 2055.
(273) Grate, J. W.; Patrash, S. J.; Abraham, M. H. *Anal. Chem.* **1995**, *67*, 2162.
(274) Grate, J. W.; Abraham, M. H.; McGill, R. A. *Handbook of Biosensors and Electronic Noses*; Kress-Rogers, E., Ed.; CRC Press: New York, 1997.
(275) McGill, R. A.; Abraham, M. H.; Grate, J. W. *Chemtech* **1994**, *24*, 27.
(276) Grate, J. W.; Patrash, S. J.; Kaganove, S. N. *Anal. Chem.* **1999**, *71*, 1033.
(277) Slater, J. M.; Paynter, J. *Analyst* **1994**, *119*, 191.
(278) Grate, J. W.; Klusty, M.; McGill, R. A.; Abraham, M. H.; Whiting, G.; Andonian-Haftvan, J. *Anal. Chem.* **1992**, *64*, 610.
(279) Grate, J. W.; Kaganove, S. N.; Bhethanabotla, V. R. *Faraday Discuss.* **1997**, *107*, 259.
(280) Grate, J. W.; Kaganove, S. N.; Bhethanabotla, V. R. *Anal. Chem.* **1998**, *70*, 199.
(281) Zellers, E. T.; Batterman, S. A.; Han, M.; Patrash, S. J. *Anal. Chem.* **1995**, *67*, 1092.
(282) Bodenhöfer, K.; Hierlemann, A.; Noetzel, G.; Weimar, U.; Göpel, W. *Anal. Chem.* **1996**, *68*, 2210.

CR980102W

INVESTIGATION OF A TECHNIQUE THAT USES BOTH ELASTIC AND ELECTROMAGNETIC WAVES TO DETECT BURIED LAND MINES

Waymond R. Scott, Jr[(1)], Christoph T. Schroeder[(1)], James S. Martin[(2)], and Gregg D. Larson[(2)]

[(1)]School of Electrical and Computer Engineering
[(2)]School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

INTRODUCTION

A system is being investigated that uses elastic waves as the primary detection mechanism to detect buried land mines [1]. The system is shown in Fig. 1. In the system, a stationary transducer, located on the surface of the soil adjacent to the search region, generates an elastic wave in the earth. The elastic wave propagates through the search region and interacts with the buried mine. This causes both the mine and the earth to be displaced. The displacement of the mine is different than that of the earth, because the mechanical properties of the mine are different than those of the earth. The radar is used to detect these displacements and, thus, the mine. The interaction of elastic waves with buried land mines is being investigated using both numerical and experimental models.

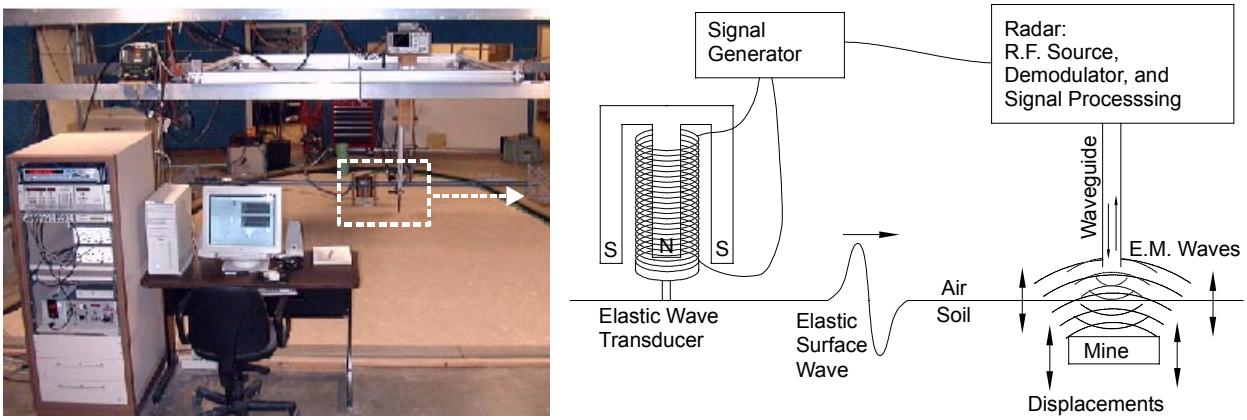


Figure 1. Photograph and schematic diagram of the experimental model.

EXPERIMENTAL AND NUMERICAL MODELS

Numerical and experimental models for the system have been constructed. The numerical model is a three-dimensional, finite-difference time-domain (FDTD) model for elastic waves traveling in the earth. The model is linear and is based on the first-order elastodynamic equations that are differenced in space and time. Pressure, shear, and Rayleigh (surface) waves are all modeled by the algorithm. A perfectly matched layer is used to absorb the waves at the edges of the model, and a free surface boundary condition is used to model the boundary between the earth and the air.

The experimental model uses an electrodynamic transducer to induce the elastic waves, a sand filled tank, a simulated mine, and a radar to measure the surface displacements. The transducer is a 20 LB moving coil shaker coupled to the sand through a narrow foot to preferentially excite surface waves. The tank is approximately 4.5 m wide, 1.5 m deep and 4.5 m long; and is filled with 50 tons of packed damp sand to simulate the earth. The radar is an 8 GHz

This work is supported in part by the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0448, by a grant from the US Office of Naval Research under contract N00014-99-1-0995, and by an equipment grant from the Intel Corporation.

continuous wave homodyne system specially designed for measuring the surface displacements. The radar is automatically scanned over the surface of the sand to measure the surface displacement field.

RESULTS

The models have been used to study the interaction of the elastic waves with a variety of different mines, simulated and inert. The results in this paper are for a simulated mine that consists of an air-filled plastic cylindrical container. The container is 9 cm in diameter and 2.2 cm in height and has thin flexible walls. This simulated mine was chosen because it is a simple structure that can be easily modeled using the current numerical code by neglecting the effects of the walls. The walls of the container are not expected to contribute significantly to the response of the mine because they are more compliant than the overlaying layer of sand. For the experiment, the simulated mine is buried in the sand and the sand is carefully re-compacted around the mine. The mine is placed 80 cm from the shaker in the center of the scan region.

A series of waterfall graphs of the displacement of the sand surface is presented in Fig. 2 for both the experimental and numerical results using two different burial depths for the simulated mine. In these graphs, the displacement is plotted as a function of time for 101 measurement points spaced in 1 cm increments away from the source in the direction of the mine. Each of the 101 time traces is shifted vertically from the previous one. The bottom trace represents the measurement point closest to the source. The region in which the mine is located is indicated in gray. Many of the discrepancies between the experimental and numerical results can be attributed to non-uniform motion of the shaker foot. The foot exhibits several sand-loaded resonances in the frequency range of interest. In addition, the depth dependence of the mechanical properties of the sand produced by the static pressure gradient is very coarsely approximated in the numerical model.

The incident pressure wave is seen to propagate toward and across the mine. The pressure wave is more apparent in the experimental data than in the model. The incident Rayleigh wave is also seen to travel toward and across the mine. Larger displacements are observed above the mine in all the data sets. These are due to a resonance of the buried mine. This resonance makes it much easier to detect the mine. In spite of the resonance, travelling waves reflected from the mine are seen to be relatively small. This indicates the difficulty that would be encountered in detecting this mine using a classical pulse-echo technique. The diameter of the mine is smaller than the wavelength of the Rayleigh wave at frequencies below 900 Hz. The resonance is spring mass like and occurs between 200 and 300 Hz. Thus, the resonance makes the mine detectable with a lower frequency seismic incident signal than would otherwise be expected. Since low frequencies attenuate more slowly in the earth, the resonance effect extends the possible search range outward from the source.

The effects of a mine resonance are not always repeatable in the experiments. This is probably due to variability in the coupling between the mine and the surrounding sand introduced when it is dug up and re-buried. Care is taken to uniformly compact the sand. However, there seems to be a long time scale cohesion of the sand that cannot be reproduced by simple wetting and compaction. Plots A and B are for burial depths of 2 and 4 cm. In both these cases, the sand in the entire scan region was tilled and repacked to make the sand more homogenous when the mine was buried. The surface displacement associated with the resonance is more pronounced for 2 cm burial depth. Plots B and C of Fig. 2 depict different results for the same mine at a 4 cm depth. The sand for case C was disturbed only above and immediately around the mine and then repacked. This approximates an actual mine burial. Case C was modeled numerically by defining a cylindrical region of earth around the mine with 20% lower wave speeds than the bulk of the medium. It can be seen from the plots that this model predicts most of the qualitative features of the data and reinforces the observation that recently disturbed volumes of soil can be more easily detected than mines. Several authors have noted this trenching effect. Interestingly, both the model and the experiment predict that the trenching effect enhances the resonant response of the mine.

Fig. 3 shows pseudo-color graphs of the displacement over the entire scan region in both the experimental and the numerical model for two different time instants. At time 1, the incident waves have not yet reached the mine. The wave fronts of the pressure and surface waves are seen to have separated in time. Small surface manifestations of head waves are discernable between these wave fronts. At time 2, both the pressure and surface incident waves have propagated beyond the mine. The circular wave fronts of scattered waves can be seen surrounding the mine location and a substantial amount of resonant motion can still be observed over the mine. Unlike the experimental data, the numerical model is able to reproduce displacements below the earth's surface. This can be seen in the cross-sectional graphs at the bottom of the figure. Here it is apparent that the surface manifestations of pressure waves have associated shear head waves propagating into the medium. It can also be seen that the dominant effect of the mine resonance is confined to the soil layer between the mine and the surface. Mode conversions can be seen to occur as both the incident pressure wave

and surface wave are scattered from the mine. There is good agreement between the model and experiment at both times depicted.

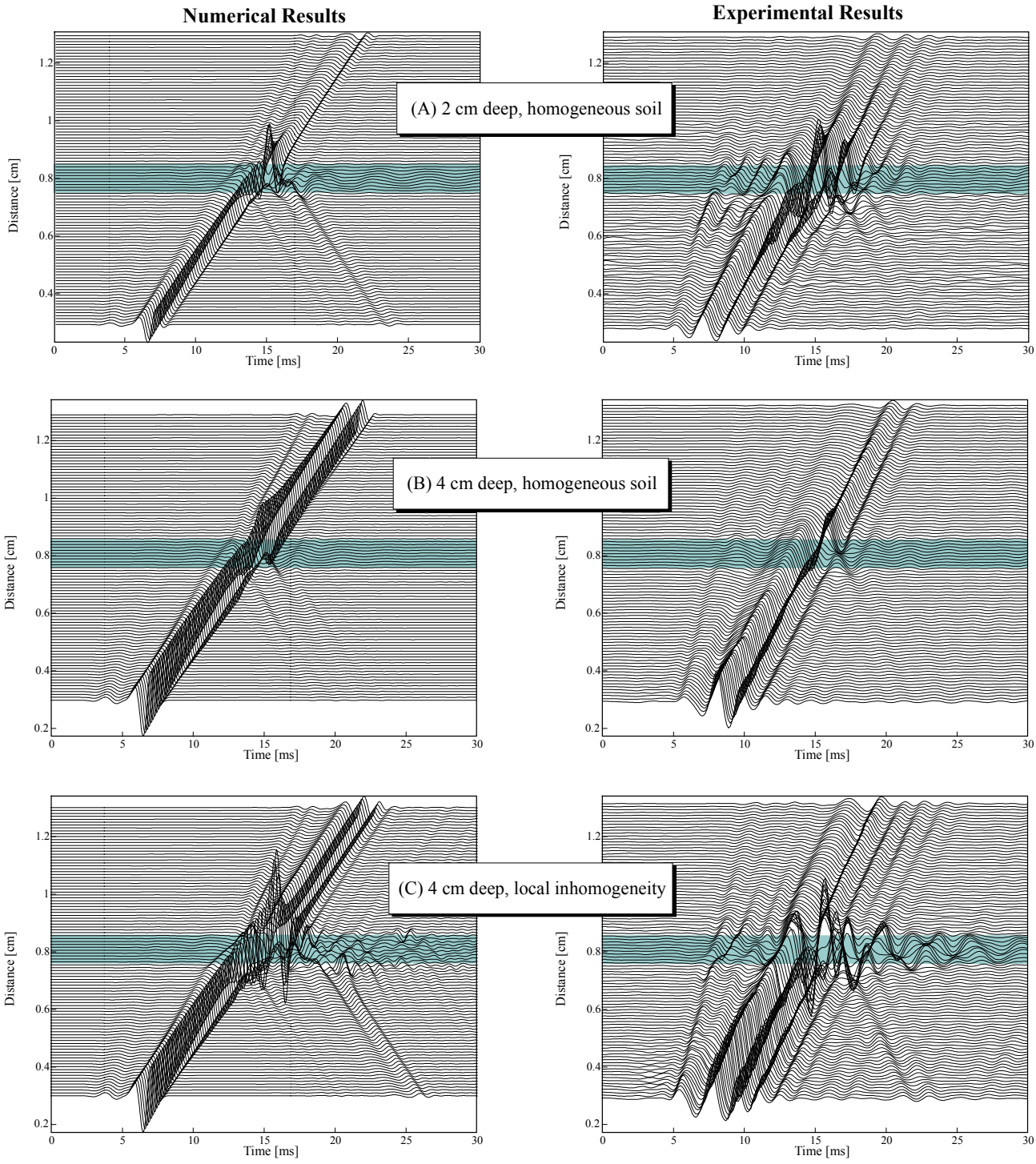


Figure 2. Waterfall graphs of the surface displacement.

CONCLUSIONS

Good agreement has been shown between experimental and numerical models for the seismic mine detection system. This agreement could be improved by eliminating resonances of the source used in the experiments and by determining the actual depth dependence of the properties of the wet compacted sand. The measured source response and depth dependence can then be incorporated into the numerical model

REFERENCES

[1] W.R. Scott, Jr. and J.S. Martin, "Experimental Investigation of the Acousto-Electromagnetic Sensor for Locating Land Mines," *Proceedings of the SPIE: 1999 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, Vol. 3710, pg. 204-14, April 1999.

[2] C.T. Schroeder and W.R. Scott, Jr., "Finite-Difference Time-Domain Model for Elastic Waves in the Ground," *Proceedings of the SPIE: 1999 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, Vol. 3710, pg. 1361-72, April 1999.

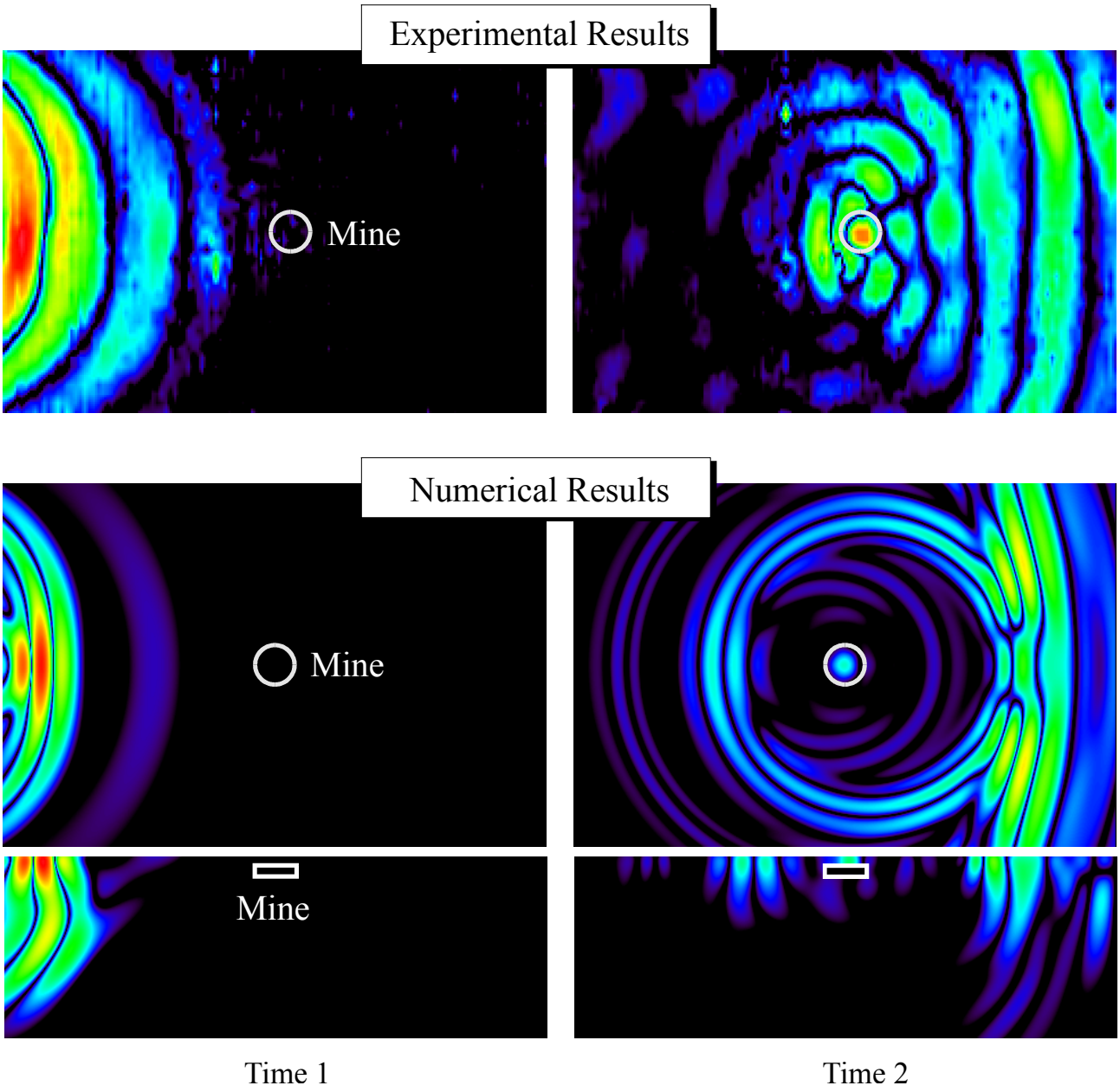


Figure 3. Pseudo-color graphs of the displacement over the entire scan region in both the experimental and the numerical model for two different time instants.

An Experimental Model of a Acousto-Electromagnetic Sensor for Detecting Land Mines

Waymond R. Scott, Jr.*
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

James S. Martin
School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0405

Introduction

A hybrid technique is presented in this paper that simultaneously uses both electromagnetic and acoustic waves in a synergistic manner to detect buried land mines. The configuration of the system currently being studied is shown in figure 1. The system consists of an electromagnetic radar and an acoustic source. An acoustic surface (Rayleigh) wave is induced in the earth by means of a stationary transducer on the surface. The transducer is placed behind the region being scanned for mines and directs the surface acoustic wave into the region being searched. The surface acoustic wave interacts with the mine and causes both the mine and the surface of the earth to be displaced. The displacement of the mine is different than the earth, because the acoustic properties of the mine are quite different than those of the earth. The displacement of the surface of the earth when a mine is present is different than when it is not present because of the waves scattered from the mine. The electromagnetic radar is used to detect these displacements and, thus, the mine. This idea has been discussed previously [1, 2], but it has not been seriously investigated.

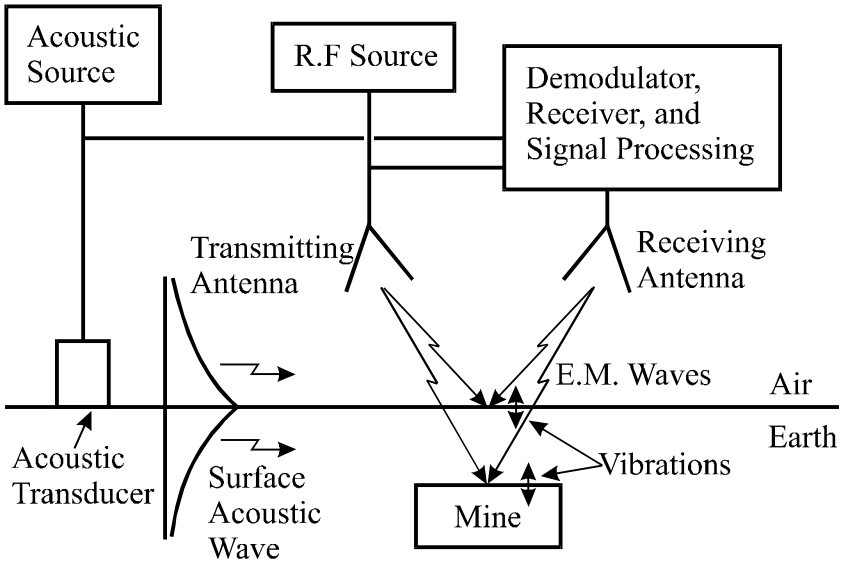


Figure 1. Acousto-electromagnetic mine detection system.

This work is supported in part by the US Army Research Office under contract DAAH04-96-1-0048.

Experimental Model

A radar has been designed and built to measure the displacements of the surface of the earth and the mine. The radar radiates electromagnetic waves that are reflected off of the surface of the earth and the mine. The reflected waves are received by the radar, and a homodyne system is used to demodulate the signals. The displacements are determined from these demodulated signals. The two biggest challenges to make this radar perform adequately for the mine detection system are 1) to make it sufficiently sensitive to be able to detect the small vibrations, and 2) to make the spot size (the area on the surface illuminated by the electromagnetic waves) sufficiently small. The radar can measure vibrations as small as 1 nm (10^{-9} m) as currently configured. To obtain this sensitivity, the radar was designed to minimize the effects of noise, such as the phase noise of the source and the electromagnetic interference from low-frequency magnetic fields. The spot size of the radar must be smaller than approximately one half of a wavelength of the acoustic waves. Currently, a small spot size is obtained by using an open-ended waveguide as the antenna for the radar. This antenna produces a sufficiently small spot size when the open end of the antenna is placed within a few centimeters of the surface. Antennas are being investigated that produce a sufficiently small spot size when they are placed farther from the surface. The radar can be configured in either a monostatic mode or in a bistatic mode as in figure 1. Currently the radar is being used in the monostatic mode to get a smaller spot size, but the radar has been used and has performed well in the bistatic mode. The radar can be operated at frequencies between 2 GHz and 8 GHz; however, all of the results presented in this work are made with the radar operating at 8 GHz. The displacements of both the surface of the sand and the mine have been measured with both the radar and an accelerometer. The measurements with the radar are good agreement with those from the accelerometer.

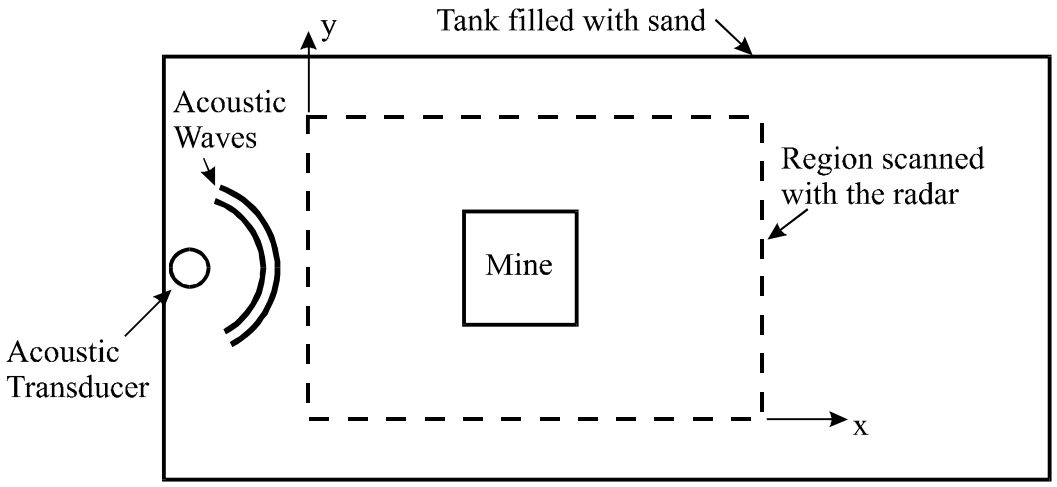


Figure 2 Top view of experimental model.

A drawing of the top view of the experimental model is shown in figure 2. The model consists of a tank that is filled with damp sand that has been packed to a relatively uniform density. The tank is approximately 120 cm wide, 120 cm deep, and 240 cm

long. A transducer is placed on the surface of the sand and is used to launch the acoustic waves into the sand. The acoustic transducer is excited with a differentiated gaussian pulse with a center frequency of 400 Hz. The transducer has been coupled to the surface so that it preferentially launches surface acoustic waves. These surface waves travel across the surface of the tank and interact with a simulated mine that is buried in the sand. The radar is used to measure the vibrations caused by the acoustic waves. An automated positioner is used to scan the radar over the surface of the sand. The region scanned (80 cm wide by 120 cm long) with the radar is also indicated on figure 2. The displacements of the surface of the sand and the mine are measured in this region as a function of time and position. The measurements are made on a uniform rectangular grid of discrete positions in the scanned region. The displacements are measured and recorded as a function of time at each of these points.

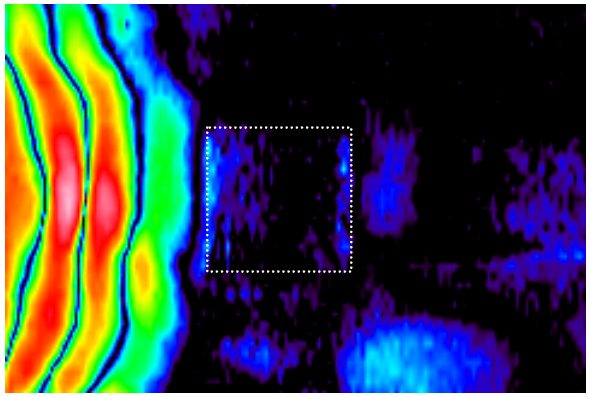
Results

Pseudo color graphs of the amplitude of the displacement of the surface are presented in figures 3 for four times. These results are for a simulated antitank mine buried in the sand. The simulated anti-tank mine made out of Acrylic plastic and is 30 cm wide, 30 cm long, and 7.5 cm height. The top of the mine is flush with the surface of the sand. The position of the mine is indicated by the dotted white line. In these graphs, the color scale goes from black to blue to green to yellow to red to white. The smallest displacements are in black and the largest displacements are in white. At time #1, the acoustic wave is seen traveling toward the mine. Note that the shape of the wave is spread out and more complex than the differentiated gaussian pulse that is the excitation for the transducer. This is partially due to the manner in which the wave is injected into the sand. Methods for injecting a more time-limited pulse are being investigated. At times #2 and #3, the wave has reached the mine, and a portion of the wave is going through the mine while the rest is going around the mine. Notice that the portion of the wave that is going through the mine is ahead of the portion that is going around the mine. This is because the mine is much stiffer than is the sand; thus, the wave appears to travel faster through the mine. The mine actually moves as if were a rigid body. The motion of the mine is essentially a rocking motion excited by the wave motion in the sand. At time #4, the wave is seen to be still going around the mine; however, a significant portion has passed through the mine. Notice that the wave that went through the mine is still ahead of the wave that went around then mine, and notice that the wave that went through the mine is smaller in amplitude than the one that went around the mine. Also notice that the displacements are smaller above the mine. The waves that are reflected from the mine can also be seen. The location and the shape of the mine are clearly evident in these graphs. Measurements have also been made with the anti-tank mine buried deeper and with simulated anti-personnel mines. The effects of the mine can also be seen in these measurements.

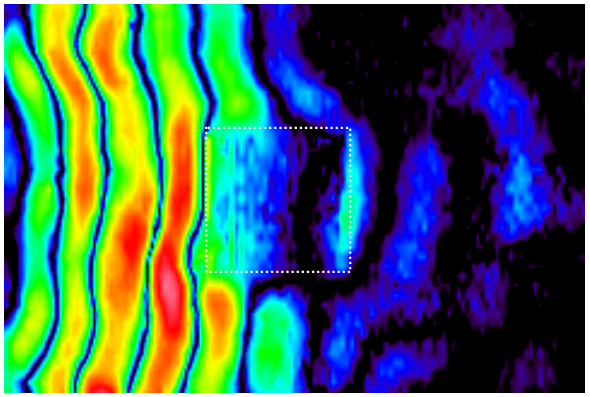
References

[1] G.S. Smith, “Summary Report: Workshop on New Directions for Electromagnetic Detection of Non-Metallic Mines,” Report for U.S. Army BRDEC and ARO, June 1992

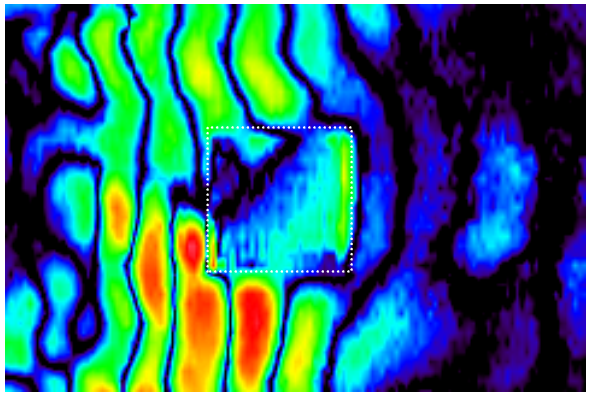
[2] C. Stewart, *Summary of Mine Detection Research*, Vol. I, pp. 172-179, Tech. Report 1636-TR, May 1960, U.S. Army Engineering Res. and Devel. Labs, Corps. of Eng., Belvoir, VA.



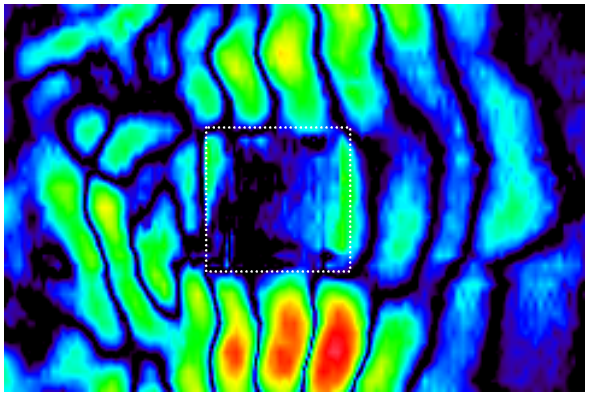
Time #1 – Acoustic pulse traveling toward the mine



Time #2 – Pulse interacting with the mine. The pulse is seen to travel faster through the mine.



Time #3 – Pulse interacting with the mine. The pulse is seen to travel faster through the mine.



Time #4 – Pulse is reflected from, transmitted through, and travels around the mine.

Figure 3 Pseudo color graphs of the amplitude of the displacement of the surface of the sand for four times when a simulated antitank mine is buried in the sand.

Environmental Influence on Microwave Radiometry for Buried Object Detection

David R. Wiggins, Baran U. Ungan, and Joel T. Johnson

ElectroScience Laboratory, The Ohio State University
1320 Kinnear Rd., Columbus, OH 43212
Phone: (614) 292-7981, Email: wiggins.38@osu.edu

Introduction

Microwave radiometers have been shown to be effective sensors for monitoring soil moisture and other geophysical data [1]-[4]. Models for geophysical medium brightness temperatures often consider horizontally stratified geometries, because exact evaluation of observed brightness temperatures is possible for this case in terms of a sum involving layered medium reflection quantities [2]. For cases involving media whose temperature is constant, the formulation reduces to "Kirchhoff's law" which relates emissivity to one minus the reflection coefficient of the layered medium. Previous studies of microwave radiometry for soil moisture remote sensing have considered both temperature and soil moisture variations with depth, and shown that these environmental effects can have important influences on soil moisture retrievals [3].

Recent studies have begun to consider the use of microwave radiometers for detecting shallow, sub-surface objects such as anti-personnel landmines [5]-[6]. Modeling studies in these references based on Kirchhoff's law and a three layer, horizontally stratified geometry (i.e. objects are infinite layers) show that significant brightness temperature contrasts can be obtained in the presence of a sub-surface object if sufficient dielectric contrast is available and if soil medium attenuation is not excessive. Reference [6] further demonstrated that use of multi-frequency brightness measurements could potentially provide detections even with low contrast or high attenuation due to the oscillatory behavior versus frequency observed in the presence of an object. Estimated environmental effects such as local surface temperature or soil moisture variations would not produce oscillatory frequency behavior (except in unusual circumstances [4]) so that detections would still be possible even in the presence of environmental "clutter".

However, previous buried object detection models have neglected variations in medium temperature (due to the use of Kirchhoff's law) and soil moisture (which was assumed constant) with depth. Since these factors can significantly impact soil mositure remote sensing, their effect on buried object detection requires consideration as well. In this paper, a multi-layer horizontally stratified emission model is coupled with a numerical solution of the heat equation and a model for water transport in the presence of a sub-surface object to estimate the importance of temperature and soil moisture variations. Results show that these factors can impact overall brightness temperatures, but that the concept of using oscillatory features in frequency swept data as an indicator of a sub-surface object remains valid.

Moisture, Thermal, and Emission Models

Figure 1 illustrates the geometry considered: a 6.8 cm thick subsurface object (modeled as a layer of constant relative permittivity $3 + i0.08$) is located 4.25 cm below an air-soil boundary. Both permittivity $\epsilon(z)$ and temperature $T(z)$ are not constant in the soil background medium, and the temperature also varies in the object layer. Specification of the medium begins with a moisture transport model [7] for evaluating soil volumetric moisture contents in the presence or absence of a sub-surface object of dimensions $25.2 \times 25.2 \times 6.8$ cm. Water transport is computed on a grid $73.5 \times 73.5 \times 59.5$ cm in the presence of a periodic "rain" forcing function. Further details of the simulation are provided in [8]. Figure 2, plot (a) illustrates the resulting moisture profiles versus depth with (taken from directly

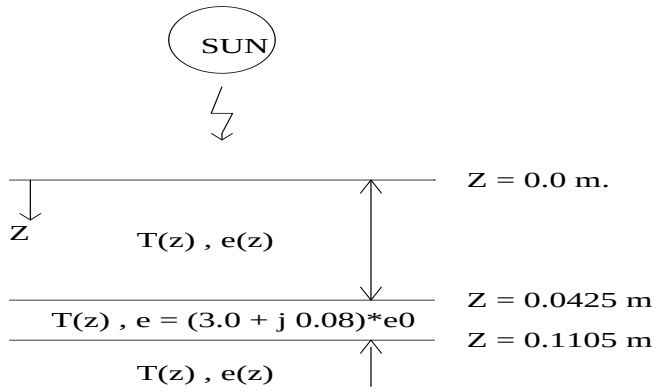


Figure 1: The stratified medium model

underneath the object) and without (taken from one of the faces of the computational domain) the object; note the increased and decreased water contents above and below the object respectively. Moisture contents versus depth then directly determine the permittivity profile $\epsilon(z)$ through the empirical soil permittivity model of [1] evaluated for sand with bulk density 1.12 g/cm^3. Obtained permittivity values ranged from $5.44 + i0.46$ to $12.74 + i2.15$; note the significant contrast with the subsurface object, although high attenuation is also obtained for this case.

Moisture contents are also used to determine the thermal conductivity and diffusivity parameters of the soil medium versus depth from the model of [9]. A one dimensional solution of the heat equation with or without the object layer is then used to determine the temperature profile versus depth as described in [10]. Parameters for solar forcing, wind convection, and other effects are as described in [10]. The heat equation is solved on a 100 layer grid up to depth 50 cm, beyond which the temperature is assumed constant. Figure 2, plot (b) illustrates typical temperature profiles obtained with and without the subsurface object at local time 15 : 42. In these simulations the subsurface object is modeled thermally as a near-insulator, allowing the relatively large temperature difference through the object layer to be maintained.

Finally, given the $\epsilon(z)$ and $T(z)$ profiles versus depth, brightness temperatures are computed from the multi-layer fluctuation-dissipation model of [2]. This model includes all potential coherent emission effects, and again is evaulated on the 100 layer grid up to depth 50 cm. Results to be illustrated consider nadir observation, so that polarization effects can be neglected. Results are compared with computations which neglect either temperature variations (by assuming the entire medium is at the surface temperature) or permittivity variations (by assuming the soil medium is all of uniform permittivity) to clarify the influence of these factors.

Results

Figure 3 compares brightness temperatures from the complete model with the object versus those assuming $T(z) = T_s = 333\ K$ or $\epsilon(z) = \epsilon$ (computed for 27% soil moisture). Results are plotted as a function of radiometer frequency from 1 to 5 GHz, and the oscillatory pattern versus frequency obtained in the presence of a subsurface object is observed. The complete model with no object is also included, and shows no strong oscillations versus frequency. For this relatively high attenuation case, the influence of temperature variations is observed to be small, since greater depths at which temperature changes are larger are not observed by the radiometer. Permittivity variations are found to have a larger effect, but the basic concept of searching for oscillatory features for object detection remains valid.

To illustrate an additional case with lower attenuation, the moisture profiles of Figure 2 are divided by 4 to obtain an average moisture content of 6.7%, and the thermal and

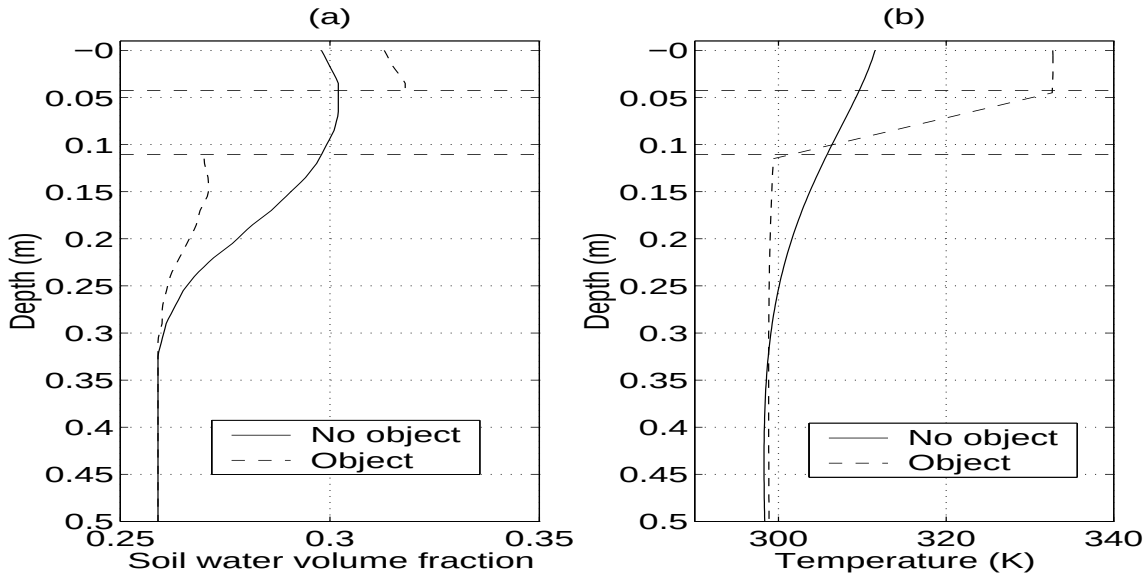


Figure 2: Moisture (a) and temperature (b) profiles with and without object

emission models re-computed. Results are illustrated in Figure 4, and show smaller oscillations due to the reduced dielectric contrast but a greater influence of both temperature and permittivity profile variations. However, the basic features with and without the subsurface object again remain similar.

The results of this study confirm that temperature and moisture profiles can influence brightness temperatures, but that these factors are unlikely to produce the oscillatory features caused by a subsurface object.

References

[1] F. T. Ulaby, R. K. Moore and A. K. Fung, *Microwave Remote Sensing: Active and Passive,* Norwood, MA: Artech House, 1986.

[2] L. Tsang, J. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing,* New York: John Wiley and Sons, 1985.

[3] Y.-A. Liou and A. W. England, "A land surface process/ radiobrightness model with coupled heat and moisture transport in soil," *IEEE Trans. Geosc. Rem. Sens.*, vol. 36, pp. 273–286, 1998.

[4] T. J. Schmugge, T. J. Jackson, P. E. O'Neill, and M. B. Parlange, "Observations of coherent emissions from soils," *Radio Sci.*, vol. 33, pp. 267-272, 1998.

[5] G. De Amici, B. Hauss, and L. Yujiri, "Detection of landmines via a passive microwave radiometer", *SPIE Aerosense meeting: Detection and Remediation Technologies for Mines and Mine-Like Targets IV,* conference proceedings, 1999.

[6] J. T. Johnson, "Theoretical study of microwave radiometry for buried object detection," *SPIE Aerosense meeting: Detection and Remediation Technologies for Mines and Mine-Like Targets V,* conference proceedings, 2000.

[7] J. Simunek, K. Huang, and M. van Genuchten, "The SWMS-3D code for simulating water flow and solute transport in three dimensional variably saturated media," *Research Report No. 139,* U. S. Salinity Laboratory, USDA, 1995.

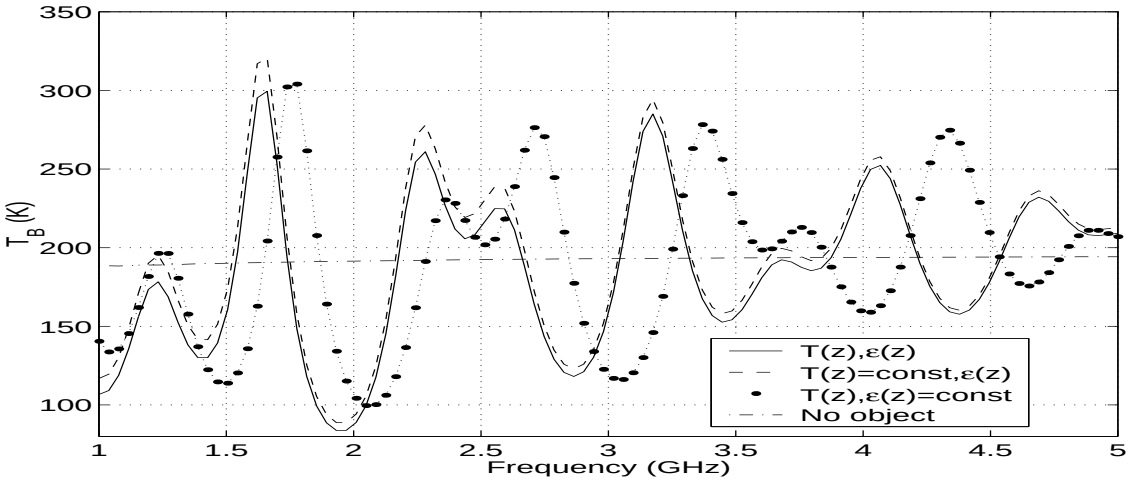


Figure 3: Brightness temperatures (T_B) vs. frequency for the high moisture case

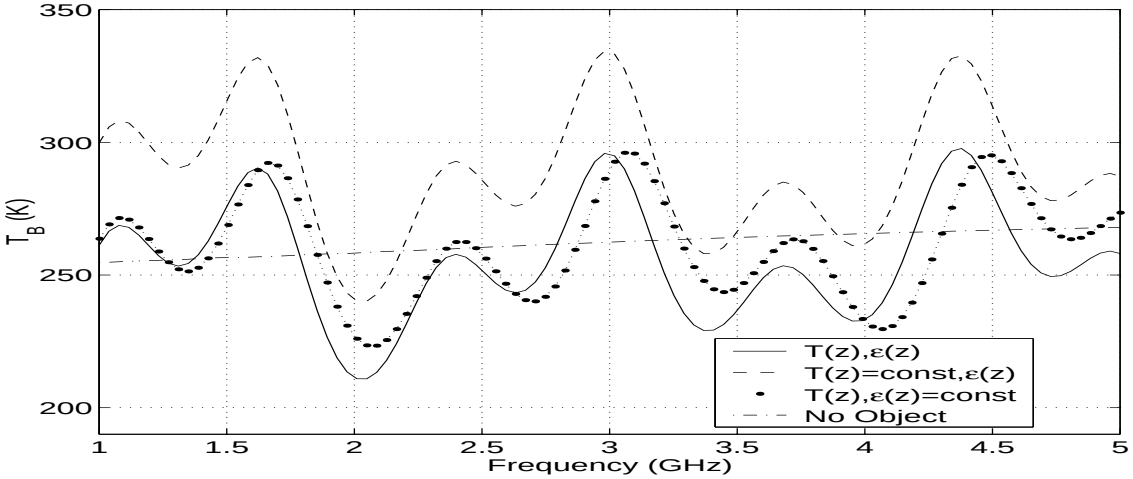


Figure 4: Brightness temperatures (T_B) vs. frequency for the reduced moisture case

[8] B. A. Baertlein and I. K. Sendur, "The role of environmental factors and mine composition in thermal IR mine signatures," submitted to *SPIE Aerosense meeting: Detection and Remediation Technologies for Mines and Mine-Like Targets VI,* 2001.

[9] P. Pregowski, W. Swiderski, R. T. Walczak, and K. Lamorski, "Buried mine and soil temperature prediction by numerical model,' *SPIE Aerosense meeting: Detection and Remediation Technologies for Mines and Mine-Like Targets V,* conference proceedings, 2000.

[10] I. K. Sendur and B. A Baertlein. "Techniques for improving buried mine detection in thermal IR imagery," *SPIE Aerosense meeting: Detection and Remediation Technologies for Mines and Mine-Like Targets IV,* conference proceedings, 1999.

4:00

4pMU7. The acoustics of a symmetric free reed coupled to a pipe resonator. James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402)

The Asian free-reed mouth organs employ symmetric free reeds mounted in resonating pipes, with the reed vibration strongly coupled to the pipe resonance. The sheng, the sho, and the khaen use a single pipe for each reed, constructed so that the pipe resonance frequency is fairly close to the natural frequency of the reed. The playing frequency is typically slightly above the resonant frequencies of both the reed and the pipe. The free-reed pipe with finger holes (known in China as the bawu) employs a pipe resonator of variable effective length in which both the pipe resonance and sounding frequency are normally well above the natural reed frequency, resulting in a striking change in tone quality. The operation of these instruments has been studied both experimentally and theoretically, with particular attention to the coupling of the reed vibration with the pipe resonator. Experimental measurements include both studies of reed vibration and impedance measurements of the pipes. In general, the experimental results can be shown to agree with the predictions of simple theoretical models.

4:15

4pMU8. The influence of upstream geometry on the activation pressure of free and restricted reed configurations. Marius O. Vermeulen, Jabus A. Wessels (Dept. of Medical Physiol., P.O. Box 19063, Tygerberg, 7505, Cape Town, South Africa, jaw@gerga.sun.ac.za), and Theodore W. von Backstrm (Univ. of Stellenbosch, Matieland, South Africa)

This paper describes a computationally efficient algorithm for the automatic calculation of reed activation pressure. The method was evaluated using both free and restricted reed configurations which were not connected to any external air column. It is shown how the reed activation is influenced by intrinsic factors such as rounded and minute burrs on the edge of the reed as well as extrinsic factors such as upstream geometry of the experimental apparatus. The introduction of constrictions in the flow path upstream from the reed resulted in marked changes in the activation pressure which was highly dependent on the length of the constriction. Expansions of the same magnitude, however, had less influence on the reed activation pressure.

4:30

4pMU9. A complexity measure for musical scales. Alpar Sevgen (Dept. of Phys., Boğaziçi Univ., Bebek 80815, Istanbul, Turkey, sevgena@boun.edu.tr)

Equally tempered scales with N semitones and M notes and interval structures $\mathbf{n}=\{n_1,n_2,\ldots,n_M\}$, where n_k is the number of semitones between the notes t_k and t_{k+1}, possess the following properties: Each distinct interval structure $\mathbf{n}$ corresponds to a multiplet of N scales. Members of a multiplet can be labeled by a set of integers $\{c\}$, modulo N, called scale labels. Each scale label is the *difference* between the number of sharps and flats occurring in that scale and is unique within the multiplet if N and M are relative primes. This labeling does not differentiate between different scale structures.To do this, *complexity* is introduced as the *sum* of the number of sharps and flats occurring in a scale. For $N=12$ and $M=7$, out of 462 possible scale structures, the major scale and its cyclical permutations, called modes, have the minimum complexity which allows the practical use of the key signatures in music. Complementary scales where notes and no notes are interchanged have the same complexity. The minimum and maximum complexity scales occupy the opposite ends of the energy spectrum under the force laws $\pm n^{\alpha}$ $(\alpha\neq 0)$, between the notes of a scale.

FRIDAY AFTERNOON, 2 JUNE 2000 SPANISH ROOM, 1:30 TO 5:15 P.M.

Session 4pPA

Physical Acoustics: Outdoor Sound Propagation and Acoustic Seismic Coupling

Gregg D. Larson, Chair
School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

Contributed Papers

1:30

4pPA1. Systematic investigation on acoustic-to-seismic responses of landmines buried in soil. James M. Sabatier and Ning Xiang (Natl. Ctr. for Physical Acoust., Coliseum Dr., Univ. of Mississippi, University, MS 38677, sabatier@olemiss.edu)

Recently, acoustic-to-seismic coupling has been successfully applied to landmine detection [Sabatier and Xiang, J. Acoust. Soc. Am. **105**, 1383 (1999); **106**, 2143 (1999)]. When airborne sound penetrates the surface of ground it is refracted towards the normal. If a landmine is buried below the surface of an insonified patch, the transmitted waves will be scattered or reflected, resulting in increased ground surface vibrational amplitudes. These distinct acoustic-to-seismic coupled vibrational changes are sensed using a scanning laser Doppler-vibrometer (LDV) device. To better understand this mine detection phenomenon, the present work is a systematic investigation of the acoustic-to-seismic response to different types of mines in different soil types and at different burial depth has been conducted. [This work is supported by U.S. Army Communications-Electronics Command.]

1:45

4pPA2. Air acoustic sensing of seismic waves. Gregg D. Larson, James S. Martin (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405), Waymond R. Scott, Jr., and Cheng Jia (Georgia Inst. of Technol., Atlanta, GA 30332)

Propagation of elastic waves in damp, compacted sand involves pressure, shear, and Rayleigh waves. The associated seismic surface displacements can be detected by sensing the acoustic pressure immediately above the surface. Propagation speeds are very low in sand. The high wave numbers of seismic displacements are, therefore, evanescent in air. Thus, the acoustic pressure can only be measured well within a seismic wavelength of the surface. Planar near-field acoustic holography techniques can then be used to back-propagate these signals and calculate surface displacements. Measurements have been made using a laboratory experimental model to investigate the potential of using this technique to detect buried land mines. The experimental model utilizes a surface-coupled transducer to generate elastic waves in a sand-filled tank, which simulates the earth. The microphone and a radar system were used to independently

measure the surface displacements. Data taken with both sensors compare well and exhibit the signature of a buried inert antipersonnel mine. For a 100–800-Hz incident pulse, the mine signature can be seen in the raw microphone data when the height of the microphone is less than 3 cm. Holographic signal-processing techniques will be investigated to increase the allowable height for the microphone. [Work supported by ARO.]

2:00

4pPA3. Electric arc source for high-frequency seismic measurement. James S. Martin, Gregg D. Larson, Peter H. Rogers (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0405, james.martin@me.gatech.edu), and Waymond R. Scott, Jr. (Georgia Inst. of Technol., Atlanta, GA 30332-0250)

An electric arc source was designed to study high-frequency seismic surface wave propagation. The noncontact nature of this source made it feasible for use in synthetic aperture transmit arrays. The transmit signal, which was not linearly controllable, was found to be predominantly in the 1- to 4-kHz band with a Gaussian spectrum. This is an octave below the simultaneously generated air acoustic pulse. The source was used to create a synthetic line array in conjunction with a stationary receiver. The experiment was conducted in a sand-filled tank. Surface wave speeds in the range of 80 m/s were measured on the resulting seismograms. Significant dispersion occurred in the propagating waveform at distances both near and far from the source. Surface wave arrivals were discernable over 1 m from the source and compressional head waves could be observed within 30 cm. The data were in good agreement with lower frequency measurements made by other techniques. The surface wave generation was studied and found to be a combination of the surface interaction of the arc and the air acoustic interaction. Strong hysteresis was observed in the first arcing event. Later, the signal was smaller but sufficiently stable for averaging. [Work supported by ARO.]

2:15

4pPA4. Seismic/electromagnetic system for landmine detection. Waymond R. Scott, Jr. (School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0250), Gregg D. Larson, James S. Martin, and Peter H. Rogers (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

A system has been designed for the detection of buried landmines. The system uses a stationary seismic source in conjunction with a movable displacement sensor that is based on an 8 GHz CW radar. The sensor measures the surface displacement by analog demodulation of the radar signal, which is reflected from the soil surface and modulated by the surface motion. The sensor is not in direct contact with the soil surface and is, therefore, capable of interrogating surface motion immediately above a buried mine. This configuration provides the dual advantage of removing half the seismic propagation path that would be encountered with a classical pulsed echo technique and detecting localized fields that would not propagate to a remote receiver. The system has been used in the laboratory to image inert antipersonnel mines and simulated antitank mines buried in damp compacted sand. Signal processing in the wave number domain provides significant improvement in the contrast between mine-related and background motion. The simplest detection cue for antipersonnel mines was found to be low-frequency resonances of their trigger mechanisms. These responded to seismic excitations with substantial local displacement. The resonances made these mines easily discernable from buried clutter such as rocks and sticks. [Work supported by ARO.]

2:30

4pPA5. A three-dimensional model for elastic waves in the ground. Christoph T. Schroeder, Kangwook Kim, and Waymond R. Scott, Jr. (School of Elec. and Computer Eng., Georgia Inst. of Technol., 777 Atlantic Dr., Atlanta, GA 30332, christoph.schroeder@ee.gatech.edu)

A three-dimensional finite-difference time-domain model for elastic waves in the ground has been developed and implemented on a massively parallel computer. The model is based on the three-dimensional equation of motion and the stress–strain relation, from which a first-order stress-velocity formulation is obtained. The boundary between the soil and the air is modeled as a free surface. A perfectly matched layer is implemented at the remaining grid edges to absorb the outward traveling waves. The numerical model has been developed as part of a project in which elastic and electromagnetic waves are used synergistically to detect buried landmines. The numerical model is being used to study the interaction of the elastic waves with the buried mines. To verify that the model accurately predicts the mine–wave interaction, the eigenfrequencies of various solid bars and plates are determined numerically and compared to analytical solutions. Currently, the model is being refined to incorporate loss within the bulk medium. Results will be shown for various landmines buried in both loss-free and lossy ground. [Work supported by ARO and ONR.]

2:45

4pPA6. Detection of land mines in fluid-saturated unconsolidated soil: Numerical modeling. Yanqing Zeng and Qinghuo Liu (Dept. of Elec. and Comput. Eng., Duke Univ., Durham, NC 27708)

Because of the strong interactions of waves with the solid grains and the fluid in the pore space, it is more appropriate to model the soil as a fluid-saturated unconsolidated material than a single-phase elastic medium. A three-dimensional finite-difference method for modeling acoustic waves propagating in fluid-saturated unconsolidated soil has been developed. Instead of the conventional elastic wave equations, Biot's equations are used for the poroelastic model. Based on the strain–stress relationship, Biot's equations are reformulated into a first-order hyperbolic system which is equivalent to strain-velocity formulation. A perfectly matched layer (PML) is used to absorb outgoing waves at the truncated boundary of an unbounded medium. The numerical method is validated by comparing numerical results to an analytical solution. Models of a land mine buried in fluid-saturated unconsolidated soil are developed. The numerical method is used to study the interaction of acoustic waves with the buried mines. Comparison of these results is made with those for a buried land mine in a conventional single-phase elastic soil model.

3:00–3:15 Break

3:15

4pPA7. Acoustically induced slow dynamics in nonlinear mesoscopic elastic materials. Alexander M. Sutin (Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030), Paul A. Johnson, and James A. TenCate (Los Alamos Natl. Lab., Los Alamos, NM 87545)

We have known about slow dynamics in rock due to continuous wave excitation drive for several years (http://www.ees4.lanl.gov/nonlinear). TenCate, Smith, and Guyer (see abstract, this meeting) have recently discovered that both the elastic modulus and the wave dissipation display log time recovery in granular solids, and that it may be thermally or mechanically induced. Much to our surprise, we have discovered that a CW or broad-frequency band acoustic source can also induce slow dynamical response. This response was observed as a variation of the ultrasonic probe wave amplitude, the resonance frequency, and Q factor after the action of a pump wave. The slow time recovery took place in materials such as powdered metals, damaged materials, concrete, and rocks as well. The variations of material properties due to the action of pump waves lead to transient amplification and an obscuration of CW probe waves. The observed behavior may be more universal than was first thought. The results have potential implications to many topics, including laboratory wave studies, earthquake strong ground motion, elastic waves emanating from a point source, damage detection, and manufacturing processes. [Work supported by Stevens and by the Department of Energy: Office of Basic Energy Sciences.]

Signal Processing of Elastic Surface Waves for Localizing Buried Land Mines*

A. Behboodian, W. R. Scott, Jr. and J. H. McClellan
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0250
behbood@ece.gatech.edu

Abstract

A system that uses elastic surface waves and electromagnetic waves has been developed to detect buried land mines. In this paper an algorithm based on local spatial frequency analysis is presented to illustrate the dispersive effect of the land mine on the acoustic wave and to estimate the wave velocity as a function of position and temporal frequency. In addition, an imaging algorithm is presented which calculates the reflected energy and generates a two-dimensional image that localizes the buried mines.

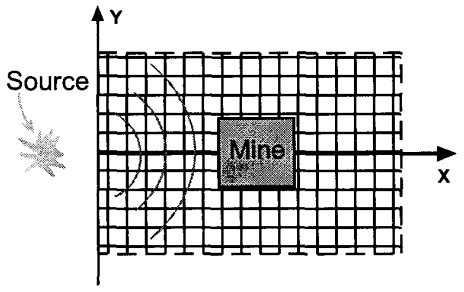


Figure 1. Top view of the experimental system.

1. Introduction

A laboratory system which uses both elastic and electromagnetic waves has been designed and constructed to locate buried mines, metallic as well as nonmetallic [3, 2]. The top view of this experimental system is shown in Fig. 1. An elastic source induces elastic waves into the ground causing both the earth and the mine to vibrate. An electromagnetic radar, designed to measure surface displacements as small as 1 nm, is used to detect these vibrations. The system takes samples on a uniform rectangular grid of discrete positions. The grid consists of 41 points in the y direction spaced 2cm apart and 121 points in the x direction spaced 1cm apart.

The characteristics of the vibrations will differ with the proximity to the mine. Mines have mechanical properties that are significantly different from soil and typical sources of clutter. The shear wave velocity is approximately 10 times higher in mines than in the ground. In addition, the structure of the mine gives rise to a structural resonance not found in other forms of clutter. In this paper signal processing algorithms have been developed to exploit these and other physical phenomena to locate buried mines. The velocity of the elastic wave is estimated locally and depth information is extracted from the data collected over the surface. By exploiting both the elastic wave reflected from the mine and the mechanical resonances induced in the mine, an imaging algorithm was developed which can localize plastic mines as small as 8cm in diameter. These plastic mines cannot be reliably detected with any of the existing methods.

*This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0048

2. Local Wave Velocity Estimation

An important physical property of the elastic surface waves is that waves with lower temporal frequencies penetrate deeper into the ground than waves with higher temporal frequencies [1]. Also the velocity of the elastic wave is higher in the mine than in the surrounding soil. Hence one can expect that waves with low temporal frequencies which penetrate deep enough to hit the buried mine travel faster in the proximity of the mine than waves with higher temporal frequencies. This is a dispersive effect of the buried mine on the elastic waves.

To compute a local estimate of velocity, we should first decompose the collected data, $s(x, y, t)$, into different temporal frequencies. This is done by Fourier transforming the

0-7803-5700-0/99/$10.00©1999 IEEE

data in the time domain. The resulting data set, $S(x, y, f)$, which depends on the spatial domains, x, y, and on the temporal frequency f is then processed separately for each temporal frequency. Depth information can be inferred from this transformed data because penetration depth is proportional to temporal frequency. However, note that while the higher temporal frequencies get affected only by the upper levels of the ground, the lower temporal frequencies get affected by both the upper levels and the lower levels of the ground. Hence the velocity which we determine for the lower temporal frequencies is not the velocity at a particular depth but a weighted average of the depth dependent velocity profile. The average velocity is frequency dependent because the penetration depth of the surface wave is frequency dependent.

To show the dispersive effect of the mine we process a slice of the transformed data, $S(x, y, f)$, at a temporal frequency of 430Hz. The wave at this temporal frequency penetrates deep enough to reach the mine which is buried 5cm under the surface. A spatial-spatial frequency analysis is performed on $S(x, y, 430)$ using a two-dimensional, sliding, rectangular window of 6cm height and 3cm width (this is 3×3 sample window). The two-dimensional Fourier transform of the windowed data is taken. The location where the maximum of the absolute value of the Fourier transform occurs is an estimate of the dominant wave number in the windowed region. Knowing the temporal frequency, one can find the localized velocity in that region by dividing the temporal frequency in rad/s by the magnitude of the wave number in rad/m. Fig. 2 shows the processing result of a data set collected over a $30 \times 30 \times 10$cm simulated mine, buried 5cm deep under the ground. The arrows in the figure depict the wavenumber vector. Only the forward going waves are shown. In this figure the source is located approximately at $(x = -30, y = 0)$. It can be seen that in the area between the source and the mine the wave propagates almost radially from the source. Over the mine the velocity of the wave increases and hence the magnitude of the wavenumber vectors become smaller. After passing the mine the wave fill in the region behind the mine. Some perturbations at the edges of the figure are mainly due to reflections of the wave off the walls of the tank in which the mine is buried.

In another processing, a two-dimensional slice of the data over the line $y = 0$ in Fig. 1 (this is $s(x, 0, t)$) was analyzed for depth information. First, we take the Fourier transform of the data in the time domain. At each temporal frequency, the spectrogram of the data is calculated in the spatial domain with a Hanning window of length 15cm (a length of 15 samples) with an overlapping of 14cm. At each window location, the spatial frequency where the maximum of the magnitude of the Fourier transform occurs is the dominant wavenumber. The velocity can hence be calculated

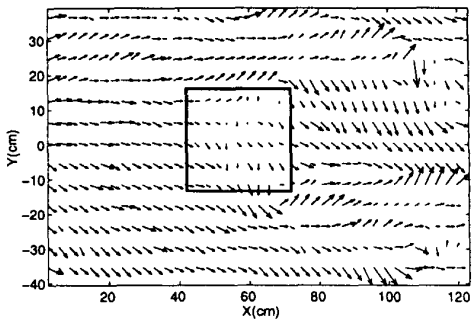


Figure 2. Wavenumber vectors at a temporal frequency of 430Hz over a buried anti-tank mine.

knowing the temporal frequency. Results over a range of temporal frequencies is shown in Fig. 3 for two different data sets. In Fig. 3(a) a simulated $30 \times 30 \times 10$cm anti-tank mine is buried 5cm under the ground whereas in Fig. 3(b) the same mine is buried 10cm under the ground. It can be seen from the figure that waves with high temporal frequencies are less affected by the presence of the mine. This is because these waves do not penetrate as much in the ground to reach the mine. Waves with medium temporal frequencies get affected the most and their velocity increases sharply over the mine. Finally, it is seen that waves with low temporal frequencies are also less affected by the presence of the mine. The reason is two fold. On one hand, we know that these waves have large wave lengths and hence have less interaction with the mine. On the other hand, these waves penetrate much deeper into the ground. Hence the presence of a 10cm thick mine does not affect the velocity of the low frequency waves as much as it affects the velocity of the medium frequency waves. Note in the figure that higher range of frequencies get less affected by the deeper mine.

3. Imaging Algorithm

In addition to processing the data to estimate the velocity of the elastic wave and infer depth information from the data, we have also developed an algorithm to image the data. The algorithm exploits two basic properties at the same time. The first property is that incident elastic wave, traveling at the surface of the ground gets reflected back upon arriving at a buried object. This property is most useful for the case of large mines. The reflection off the small mines such as anti-personnel mines is small and hard to distinguish from the background noise. Another fundamental property of mines is their structural resonance. Elastic

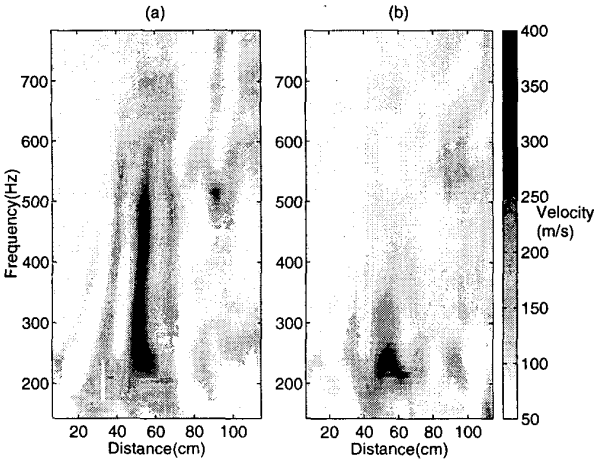


Figure 3. Temporal-Spatial analysis of collected data over y = 0. (a) 5cm deep. (b) 10cm deep.

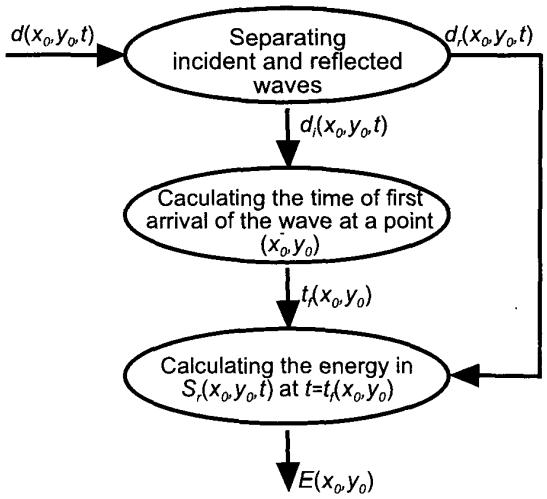


Figure 4. The three steps involved in formation of an image.

waves, upon arriving at the mine, cause the mine to resonate [2]. The resonance not only enables us to locate small mines, but also gives us the ability to distinguish between clutter such as stones and small mines. An explanation of how the imaging algorithm works comes next.

One way to define a reflector is that it is located at a point where the first arrival of the incident wave is time coincident with the reflected wave. The resonance also starts at the first arrival of the wave to the mine. The basic idea behind our algorithm is to calculate the time it takes for the wave to get to a point on the surface and then calculate the energy in the reflection and resonance near that point. To exploit this idea we proceed in three steps depicted in Fig. 4. In the first step, we separate the incident and reflected waves. The separation process can readily be explained in the $\omega - k$ domain. In two dimensions, the incident wave has energy in the first and third quadrants in the $\omega - k_x$ domain while the reflected wave has energy in the second and fourth quadrant. Fig. 5 shows the frequency transform of the data collected over the line $y = 0$ in Fig. 1. Obviously the energy of the reflected wave is less than the energy of the incident wave which is also seen from the figure. The first and third quadrants are filtered out leaving us with the reflected wave.

In the second step, we find the time $t_f(x, y)$, when the incident wave first arrives at a point (x, y) on the surface. This requires us to have an estimate of the wave velocity. Here, we use an estimate of the bulk velocity by simply measuring the slope of the traveling wavelet in the $x - t$ domain of the collected data. The $x - t$ domain data is shown in Fig. 6 with a line fitted to the incident wave. The processes involved in the last step of the imaging algorithm are shown in Fig. 7. For each point (x_0, y_0) on the surface we calculate the energy of the reflected wave near $t_f(x_0, y_0)$. This energy is obtained by calculating the energy of a windowed portion of the time signal for that location. The window is centered at the time it takes for the traveling wave to reach this location, which was calculated in the previous step. A triangular window is used. The energy calculated is also compensated for the energy loss by multiplying the result for each point by the squared distance of the source to that point. Results of the imaging algorithm is shown in Fig. 8 and Fig. 9. In Fig. 8 the simulated anti-tank mine is buried 5cm under the ground centered at $x = 55, y = 0$. Here the mine does not have a structural resonance and the imaging algorithm is mainly exploiting the reflection of the wave off the mine. An actual mine will also have structural resonances which should improve the image. In Fig. 9 a small mine (8cm in diameter) is buried about 2cm deep in the ground centered at $x = 50, y = 0$ along with rocks and sticks. Here the reflection is small and the algorithm is basically locating the energy of the resonance which is also a signature to distinguish the mine from the clutter.

4. Conclusion

Signal processing algorithms have been developed to estimate the localized velocity of the elastic wave, inferring depth information and imaging the data collected by an electro-acoustic system. Examples of the algorithms have been presented. Using the imaging algorithm, one can lo-

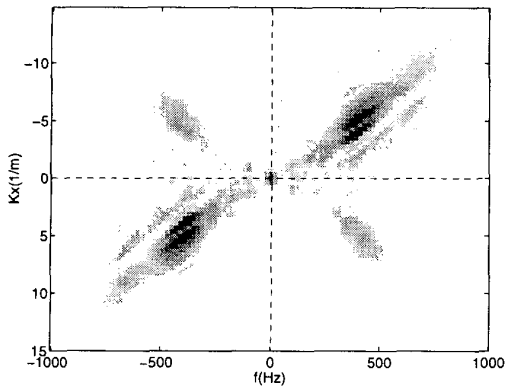


Figure 5. Two-dimensional Fourier transform of the data collected along $y = 0$.

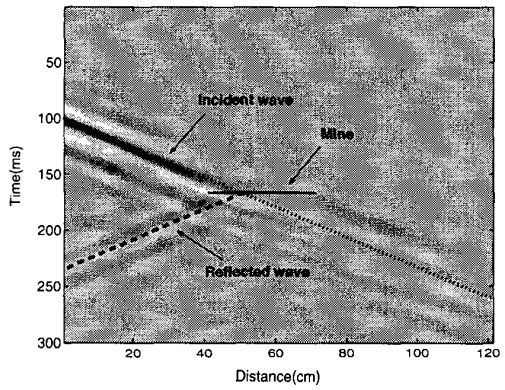


Figure 6. Data collected at $y = 0$.

calize small mines and distinguish them from the clutter. Research in increasing the resolution of the images by a better estimation of different parameters is currently being pursued.

References

[1] K. F. Graph. *Wave Motion in Elastic Solids.* Dover Publications, 1975.

[2] W. R. Scott, Jr. and J. S. Martin. Experimental investigation of the acousto-electromagnetic sensor for locating land mines. *Proceedings of the SPIE: 1999 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls, Orlando, FL,* 3710:204–214, April 1999.

[3] W. R. Scott, Jr., C. Schroeder, and J. S. Martin. A hybrid acoustic/electromagnetic technique for locating land mines. *Proceedings of the 1998 International Geoscience and Remote Sensing Symposium,Seattle, Washington,* pages 216–218, July 1998.

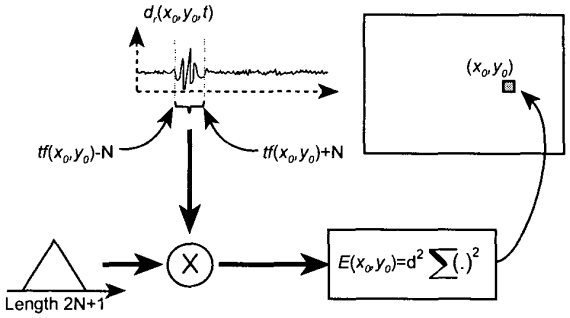


Figure 7. Process involved in the last step of the imaging algorithm.

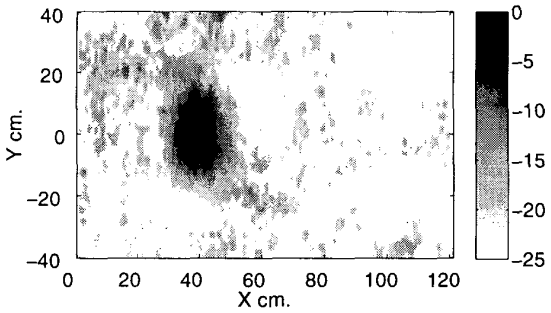


Figure 8. A 25dB image of an anti-tank mine obtained by the imaging algorithm.

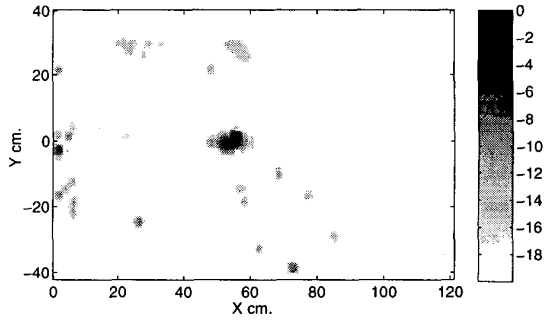


Figure 9. A 20dB image of an anti-personnel mine obtained by the imaging algorithm.

ELSEVIER

Sensors and Actuators B 82 (2002) 54–74

SENSORS AND ACTUATORS B CHEMICAL

www.elsevier.com/locate/sensorb

Exploitation of spatiotemporal information and geometric optimization of signal/noise performance using arrays of carbon black-polymer composite vapor detectors

Shawn M. Briglin[a], Michael S. Freund[a], Phil Tokumaru[b], Nathan S. Lewis[a,*]

[a]*Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA*
[b]*AeroVironment Inc., Monrovia, CA 91016, USA*

Accepted 22 October 2001

Abstract

We have investigated various aspects of the geometric and spatiotemporal response properties of an array of sorption-based vapor detectors. The detectors of specific interest are composites of insulating organic polymers filled with electrical conductors, wherein the detector film provides a reversible dc electrical resistance change upon the sorption of an analyte vapor. An analytical expression derived for the signal/noise performance as a function of detector volume implies that there is an optimum detector film volume which will produce the highest signal/noise ratio for a given carbon black-polymer composite when exposed to a fixed volume of sampled analyte. This prediction has been verified experimentally by exploring the response behavior of detectors having a variety of different geometric form factors. We also demonstrate that useful information can be obtained from the spatiotemporal response profile of an analyte moving at a controlled flow velocity across an array of chemically identical, but spatially nonequivalent, detectors. Finally, we demonstrate the use of these design principles, incorporated with an analysis of the changes in detector signals in response to variations in analyte flow rate, to obtain useful information on the composition of analytes and analyte mixtures. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Sensor arrays; Spatiotemporal effects; Conducting polymer vapor sensors

1. Introduction

Arrays of vapor detectors that rely on the partially selective interaction of vapors with various polymers have received significant attention in the recent literature. Detector types of interest include carbon black-insulating polymer composites [1], conducting organic polymers [2–4], polymer-coated quartz crystal microbalances (QCM) [5], polymer-coated surface acoustic wave (SAW) devices [6,7], polymer-coated capacitors [8], and arrays of dye-impregnated polymeric beads or coated optical fibers [9–11]. The responses of such sorption-based detectors depend primarily on the partition coefficient of the gaseous analyte into the polymer [12]. Arrays of detectors, in which each element contains a chemically different polymer, have been demonstrated to allow discrimination between various vapors based on the differences in response patterns produced by the detector array [13].

In most studies to date, the detectors in such an array are placed in nominally spatially equivalent positions relative to the analyte flow path [1,11,14]. In such a configuration, any spatiotemporal differences between detectors are minimized, and the array response pattern is determined by the differing physicochemical responses of the various detectors towards the analyte of interest. The variations in analyte sorption amongst various detectors thus determines the resolving power of the detector array and determines the other performance parameters of such systems. In this work, we have deliberately placed detectors in spatially nonequivalent positions relative to the flow path of the sampled analyte. We demonstrate that the spatiotemporal response properties of such an array can be used advantageously to obtain information on the identity of analyte vapors and also to produce information on the composition of analyte mixtures.

Additionally, in most studies of detector arrays to date, the form factor of the individual detectors is constrained by factors related to the mode of signal transduction. For example, most film-coated QCM devices must have specified dimensions so that a resonant bulk acoustic wave

* Corresponding author. Tel.: +1-626-395-6335; fax: +1-626-795-7487.
E-mail address: nslewis@caltech.edu (N.S. Lewis).

0925-4005/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0925-4005(01)00991-1

can be maintained in the quartz crystal transducer element [15,16]. Similarly, the geometry of SAW devices is constrained by the need to sustain a Rayleigh wave of the appropriate resonant frequency at the surface of the transducer crystal [15]. Each detector in a QCM or SAW array typically has an identical area and form factor; consequently, the array response is based solely on the different polymer/analyte sorption properties of the differing detector films. Although in principle these types of devices could be constructed with a range of form factors, relatively little attention has been focused on varying the form factors of the detector to optimize the signal/noise ratio (S/N) for a particular analyte. Recent work in our laboratories has focused on the use of chemically sensitive vapor detectors comprised of regions of electrical conductors interspersed amongst regions of insulating organic polymers [1]. The swelling of these films upon sorption of an analyte vapor produces a readily measured, dc electrical resistance change. Spray-coating deposition techniques using masked substrates permits the fabrication of such chemiresistor-type vapor detectors in virtually any geometry where the film can bridge two electrically conducting contact leads [17]. This freedom to explore various form factors allows convenient exploration of the geometrical aspects of sorption-based vapor detector design.

We demonstrate herein that different form factors of a given detector film in conjunction with specific types of analyte flow paths can provide very different detection performance for different types of analyte vapors. An analytical expression has been derived to predict the optimum volume of a detector film as a function of the sample volume and the analyte/polymer partition coefficient. Under certain conditions, detectors of very small areas are expected to have the best S/N performance, whereas for other conditions, relatively large detector areas are optimal. These predictions have been verified through measurements of the response properties of conducting polymer composite chemiresistor vapor detectors. We also demonstrate that, based on these principles, the use of an array of detectors that are nominally identical chemically, but which have different form factors relative to the analyte flow path, can provide useful information on the composition and identity of an analyte vapor. Finally, we report S/N data that allow comparisons between the detection limits of several polymer/analyte combinations using two different modes of signal transduction: frequency shifts in SAW devices and dc electrical resistance changes in composites of carbon black and insulating organic polymers.

2. Theoretical considerations

2.1. Dependence of the noise power on the area of a carbon black-polymer composite vapor detector

At open circuit, resistors exhibit voltage fluctuations whose power spectrum is constant as the frequency is varied. These fluctuations are known as Johnson noise. The root mean squared (rms) noise voltage density of this Johnson noise, V_{JN}, is related to the resistance, R, of a resistive detector as follows:

$$V_{JN} = (4kTRB)^{1/2} \quad (1)$$

where k is Boltzmann's constant, T the temperature (in K), and B is the bandwidth [18]. This Johnson noise is the fundamental lower limit on the noise of any device of resistance R, and its magnitude is independent of the volume or of other fabrication-dependent properties of the resistor. However, when current flows through most types of resistive materials, a voltage fluctuation is observed with a power spectral density that displays an inverse dependence on frequency. This additional noise, which is typically of the form $1/f^{\gamma}$, where $\gamma = 1 \pm 0.1$, is designated $1/f$ noise [19,20].

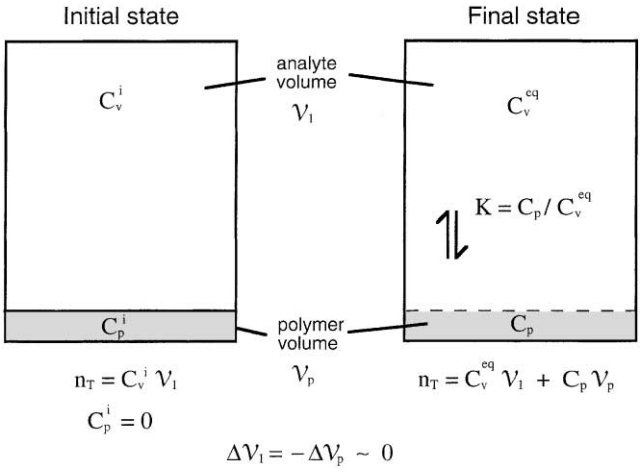
Even for a series of resistors that are fabricated by an identical process, the magnitude of the $1/f$ noise depends on the volume, $\mathcal{V}$, of the resistor. When the correlation length of the resistive particle network is small compared to the physical length scale of interest, the $1/f$ noise of a resistance-based detector is expected to be proportional to $\mathcal{V}^{-(1/2)}$ [21]. For a given film thickness, this implies that the total noise of a resistive detector scales as $A^{-(1/2)}$, where A is the total area of the detector film between the electrical contact leads. This dependence requires that the magnitude of the $1/f$ noise, in the frequency window of the measurement, is much greater than the magnitude of the Johnson noise, so that the total noise is dominated by the $1/f$ contribution. As a consequence of Ohm's law, the power spectral density, $S_n(V)$, of the $1/f$ resistance noise scales with the square of the bias voltage, V_b, applied to the resistor. The quantity of fundamental interest in characterizing the noise of a resistive detector element is thus

$$S_n = \frac{S_n(V_b)}{V_b^2} \quad (2)$$

where S_n is the relative noise power spectral density and V_b is the biasing voltage [21,22]. In contrast to the Johnson noise, the level of the $1/f$ noise in carbon black-polymer composite resistors varies with many factors, including the structure of the carbon black, its volume fraction in the composite, the type of insulator, the resistivity of the composite, and the method of resistor preparation [21,23].

2.2. Dependence of signal/noise on the area of a carbon black composite chemiresistor

Given the above expectations for the scaling of the noise power of a chemically sensitive resistor with the volume of the detector film, we now consider how the signal produced by sorption of an analyte will depend on the volume of the detector film. Consider introducing a fixed quantity of an analyte into a sample chamber of total volume $\mathcal{V}_1$ to produce an initial analyte concentration C_v^i in the vapor



Scheme 1. Equilibration between a finite volume of sampled analyte and a finite volume of a sorption-based vapor detector film. Before equilibration, the initial analyte at a concentration C_v^i is introduced into a total headspace volume $\mathcal{V}_1$. Equilibration of analyte between the vapor and polymer phases results in an equilibrium concentration of analyte in the polymer phase, C_p, and an equilibrium concentration of analyte in the vapor phase, C_v^{eq}, such that $C_p/C_v^{eq} = K$, where K is the polymer/gas partition coefficient. The headspace volume $\mathcal{V}_1$ is essentially constant before and after analyte sorption provided that the change in volume of the polymer phase due to analyte sorption, $\Delta\mathcal{V}_p$, is small compared to the initial headspace volume $\mathcal{V}_1$.

phase (Scheme 1). This analyte could either be introduced as a pulse of concentrated analyte into the volume $\mathcal{V}_1$ or by introducing a sampled volume of analyte in conjunction with a dead volume of carrier gas in the sampling path such that initially after the sampling process has been completed, an analyte concentration C_v^i is present in a total headspace volume $\mathcal{V}_1$. Assuming that no analyte is present initially in either the background gas or the polymer, the total number of moles of analyte available for sorption into the polymer is therefore $n_T = C_v^i \mathcal{V}_1$. Sorption of the analyte into a polymer of volume $\mathcal{V}_p$ will proceed with a polymer/gas partition coefficient, $K = C_p/C_v^{eq}$, where C_p is the concentration of analyte in the polymer phase, C_v^{eq} is the concentration of the analyte in the vapor phase, and both concentrations refer to the situation after equilibrium has been reached.

Assuming that the change in volume of the polymer phase due to analyte sorption, $\Delta\mathcal{V}_p$, is small compared to the value of the initial headspace volume $\mathcal{V}_1$ implies that $\mathcal{V}_1$ also equals the headspace volume after equilibrium has been reached (for typical detector film thicknesses of 0.2–1.0 μm, and for typical headspace thicknesses of greater than 0.1 cm, even 100% increases in film thickness due to sorption-induced film swelling will produce a negligible change in the headspace volume). Under these conditions, conservation of mass of analyte implies that

$$n_T = \mathcal{V}_1 C_v^i = \mathcal{V}_p C_p + \mathcal{V}_1 C_v^{eq} \tag{3}$$

Hence

$$n_T = \frac{\mathcal{V}_p C_p + \mathcal{V}_1 C_p}{K} \tag{4}$$

or

$$C_p = \frac{n_T}{(\mathcal{V}_p + (\mathcal{V}_1)/K)} \tag{5}$$

We further assume that the signal, S, obtained due to sorption of analyte into the polymer is linearly related to the sorbed analyte concentration through a sensitivity factor, X_1, for each analyte/polymer combination [13]:

$$S = X_1 C_p \tag{6}$$

In the limit where the $1/f$ noise dominates the total noise of a chemically sensitive resistor, the measurement noise, N, scales as $\mathcal{V}^{-(1/2)}$ (vide supra). Hence one can write

$$N = X_2 \mathcal{V}_p^{-(1/2)} \tag{7}$$

where X_2 is a constant that is independent of the film volume.The S/N is therefore

$$\frac{S}{N} = \frac{X_1 C_p}{X_2 \mathcal{V}_p^{-(1/2)}} \tag{8}$$

Substituting for C_p from Eq. (5) produces

$$\frac{S}{N} = \left(\frac{X_1}{X_2}\right) n_T \left[\mathcal{V}_p^{1/2} + \left(\frac{\mathcal{V}_1}{K}\right)\mathcal{V}_p^{-(1/2)}\right]^{-1} \tag{9}$$

Multiplying both the numerator and denominator of the right hand side of Eq. (9) by $(K/\mathcal{V}_1)^{1/2}$ yields

$$\frac{S}{N} = \left(\frac{X_1}{X_2}\right) n_T \left(\frac{K}{\mathcal{V}_1}\right)^{1/2} \times \left[\left(\frac{\mathcal{V}_1}{K}\right)^{-(1/2)} \mathcal{V}_p^{1/2} + \left(\frac{\mathcal{V}_1}{K}\right)^{1/2} \mathcal{V}_p^{-(1/2)}\right]^{-1} \tag{10}$$

With the substitution $\mathcal{X} = \mathcal{V}_p K/\mathcal{V}_1$, Eq. (10) becomes

$$\frac{S}{N} = \left(\frac{X_1}{X_2}\right) n_T \left(\frac{K}{\mathcal{V}_1}\right)^{1/2} [\mathcal{X}^{1/2} + \mathcal{X}^{-(1/2)}]^{-1} \tag{11}$$

This function is maximized when $\mathcal{X} = 1$, i.e. when $K\mathcal{V}_p/\mathcal{V}_1 = 1$, which implies that

$$\mathcal{V}_p = \frac{\mathcal{V}_1}{K} \tag{12}$$

at maximal S/N ratio.

When $\mathcal{V}_p = \mathcal{V}_1/K$, Eqs. (3) and (4) yield $C_v^{eq}\mathcal{V}_1 = (1/2)n_T$ and $C_p\mathcal{V}_p = (1/2)n_T$. In other words, for a finite quantity of sampled analyte, the maximal S/N ratio is obtained when the detector volume equals the headspace volume $\mathcal{V}_1$ divided by the polymer/gas partition coefficient. This produces a situation in which equal numbers of moles of analyte are present in the polymer and vapor phases after equilibrium has been attained. In practice, the film thickness of the detector is typically as small as possible to minimize the time constant for sorbtion/desorbtion of analyte. Hence, at constant, minimized film thickness, Eqs. (9) and (12) imply that there is an optimum detector film area for a given

headspace volume and a given initial headspace analyte concentration. Smaller detector areas than this optimum value fail to exhibit optimally low noise, while larger detector areas result in the sorption of the fixed number of moles of analyte into too large of a polymer volume and therefore produce a reduced magnitude of signal after equilibrium has been reached. Another consequence of the analysis presented above is that the different response properties of a set of detectors having a common polymer sorbent layer, but having different form factors, can provide information on the value of K, if $\mathscr{V}_1$ is known and/or is held constant during the experiment. Below we evaluate the validity of these predictions for sorption-based detectors fabricated using carbon black-filled chemiresistors as exemplary systems.

3. Experimental

3.1. Materials

Poly(ethylene-co-vinyl acetate) with 25% acetate (PEVA), and poly(caprolactone) (PCL) were purchased from Scientific Polymer Products. The solvents *n*-hexane, *n*-decane, *n*-dodecane, *n*-tridecane, and *n*-hexadecane were purchased from Aldrich, while toluene and methanol were purchased from EM Science. All solvents were used as received.

3.2. Fabrication of substrates and detector films

3.2.1. Detector film fabrication

The carbon black-polymer composite suspensions used to form the detector films were prepared by dissolving 160 mg of polymer in toluene, followed by addition of 40 mg of carbon black (Cabot Black Pearls 2000) [1]. The mixtures were sonicated for 10 min and were then sprayed in several lateral passes using an airbrush (Iowata HP-BC) held at a distance of 10–14 cm from the substrate.

3.2.2. Substrates for noise measurements

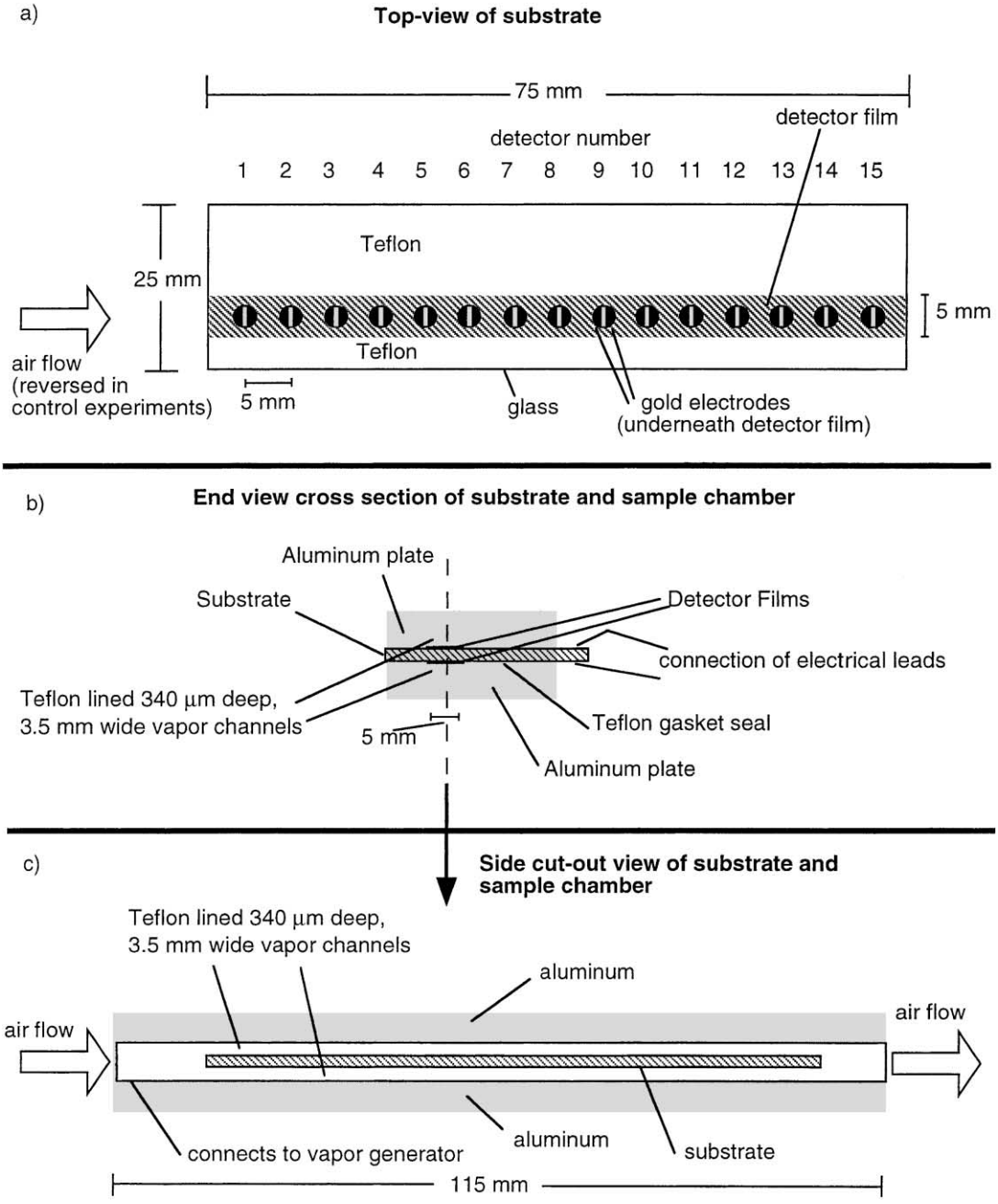
For measurements of the noise properties of the detector films, glass microscope slides were coated with a 50 nm thick layer of Au on top of a 15–30 nm thick layer of Cr, in a pattern that produced rectangular gaps between two parallel metal contact regions. The ratio of the rectangular edge length to the gap length was 8:1, and this aspect ratio was held constant as the area of the gap was varied. After film deposition, this procedure resulted in detector films of similar resistance values that had systematically varying film volumes. Carbon black composite films containing either PEVA or PCL, and having areas of 0.080, 0.30, 1.2, 1.3, 5.0, 33.0, and 132 mm^2, with resistance values ranging from 70 to 160 kΩ, were then deposited onto these substrates. The resulting detector film thicknesses, which were between 180 and 300 nm for the PEVA films and between 60 and 120 nm for the PCL films, were measured with a Sloan Dektak model 3030 profilometer.

3.2.3. Substrates for investigation of the vapor response of chemically equivalent, spatially nonequivalent detectors

Two additional types of substrates were used for investigation of the spatiotemporal and geometric aspects of the chemiresistive vapor detectors. In the first set of experiments (Scheme 2), a series of parallel Cr/Au contacts was formed on each side of 75 mm × 25 mm glass slides. These contact electrodes were 1.8 mm long and were separated by a gap of 0.4 mm. Each pair of electrodes, which defined the contacts for an individual detector, was spaced 5 mm apart, permitting formation of 15 individual detectors on each side of the glass slide. The area surrounding the electrodes was coated with a thin layer of Teflon.

Both sides of the substrate were masked, with the exception of a 5 mm × 75 mm rectangular region on each side of the substrate that was centered on the row of electrical contacts used to form the detectors. Through this mask, carbon black-PEVA composites were sprayed onto one side of the glass microscope slide and carbon black-PCL composites were sprayed onto the other side of the glass slide. After spraying, the carbon black-polymer films covered the entire length of these substrates (Scheme 2). Two such substrates were prepared. On the first substrate, the resulting detectors had resistance values that ranged from 60–160 kΩ on the side sprayed with a PCL-carbon black composite and from 140–180 kΩ on the side sprayed with a PEVA-carbon black composite. The ranges on the second substrate were 70–110 kΩ on the side sprayed with the PCL-carbon black composite and 170–260 kΩ on the side sprayed with a PEVA-carbon black composite.

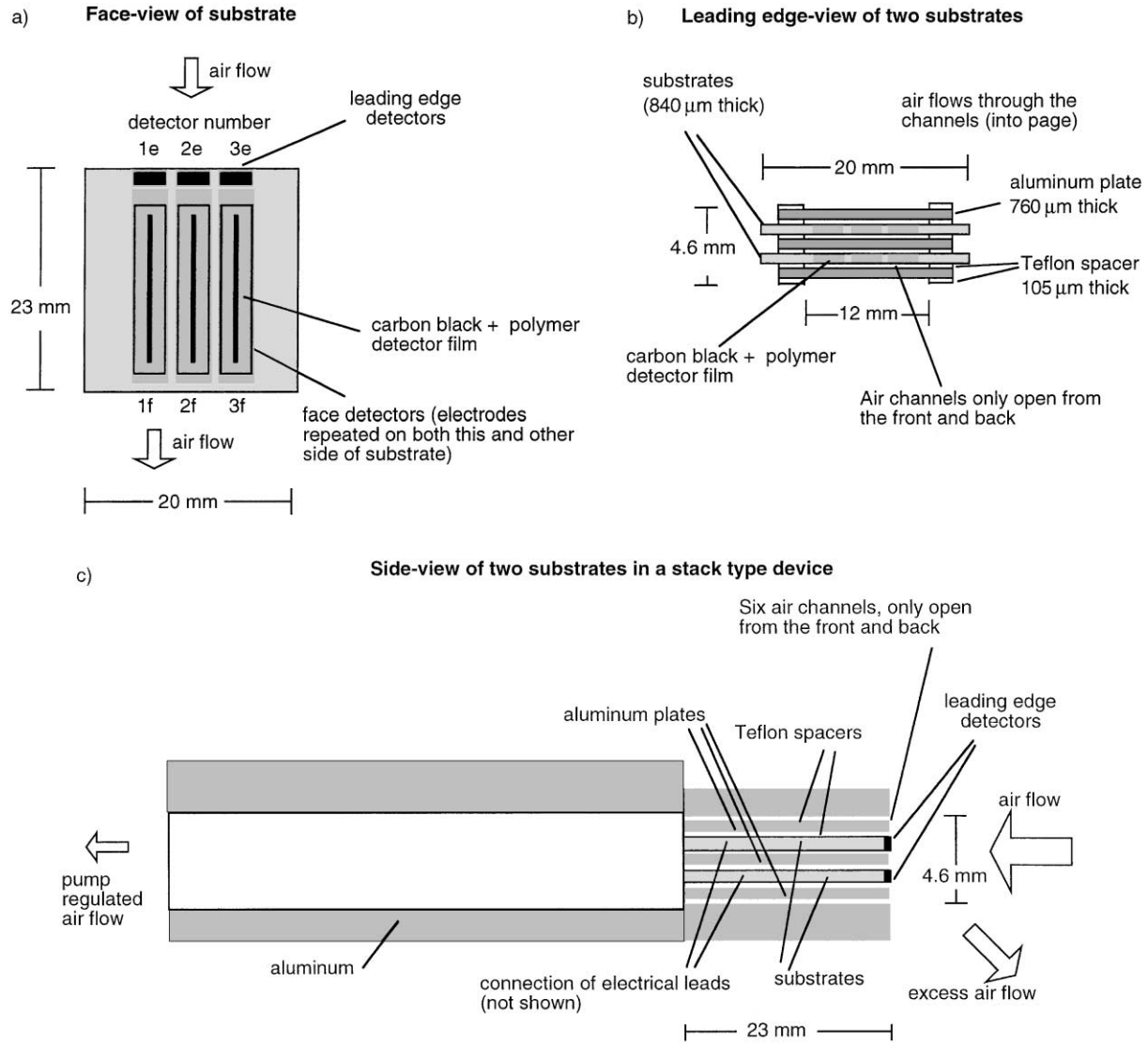
A low volume vapor sample chamber was custom fabricated for the vapor response experiments (Scheme 2). The detector substrate was placed between two pieces of Al, each of which had a recess 3.5 mm wide and 400 μm in depth machined along its length. Prior to assembly, a thin piece of Teflon tape was smoothed over the surface of the Al pieces and into the channel, effectively lining the top and the sides of the channel with an ≈60 μm thick layer of Teflon. This Teflon prevented contact between the analyte and the Al and also formed an airtight gasket between each Al piece and the substrate. Assembly of the Al pieces and the substrate created one shallow channel above the substrate and one shallow channel below the substrate, with each channel being 340 μm deep (400 μm channel depth minus 60 μm thickness of Teflon insulation) and 3.4 mm wide (3.5 mm machined width minus 2 mm × 0.06 mm thickness of Teflon insulation). Each channel spanned the entire length of the row of 15 detectors on its corresponding side of the substrate. The 3.4 mm width of the channel bounded the gas flow into a region that was less than the width of the detector film that had been sprayed onto the substrate. Hence, for the entire length of the channel, the detector film completely coated the substrate in the region adjacent to the channel.



Scheme 2. Layout of a substrate and associated sample chamber used to fabricate a linear array of detectors having a defined headspace and analyte flow configuration. (a) Top view of substrate: a carbon black-PCL composite detector film was sprayed through a mask into a rectangular region of one side of the substrate, and a carbon black-PEVA film was sprayed into an identical pattern onto the other side of the substrate. Each polymer film coated a row of 15 pairs of Au/Cr contacts, creating 15 detectors on each side of the substrate. (b) End view cross section of substrate and sample chamber: a two part Al chamber was then clamped onto both sides of the completed substrate, creating a Teflon-lined channel 340 μm deep and 3.4 mm wide that extended along each row of detectors for the entire length of both sides of the substrate. The wall of the channel that was adjacent to the substrate (i.e. the only wall that was not Al-coated Teflon) was fully spanned by the detector film. (c) Side cut-out view of substrate and sample chamber: air flow was then directed either from low to high detector position or from high to low detector position, as desired, by connecting the substrate/chamber assembly to the output of a vapor generator in conjunction with flow control equipment.

3.2.4. Substrates for investigation of the vapor response of geometrically optimized detector films

A second set of experiments (Scheme 3) used rectangular 20 mm × 23 mm substrates that were fabricated by a commercial vendor (Power Circuits, Santa Ana, CA) using standard printed circuit board technology. Each of these substrates had electrical contacts deposited in a pattern that created a total of six detectors. Three detectors were located on the face of the substrate and three on the edge of the substrate. The three leading edge detectors were formed on



Scheme 3. Schematic of the printed circuit board substrates used to fabricate two types of detectors with different film areas. (a) Face view of substrate: each detector contained three detectors on the edge of the substrate (positions 1e, 2e, 3e in the scheme) and three detectors on each face of the substrate (detectors on the face shown are in positions 1f, 2f, and 3f). Detectors in corresponding face positions on the two sides of a substrate were wired in parallel to form one bigger detector film. (b) Leading edge view of two substrates: two of these substrates were assembled into a stack structure constructed to direct the flow of an air stream containing an analyte vapor over first the small detectors on the edge of the substrate and subsequently over the larger detectors on the face of the substrate. The stack consisted of Al plates alternating with detector-coated substrates, each interspaced with Teflon spacers. (c) Side view of two substrates in a stack type device: the stack was then inserted into an assembly that connected to a controlled flow from a vapor generator. Vapor flow through the assembly was controlled by connecting the output of the assembly to a regulated pump. The stack along with the Al walls of the assembly formed a total of six channels for vapor flow through the stack, first encountering the three detectors on the leading edge of each of the two substrates, then encountering the three two-sided detectors on the faces of each substrate. One substrate in an assembly was formed by spray deposition of a PEVA-carbon black composite detectors and the other substrate was formed by spray deposition of PCL-carbon black composite detectors.

the 840 μm thick edge of the substrate between parallel contacts that were located on each face of the circuit board. These detectors were located in positions 1e, 2e and 3e in Scheme 3. The 20 mm × 23 mm faces of the circuit board supported the three larger detectors, each of which had dimensions of 2 mm × 15 mm (positions 1f, 2f, and 3f in Scheme 3). The electrodes that formed face detectors in the same location on the top and bottom of each substrate were wired together in parallel (i.e. the leads to detector 1f on the top face were connected in parallel to the leads that addressed detector 1f on the bottom face of the substrate). On each substrate this arrangement therefore produced three face detectors, each having a total film area of 60 mm^2 (2 × 2 mm × 15 mm).

Six total substrates of this type were prepared. Three of these substrates were prepared by spraying PEVA-carbon black films onto the edge and face detectors of the substrates, and three by spraying PCL-carbon black films onto the edge and face detectors of the substrates. To prevent current leakage between adjacent detectors, the films of the all

individual detectors were isolated from each other by masking during spraying to produce a narrow (1 mm wide) gap in the detector film between adjacent detectors. Each of the six substrates was sprayed from an independently prepared suspension of carbon black and polymer, but both faces and the leading edge of a given substrate were sprayed from the same suspension. The two faces of a substrate were coated with a film of approximately the same resistance, to create films of similar thickness on each side of a given substrate.

One substrate sprayed with a PEVA-carbon black composite and one sprayed with a PCL-carbon black composite were then assembled into a stack that also contained 760 μm thick Al plates and 105 μm thick Teflon spacers. This assembly created a set of small channels, each of dimensions 0.105 mm × 12 mm × 23 mm, that permitted vapor to be drawn over each set of face detectors. The Teflon spacers served as the side walls for each channel. The assembled stack was 4.59 mm high ($2 \times 0.840 + 3 \times 0.760 + 6 \times 0.105$). Three separate stack assemblies of this type were built.

The stack assemblies were fitted into an Al chamber that had an open front and a tube connector on the back (away from the leading edge detectors). This tube connector was piped to a vacuum pump through a combination airflow meter and regulator (Cole Parmer). Each of the three stack assemblies used in this experiment contained six total channels formed collectively between the two substrates, the three Al plates, and the two walls of the chamber. Hence the volumetric flow of sampled gas through each individual channel was 1/6 of the volumetric flow of sample gas through the entire stack assembly.

3.3. Spectral noise measurements

A standard method was used to determine the noise of the detector films [21,24]. Briefly, the films were placed into a metal box and were biased with a stack of batteries (18 V total) that was connected in series to a 1 MΩ resistance. The 1 MΩ low-noise resistance was formed from ten 100 kΩ wire-wound resistors (Newark Electronics) that were soldered together in series. The bias voltage across the detector film was ac coupled to an SR560 wide-band low-noise voltage preamplifier (Stanford Research Systems), and the output of the preamplifier was sent to an SR785 dynamic signal analyzer (Stanford Research Systems). Using an average of 100 measurements, a power spectral density from 1 to 800 Hz was collected for each film. Data collection occurred over a period of in excess of 100 s for each noise spectral power measurement. These spectra were divided by the square of the bias voltage applied to the chemiresistor, V_b^2, to yield the relative power spectral density S_n for each detector film [21].

A control experiment was performed to evaluate whether film-substrate contacts dominated the observed noise properties of the detectors. Two composite films of approximately the same thickness, film area, and resistance were fabricated, with one film deposited in five 0.38 mm gaps between ten parallel 5.0 mm wide Cr/Au electrical contact pads, and the other film deposited across only one 2.0 mm gap between two parallel 5.0 mm wide Cr/Au contact pads. The additional film/substrate contacts produced no change in the relative noise power of the films, suggesting that the measured noise resulted primarily from the properties of the bulk detector film as opposed to the properties of the film electrode contacts. The properties of commercial, low noise, wire-wound resistors that had resistances similar to those of the carbon black composite films were also measured. The much lower noise values observed for these wire-wound resistors, which are known to exhibit little or no 1/*f* noise, confirmed that the Johnson noise of the resistors plus any additional amplifier noise of the experimental set-up was much lower than the 1/*f* noise observed for the carbon black composite films. No correction for the amplifier noise was therefore performed in analysis of the noise data of the carbon black composite detector films.

3.4. Vapor flow apparatus

An automated flow system was used to deliver pulses of a diluted stream of solvent vapor to the detectors [14]. The carrier gas was oil-free air obtained from the house compressed air source (1.10 ± 0.15 parts per thousand (ppth) of water vapor) controlled with a 28 ml min^{-1} or a 625 ml min^{-1} mass flow controller (UNIT). To obtain the desired concentration of analyte in the gas phase, a stream of carrier gas controlled by a 625 ml min^{-1} or a 60 ml min^{-1} mass flow controller was passed though one of five bubblers. Saturation of the gas flow through the bubbler of interest was confirmed with a flame ionization detector (model 300 HFID, California Analytical Instruments Inc.). The saturated gas stream was then mixed with background air to produce the desired analyte concentration while maintaining the total air flow at the desired value for the linear flow chamber experiments (Scheme 2) and at a constant value of 2 l min^{-1} for the geometrically optimized detector experiments (Scheme 3).

For detectors in the linear flow chamber, the air flow was connected directly to the channel adjacent to the row of detectors. To produce the low flow rates required by this experiment, the analyte-containing vapor was generated at higher flow rates, and a constant 200 ml min^{-1} was subtracted with a flow-regulated pump, permitting the difference to flow into the detector chamber. This flow was then divided into the two equally sized openings of the two channels in the Scheme 2 chamber. The volumetric flow rates quoted below reflect the volumetric flow rate in each separate gap between the detector substrate and the Teflon-lined Al block.

For detectors arranged in the stack assembly of Scheme 3, a constant output of 2 l min^{-1} from the vapor generator was directed at the front end of the sampling device through use of a Teflon tube that was slightly larger in diameter than

the opening of the stack device. Vapor flow through the channels in the stack assembly was maintained at a volumetric flow rate of 75 ml min^{-1}, i.e. 12.5 ml min^{-1} per channel. The excess flow of 1.925 l min^{-1} flowed away from the stack device without proceeding through the channels or over the face sensors.

All exposed parts of the flow system were constructed from Teflon, stainless steel, or Al. The temperature during data collection was approximately 294 K, and the temperature was passively controlled by immersing the solvent bubblers into large tanks of water. For the linear row of detectors (Scheme 2), vapor presentations were 300 s in duration, and analyte exposures were separated in time by at least 75 min to minimize any possible influence of the previous exposure. The analyte was delivered at a constant activity of $P/P^\circ = 0.10$, where P is the partial pressure and P° is the vapor pressure of the analyte. For experiments with geometrically optimized detectors (Scheme 3), the vapor presentations were 240 s in duration, separated in time by 25 min, and were conducted at a fixed analyte activity of $P/P^\circ = 0.050$. Flow experiments were performed separately on each of the three separate stack assemblies. Each stack assembly received 10 exposures to each of four analytes, and the order of these 40 total presentations was randomized with respect to the analyte identity and with respect to replicate exposures to a given analyte. A different randomized analyte presentation order was used for each of the three stack assemblies. A personal computer running programs developed with LabVIEW 5.0 controlled both the flow system and the data acquisition apparatus.

3.5. Direct current resistance measurements

Direct current resistance data were collected using a Keithley 2002 multimeter and a Keithley 7001 multiplexer. Shielded, twisted pair cables were used, and each resistance value was integrated over 2 or 10 power line cycles to reject 60 Hz pick-up. Data were processed using a program written in Microsoft Excel Basic. The relative differential resistance change, $\Delta R_{final}/R_b$, was calculated for each detector, where R_b is the baseline resistance averaged over approximately 20 s prior to vapor presentation, and ΔR_{final} is the differential resistance change relative to R_b. The value of ΔR_{final} was evaluated over a period of approximately 20 s at a fixed time after initiating the vapor presentation. This time varied between the different types of experiments, either from 40 to 60, 200 to 220 or 240 to 260 s after the start of the vapor presentation. No averaging was performed to collect the data represented in the figures which show the response characteristics of the detectors as a function of time. For ease of visualization on a common graph of the different absolute responses of the various detector/analyte combinations, the $\Delta R/R_b$ data in some figures have been normalized. In these figures, data were normalized by the mean response value, $(\Delta R/R_b)_j$, of the detector in the physical position j for each set of identical exposures (i.e. for exposures to a common analyte, or for exposures to a common analyte at a common flow rate, as specified). The value for j was always chosen as the position of the detector to first physically encounter the analyte.

The rms noise, N_{rms}, of a detector was measured as the standard deviation of the data points obtained from the multimeter in the period immediately prior to each vapor presentation, divided by the average resistance value of the multimeter data points produced over that same measurement period. The period used to measure this baseline noise was equal to the time elapsed between determination of the baseline resistance and the determination of the differential resistance change upon analyte exposure. This ensured that the signals were measured in the same bandwidth as the noise. The multimeter was used to determine both the signal and noise values for this calculation, because it was desirable to measure the signal and noise of the detectors using the same instrumental apparatus (i.e. the N in S/N is N_{rms}). The values of the S/N were calculated independently for each separate presentation of analyte to each detector. For the multimeter measurement of the noise of the films of different sizes described above, the same analysis was used, except the noise was calculated over an interval of only 20 s, and five of these values, separated in time by 100 s, were averaged to generate N_{rms}. Unlike the values for S_n, which is a measure of the noise power, these noise values, N_{rms}, were first squared to yield N_{rms}^2 prior to plotting them against film volume.

3.6. Determination of polymer/gas partition coefficients

Quartz crystal microbalance measurements were performed on pure films of both PEVA and PCL at 294 K using 10 MHz resonant frequency quartz crystals and a measurement apparatus that has been described previously [25]. Twenty vapor presentations, each 120 s in duration and separated in time by 15 min, were performed at each of four concentrations ($P/P^\circ = 0.010$, 0.030, 0.050, 0.10) of *n*-hexane and of methanol. The order of vapor presentation was randomized with respect to analyte identity, analyte concentration, and repetition of conditions. The frequency shifts of the polymer-coated QCM crystals arising from deposition of the polymer film, $\Delta f_{polymer}$, were recorded as the difference in the resonant frequency of the crystal before and after deposition of the polymer film. The frequency change upon exposure to analyte vapor, $\Delta f_{analyte}$, was calculated as the difference in the resonant frequency of the film-coated crystal during exposure to the specific analyte vapor relative to the baseline resonant frequency of the film-coated crystal in background air. The baseline frequency was taken as the mean frequency value obtained for the film-coated crystal during a 30 s period immediately prior to exposure to the analyte, and the frequency during exposure to analyte vapor was taken to be the mean frequency value observed between 80 and 110 s after the vapor exposure had been initiated.

4. Results

4.1. Noise spectral density measurements for carbon black composite vapor detectors

Fig. 1 displays the noise power spectral density, $S_n(V_b)$, between 1 and 800 Hz for a set of carbon black composite thin film detectors as a function of the area covered by the composite between the electrical contact pads. The electrode contact dimensions in these experiments were scaled such that the resistance (≈100 kΩ) was approximately constant as the film area was varied. Any variation in the noise thus arose from the film area and not from a variation in response of the preamplifier to different absolute input resistance values. An additional advantage of maintaining a constant aspect ratio for the different volume films is to reduce the variation in the noise that has been observed in some thick-film resistors of different aspect ratios [26]. Fig. 1 also displays the power spectral density for a commercial, low noise, wire-wound resistor.

The power spectral density of the carbon black-polymer thin film composites was well fit to a function of the form $S_n(V_b) \propto 1/f^{\gamma}$ with an exponent of $\gamma = 1.1$. Some deviation from the 1/*f* behavior was observed at very low frequencies (<5 Hz), but this deviation may have resulted from the mechanical contacts used to make connections to the Au/Cr/glass substrates. The noise power spectral density of the wire-wound resistor was much lower than the 1/*f* noise of any of the detector films at the frequencies investigated in this study.

Fig. 2 depicts the value of the $S_n \times f$ product for carbon black composite detectors fabricated from PEVA or from PCL as a function of the volume of the detector film. For these comparisons, the data were taken as the value of S_n at 10 Hz to avoid the lower frequency contact noise. These values are directly comparable, because they were taken at the same frequency, but the $S_n \times f$ product was displayed, because it is essentially independent of frequency for the 1/*f* region above about 5 Hz in frequency. Also shown are the square of the noise values, N^2_{rms}, derived from analysis of the standard deviation of the baseline resistance values versus time as determined on these same films using the multimeter. The detector films used in these experiments were all approximately the same thickness, but the film volume data were calculated using the actual thickness values determined from profilometry measurements of the thickness of each detector film.

The N^2_{rms} and $S_n \times f$ values decreased approximately linearly with the film volume, $\mathcal{V}$, with a plot of $S_n \times f$ versus $\mathcal{V}$ for PEVA-containing carbon black composites

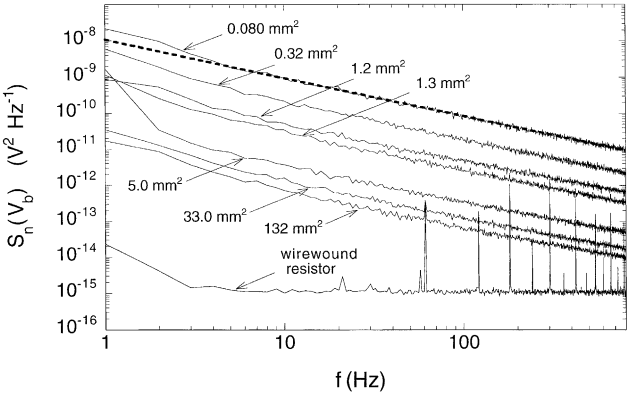


Fig. 1. Power spectral density of the noise, $S_n(V_b)$, vs. frequency, *f*, for seven poly(ethylene-co-vinyl acetate), 25% acetate (PEVA)-carbon black composite detector films of varying area. The dimensions of the rectangularly shaped regions bridged by polymeric composite between the electrical contact pads were (in mm): 0.10×0.80, 0.20×1.60, 0.38×3.05, 0.40×3.20, 0.79×6.3, 2.03×16.3, 4.06×32.5. The PEVA-carbon black composite films were ≈230 nm in thickness as determined by profilometry. The dashed line indicates a fit of one such plot to a function of the form $S_n(V_b) = 1 \times 10^{-8}/f^{1.054}$. Also shown are data for a wire-wound, low noise, 70 kΩ resistor.

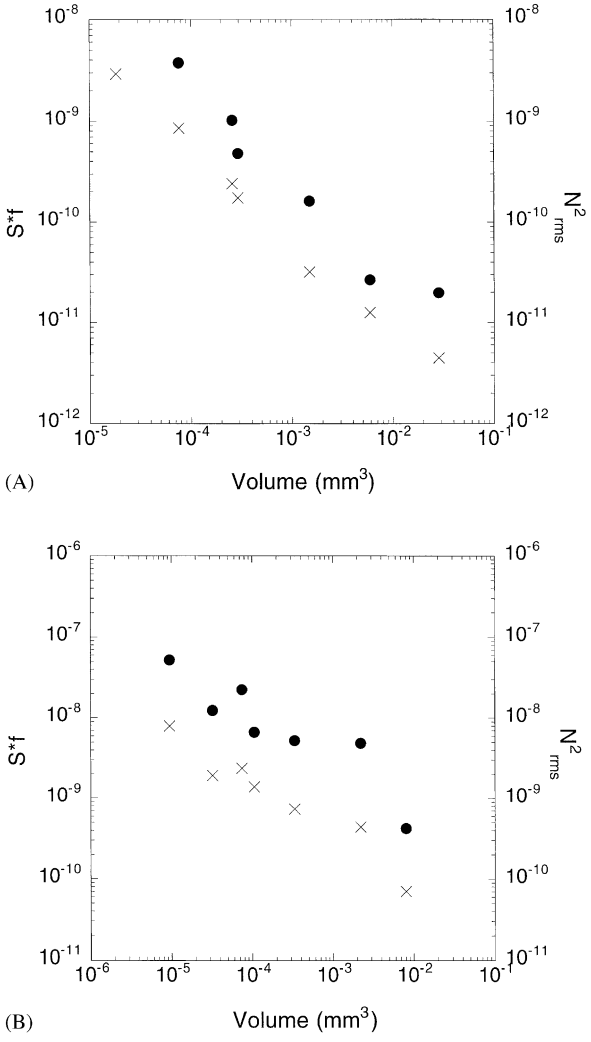


Fig. 2. Values of $S_n \times f$ (crosses) at 10 Hz and N^2_{rms} (filled circles) vs. volume for carbon black composite detectors fabricated from (a) PEVA and (b) PCL. The PEVA-carbon black composite films were ≈230 nm in thickness and the PCL-carbon black composites were ≈80 nm in thickness as determined by profilometry.

having a slope of −0.95 ($R^2 = 0.989$) and a plot of N^2_{rms} versus $\mathscr{V}$ having a slope of −0.91 ($R^2 = 0.964$). For the PCL-containing carbon black composite films, the slope of $S_n \times f$ versus $\mathscr{V}$ was −0.60 ($R^2 = 0.933$), whereas the slope of N^2_{rms} versus $\mathscr{V}$ was −0.58 ($R^2 = 0.833$). It is difficult to perform a quantitative comparison between the $S_n \times f$ and N^2_{rms} values, due to the impedance mismatch between the input amplifier of the multimeter and the resistive load of the detector, the variable bandwidth of the multimeter during various resistance readings, and other well-known electronic circuit considerations [27]. However, the inverse dependence of the N^2_{rms} value on the volume of the detector film is clearly seen in both sets of measurements. Deviations from a strictly linear dependence of the relative noise power on $\mathscr{V}$ with a slope of −1 have been observed previously for polymer film resistors, and have been explained by factors arising from the film-electrode contacts, inhomogeneities in film composition, and/or variability in film thickness over the measured detector area [23,26]. The deviations that we observed may also have resulted from properties related to the relatively thin nature of the films used in this study.

4.2. Determination of polymer/gas partition coefficients

For a given volume of sampled analyte, the detector volume that will produce optimum S/N performance for a specific polymer/analyte combination can be calculated from Eq. (12) if the polymer/gas partition coefficient is known. Accordingly, data for the partition coefficients of hexane and methanol into PCL and PEVA were determined using QCM measurements. Fig. 3a displays the QCM frequency shifts measured for PEVA films exposed to hexane or methanol, while Fig. 3b displays the frequency shifts measured for PCL films exposed to these same analytes.

The frequency shifts of the polymer-coated QCM crystals arising from deposition of the polymer film, $\Delta f_{polymer}$ and from sorption of the analyte vapor, $\Delta f_{analyte}$, were in total much less than 2% of the resonant frequency of the uncoated crystal. Under such conditions, prior work has concluded that mechanical losses are minimal and that the frequency shifts are predominantly due to changes in mass uptake [28], which can be calculated from the Sauerbrey equation [28,29]. Polymer/gas partition coefficients were therefore calculated by fitting a line with a forced zero intercept through the $\Delta f_{analyte}$ versus concentration data for each polymer/analyte combination. The slopes of these lines were −4.36 ($R^2 = 0.9988$) and −0.910 ($R^2 = 0.9995$) for hexane and methanol, respectively, sorbing into PEVA, and were −0.612 ($R^2 = 0.9977$) and −0.930 ($R^2 = 0.9995$) for hexane and methanol, respectively, sorbing into PCL. The slopes of the resulting lines were converted into partition coefficients using

$$K = \frac{10^6 \rho \mathscr{R} T m}{M_w \, \Delta f_{polymer} P_{atm}} \quad (13)$$

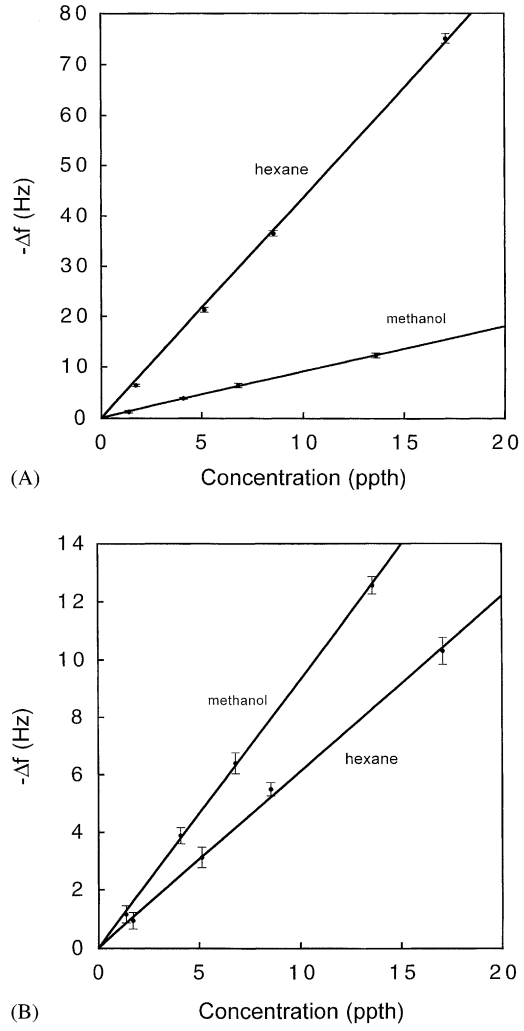


Fig. 3. Differential frequency changes, $-\Delta f_{analyte}$, of QCM coated with (a) PEVA and (b) PCL polymer films (no carbon black) during exposure to hexane at $P/P^\circ = 0.010$, 0.030, 0.050, and 0.10 (1.7, 5.1, 8.5, 17 ppth) and methanol at $P/P^\circ = 0.010$, 0.030, 0.050, and 0.10 (1.3, 4.1, 6.8, 14 ppth), where P is the partial pressure of analyte and P° is the vapor pressure of the analyte at 294 K. Each data point represents an average of 20 $\Delta R/R_b$ responses, and the error bars indicate ±1 S.D. around the mean. The frequency shifts corresponded to decreases in frequency upon exposure to analyte. Lines were fitted through these points with a forced zero intercept. The slopes of these lines were (a) hexane: 4.36 ($R^2 = 0.9988$); methanol: 0.910 ($R^2 = 0.9995$); (b) hexane: 0.612 ($R^2 = 0.9977$); methanol: 0.930 ($R^2 = 0.9995$). The frequency shifts due to coating the crystal with the polymer were −6835 Hz for PEVA and −4355 Hz for PCL.

where $\mathscr{R}$ in the ideal gas constant (l atm mol^{-1} K^{-1}), ρ the density (g ml^{-1}) of the polymer, T the temperature (K), m the slope of $\Delta f_{analyte}$ versus concentration (Hz/ppth in air), M_w the molecular weight (g mol^{-1}) of the analyte, $\Delta f_{polymer}$ (Hz) the frequency shift corresponding to deposition of the polymer, and P_{atm} is the atmospheric pressure (atm). The partition coefficients for each analyte/polymer combination are displayed in Table 1.

Partition coefficients for the lower vapor pressure analytes, dodecane and hexadecane, were difficult to measure, because these very low vapor pressure analytes adsorbed to

Table 1
Responses, noise, and S/N for two types of polymer-carbon black composite detectors in the configuration of Scheme 3[a]

Analyte	Vapor pressure of pure analyte		Logarithm of partition coefficient (log K)[b]		Stack assembly	$\Delta R/R_b \times 100$				N_{rms}				S/N			
	P° (Torr)	PPM[c]	PCL	PEVA		PCL		PEVA		PCL		PEVA		PCL		PEVA	
						Edge[d]	Face	Edge	Face	Edge	Face	Edge	Face	Edge	Face	Edge	Face
Hexane	1.28×10^2	1.71×10^5	1.65	2.23	A	1.4 ± 0.2	1.07 ± 0.03	3.3 ± 0.1	3.5 ± 0.6	$(1.5 \pm 0.7) \times 10^{-3}$	$(1.9 \pm 0.5) \times 10^{-4}$	$(5 \pm 1) \times 10^{-4}$	$(8 \pm 2) \times 10^{-5}$	13 ± 7	60 ± 14	73 ± 20	460 ± 200
					B	1.1 ± 0.4	0.77 ± 0.04	3.6 ± 0.3	2.5 ± 0.1	$(2 \pm 1) \times 10^{-3}$	$(3.2 \pm 0.8) \times 10^{-4}$	$(9 \pm 2) \times 10^{-4}$	$(1.3 \pm 0.3) \times 10^{-4}$	5 ± 2	26 ± 9	44 ± 12	200 ± 45
					C	1.3 ± 0.2	1.17 ± 0.08	2.8 ± 0.3	2.4 ± 0.1	$(1.2 \pm 0.6) \times 10^{-3}$	$(1.8 \pm 0.2) \times 10^{-4}$	$(4 \pm 2) \times 10^{-4}$	$(2.7 \pm 0.9) \times 10^{-4}$	23 ± 23	100 ± 60	77 ± 25	100 ± 37
					Mean	1.3	1.0	3.2	2.8	2×10^{-3}	2.3×10^{-4}	6×10^{-4}	1.6×10^{-4}	14	64	65	260
Methanol	1.02×10^2	1.36×10^5	2.26	1.98	A	2.4 ± 0.2	2.7 ± 0.1	2.0 ± 0.4	2.1 ± 0.5	$(1.4 \pm 0.8) \times 10^{-3}$	$(2.0 \pm 0.5) \times 10^{-4}$	$(5 \pm 1) \times 10^{-4}$	$(9 \pm 2) \times 10^{-5}$	23 ± 12	140 ± 42	42 ± 10	270 ± 120
					B	3.3 ± 0.5	2.4 ± 0.2	1.8 ± 0.3	1.61 ± 0.08	$(3 \pm 1) \times 10^{-3}$	$(3.0 \pm 0.6) \times 10^{-4}$	$(9 \pm 2) \times 10^{-4}$	$(1.5 \pm 0.3) \times 10^{-4}$	14 ± 5	80 ± 16	21 ± 5	120 ± 25
					C	2.6 ± 0.8	2.8 ± 0.2	1.1 ± 0.2	1.2 ± 0.1	$(1.2 \pm 0.8) \times 10^{-3}$	$(1.3 \pm 0.7) \times 10^{-4}$	$(4 \pm 2) \times 10^{-4}$	$(2.6 \pm 0.9) \times 10^{-4}$	33 ± 22	260 ± 110	28 ± 11	50 ± 17
					Mean	2.8	2.6	1.6	1.6	2×10^{-3}	2.1×10^{-4}	6×10^{-4}	1.6×10^{-4}	23	160	30	140
Dodecane	9.71×10^{-2}	1.29×10^2	4.77[e]	5.35[e]	A	1.6 ± 0.2	1.16 ± 0.03	3.7 ± 0.1	3.6 ± 0.6	$(1.3 \pm 0.6) \times 10^{-3}$	$(2.0 \pm 0.4) \times 10^{-4}$	$(5 \pm 1) \times 10^{-4}$	$(9 \pm 0.3) \times 10^{-5}$	15 ± 7	60 ± 13	76 ± 21	440 ± 170
					B	1.2 ± 0.4	0.88 ± 0.07	3.8 ± 0.3	2.6 ± 0.1	$(3 \pm 1) \times 10^{-3}$	$(3.2 \pm 0.9) \times 10^{-4}$	$(9 \pm 2) \times 10^{-4}$	$(1.4 \pm 0.2) \times 10^{-4}$	5 ± 2	30 ± 10	45 ± 14	190 ± 35
					C	1.6 ± 0.2	1.25 ± 0.04	3.4 ± 0.1	1.3 ± 0.2	$(1.2 \pm 0.8) \times 10^{-3}$	$(9 \pm 5) \times 10^{-5}$	$(4 \pm 2) \times 10^{-4}$	$(2.5 \pm 0.7) \times 10^{-4}$	32 ± 32	150 ± 64	100 ± 41	54 ± 21
					Mean	1.5	1.1	3.6	2.5	2×10^{-3}	2.1×10^{-4}	6×10^{-4}	1.6×10^{-4}	17	81	73	230
Hexadecane	9.11×10^{-4}	1.21	6.70[e]	7.35[e]	A	0.3 ± 0.2	0.01 ± 0.09	0.26 ± 0.09	0.01 ± 0.01	$(1.4 \pm 0.9) \times 10^{-3}$	$(1.9 \pm 0.3) \times 10^{-4}$	$(5 \pm 1) \times 10^{-4}$	$(8 \pm 3) \times 10^{-5}$	3 ± 2	1 ± 1	6 ± 2	2 ± 2
					B	0.3 ± 0.3	0.02 ± 0.03	0.4 ± 0.1	0.02 ± 0.04	$(2 \pm 2) \times 10^{-3}$	$(3.1 \pm 0.9) \times 10^{-4}$	$(9 \pm 2) \times 10^{-4}$	$(1.4 \pm 0.3) \times 10^{-4}$	2 ± 1	1 ± 1	5 ± 2	1 ± 3
					C	0.3 ± 0.2	0.03 ± 0.03	0.3 ± 0.1	0.04 ± 0.07	$(1.1 \pm 0.7) \times 10^{-3}$	$(1.2 \pm 0.6) \times 10^{-4}$	$(4 \pm 2) \times 10^{-4}$	$(2.3 \pm 0.7) \times 10^{-4}$	5 ± 4	4 ± 4	8 ± 3	2 ± 3
					Mean	0.3	0.02	0.3	0.03	2×10^{-3}	2.1×10^{-4}	6×10^{-4}	1.5×10^{-4}	3	2	6	2

[a] Data were averages of 10 randomized presentations of the four analytes each at $P/P^\circ = 0.050$, across three copies of each of the two detector types, with each value representing 30 vapor/polymer interactions. The experiment was repeated for three independently prepared stack assemblies (A, B and C). The data represent responses after 200 s of exposure to analyte.

[b] Determined from QCM measurements on unfilled polymers.

[c] Vapor pressure of analyte expressed in ppm of air at 294 K.

[d] Edge refers to the leading edge sensors and face refers to the face sensors depicted in Scheme 3. The uncertainties are expressed as 95% confidence intervals.

[e] Values were estimated based on measurements of K for hexane and correction for the differences in vapor pressure between hexane and the alkane of interest assuming constant activity coefficients for the sorption of the alkanes into a given polymeric phase.

the walls of the chamber and required long times as well as high analyte volumes to reach true equilibrium conditions. Instead, the values for these analytes were estimated by multiplying the measured polymer/gas partition coefficients for hexane by the ratio of the vapor pressures of dodecane and hexadecane relative to that of hexane [30]. This is a good approximation provided that the activity coefficients do not vary significantly for sorption of these three alkanes into the polymers of interest. As displayed in Table 1, the polymer/gas partition coefficients varied from measured values of 10^2 for hexane and methanol to values of over 10^7 estimated for the lowest vapor pressure analyte, hexadecane.

The wide difference in vapor pressures between the analytes of concern is expected to have a significant influence on the physical array design for optimization of the S/N ratio as given by Eq. (9). In a chamber of headspace thickness of 1.0×10^{-2} cm, with a detector film thickness of 1.0×10^{-4} cm, the optimum detector area for a 1.0 cm^3 volume of an analyte sample for which the analyte polymer/gas partition coefficient is 1.0×10^2 is 1.0 cm^2. In contrast, for the same sampled volume, headspace thickness, and detector film thickness, a detector area of only 1.0×10^{-5} cm^2 produces maximum S/N performance for an analyte having a polymer/gas partition coefficient of 1.0×10^7. The implications of this wide variation in polymer/gas partition coefficient for optimizing the S/N performance of sorption-based vapor detectors are explored in detail below.

4.3. Spatiotemporal response data from linear arrays of carbon black-polymer composite chemiresistive vapor detectors

The responses of an array of carbon black-polymer composite vapor detectors were investigated as a function of position relative to the location of analyte flow injected into the detection chamber. The pattern of the contacts beneath the film of carbon black-polymer composite produced an array of chemiresistive detectors that were arranged in a linear geometry, parallel to the analyte flow path, and which were spaced at 5 mm intervals downstream from the location of flow injection (Scheme 2). The headspace volume was defined by the 3.4 mm width, 340 μm depth, and 75 mm total length of this channel over the detector film. The area of the carbon black-polymer composite film spanned the entire length of the substrate and was sufficiently wide to ensure that the entire region of the substrate in contact with this vapor channel was coated with the detector film (Scheme 2). Hence, in many respects this experimental apparatus is analogous to probing the spatiotemporal distribution of analyte in the sorbent phase after injection of a sample onto a gas chromatography column or to ascertaining spectroscopically the position of analyte in a thin layer chromatography experiment as a function of time.

Fig. 4 displays data collected for the array exposed in this configuration at a fixed, low carrier gas flow rate of three analytes of differing vapor pressure (hexane, dodecane, and tridecane) to a series of PEVA-carbon black composites. The data are the relative differential resistance values measured in a 20 s period after 240 s of continuous exposure to the various analytes of interest. The analyte exposures used to produce these data were randomized with respect to analyte identity and with respect to the five replicate exposures of each analyte at the concentration of interest.

For high vapor pressure analytes, the detectors all produced nominally identical, time-independent, responses to the analyte in the final 20 s of this 260 s exposure period. For example, the standard deviation of the mean response to hexane at $P/P^\circ = 0.10$ for the 15 nominally identical detectors was less than 5% of the mean $\Delta R/R_b$ response value for this detector/analyte combination. This degree of reproducibility is consistent with prior reports that have evaluated the reproducibility of the response of carbon black-polymer composite detectors [1].

In contrast, during 260 s exposures to low vapor pressure analytes such as tridecane, the $\Delta R/R_b$ values observed from the detectors to first encounter the vapor stream were much higher than $\Delta R/R_b$ values observed for detectors located at positions remote from the injection location. The position-related variation in $\Delta R/R_b$ in response to the low vapor pressure analytes was clearly much greater than the standard deviation of the $\Delta R/R_b$ value observed for replicate exposures to any of the analytes investigated. The trend was systematic in that the detectors closest to the analyte injection position displayed the highest $\Delta R/R_b$ values, the response decreased monotonically with position from the location of analyte injection, and the magnitude of the effect increased as the vapor pressure of the analyte decreased. Furthermore, for the low vapor pressure analytes, the change in mean response versus detector position far exceeded the standard deviation of the mean responses observed for these same detectors when exposed, in the identical apparatus, to analytes having high vapor pressures. To conclusively prove that the effect was associated with the geometry of the flow system relative to the position of the detectors in the chamber, and not with any physicochemical inequivalence in the detectors themselves, the position of analyte injection was changed such that the flow proceeded in the opposite direction through the chamber, with analyte first encountering detector number 15 and finally encountering detector number 1 in Scheme 2. The same analytes were used and the order of presentation was again randomized with respect to solvent identity and with respect to the five replicate exposures to each analyte; however, the exposure order was the same as that used when the flow proceeded from low to high detector number. As shown in Fig. 4, the detectors again provided essentially equivalent responses when exposed to high vapor pressure analytes at a volumetric flow rate of 6 ml min^{-1}. For low vapor pressure analytes, the highest $\Delta R/R_b$ values were again observed from the detectors that first physically encountered the vapor stream.

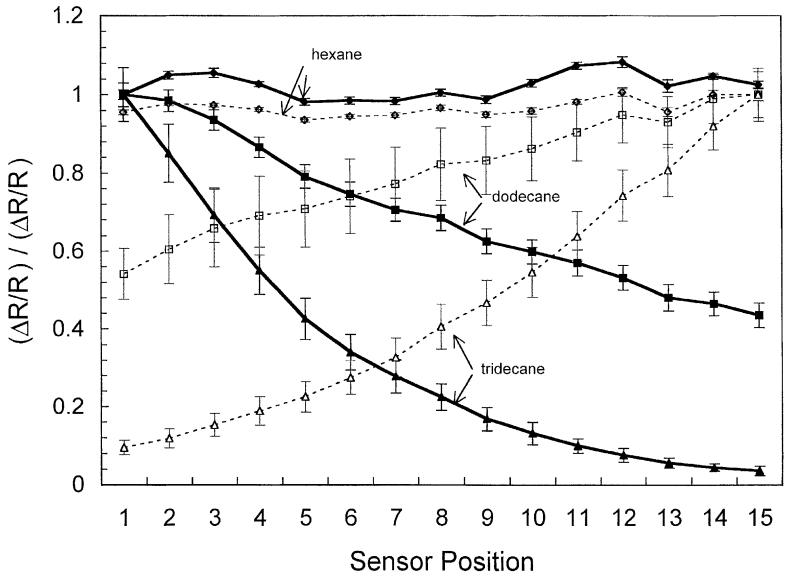


Fig. 4. Normalized relative differential resistance responses, $(\Delta R/R_b)/(\Delta R/R_b)_j$, of carbon black-PEVA composite vapor detectors exposed to three analytes: a high vapor pressure analyte (hexane), a moderately low vapor pressure analyte (dodecane), and a low vapor pressure analyte (tridecane, vapor pressure of 3.9×10^{-2} Torr at 294 K), each at a constant activity of $P/P^\circ = 0.10$ and at a volumetric flow rate of 6 ml min^{-1}. Each point represents an average of five responses, and the error bars indicate ± 1 S.D. from the mean. Data were obtained from $\Delta R/R_b$ values observed between 240 and 260 s after initiation of exposure to the analyte of interest. The detectors were arranged in a linear geometry as depicted in Scheme 2. For ease of visualization on a common graph of the different absolute responses of the various detector/analyte combinations, the data in this figure have been normalized relative to the mean response of the first detector that physically encountered the analyte. The solid lines indicate responses when the analyte flowed in the direction from the leftmost detector (corresponding to the detector with the lowest numbered position) to rightmost detector. These data (and associated standard deviations) were normalized to the mean response value of the detector in position 1 in the array ($j = 1$) for the five exposures to the analyte of interest. The normalization constants (values by which the data were multiplied for display on the plot) are: 10.8, 16.7, and 32.1, for hexane, dodecane, and tridecane, respectively. The dashed lines indicate responses recorded when the same row of detectors was exposed to vapor flowing in the opposite direction through the detector chamber; consequently, these data (and associated standard deviations) were normalized to the mean response value of the detector in position 15 in the array ($j = 15$) to the five exposures of the analyte of interest. Normalization constants for these data are: 10.4, 15.3, and 30.2, for hexane, dodecane, and tridecane, respectively. Analytes were delivered in a random order during an individual experiment, but the exposure order when analyte entered from the right was identical to the exposure order for experiments in which analyte entered the array from the left of the substrate as displayed in Scheme 2.

The effect of sorption of low vapor pressure analytes into the composite vapor detectors arranged in the linear geometry of Scheme 2 was also evident in the time-dependent response of the different detectors in the array. Fig. 5 displays the $\Delta R/R_b$ values of the detectors at various times during the course of exposure to hexane or dodecane, respectively. These exposures were performed in the same, low flow rate vapor response, experiment described above in which the analyte first encountered detector number 1 in Scheme 2. No averaging time window was employed in this representation of the data so that the response at a variety of times could be displayed. As shown in Fig. 5, at this flow rate all of the detectors reached a steady-state response within 20 s during exposure to hexane, but no detector reached a steady-state response during the first 240 s of exposure to the lower vapor pressure analyte, dodecane.

Position-related differences in the spatiotemporal response to different analytes are also apparent through examination of the temporal response of a single detector. Fig. 6 shows resistance versus time data for exposure of a PEVA-carbon black composite to hexane (at $P/P^\circ = 0.10$) followed immediately by exposure to a mixture of hexane and dodecane (each at $P/P^\circ = 0.10$). These data were obtained at a relatively low carrier flow velocity (6 ml min^{-1}) on a PEVA-carbon black detector located at position 7 in Scheme 2. As shown in Figs. 4–6, under these conditions, the different analytes can be distinguished based on their characteristic temporal responses on the detectors that arise from the interactions with the analyte flow in the detector chamber.

Fig. 7 displays similar data, collected on a different substrate, as a function of analyte flow velocity. Data presented are for two analytes, one having a high vapor pressure (hexane) and the other having a low vapor pressure (dodecane), both exposed to either PEVA-carbon black (Fig. 7a) or to PCL-carbon black (Fig. 7b) composite detector films. For each flow rate, hexane and dodecane were alternately presented to the detectors. This procedure was repeated for each of five flow rates, proceeding sequentially from the lowest volumetric flow rate to the highest volumetric flow rate. This 10 exposure protocol was then repeated in its entirety four times, producing 50 total exposures of analyte to the detectors. For high vapor pressure analytes (i.e. analytes with relatively small polymer/gas partition coefficients), all of the detectors exhibited essentially the same $\Delta R/R_b$ response values in the 20 s period after 240 s of analyte exposure

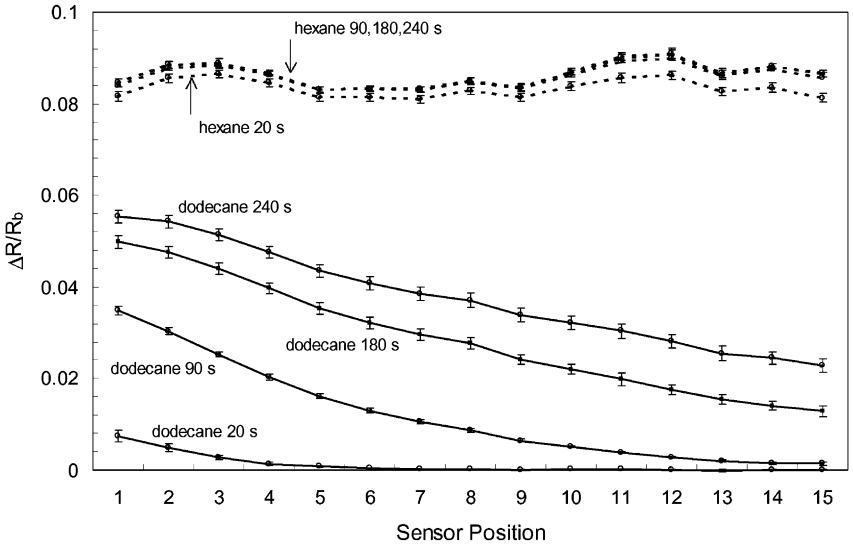


Fig. 5. Relative differential resistance responses, ($\Delta R/R_b$), of carbon black-PEVA composite vapor detectors configured as indicated in Scheme 2 for each detector position as a function of time during exposure to analyte vapor. These detectors were exposed to two analytes: a high vapor pressure analyte (hexane), indicated with a dashed line, and a moderately low vapor pressure analyte (dodecane), indicated with a solid line. Each analyte was delivered in random order at a constant activity of $P/P^\circ = 0.10$ and at a volumetric flow rate of 6 ml min^{-1}. The analyte flow first encountered detector 1 in Scheme 2. Data for each analyte are displayed as $\Delta R/R_b$ values calculated at the specific times (20, 90, 180, 240 s) after initiation of the vapor exposure of interest. Each point represents an average of five values produced by independent exposures to the analyte of interest, and the error bars indicate ±1 S.D. from the mean.

at all tested flow rates, regardless of the position of the detector relative to the point of analyte injection. This is expected because the analyte sorption process determines the steady-state value of $\Delta R/R_b$ [30], and because all of the detectors experienced essentially identical concentrations of analyte under such conditions.

Low vapor pressure analytes (i.e. analytes with large polymer/gas partition coefficients), however, produced different behavior. At high flow rates, all detectors produced essentially identical $\Delta R/R_b$ signals in the 20 s period after 240 s of analyte exposure, further confirming that the concentration of the analyte in proximity to each detector was similar and that the detectors themselves were very similar in response properties. However, at lower flow rates, lower $\Delta R/R_b$ values were observed in the 20 s period after 240 s of analyte exposure for the detectors to last encounter the vapor stream. To confirm that this effect was due to the physical location of the detector relative to the position of analyte flow injection, the direction of analyte flow in the chamber was again reversed and data were recollected for the entire sequence of analyte exposures. The lowest $\Delta R/R_b$ responses were again observed for detectors that were located farthest from the position of analyte injection. For both analyte injection positions, the responses of detectors located at positions remote from the position of analyte injection were still increasing at the end of the 260 s analyte exposure period.

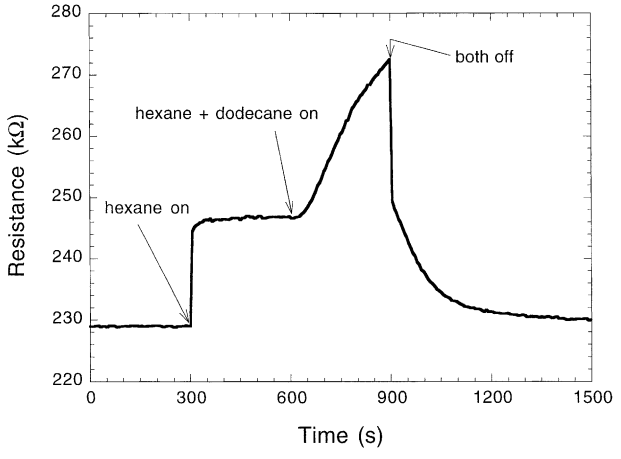


Fig. 6. Resistance response versus time for a PEVA-carbon black composite detector (located at position 7 in Scheme 2) exposed to hexane at $P/P^\circ = 0.10$ from 300 to 600 s, and then to a mixture of both hexane at $P/P^\circ = 0.10$ and dodecane at $P/P^\circ = 0.10$ from 600 to 900 s.

At times shorter than the period required to produce time-independent responses on all of the detectors in the array, the concentration of the low vapor pressure analyte stream is clearly depleted by sorption into the first region of polymer composite film that it encounters, and the analyte concentration in the boundary layer that is exposed to the film is decreased further as the gas flow progresses along the length of the polymer composite. For analytes of low vapor pressure, all detectors produced essentially identical responses at high flow rates after 240 s of exposure time, whereas at sufficiently low flow rates different responses were still observed after this exposure time for detectors located in different positions relative to the position of analyte injection into the chamber. In this transitional region of behavior, analysis of the relative signal strengths of the detectors in the array can provide information on the partition coefficient of the analyte into the polymer film of interest. Fig. 4 shows this effect quite clearly for hexane, dodecane, and tridecane.

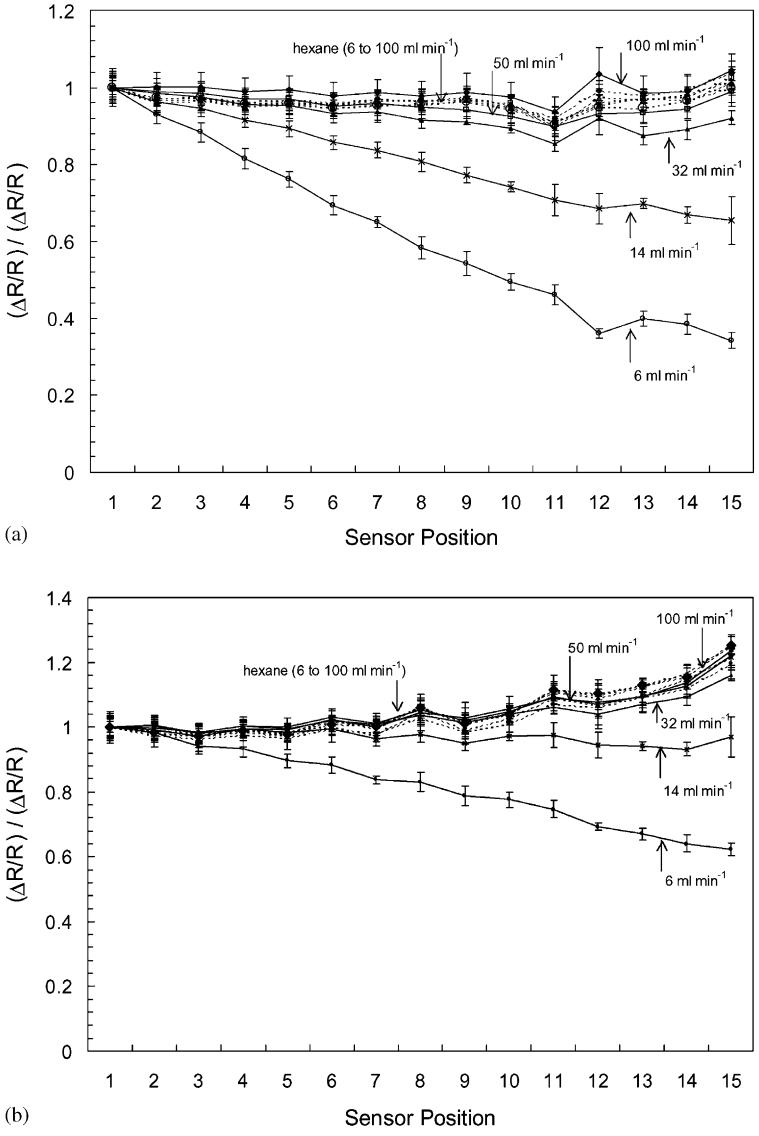


Fig. 7. Normalized relative differential resistance responses, $(\Delta R/R_b)/(\Delta R/R_b)_j$, of (a) carbon black-PEVA composite vapor detectors and (b) carbon black-PCL detectors to hexane (dashed lines) and dodecane (solid lines) at a constant activity of $P/P^\circ = 0.10$ in an air background. Data are averages of five responses, and the error bars indicate ± 1 S.D. values around the mean. Data were obtained from $\Delta R/R_b$ values observed between 240 and 260 s after initiation of exposure to the analyte of interest. For each flow rate, hexane and dodecane were alternately presented to the detectors. This procedure was repeated for each of five flow rates, proceeding sequentially from the lowest volumetric flow rate to the highest volumetric flow rate. This 10 exposure protocol was then repeated four times, producing 50 total exposures of analyte to the detectors. The detectors were arranged in a linear geometry as depicted in Scheme 2, and the analyte flowed from the leftmost detector (corresponding to the detector with the lowest numbered position) towards the rightmost detector. The volumetric flow rate was varied in five values between 6 and 100 ml min^{-1}. The data (and associated standard deviations) collected for each flow rate were separately normalized to the mean response value at that flow rate of analyte for the detector in position 1 ($j = 1$) of the array. Normalization constants (volumetric flow rate in ml min^{-1}) for the hexane data of (a) are: 10.8 (6); 10.5 (14); 10.0 (32); 10.3 (50); 11.1 (100). Normalization constants (volumetric flow rate in ml min^{-1}) for the dodecane data of (a) are: 26.7 (6); 20.4 (14); 16.8 (32); 16.5 (50); 16.8 (100). Normalization constants (volumetric flow rate in ml min^{-1}) for the hexane data of (b) are: 45.7 (6); 44.9 (14); 45.0 (32); 46.0 (50); 49.6 (100). Normalization constants (volumetric flow rate in ml min^{-1}) for the dodecane data of (b) are: 79.5 (6); 67.0 (14); 59.2 (32); 60.5 (50); 64.4 (100).

4.4. Flow system experiments with a signal/noise enhancement targeted towards an analyte's vapor pressure

The data presented above indicate that the noise decreases approximately as the square root of the detector area. Thus, for sufficient headspace volumes and quantities of sampled analyte such that the concentration of analyte sorbed into the polymer film remains constant as the detector area increases (as given by $K = C_p/C_v^{eq}$), an increased detector area will produce no change in the magnitude of the steady-state signal, a reduced value of the noise, and hence an increase in S/N ratio. However, for finite duration pulses of low vapor pressure compounds injected at low flow rates onto polymer films that have large polymer/gas partition coefficients,

analyte sorption will only effectively occur onto the subset of detectors that are encountered initially by the analyte flow. In this situation, increasing the detector area decreases the S/N ratio and additionally masks the spatiotemporal dependence of analyte sorption that can be used to discriminate between analytes of differing polymer/gas partition coefficients (Figs. 4–7). In this section, we describe the results of experiments designed to exploit both aspects of these properties of detector/analyte/flow interactions. To investigate this trade-off between detector S/N and detector area, detector arrays arranged as depicted in Scheme 3 were exposed to various analytes of interest. In this configuration, a detector film was deposited onto the edge of a printed circuit board substrate, and two other detector films of nominally identical composition were then deposited onto the two faces of the substrate. The face detector serves in essence as one large collection of detectors arranged linearly as in Scheme 2, thereby inherently averaging the responses, and providing reduced noise, for analytes with small polymer/gas partition coefficients. In contrast, the edge detector has a small area so that it can provide enhanced S/N performance for analytes with large polymer/gas partition

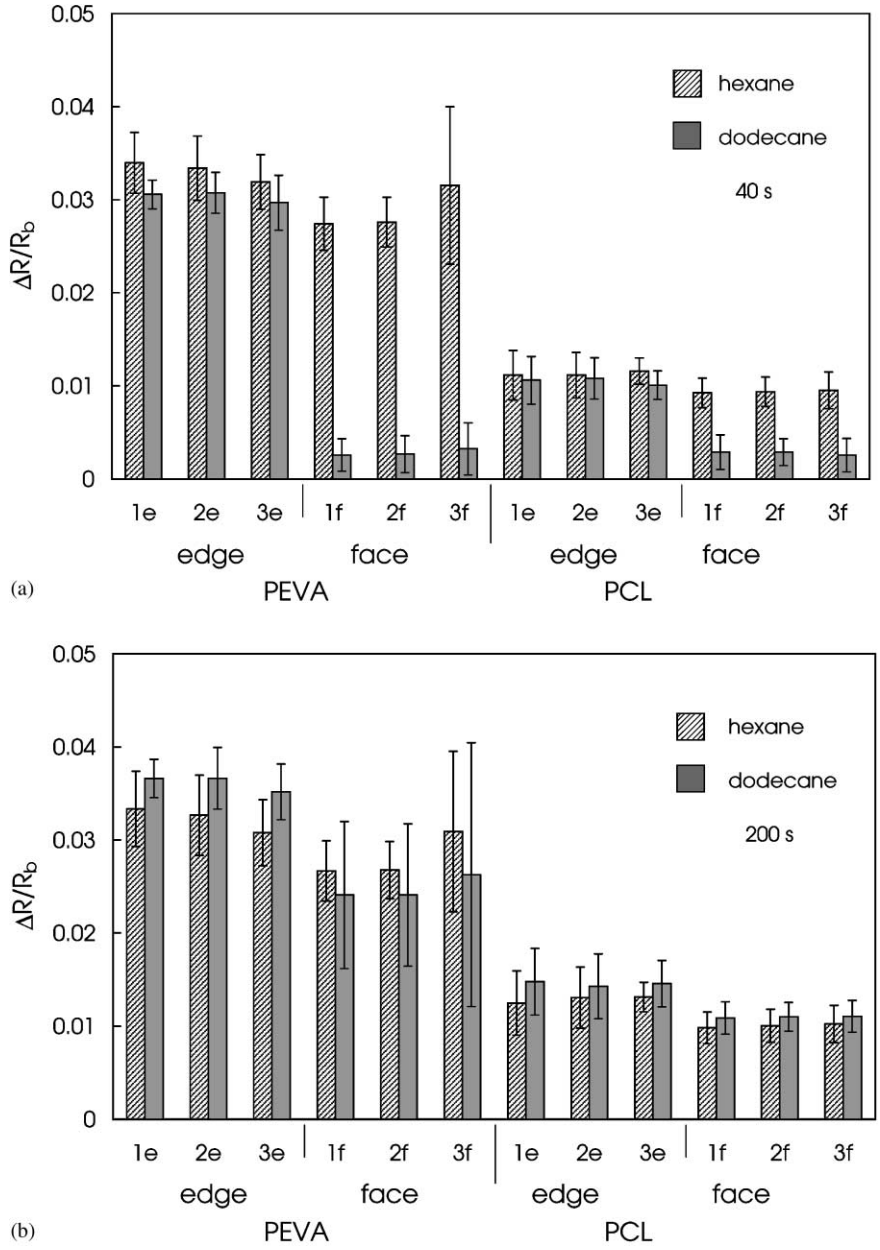


Fig. 8. Relative differential resistance responses to hexane (diagonal lines) and dodecane (gray) after (a) 40 s and (b) 200 s of carbon black composite detectors on the edge and face positions of the substrates arranged in the stacked configuration of Scheme 3. Data are averages of 30 $\Delta R/R_b$ responses, representing 10 responses to detectors of the same position on each of the three stack assemblies. The error bars indicate ± 1 S.D. value around the mean response value of these 30 detector/analyte combinations. Each analyte was delivered at $P/P^\circ = 0.050$ in a laboratory air background.

coefficients. Two such substrates were oriented so that the analyte flow encountered the leading edge of each detector first, and a component of this flow subsequently flowed along the faces of the substrate. One substrate had one polymer type forming its detectors and the other substrate had a separate, different carbon black-polymer composite material forming all of its detectors. The gaps between the substrates and the adjacent Al plates were sufficiently thin to insure that the flow would proceed in the desired direction. The entire experimental procedure and data collection were fully repeated three independent times, each time with two independently prepared substrates that were assembled into the configuration of Scheme 3.

The reported $\Delta R/R_b$ responses, N_{rms} values, and S/N values (Table 1) for each stack assembly are averages over the three detectors of the same geometry (face or edge) on a single substrate for 10 exposures to a given analyte. In Table 1, the results of the experiments on the three independently prepared stack devices are displayed separately. The average responses to high vapor pressure analytes (hexane and methanol) on the face detectors were between 75 and 100% of the magnitude of the responses on the edge detectors, while the lowest vapor pressure analyte, hexadecane, produced responses on the face detector that were all less than 15% of the values observed on the edge detectors (Table 1). This difference was much greater than the standard deviation of the responses of either all of the face detectors or all of the edge detectors on given substrate to an exposure to the analyte of interest.

The detector films on the leading edge of the substrate had 1/24 the area of the films on the face of the detectors, and therefore exhibited higher noise levels than the detectors on the face of the substrate. Noise values, N_{rms}, in the dc resistance readings measured using the multimeter were on average eight times higher for the PCL edge detectors than for the PCL face detectors, and were on average four times higher for the PEVA edge detectors than the PEVA face detectors (Table 1). The high vapor pressure analytes produced similar $\Delta R/R_b$ values on both detector types when exposed to methanol or hexane, hence the face detectors exhibited S/N ratios that reflected the decrease in noise produced by large volume detector films (Table 1). For 200 s exposures to hexane, S/N values were six times higher for PCL face detectors and were four times higher for PEVA face detectors than for the corresponding edge detectors. In contrast, for 200 s exposures to hexadecane, the analyte with the lowest vapor pressure, the S/N values were about twice as high on the leading edge detectors as on the face detectors. Clearly, the different geometric form factors and interactions with the analyte flow streamlines produced different performance characteristics from a S/N viewpoint for these different types of detectors.

The temporal evolution of the detector response properties can also be exploited to differentiate between analytes. As shown in Fig. 8a and b, the responses of the face and edge detectors to hexane were similar after 40 s of vapor presentation, and remained similar after 200 s. These hexane responses are similar in magnitude to the signals for dodecane after 200 s (Fig. 8b), and the two analytes could not easily be distinguished based on these data alone. However, the responses for these two analytes are clearly separable at 40 s (Fig. 8a), when the hexane has fully equilibrated with the given polymer film area, but the dodecane is still being depleted from the analyte sample due to its very high polymer/gas partition coefficient. The separation of these analytes as a function of time therefore demonstrates an increase in the resolving power attainable through the use of such spatiotemporal response information in conjunction with a spatially ordered array of vapor detectors.

5. Discussion

5.1. Detection limits of chemiresistor-based vapor detectors

We have previously reported that the steady-state $\Delta R/R_b$ values for various carbon black-polymer composite chemiresistors are linearly dependent on analyte concentration over a range of analyte/detector combinations and analyte concentrations [31]. The noise measurements reported herein, in conjunction with the previously reported dependence of $\Delta R/R_b$ on the partial pressure of the analyte [30] and the analyte/polymer sensitivity factors that can be deduced from such plots, allow estimation of the detection limits for various analyte/carbon black composite detector combinations. Two limiting cases will be considered: (a) high vapor pressure analytes, which have relatively small partition coefficients for sorption into the carbon black composite detectors and (b) low vapor pressure analytes, which generally sorb strongly and exhibit very large polymer/gas partition coefficients into the polymers of concern.

When the polymer/gas partition coefficient is relatively small, sufficient analyte will, in general, be present in the sampled volume to produce the equilibrium volume swelling of the entire available detector area. In this situation, too little detector volume is generally present to satisfy the optimum detector volume as given by Eq. (12). At constant film thickness, the steady-state $\Delta R/R_b$ value of a given carbon black-polymer composite is directly related to the swelling of the film. Thus, a given analyte concentration should produce the same steady-state $\Delta R/R_b$ signal in the film regardless of the area of such a detector.

Under these conditions, the scaling of the S/N (in a given measurement bandwidth) with detector area is determined by the dependence of the noise on detector area. As discussed above, the background noise of the carbon black composite chemiresistors at low measurement frequencies scales as $A^{-(1/2)}$. The S/N, and thus the detection limits of a particular carbon black-polymer composite detector towards a given analyte, therefore scale as $A^{1/2}$. The use of a detector film having the largest practical volume possible (up to the

Table 2
LOD for carbon black-polymer composite detectors and polymer film SAW detectors

	Polymer	LOD (μg/l)			
		Benzene	Cyclohexanone	Hexane	Nonane
Carbon black composite[a]	PEVA	1.8×10	1.5	4.0×10	1.3
	PCL	5.2×10^2	4.5×10	1.3×10^3	4.8×10
SAW[b]	Poly[bis(cyanoallyl)siloxane]	4.0×10^2	1.5×10	5.3×10^3	5.7×10^2
	Poly(methylphenylsiloxane)	3.0×10^2	1.4×10	1.5×10^3	1.1×10^2
	Poly(phenyl ether)	2.2×10^2	1.3×10	9.9×10^2	7.9×10
	Poly(isobutylene)	2.6×10^2	3.2×10	3.5×10^2	1.9×10

[a] Carbon black composite LOD are calculated from the slopes of $\Delta R/R_b$ vs. P/P° at 294 K in [31] using 3 S.D. noise values for 1 cm^2 of the same film type at average experimental film thickness values of 230 nm for PEVA and 80 nm for PCL.
[b] SAW values are taken from [32] for 158 MHz SAW oscillators at 298 K.

limit of optimum volume indicated by Eq. (12), or the volume at which the $1/f$ noise, for the measurement bandwidth, falls below the Johnson noise and the total noise no longer exhibits a dependence on volume) is thus the optimum detector design under such conditions.

Table 2 reports the S/N values and deduced limits of detection (LOD) for representative carbon black-polymer composite detectors with various vapor analytes, for 1 cm^2 of detector area. Table 2 also reports representative values taken from the literature for selected polymer-coated SAW vapor detectors [32]. For the given area, the detection limits are comparable for both types of signal transduction, although the carbon black composites exhibit somewhat higher sensitivity than the SAW devices for the analyte/polymer combinations chosen for comparison. We have only reported LOD as opposed to limits of classification; the former quantity depends only on the properties of the analyte/detector combination, while the latter quantity also depends on the test set of analytes presented to the array as well as on the algorithms used to perform the classification. In one particular situation evaluated in the literature, the limit of classification of an analyte was shown to be within a factor of three of the LOD of that same analyte, indicating that the limit of classification is likely to be on the same order of magnitude as the LOD, at least for some tasks [33].

In the limit where the analyte exhibits a very strong sorption into the polymer film of the carbon black composite detector, the S/N optimization methodology is quite different. As given in Eq. (5), the sorption process under such conditions will be limited by the amount of analyte in the sampled volume. The $\Delta R/R_b$ signal of the detector is proportional to the swelling of the detector film [25], so increasing the detector area will reduce the signal (by diluting the fixed amount of sorbed analyte into a correspondingly larger volume of polymer). As long as the swelling is linearly dependent on the concentration of analyte sorbed into the polymer [25], this dilution will produce a linear decrease in the $\Delta R/R_b$ signal with increased detector volume. Because the noise scales as $A^{-(1/2)}$ (at constant film thickness), the S/N under such conditions scales as $A^{-(1/2)}$ and small detector areas are favored. In fact, the design goal under such conditions is to insure that the most analyte is sorbed into the least area of detector film, and signals should only be acquired from the limited, highly analyte-swollen, portion of the detector. For example, this design principle is appropriate for applications in which 2,4-dinitrotoluene, a vapor component above-buried land mines [17,34,35], or other low vapor pressure analytes, are being detected. This principle is exemplified in the detector arrangement of Scheme 3.

This relationship also has implications for sample chamber design of vapor detector arrays. For example, assume that the analyte headspace is comprised of a vertical column equal in area to the area of the detector film, and that the detector film thickness is 1.0×10^{-4} cm. For $K = 1.0 \times 10^2$, the sorbed analyte would come to equilibrium in a closed system with the vapor phase analyte that is contained in a headspace thickness of 1.0×10^{-2}. Increasing the thickness of the headspace would simply provide more analyte than is needed to attain the optimal S/N ratio for the detector response and would require introduction of more sample into the headspace chamber. Alternatively, larger detector areas could be used advantageously to obtain improved S/N ratios from the increased number of analyte molecules available in a thicker (closed system) headspace chamber. In contrast, for $K = 1.0 \times 10^7$, a 1.0×10^{-4} cm thick detector film would sorb essentially all of the analyte from a closed system having a 1000 cm thick headspace. A 2.6 cm^2 area of such a detector film could sorb essentially all of the analyte in a 3.0×10^{-2} cm thick headspace (cf. Scheme 2) that is supplied at a continuous volumetric flow rate of 10 cm^3 min^{-1} for a period of 260 min. For shorter analyte injection times (at constant analyte flow rate), smaller detector areas are optimal, because otherwise the fixed amount of analyte is distributed into too large a detector area, thereby diminishing the magnitude of the signal and deleteriously affecting the S/N ratio of the detector.

Given the relationships reported previously between the mass loading of analyte and the $\Delta R/R_b$ values for carbon black composite vapor detectors [25], in conjunction with the background noise levels reported herein, detection limits can be evaluated in the high sorption/low analyte vapor

pressure regime. At a noise level of ≈10 ppm, and with a $\Delta R/R_b = 0.10$ produced at a mass loading of 5.0 mg of analyte sorbed into 1 cm^2 of polymer [25], the computed 3 S.D. detection limit of a PCL-carbon black composite is 1.5 ng cm^{-2}. This value can only be reached in practice if an efficient sampling and delivery system is available, such that the full amount of the sampled analyte can be delivered effectively to the 1 cm^2 area of the detector film. Of note is that the detection limit scales inversely with the film area and linearly with the efficiency of delivery analyte to the sampled film area.

In the intermediate sorption/partition coefficient regime, an optimum detector volume clearly exists for which the S/N, and therefore the detection limit performance, of a particular analyte/polymer combination is maximized. This detector volume, and consequently the optimum film area, depends on only the analyte/polymer partition coefficient and the sampled analyte volume available for sorption into the detector, and is readily calculated from Eq. (12). The S/N can therefore be optimized for different vapor pressure analytes through control over the form factor of the detector film, as demonstrated herein both theoretically and experimentally.

5.2. Geometric considerations of the detector for optimum signal/noise performance with a fixed sample volume

The dependence of optimum detector area on the analyte/polymer partition coefficient can also be used advantageously in the classification of analytes and analyte mixtures. In such a system, analytes with a high polymer/gas partition coefficient (generally analytes with low vapor pressures) would be sorbed into the smallest detector area possible, producing the largest signal and therefore the largest S/N ratio for that particular analyte/polymer/sampler combination. Higher vapor pressure analytes are, in turn, detected with higher S/N performance at detectors having larger film areas. In fact, an array of contacts spaced exponentially along a polymer film could be used advantageously to gain information on the sorption coefficients of analytes into polymers, and therefore can provide additional classification information on analytes and analyte mixtures relative only to equilibrium $\Delta R/R_b$ values on a detector film having a single, fixed form factor for all analytes. A relatively simple demonstration of this principle was performed herein, in which two analytes were not readily distinguishable based on their responses on a single type of detector located on the edges of the substrates, but the analytes were clearly distinguishable when information on the relative response values of detectors located on the edge versus the face of the substrate was used in the data analysis. Additional information is available if the analyte flow rate is also varied over the detector array.

Variation in the geometric form factor of detectors also could potentially have practical advantages in the implementation of instruments based on arrays of vapor detectors. Although information similar to that produced by a collection of spatiotemporally arrayed detectors could in principle be obtained from an analysis of the time response of a collection of detectors that are equivalent both geometrically and with respect to the point of analyte injection, the spatiotemporal implementation discussed above has the advantage that analytes are detected on films that have nearly optimal S/N for the analyte of interest. In addition, electronically referencing the response of a face detector to that of an edge detector, for example, would allow nulling of the response to a high vapor pressure analyte and subsequent amplification of only those signals arising from low vapor pressure analytes.

Another feature of note is the possible relationship of this type of detector design to the biological olfactory system. Sobel et al. have recently reported differences in human perception of binary odorant mixtures that contained an odorant having a high vapor pressure and an odorant having a low vapor pressure [36]. The perceptual changes were shown to be produced by differences in flow rate between the two nostrils of the human subjects. Although the mixture contained fixed concentrations of each odorant, the subjects perceived the mixture to be enriched in the lower vapor pressure odorant when sampled through the higher flow rate nostril, and the same mixture was perceived to be enriched in the higher vapor pressure odorant when sampled through the lower flow rate nostril. The perceived responses changed when the flow rates in the nostrils were naturally interchanged due to normal physiological processes. The authors concluded that the spatiotemporally dependent responses of olfactory receptors are useful to humans in resolving certain odor mixtures and in obtaining additional information on the composition of odorants [36]. The relatively primitive system investigated herein demonstrates an analogous principle in a non-biological array of broadly cross-reactive vapor detectors. Thus a differential measurement of the response of two conducting polymer composite detector arrays, sampling the same analyte at different injected flow rates, might provide an interesting platform for evaluating the degree to which spatiotemporal response information can be used to obtain additional classification information in odor detection.

6. Conclusions

The dependence of the relative power spectral density on the volume of carbon black-polymer composite vapor detectors was of the form $S_n \propto 1/\mathcal{V}^n$, with $n = 1$ for PEVA-carbon black detectors and $n = 0.6$ for PCL-carbon black detectors in the frequency range of 1–800 Hz. Analytes with moderate polymer/gas partition coefficients produce the same $\Delta R/R_b$ response values on detectors of constant film thickness but of different area, so under these conditions the S/N is optimized for detectors of very large area. In contrast,

for finite quantities of injected sample, analytes with high polymer/gas partition coefficients produce much larger $\Delta R/R_b$ values on detectors of small area that are positioned to best sample the injected analyte flow. For such detector/analyte combinations, detectors of small area will exhibit significantly better vapor detection sensitivity. Manipulation of the geometric form factor of carbon black composite vapor detectors thus provides a facile method for optimizing the S/N performance for a particular detector/analyte combination of interest. An array of nominally identical polymer-carbon black detectors arranged linearly relative to the analyte flow path produces different spatiotemporal response patterns for analytes having different polymer/gas partition coefficients. Analytes with moderate polymer/gas partition coefficients produce the same signals on all detectors over a range of flow rates, whereas before steady state is reached on all of the detectors, analytes with very large polymer/gas partition coefficients produce signals that are highly dependent on the analyte flow rate and the spatial position of the detector in the array. Such a configuration produces useful information on the composition of binary analyte mixtures and adds classification information to an array of compositionally different conducting polymer composite vapor detectors.

Acknowledgements

We acknowledge NASA, DOE, the NIH, and an Army MURI for their generous support of this work.

References

[1] M.C. Lonergan, E.J. Severin, B.J. Doleman, S.A. Beaber, R.H. Grubbs, N.S. Lewis, Array-based vapor sensing using chemically sensitive, carbon black-polymer resistors, Chem. Mater. 8 (1996) 2298–2312.
[2] H.V. Shurmer, P. Corcoran, J.W. Gardner, Integrated arrays of gas sensors using conducting polymers with molecular-sieves, Sens. Actuators B 4 (1991) 29–33.
[3] T.C. Pearce, J.W. Gardner, S. Friel, P.N. Bartlett, N. Blair, An electronic nose for monitoring the flavor of beers, Analyst 118 (1993) 371–377.
[4] M.S. Freund, N.S. Lewis, A chemically diverse conducting polymer-based electronic nose, Proc. Natl. Acad. Sci. U.S.A. 92 (1995) 2652–2656.
[5] J.M. Slater, E.J. Watt, An examination of ammonia poly(pyrrole) interactions by piezoelectric and conductivity measurements, Analyst 116 (1991) 1125–1130.
[6] J.W. Grate, S.J. Martin, R.M. White, Acoustic-wave microsensors. 2, Anal. Chem. 65 (1993) A987–A996.
[7] J.W. Grate, S.J. Martin, R.M. White, Acoustic-wave microsensors. 1, Anal. Chem. 65 (1993) A940–A948.
[8] C. Cornila, A. Hierlemann, R. Lenggenhager, P. Malcovati, H. Baltes, G. Noetzel, U. Weimar, W. Gopel, Capacitive sensors in CMOS technology with polymer coating, Sens. Actuators B 25 (1995) 357–361.
[9] S.R. Johnson, J.M. Sutter, H.L. Engelhardt, P.C. Jurs, J. White, J.S. Kauer, T.A. Dickinson, D.R. Walt, Identification of multiple analytes using an optical sensor array and pattern recognition neural networks, Anal. Chem. 69 (1997) 4641–4648.
[10] T.A. Dickinson, J. White, J.S. Kauer, D.R. Walt, A chemical-detecting system based on a cross-reactive optical sensor array, Nature 382 (1996) 697–700.
[11] J. White, J.S. Kauer, T.A. Dickinson, D.R. Walt, Rapid analyte recognition in a device based on optical sensors and the olfactory system, Anal. Chem. 68 (1996) 2191–2202.
[12] G. Harsanyi, Polymer Films in Sensor Applications, Technomic, Lancaster, 1995, pp. 136–140.
[13] K.J. Albert, N.S. Lewis, C.L. Schauer, G.A. Sotzing, S.E. Stitzel, T.P. Vaid, D.R. Walt, Cross-reactive chemical sensor arrays, Chem. Rev. 100 (2000) 2595–2626.
[14] B.J. Doleman, R.D. Sanner, E.J. Severin, R.H. Grubbs, N.S. Lewis, Use of compatible polymer blends to fabricate arrays of carbon black-polymer composite vapor detectors, Anal. Chem. 70 (1998) 2560–2564.
[15] E. Benes, M. Groschl, W. Burger, M. Schmid, Sensors based on piezoelectric resonators, Sens. Actuators A 48 (1995) 1–21.
[16] J.W. Grate, M. Klusty, Surface acoustic-wave vapor sensors based on resonator devices, Anal. Chem. 63 (1991) 1719–1727.
[17] S.M. Briglin, M.C. Burl, M.S. Freund, N.S. Lewis, A. Matzger, D.N. Ortiz, P. Tokumaru, Progress in use of carbon black-polymer composite vapor detector arrays for land mine detection, Proc. SPIE Int. Opt. Eng. 4038 (2000).
[18] T.H. Wilmshurst, Signal Recovery from Noise in Electronic Instrumentation, Adam Hilger, Boston, 1985, p. 60.
[19] J.R. Larry, R.M. Rosenberg, R.O. Uhler, Thick-film technology: an introduction to the materials, IEEE Trans. Comp. Hybrids Manufact. Technol. 3 (1980) 211–225.
[20] M.B. Weissman, 1/*f* Noise and other slow, nonexponential kinetics in condensed matter, Rev. Mod. Phys. 60 (1988) 537–571.
[21] A. Dziedzic, A. Kolek, 1/*f* Noise in polymer thick-film resistors, J. Phys. D: Appl. Phys. 31 (1998) 2091–2097.
[22] J.H. Scofield, D.H. Darling, W.W. Webb, Exclusion of temperature fluctuations as the source of 1/*f* noise in metal films, Phys. Rev. B 24 (1981) 7450–7453.
[23] S.L. Fu, M.S. Liang, T. Shiramatsu, T.S. Wu, Electrical characteristics of polymer thick film resistors. I. Experimental results, IEEE Trans. Comp. Hybrids Manufact. Technol. 4 (1981) 283–288.
[24] M.J. Deen, S. Rumyantsev, J. Orchard-Webb, Low frequency noise in heavily doped polysilicon thin film resistors, J. Vac. Sci. Technol. B 16 (1998) 1881–1884.
[25] E.J. Severin, N.S. Lewis, Relationships among resonant frequency changes on a coated quartz crystal microbalance, thickness changes, and resistance responses of polymer-carbon black composite chemiresistors, Anal. Chem. 72 (2000) 2008–2015.
[26] A. Peled, R.E. Johanson, Y. Zloof, S.O. Kasap, 1/*f* Noise in bismuth ruthenate based thick-film resistors, IEEE Trans. Comp. Package Manufact. Technol. A 20 (1997) 355–360.
[27] P. Horowitz, W. Hill, The Art of Electronics, 2nd Edition, Cambridge University Press, New York, 1989, pp. 449–454.
[28] C. Lu, A.W. Czanderna, Applications of Piezoelectric Quartz Crystal Microbalances, Elsevier, New York, 1984, pp. 251–253.
[29] D.A. Buttry, in: A.J. Bard (Ed.), Electroanalytical Chemistry: A Series of Advances, Vol. 17, Marcel Dekker, New York, 1991, Chapter 17, pp. 2–82.
[30] B.J. Doleman, E.J. Severin, N.S. Lewis, Trends in odor intensity for human and electronic noses: relative roles of odorant vapor pressure vs. molecularly specific odorant binding, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 5442–5447.
[31] E.J. Severin, B.J. Doleman, N.S. Lewis, An investigation of the concentration dependence and response to analyte mixtures of carbon black/insulating organic polymer composite vapor detectors, Anal. Chem. 72 (2000) 658–668.
[32] S.J. Patrash, E.T. Zellers, Characterization of polymeric surface-acoustic-wave sensor coatings and semiempirical models of sensor responses to organic vapors, Anal. Chem. 65 (1993) 2055–2066.

[33] E.T. Zellers, J. Park, T. Hsu, W.A. Groves, Establishing a limit of recognition for a vapor sensor array, Anal. Chem. 70 (1998) 4191–4201.

[34] M. la Grone, C. Cumming, M. Fisher, D. Reust, R. Taylor, Landmine detection by chemical signature: detection of vapors of nitroaromatic compounds by fluorescence quenching of novel polymer materials, Proc. SPIE Int. Opt. Eng. 3710 (1999) 409–420.

[35] V. George, T.F. Jenkins, D.C. Leggett, J.H. Cragin, J. Phelan, J. Oxley, J. Pennington, Progress on determining the vapor signature of a buried land mine, Proc. SPIE Int. Opt. Eng. 3710 (1999) 258–269.

[36] N. Sobel, R.M. Khan, A. Saltman, E.V. Sullivan, J.D.E. Gabrieli, Olfaction—the world smells different to each nostril, Nature 402 (1999) 35–37.

ELSEVIER

Sensors and Actuators B 72 (2001) 149–159

SENSORS AND ACTUATORS B CHEMICAL

www.elsevier.nl/locate/sensorb

Assessing the ability to predict human percepts of odor quality from the detector responses of a conducting polymer composite-based electronic nose

Michael C. Burl[b], Brett J. Doleman[a], Amanda Schaffer[a,b], Nathan S. Lewis[a,*]

[a] *127-72 Noyes Laboratory, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA*
[b] *126-347 Jet Propulsion Laboratory, 4800 Oak Grove Blvd, Pasadena, CA 91109, USA*

Received 30 May 2000; received in revised form 25 August 2000; accepted 2 September 2000

Abstract

The responses of a conducting polymer composite ''electronic nose'' detector array were used to predict human perceptual descriptors of odor quality for a selected test set of analytes. The single-component odorants investigated in this work included molecules that are chemically quite distinct from each other, as well as molecules that are chemically similar to each other but which are perceived as having distinct odor qualities by humans. Each analyte produced a different, characteristic response pattern on the electronic nose array, with the signal strength on each detector reflecting the relative binding of the odorant into the various conducting polymer composites of the detector array. A ''human perceptual space'' was defined by reference to English language descriptors that are frequently used to describe odors. Data analysis techniques, including standard regression, nearest-neighbor prediction, principal components regression, partial least squares regression, and feature subset selection, were then used to determine mappings from electronic nose measurements to this human perceptual space. The effectiveness of the derived mappings was evaluated by comparison with average human perceptual data published by Dravnieks. For specific descriptors, some models provided cross-validated predictions that correlated well with the human data (above the 0.60 level), but none of the models could accurately predict the human values for more than a few descriptors. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Electronic nose; Conducting polymer composites; Odor quality; Human perception; Olfaction; Regression models; Feature selection

1. Introduction

Arrays of broadly responsive vapor detectors are attracting increasing interest as ''artificial noses'' [1–3]. Like the receptors in the mammalian olfactory system [4], each detector in an ''artificial nose'' responds to more than one analyte, and each analyte elicits a response from more than one detector [1–3]. Pattern recognition algorithms are then used to classify, identify, and in some cases quantify, an analyte in the vapor phase. One motivation for studying such arrays is eventually to learn enough about the process of olfaction to construct a man-made, functional analogue of a mammalian olfactory system [5].

Perhaps the ultimate challenge for an artificially-constructed olfactory system is to mimic faithfully the mapping of an odorant-induced detector response pattern to the quality of an odor, e.g. to its ''minty-ness'', as perceived by a human. This task is difficult because the human olfactory system is highly nonlinear in many respects. For example, perceived odor intensity is a nonlinear function of analyte concentration [6]. In addition, qualitatively different human percepts are often produced by varying the concentration of a given odorant. Cross-adaptation, masking, and other processes involved with the human perception of odor mixtures [7] further complicate the signal processing involved in olfaction [8]. A further level of complexity results because humans are variable genetically in their perception of many odorants [9]. Thus, any static, generically-constructed, artificial olfactory device could at best capture some average human perceptual processes for a representative set of odorants. Of course, this does not eliminate the possibility of a ''trainable'' device that could be tuned to match the perceptual profile of a specific individual, however, developing such a system poses yet another set of challenges.

The artificial nose implementation that was used in this study consists of an array of conducting polymer composites, in which each detector material of the array has regions of a conductive material interspersed into regions of an insulating organic polymer [10,11]. The conductive material is typically carbon black although it could also be an

* Corresponding author. Tel.: +1-626-395-6335; fax: +1-626-795-7487.
E-mail address: nslewis@caltech.edu (N.S. Lewis).

0925-4005/01/$ – see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0925-4005(00)00645-6

inorganic metal or an organic electrically-conductive polymer. Sorption of an odorant into the polymer produces a swelling of the polymer film, which in turn leads to an increase in the dc electrical resistance of the detector. The electrical output signals from an array of such detectors are then transferred to a central processing unit for odorant analysis and classification. This implementation of an electronic nose was chosen for study because it is readily investigated experimentally [11], allows inclusion of a chemically diverse set of detectors [12], has been shown to parallel mean human olfactory detection threshold behavior for several classes of organic vapors [10], and has been shown in selected test cases to parallel human and monkey olfaction in the positive correlation between its discrimination ability and the chemical dissimilarity between members of a pair of odorants [13].

The specific focus of the current study was to investigate whether the responses of an array of such detectors could be used to predict accurately the perceived quality of an odorant as reported by human panelists. Only chemically pure single-component analytes were investigated, due to the further complications described above relating to the human olfactory perception of odorant mixtures. Data on human perception of odor quality for a variety of odorants were obtained from tabulations available in the literature [14]. Electronic nose responses were collected for a selected subset of the same compounds. The odorants investigated in our work included molecules that are chemically similar but which are perceived as being different by humans, as well as molecules that are chemically quite distinct from each other. Successful odor quality prediction is critical not only for meeting the intellectual challenge of constructing an artificial olfactory system, but also for many industrial quality control applications of an artificial nose in which product assessments (good/bad) must be made with respect to human perception rather than with respect to changes in the chemical composition of the odors of concern [15–17].

2. Experimental

2.1. Chemicals and data collection

Twenty-one odorants (Table 1) were evaluated in this work. All chemicals were obtained from Aldrich Chemical Corp. and were used as received. Sets of chemically homologous odorants (for example, a series of straight-chain alcohols, a series of aliphatic esters, a series of straight-chain aliphatic acids, a series of benzene derivatives, etc.) were chosen such that the odors were associated with common, but not identical, human odor descriptors both within a set and between sets of odorants. A total of 20 insulating polymers were used to form the carbon black/polymer composite detectors in the electronic nose (Table 2). Detectors were fabricated as described previously [12].

Table 1
Odorants used in this study

1	1-Butanol
2	1-Hexanol
3	1-Heptanol
4	1-Octanol
5	Ethyl propionate
6	Ethyl butyrate
7	Propyl butyrate
8	Amyl butyrate
9	Isopentyl acetate
10	Pentanoic acid
11	Hexanoic acid
12	Toluene
13	Anisole
14	Phenyl ethanol
15	Phenyl acetylene
16	Tetrahydrothiophene
17	Thiophene
18	Butanoic acid
19	Pyridine
20	Citral
21	Limonene

All odorant exposures were performed using a computer-controlled vapor generation and control system that regulated the identity, concentration, exposure time, and flow rate of the analyte above the detectors [18]. The experimental protocol for each odorant exposure was 5 min of clean air flow, followed by 5 min of air flow containing the odorant at a partial pressure corresponding to 5% of its vapor pressure, followed by another 5 min of clean air flow. The detectors were exposed to each odorant a minimum of 10 times. Analyte identities were varied in random order between each exposure.

Table 2
Polymers contained in the detectors of the carbon black/polymer composite electronic nose array

Detector	Polymer
1	Poly(4-vinyl phenol)
2	Poly(*N*-vinylpyrrolidone)
3	Poly(sulfone)
4	Poly(methyl methacrylate)
5	Poly(caprolactone)
6	Poly(ethylene-*co*-vinyl acetate), 82% ethylene
7	Poly(ethylene oxide)
8	Poly(ethylene)
9	Poly(vinylidene fluoride)
10	Poly(ethylene glycol)
11	Poly(vinyl acetate)
12	Poly(styrene)
13	Poly(butadiene)
14	Poly(styrene-*co*-allyl alcohol)
15	Poly(α-methylstyrene)
16	Hydroxypropyl cellulose
17	Poly(styrene sulfonic acid)
18	Poly(carbonate bisphenol A)
19	Poly(epichlorohydrin)
20	Poly(styrene-*co*-butadiene)

Only the steady-state response data were used in analysis of the electronic nose array signals. Specifically, the data were reduced to produce a $\Delta R/R_b$ value for each detector, where R_b is the drift-corrected baseline response of the detector during the analyte exposure period and ΔR is the steady-state differential resistance response of the detector with respect to the value of R_b. The data for each exposure were then expressed as a response vector, with each component of the vector corresponding to the $\Delta R/R_b$ value of a particular detector. The results from individual exposures to a given odorant were averaged to produce a single twenty-dimensional vector that described the response of the detector array to each odorant. The electronic nose measurements were, thus, reduced to a 21×20 matrix, $\boldsymbol{M}$, whose rows corresponded to different odorants and whose columns corresponded to different detectors.

The measurement data can be visualized to some extent by performing principal components analysis (PCA) on the raw twenty-dimensional measurement space and projecting the data onto the two leading principal component directions. Fig. 1 shows all of the odorants in this PCA space. The numerical label next to each point can be translated into a chemical name using Table 1. Note, for example, that points 1–4, which correspond to straight chain aliphatic alcohols, are well-clustered in the principal components space.

2.2. Perceptual odor quality

Perceptual odor quality values for humans were obtained from Dravnieks' *Atlas of Odor Character Profiles*. [14]. In Dravnieks' study, over 100 people of both sexes, spanning a wide range of ages, and including smokers as well as non-smokers, were evaluated. The rationale for using such a diverse group of panelists was apparently to insure that the reported percepts would be consistent with those of the population at large. Each participant was asked to smell a collection of odorants and was instructed to assign a score from 0 through 5 to each of 146 different descriptors (adjectives) that are used in the English language to describe odors. For example, a panelist could give an odorant a score of 3 in the ''etherish, anesthetic'' category, 4 in the ''minty'' category, and 0 in each of the remaining categories. Scores are intended to reflect the degree to which the panelist believes that a descriptor is appropriate for a given odorant, with a value of 0 meaning not appropriate. As described by Dravnieks, care was exercised in the experimental procedure to insure that artificial biases were not introduced into the results. Of the 146 descriptors considered by Dravnieks, the seventeen descriptors listed in Table 3 were selected for use in our study based on the frequency and extent to which they were used by the panelists to describe our selected set of test odorants (Table 1).

For the purposes of our study, a limitation with Dravnieks' *Atlas* is that only averages across the entire group of panelists are provided, so score profiles for individual participants are not available. Also, the variance (or the distribution) of scores given to a particular odorant–descriptor pair was not reported. Instead, the available data for each odorant–descriptor pair consist of two quantities: percentage of usage and percentage of applicability. The usage indicates

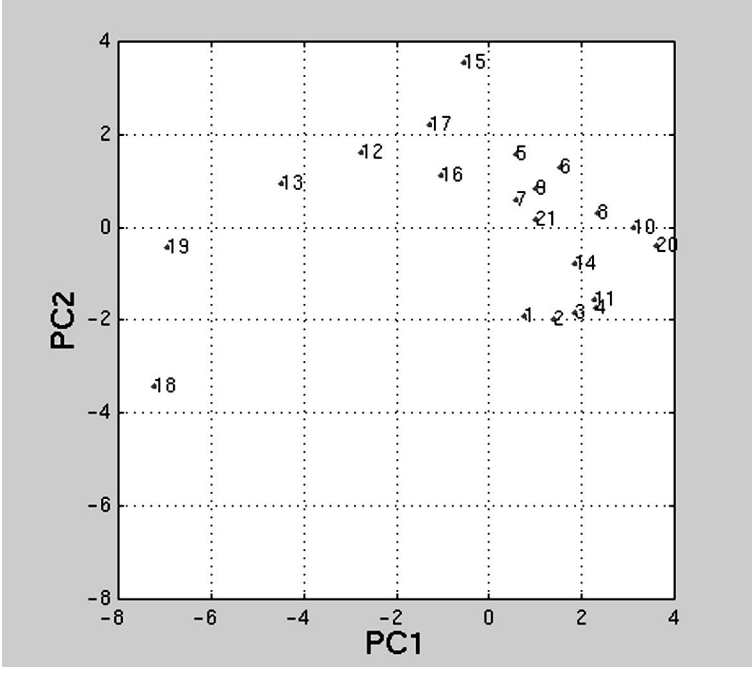


Fig. 1. Conducting polymer composite detector response data for the 21 odorants in two-dimensional principal component space. The numerical labels associated with each point correspond to the analytes listed in Table 1.

Table 3
Scent descriptors used in this study[a]

1	Fruity (citrus)
2	Fruity (non-citrus)
3	Floral
4	Minty
5	Etherish
6	Gasoline
7	Sharp, pungent
8	Oily
9	Putrid, foul, decayed
10	Woody
11	Sweet
12	Herbal
13	Musty
14	Medicinal
15	Sour
16	Paint
17	Sweaty

[a] From [14].

the percentage of panelists who assigned a non-zero score to the descriptor for the given odorant. This quantity can be interpreted as a probability value. The percentage of applicability is the geometric mean of the usage and the average score level assigned by the panelists.

Dravnieks suggests that the percentage of applicability is "the most equitable indicator of the descriptor applicability" to human perception. However, we have observed that the usage, score level, and applicability are all highly correlated. For example, Fig. 2 shows a plot of the score level versus usage for the chemicals and descriptors used in our study. A clear functional relationship is apparent between the score level and usage. However, because the relationships between the quantities are nonlinear, it is not clear which, if either, quantity will be more easily predicted from the electronic nose measurements. Feature vectors for the electronic nose response data and values of the organoleptic descriptors for the 21 odorants of Table 1 are available at http://www-aig.jpl.nasa.gov/mls/home/burl/data/odor_-quality/.

2.3. Data analysis

The goal is to assess the degree to which the response of the electronic nose to a given odorant can quantitatively predict the usage, score level, and/or applicability that would be provided on average by a group of human panelists for each of the seventeen descriptors. The different approaches that have been explored towards this goal are described below.

2.3.1. Standard regression and nearest-neighbor approaches

2.3.1.1. Standard linear regression. With the electronic nose measurements expressed as a 21×20 matrix $\boldsymbol{M}$ and the human panelist ratings for a particular descriptor/quantity expressed as a 21×1 vector $\boldsymbol{h}$, the problem reduces to finding a vector-valued function f such that the prediction $\boldsymbol{h}_p = f(\boldsymbol{M})$ is approximately equal to $\boldsymbol{h}$. Of the many possible classes of functions f, we restricted our attention to linear ($f(\boldsymbol{M}) = \boldsymbol{M}\boldsymbol{w}$) and affine ($f(\boldsymbol{M}) = \boldsymbol{M}\boldsymbol{w} + \boldsymbol{w}_0$) models, as well as "clipped" linear and affine models

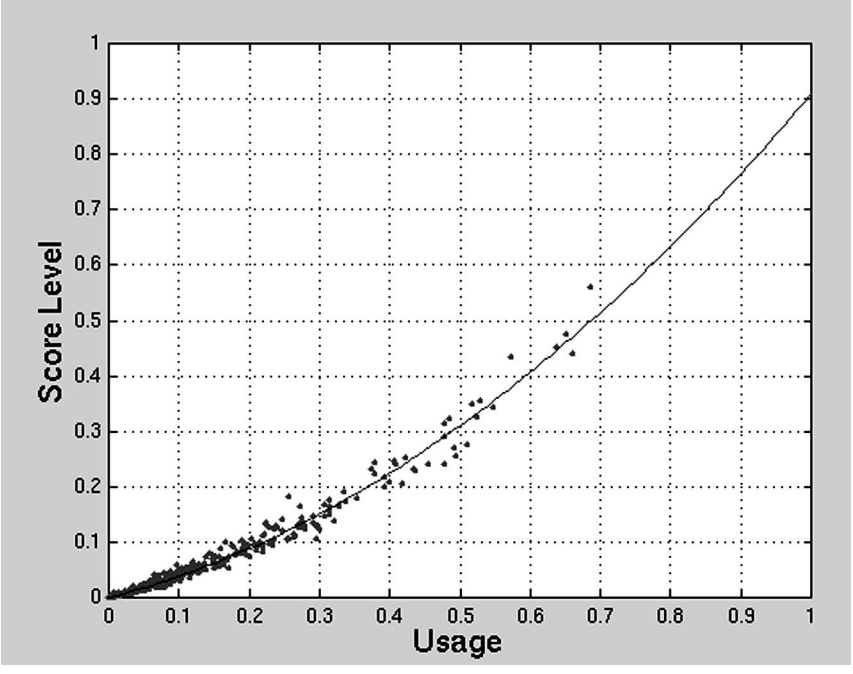


Fig. 2. Plot of the average human-assigned score level vs. percent usage for each chemical/descriptor pair in our study. The two quantities are clearly highly correlated, with s approximately equal to $0.33u + 0.58u^2$.

that incorporate a nonlinearity on the output to confine predictions to the range [0, 1], the same range as for the human perceptual data. Note that the affine form can be reduced to the linear case by augmenting the measurement matrix $\boldsymbol{M}$ with an additional column of ones and increasing the dimensionality of the weight vector by one. Hence, the linear form can be focused on without loss of generality.

The weight vector that minimizes the mean squared error between the predictions and the targets is well known:

$$\boldsymbol{w} = (\boldsymbol{M}'\boldsymbol{M})^{-1}\boldsymbol{M}'\boldsymbol{h} \quad (1)$$

Given a 1×20 measurement vector $\boldsymbol{m}$ for an odorant whose quality descriptors are to be predicted, the regression model predicts: $\boldsymbol{h}_p = \boldsymbol{m}\boldsymbol{w}$. Note that for numerical calculations the weight vector $\boldsymbol{w}$ is generally obtained by solving the equation $(\boldsymbol{M}'\boldsymbol{M})\boldsymbol{w} = \boldsymbol{M}'\boldsymbol{h}$ via LU decomposition and backsubstitution, rather than by computing the inverse of $(\boldsymbol{M}'\boldsymbol{M})$ [19].

2.3.1.2. Nearest neighbor models. A different type of predictor is based on the idea of nearest neighbors. If two odorants have similar electronic nose patterns, then it might be expected that the descriptors provided by human observers for these two odorants would also be very similar. The nearest neighbor model makes a prediction for a test odorant having an electronic nose signature $\boldsymbol{m}$ by first finding the row in $\boldsymbol{M}$ that is most similar to $\boldsymbol{m}$ (e.g. using a Euclidean distance metric). The human perceptual value for this nearest neighbor is then taken as the prediction for the test odorant. The reference library of electronic nose measurements imparts a partitioning of the measurement space into distinct regions in which a single library example will be the nearest neighbor of any example that falls into the region. The predictions for any example falling in this region will be the same as for the library example.

2.3.2. Basic testing procedure

A leave-one-out cross-validation procedure was used to evaluate the performance of the predictors. In this approach, the electronic nose signals for one odorant were withheld for use in testing. The remaining $N - 1(=20)$ odorants were then used to train the model (i.e. to determine the weight vector $\boldsymbol{w}$ in the regression approach or to serve as the reference library in the nearest neighbor approach). This process was repeated N times, with each odorant serving a turn as the holdout example. The predictions of both models for all odorants were then compared to the actual human target data for the odorants of interest.

3. Results

3.1. Performance of standard regression and nearest neighbor approaches

Fig. 3a shows the correlation coefficients between the predictions of the clipped linear regression model and the odor targets. The integer-valued x-coordinates (values from 1 to 17) represent the particular descriptors indicated in Table 3. The y-axis shows the correlation coefficient achieved for each descriptor/quantity (u = usage, s = score, a = applicability). Overall, the regression predictor performed poorly. Only 3 of the 17 descriptors (''floral'', ''sour'', and ''paint'') have at least one quantity predicted above the 0.60 level. The median correlation coefficient across all descriptors/quantities for this model is 0.21. The predictability of usage, score, or applicability were all fairly similar, having median values of 0.21, 0.21, and 0.22, respectively.

Fig. 3b presents the correlation coefficients between the predictions of the nearest neighbor model and targets. Overall, this predictor also performed poorly. Only 3 of the 17 descriptors (''fruity non-citrus'', ''woody'', and ''sweet'') have at least one quantity predicted above the 0.60 level. It is interesting to note that the descriptors that are predicted well by the regression model do not intersect with the descriptors that are predicted well by the nearest-neighbor model. The median correlation coefficient for the nearest neighbor model across all descriptors/quantities is 0.25.

Several variations in the preprocessing of the electronic nose measurements were explored, including normalization of each chemical signature to remove concentration information, and auto-scaling the detector values to remove means and equalize variances. Concentration normalization resulted in a slight increase in the performance of the regression model (median = 0.32) and a modest decrease in the performance of the nearest neighbor model (median = 0.07; only one descriptor above 0.60). Auto-scaling does not affect the regression model, but resulted in degraded performance for the nearest neighbor model (median = 0.07).

3.2. More sophisticated regression and feature selection approaches

The poor predictive abilities of the standard linear regression and nearest neighbor models should not be surprising given that the number of examples (odorants) is comparable to the number of dimensions (detectors). In fact, under the leave-one-out cross validation procedure, 20 examples and 20 dimensions were present. Assuming that the 20×20 measurement matrix is non-singular, the regression model, thus, provides a unique weight vector that *exactly* maps the training measurements to the training targets. However, since the measurements (and targets) include noise, this will clearly lead to overfitting and poor generalization on new examples. There are several approaches available for such situations: (1) ridge regression to improve the conditioning of the $\boldsymbol{M}'\boldsymbol{M}$ matrix; (2) principal components regression to orthogonalize the measurements and discard noise; (3) partial least squares regression (PLS), also used to ignore redundant detector measurements and to discard

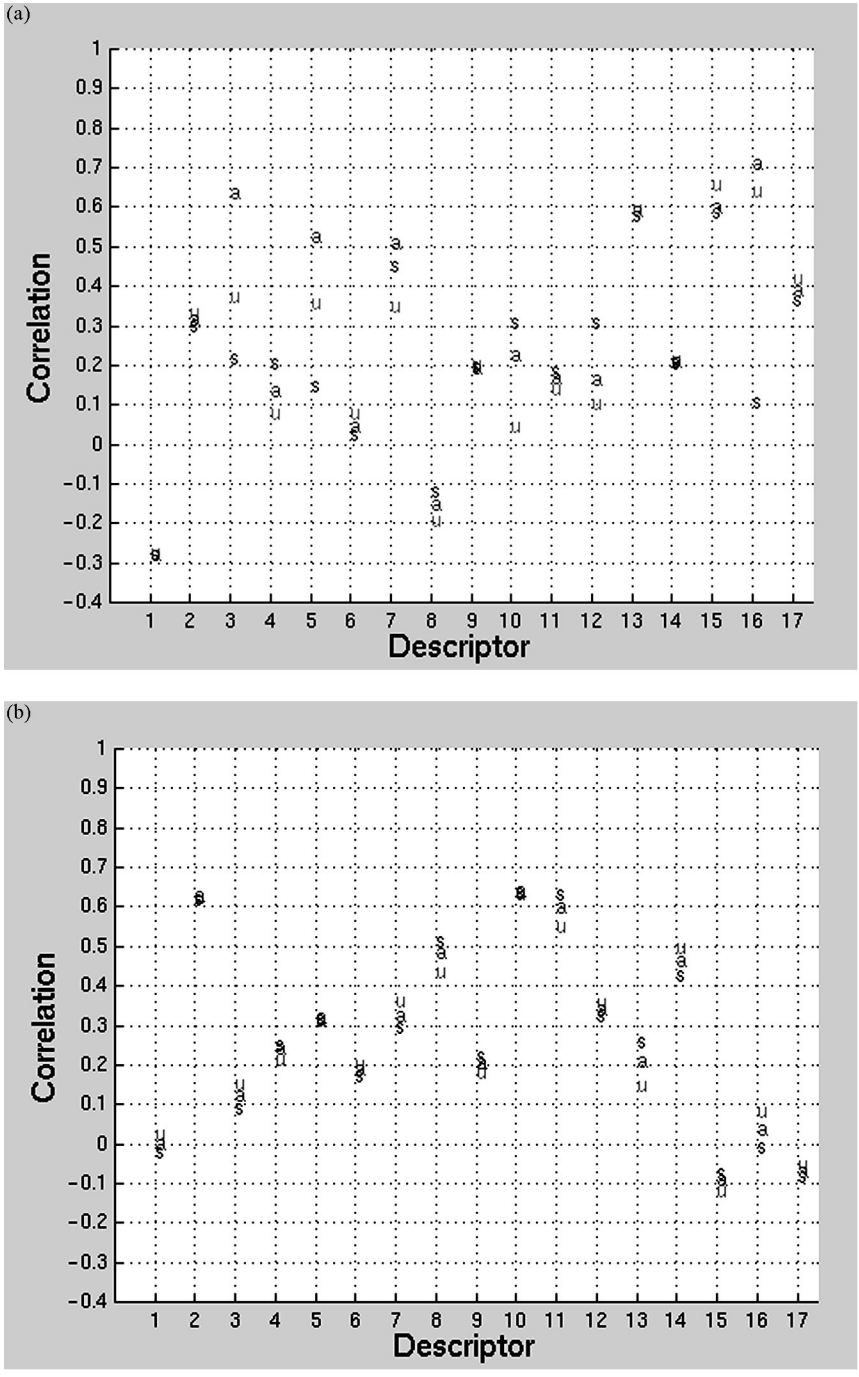


Fig. 3. (a) Correlation between clipped linear regression predictions and human target values. The symbols "*u*", "*s*", and "*a*" represent usage, score, and applicability, respectively. Only descriptors 3, 15, and 16 (floral, sour, paint) are predicted above the 0.6 level. (b) Correlation between nearest neighbor predictions and human target values. Only descriptors 2, 10, and 11 (fruity non-citrus, woody, sweet) are predicted above the 0.6 level.

noise; and (4) feature subset selection algorithms to reduce the number of features used by the predictor.

3.2.1. Ridge regression

One view of standard regression is that the matrix $\boldsymbol{M}'\boldsymbol{M}$ is a covariance matrix (if the mean value of each detector is first removed from $\boldsymbol{M}$). The standard regression solution, therefore, attempts to estimate the covariance between detectors based on data from a very limited number of examples. Such estimates can often be improved by shrinking toward the identity matrix [20]. Mathematically, $\boldsymbol{M}'\boldsymbol{M}$ is replaced by $(1-\gamma)\boldsymbol{M}'\boldsymbol{M} + \gamma\,\mathrm{tr}[\boldsymbol{M}'\boldsymbol{M}]/\mathrm{ns}\boldsymbol{I}$, where ns is the number of detectors and γ a free parameter that controls the amount of regularization.

To select the proper value of γ, the following nested cross-validation procedure was used:

```
for x = 1: nx (% holdout example x for final testing)
  % Perform selection of γ*
  for γ = 0: dγ:_1
    for y = 1: nx−1
        Perform cross-validated evaluation with
        parameter γ using all examples excluding x.
    end
  end
  Choose γ* = _best γ_value based on inner loop
  results.
  Train regularized regression model using γ* and all
  examples excluding x.
  Use regularized regression model to predict human
  value for example x.
end
Compare predicted values with targets.
```

Experimentally, a value of $d\gamma = 0.01$ was used. Due to the small number of examples available in the inner loop, the γ loop was started at $\gamma = d\gamma$ rather than at 0. The inner loop cross-validation generally selected γ^* values of 0.01 or 0.02, with one value as high as 0.04; however, the outer loop performance using the γ^* regularized regression model turned out to be slightly worse than using the standard regression model with $\gamma = 0$.

3.2.2. Principal components regression

When the measurements from different detectors are highly correlated or are noisy, the presence of the inverse in Eq. (1) often precludes obtaining a good weight vector through standard linear regression. One approach to resolve this problem is to perform a principal components analysis on $\boldsymbol{M}'$ to determine the directions that have the most variance. The data are then projected onto this reduced dimensional subspace, and directions with smaller variance are presumed to correspond to noise and are discarded. The target values are then predicted from the projected subspace rather than from the original data. In the chemometrics literature, this approach is known as principal components regression (PCR), and the projected data are commonly referred to as the ''score matrix''. In many cases, PCR provides an alternative solution to the regression problem that may be better-behaved than standard regression.

For the data of concern in this work, however, the PCR approach performed poorly. In fact, the median cross-validated correlation coefficient across all descriptors/quantities was worse than for the raw 20-dimensional data for all values of k, except $k = 18$, for which the result obtained using PCR was better by a small amount than that obtained using standard regression.

3.2.3. Partial least squares regression

Partial least squares regression (PLS) is another method that provides an alternative solution to the regression problem. The PLS method is similar to PCR, except that both the target vector and the measurements are used to determine a lower-dimensional subspace from which the predictions will be made. Determination of the subspace is accomplished through an iterative procedure as described in the literature [21].

The same leave-one-out cross-validation testing scheme described above was used to evaluate the effectiveness of PLS regression in predicting the values of the human-provided descriptors from the conducting polymer composite vapor detector measurements. On average, the PLS predictor performed slightly better than standard regression, producing median correlation values for usage, score, and applicability of 0.30, 0.29, and 0.30, respectively. However, as shown in Fig. 4, only one descriptor (#10, ''woody'') was predicted above the 0.60 level.

3.2.4. Feature subset selection

A different approach to possibly improve predictions based on models derived from limited data is to consider subsets of the raw detectors. For 20 detectors, there are $n\text{fs} = (2\hat{}20) - 1$ (slightly over 1 million) possible subsets. Various feature subset selection algorithms that use heuristics to guide the search for good subsets have been developed in the machine learning and pattern recognition communities [22,23]. In our study, however, we have avoided the use of such techniques and instead have used large amounts of computation to exhaustively evaluate *every* subset of features. By doing so, we avoid the uncertainty inherent in not knowing whether the search heuristics lead to significant degradations in achievable performance. In other words, exhaustive subset evaluation is an academic approach that will enable us to study the prediction problem without adding confusion from approximations. However, for significantly larger array sizes and/or with limited computational resources, the heuristic approaches may be the only feasible way to proceed.

To evaluate a particular subset of features, the following nested cross-validation procedure was considered:

```
for x = 1: nx (% holdout example x for final testing)
  % Perform feature selection
  for fs = 1: nfs (%)
    for y = 1: nx−1 (%)
      Perform cross-validated evaluation of feature
      set f using all examples excluding x.
    end
  end
  Choose feature set fs* based on some selection
  criteria.
  Train regression model using feature set fs* and all
  examples excluding x.
  Use regression model to predict human value for
  example x based on feature set fs*.
end
Compare predicted values with targets.
```

Clearly, the computational requirements of this procedure are quite demanding. To make the execution feasible, the

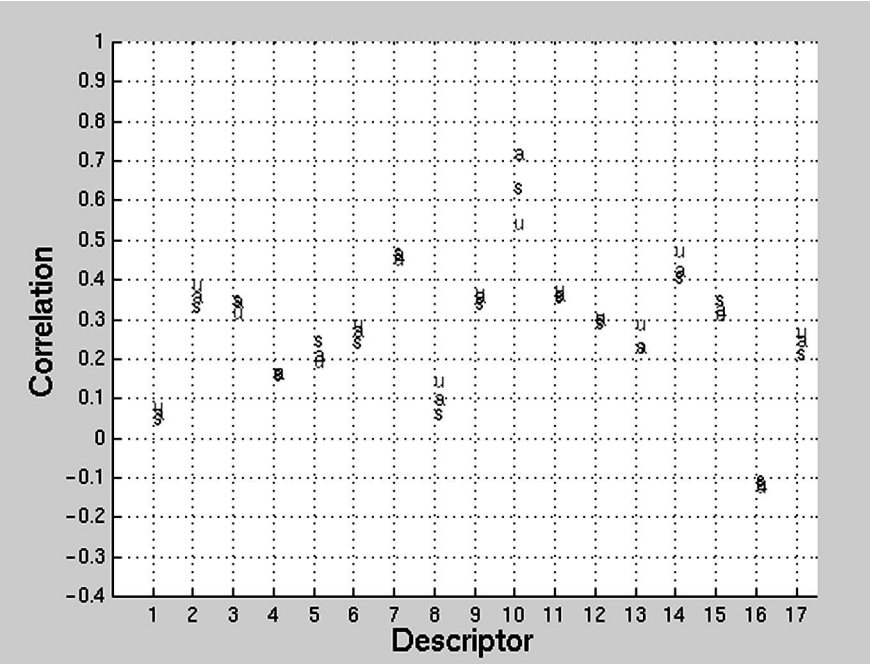


Fig. 4. Correlation between partial least squares predictions and human target values. The symbols "*u*", "*s*", and "*a*" represent usage, score, and applicability, respectively. Only descriptor 10 (woody) is predicted above the 0.6 level.

outer *x*-loop was parallelized across a dozen or so SUN Ultra 60 (dual CPU) and SUN Ultra 2 (single CPU) workstations, with each processor handling one pass through the body of the loop. Results were saved to disk at the end of the fs-loop and later merged.

The inner-most cross-validation loop (the *y*-loop) requires the computation of $(\boldsymbol{M}'_{xy}\boldsymbol{M}_{xy})^{-1}\boldsymbol{M}'_{xy}\boldsymbol{H}_{xy}$, where $\boldsymbol{M}_{xy}$ consists of all the measurements excluding examples *x* and *y* and $\boldsymbol{H}_{xy}$ consists of the corresponding human values. $\boldsymbol{H}_{xy}$ is a matrix with each column containing the human data for a particular descriptor/quantity; thus $\boldsymbol{H}_{xy}$ has $51 = 17 \times 3$ columns. Each pass through the *y*-loop appears to require an inversion of the matrix $\boldsymbol{M}'_{xy}\boldsymbol{M}_{xy}$, requiring work proportional to the third power of the number of features involved. However, the Sherman–Morison–Woodbury formula, which specifies how a rank one correction to a matrix affects the inverse, can be applied so that only one inversion is required, rather than $nx-1$ inversions. The main idea is to compute the inverse of $\boldsymbol{M}'_x\boldsymbol{M}_x$ (measurements excluding example *x*) once before the *y*-loop and then note that $\boldsymbol{M}'_{xy}\boldsymbol{M}_{xy}$ is related to $\boldsymbol{M}'_x\boldsymbol{M}_x$ by a rank one correction involving example y. A correction can also applied to the quantity $\boldsymbol{M}'_x\boldsymbol{H}_x$ to obtain $\boldsymbol{M}'_{xy}\boldsymbol{H}_{xy}$ with less work than recomputing on each pass through the loop. This procedure greatly reduced the computational complexity.

A number of possible selection criteria could be applied at the conclusion of the loop over feature sets. The most straightforward criterion is to choose for each descriptor/quantity the feature set fs^* that results in the best inner-loop (*y*-loop) cross-validated correlation. Other inner-loop assessment metrics such as mean square error, maximum absolute error, and maximum relative error can also be considered. Intuition suggests, however, that models based on a smaller number of features may generalize better to new examples, so it seems reasonable to bias the feature selection strategy towards feature sets having smaller numbers of features. For example, rather than choosing the feature set providing the best inner loop performance, the best feature set having exactly (or at most) a predetermined number of features was chosen. Similar biasing rules such as selecting the smallest-sized feature set whose performance on the inner loop exceeds a given threshold, or whose performance on the inner loop is "close enough" to the best performance achieved by any feature set, may also be considered.

Table 4 shows the results of the feature selection experiment. For each iteration of the outer loop, the inner loop cross-validated correlation scores were used to select the best set of *k* detectors. Thus, *nx* best detector sets could potentially be identified for a given *k* value. The median inner-loop correlation values over these best detector sets are shown in column 2 of the table, where the median is taken over all descriptors/quantities as well as over the best detector sets. The score of the best inner loop models when applied to the outer loop holdout examples is shown in column 3 (median across all descriptors/quantities).

The best cross-validated correlation score that could have been achieved by a single feature set of size *k* is shown in column 4. Note that these scores were obtained by evaluating all feature sets of size *k* on the outer loop task; then the set that produced the best results was chosen and its results were reported in column 4 (i.e. feature subset selection was based on the test set).

Table 4
Results of feature selection experiments

Number of detectors (k)	Median inner loop score	Outer loop score	Idealized outer loop score
1	0.37	0.19	0.40
2	0.48	0.10	0.50
3	0.55	0.10	0.57
4	0.61	0.10	0.63
5	0.65	0.13	0.65
6	0.70	0.12	0.74
7	0.74	0.11	0.77
8	0.76	0.13	0.77
9	0.77	0.15	0.78
10	0.82	0.21	0.82
11	0.86	0.25	0.87
12	0.88	0.30	0.90
13	0.90	0.26	0.90
14	0.91	0.25	0.91
15	0.92	0.26	0.88
16	0.93	0.17	0.86
17	0.96	0.21	0.82
18	0.99	0.08	0.84
19	0.39	0.09	0.63
20	0.02	0.18	0.25

The fact that the column 4 numbers are large is interesting, because it indicates that certain linear regression models using subsets of features can accurately predict the human data in cross-validation tests. For example, Table 5 shows the smallest models that yielded cross-validated correlation above 0.6 for the *usage* quantity of each descriptor. However, in Table 4 the large margin between column 2 and column 3 casts doubt on whether the models that produced column 4 will generalize well to new data. More specifically, column 2 represents the expected performance based on the models that worked best in the inner-loop cross-validation (train on $nx-2$ examples, test on 1). Column 3 indicates the performance of these ''good'' inner-loop models when applied to the outer-loop holdout example. Clearly, the performance of the models was highly degraded under these conditions. Even though a good inner-loop model could be obtained, the existence of such a model did not insure that the same model would work well in the outer-loop cross-validation. Analogously, some models worked well on the outer-loop cross-validation, but this does not insure that the same models will work well for new data.

Table 5
Smallest models yielding a cross-validated correlation above 0.6 for the descriptor usage quantity

	Descriptor	k	Model[a]
1	Fruity citrus	15	5, 7, 11, 16, 19
2	Fruity non-citrus	5	5, 12, 13, 15, 18
3	Floral	8	2, 3, 5, 10, 12, 14, 17, 19
4	Minty	9	2, 4, 7, 9, 10, 11, 14, 18, 20
5	Etherish	4	1, 5, 6, 13
6	Gasoline	3	1, 5, 13
7	Sharp, pungent	2	4, 14
8	Oily	2	16, 20
9	Putrid, foul, decayed	5	1, 6, 15, 18, 20
10	Woody	4	11, 13, 15, 19
11	Sweet	4	7, 12, 15, 17
12	Herbal	2	6, 14
13	Musty	2	6, 20
14	Medicinal	7	3, 4, 5, 7, 8, 11, 12
15	Sour	6	1, 2, 5, 15, 16, 18
16	Paint	4	3, 6, 13, 14
17	Sweaty	2	2, 20

[a] Underlined values indicate excluded detectors.

4. Discussion

This study considered single-component odorants consisting of simple organic vapors without significant aroma activity (arguably the simplest case). For specific descriptors, some models provided cross-validated predictions that correlated well with the human data (above the 0.60 level), but none of the models could accurately predict the human values for more than a few descriptors.

The models based on feature subset selection were especially intriguing. Relatively small subsets of detectors (e.g. as listed in Table 5) in some cases provided good cross-validated predictions of most of the human descriptors. However, because these subsets could not be identified through a rigorous model selection procedure, the results may not generalize to new data. Further evidence for this conclusion is given by the model selection procedure itself. The large discrepancy between the inner loop cross-validated correlation values and the resulting outer loop cross-validated correlation values (i.e. comparison of column 2 and column 3 in Table 4) indicates that at least some ''good'' inner-loop models perform poorly on new data.

The term ''overfitting'' is typically applied to describe the situation in which models or parameters are adjusted to fit a training set to the best degree possible. In our case, the feature subset selection procedure uses the criteria ''find the best subset of features smaller than size k'', which provides a bias toward smaller models, i.e. the decision as to which feature subset to choose is not based solely on optimizing the predictions to the targets. Hence, we are not strictly overfitting during the subset selection. The poor generalization results, however, may indicate that the bias is inadequate to lead the model selection procedure to good subsets that will generalize.

Once a feature subset is selected, the standard regression solution attempts to find the best set of weights that fits the training set, i.e. it is overfitting. It does not appear that this weight overfitting is catastrophic, however, because it also is present for column 4 of Table 4.

There are several plausible explanations for why human perception of odor quality could not be predicted reliably from the conducting polymer composite electronic nose

signals. First, the number and diversity of receptors in the artificial nose was significantly (1–2 orders of magnitude) more limited than in humans. Second, linear and affine models may be too simplistic to enable human percepts of even single component organic vapor odorants to be predicted well from conducting polymer composite electronic nose signals. This may reflect the fact that these data models have no physically-based relationship to the neuronal connections and signal processing involved in olfactory perception. A third possibility is that the form of the data models may be adequate for the limited task considered, but the parameters of these models could not be estimated reliably enough from the amount of data available. A fourth possibility is that the feature subset selection procedure did not have enough data to identify feature subsets that would provide good generalization ability.

In any case, it is clear that the situation would be significantly worse for odorants that are mixtures of pure compounds, because human odor perception of mixtures is often not linearly related to the mole fraction of the individual components of the mixture. Similarly, for mixtures of odorants that contain aroma-active compounds which are detectable at very low concentration levels in the human nose but which would not even produce signals above the noise level of the current conducting polymer composite detectors, no correlation would be expected between the conducting polymer composite vapor detector response and that of human odor perception.

Given the complex, nonlinear characteristics of human olfaction, it is not surprising that the analogy between the electronic nose and the human olfactory system only extends to the design principle that both systems utilize arrays of broadly responding, cross-reactive detectors. Hence, the primary contribution of the present study is to advance the idea of formulating, developing, and testing models of olfactory perceptual processing using artificial data sets such as those generated by the electronic nose in place of spike train data measured only with great difficulty on biological olfactory receptors. In this respect, one advantage of the conducting polymer composite-based electronic nose used in this study is that characteristic response patterns that are essentially independent of the concentration of the analyte presented to the detectors can be obtained. Thus, one degree of freedom (choice of analyte concentration) can be eliminated, unlike the situation for metal-oxide detectors, dye-impregnated polymers on optical fibers, and other detectors that exhibit response patterns that are a function of concentration of the analyte of interest [3]. A secondary contribution is that the results based on simple numerical models provide a benchmark against which more sophisticated models can be judged, and strongly suggest that significant system architecture changes and highly increased data analysis algorithm complexity are needed before any artificial model could provide a robust predictive ability for human odor quality descriptors on even a simple set of test odorants. The correlation between the predictions for certain odor descriptors and the human perceptual data are definitely interesting, but more extensive experiments would be required to assess the validity of these predictors for other odorants.

Acknowledgements

We acknowledge a MURI supported through the Army Research Office for support of this work, and thank W. Cain of UCSD and D. DeCoste for numerous helpful discussions.

References

[1] H.V. Shurmer, An electronic nose — a sensitive and discriminating substitute for a mammalian olfactory system, IEEE Proc.-G Circ. Dev. Syst. 137 (1990) 197–204.
[2] J.W. Gardner, P.N. Bartlett, A brief-history of electronic noses, Sens. Actuators B 18 (1994) 211–220.
[3] K.J. Albert, N.S. Lewis, C.L. Schauer, G.A. Sotzing, S.E. Stitzel, T.P. Vaid, D.R. Walt, Cross-reactive chemical sensor arrays, Chem. Rev. 100 (2000) 2595–2626.
[4] B. Malnic, J. Hirono, T. Sato, L.B. Buck, Combinatorial receptor codes for odors, Cell 96 (1999) 713–723.
[5] T.C. Pearce, Computational parallels between the biological olfactory pathway and its analogue 'The Electronic Nose'. 2. Sensor-based machine olfaction, Biosystems 41 (1997) 69–90.
[6] F.T. Schiet, W.S. Cain, Odor intensity of mixed and unmixed stimuli under environmentally realistic conditions, Perception 19 (1990) 123–132.
[7] D.G. Laing, H. Panhuber, M.E. Willcox, E.A. Pittman, Quality and intensity of binary odor mixtures, Physiol. Behav. 33 (1984) 309–319.
[8] J.D. Pierce Jr., X.N. Zeng, E.V. Aronov, G. Preti, C.J. Wysocki, Cross-adaptation of sweaty-smelling 3-methyl-2-hexenoic acid by a structurally-similar, pleasant-smelling odorant, Chem. Senses 20 (1995) 401–411.
[9] S. Ayabe-Kanamura, I. Schicker, M. Laska, R. Hudson, H. Distel, T. Kobayakawa, S. Saito, Differences in perception of everyday odors: a Japanese–German cross-cultural study, Chem. Senses 23 (1998) 31–38.
[10] B.J. Doleman, E.J. Severin, N.S. Lewis, Trends in odor intensity for humans and electronic noses: relative roles of odorant vapor pressure vs. molecularly specific odorant binding, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 5442–5447.
[11] M.C. Lonergan, E.J. Severin, B.J. Doleman, S.A. Beaber, R.H. Grubbs, N.S. Lewis, Array-based vapor sensing using chemically sensitive, carbon black-polymer resistors, Chem. Mater. 8 (1996) 2298–2312.
[12] B.J. Doleman, M.C. Lonergan, E.J. Severin, T.P. Vaid, N.S. Lewis, Quantitative study of the resolving power of arrays of carbon black-polymer composites in various vapor-sensing tasks, Anal. Chem. 70 (1998) 4177–4190.
[13] B.J. Doleman, N.S. Lewis, Comparison of odor detection thresholds and odor discriminablities of a conducting polymer composite electronic nose vs. mammalian olfaction, Sens. Actuators B 72 (2001) 41–50.
[14] A. Dravnieks, Atlas of Odor Character Profiles, American Society for Testing and Materials, Baltimore, 1985.
[15] C. Di Natale, A. Macagnano, R. Paolesse, R. Mantini, E. Tarizzo, F. Sinesio, et al., Electronic nose and sensorial analysis: comparison of performance in selected cases, Sens. Actuators B: Chem. 50 (1998) 246–252.

[16] B.-I.L. Willing, A. Brunders, I. Lundstrom, Odour analysis of paperboard, the correlation between human senses and electronic sensors using multivariate analysis, Packaging Technol. Sci. 11 (1998) 59–67.
[17] Y. Blixt, E. Borch, Using an electronic nose for determining the spoilage of vacuum-packaged beef, Int. J. Food Microbiol. 46 (1999) 123–134.
[18] E.J. Severin, B.J. Doleman, N.S. Lewis, An investigation of the concentration dependence and response to analyte mixtures of carbon black-insulating organic polymer composite vapor detectors, Anal. Chem. 72 (2000) 658–668.
[19] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C, Cambridge University Press, Cambridge, 1988.
[20] J.H. Friedman, Regularized discriminant analysis, J. Am. Stats. Assoc. 84 (405) (1989) 165–175.
[21] P. Geladi, B.R. Kowalski, Partial least squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[22] F. Ferri, P. Pudil, M. Hatef, J. Kittler, Comparative study of techniques for large-scale feature selection. IV. Pattern Recognition in Practice, in: E.S. Gelsema, L.N. Kanal (Eds.), Elsevier, Amsterdam, 1994, pp. 403–413.
[23] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, Oxford, 1995.

Classification of Landmine-Like Metal Targets Using Wideband Electromagnetic Induction

Ping Gao, *Student Member, IEEE,* Leslie Collins, *Member, IEEE,* Philip M. Garber, *Student Member, IEEE,* Norbert Geng, *Member, IEEE,* and Lawrence Carin, *Senior Member, IEEE*

***Abstract*—In our previous work, we have shown that the detectability of landmines can be improved dramatically by the careful application of signal detection theory to time-domain electromagnetic induction (EMI) data using a purely statistical approach. In this paper, classification of various metallic landmine-like targets via signal detection theory is investigated using a prototype wideband frequency-domain EMI sensor. An algorithm that incorporates both a theoretical model of the response of such a sensor and the uncertainties regarding the target/sensor orientation is developed. This allows the algorithms to be trained without an extensive data collection. The performance of this approach is evaluated using both simulated and experimental data. The results show that this approach affords substantial classification performance gains over a standard approach, which utilizes the signature obtained when the sensor is centered over the target and located at the mean expected target/sensor distance, and thus ignores the uncertainties inherent in the problem. On the average, a 60% improvement is obtained.**

***Index Terms*—Bayes procedures, electromagnetic induction, object detection, pattern classification.**

I. Introduction

A PERSISTENT problem with traditional narrowband EMI sensors involves not just detection of metal objects, but discrimination of targets from clutter. In most fielded sensors, the energy in the output of such sensors is calculated, and a decision regarding the presence or absence of a target is made using this statistic [1]. This approach leads to excessively large false alarm rates. When each piece of buried metal must be excavated in order to determine whether it is a target of interest, significant costs are incurred both due to lost time and costs associated with digging. The false alarm issue is particularly problematic in real world landmine-detection scenarios. In order to facilitate the discrimination of targets of interest from other pieces of metal, several modifications to traditional EMI sensors have been considered [1]–[8]. For instance, the late time EMI fields are characterized by an exponential decay in the time-domain [2], [3], [7], [8]. The decay rate has been used for target identification, because it strongly depends on the target conductivity, permeability, shape, and orientation. Alternatively, a promising approach is to operate the sensor in the frequency-domain by utilizing wideband excitation. The frequency dependence of the induced fields excited by buried conducting targets can then be exploited by a detector.

A second problem that besets statistical algorithms is the need for adequate training data. A wide range of targets and clutter signatures must be obtained at all possible object/sensor orientations. Classically, this problem has been addressed by using data measured in the field [1], [9], [10]. However, a lack of sufficient training data can severely degrade performance [1], [11]. In general, it is difficult if not impossible to obtain such data for all possible object/sensor orientations. An alternative is to train the algorithms with data produced by computational models, which have only recently become general enough to consider such problems [8].

In this paper, we consider the problem in which we assume an object has been detected, and a decision as to "target" or "clutter" is required. In this approach, we use the complex frequency-dependent EMI response as a signature. A full-wave model is developed for the wideband EMI response of targets, specialized to the case of a body of revolution, thus obviating the need for training on field-collected data. Subsequently, a Bayesian classification algorithm is developed, which incorporates the wave model and the target/sensor position uncertainty. Substantial improvements are achieved via this approach over a processor, which ignores the orientation uncertainties.

This paper is organized as follows. In Section II, we describe a new prototype wideband frequency-domain EMI sensor, the GEM-3 [9]. In Section III, we discuss a model that calculates the wideband EMI responses. A model-based Bayesian approach for discriminating targets is discussed in Section IV. The process used to generate the simulated data and the experiment performed to collect the measured data are described in Section V. Next, the results from both simulated and measured data are shown. Finally, we summarize our major findings based on these results.

II. Sensor Overview

When operating an EMI sensor in the frequency-domain, it has been shown that the frequency-dependent induced fields can differ significantly depending on the target shape and conductivity [12]. This variability may be exploited to enhance discrimination performance. Therefore, data from a prototype wideband EMI sensor, the GEM-3, developed by Geophex Ltd., was selected for this analysis. The validity of a numerical model that predicts the wideband EMI responses (discussed in the next section) can be tested using data collected with the GEM-3. Furthermore, a decision-theoretic discrimination algorithm can be applied to both simulated data generated based on the model

Manuscript received November 23, 1998; revised June 15, 1999. This work was supported by the Army Research Office, Research Triangle Park, NC, under Grants DAAH04-96-1-0448 (Demining MURI) and DAAG55-98-1-0340.

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA (e-mail: lcollins@ee.duke.edu).

Publisher Item Identifier S 0196-2892(00)02849-7.

0196–2892/00$10.00 © 2000 IEEE

predictions and real data measured with the sensor. In this section, the sensor operating principles are briefly described.

The design of the transmitting coils of the GEM-3, two concentric circular coils, results in a magnetic cavity in the center zone of the two coils, i.e., an area where the primary magnetic flux vanishes. A small receiving coil is located in this magnetic cavity [13]. Therefore, the receiver can sense a weak, secondary field returned from the earth and any buried objects. By using two transmitting coils connected in series, the coils carry the same amount of current. However, current flows in opposite directions. The magnetic cavity is created by carefully choosing the radius of the two coils and the number of turns of the coils. The sensor records the real and imaginary parts (in-phase and quadrature) of the induced complex voltage at the receiving coil, relative to that on the transmitting coils.

Instead of using a pulse excitation, such as that used by time-domain EMI systems, the transmitting coils of the wideband frequency-domain EMI sensor transmit a continuous, complex waveform consisting of multiple frequencies predefined by the operator [9], [13]. Thus, the sensor is only subject to the noise at the frequencies of interest, not within the whole frequency band, as is the case for time-domain EMI sensors. Therefore, frequency-domain EMI sensors can operate at much higher SNR's than time-domain systems. In addition to the improved SNR, theoretical calculations and experimental data have shown that the frequency-domain EMI signatures differ significantly across targets [12], which provides the underlying physical mechanisms important for discriminating, identifying, or classifying targets.

III. Model for Wideband Frequency-Domain EMI Responses

In this paper, a model-based Bayesian decision-theoretic approach is investigated to discriminate four manmade metal targets of different shapes, sizes, and metal types under conditions where the target/sensor relative position is unknown. In order to model the signature of these targets, a method of moment (MoM) analysis is used to predict the theoretical response from the target [12], [14]. The calculation provides the theoretical induced voltage (magnitude and phase, or in-phase and quadrature components) for each target and frequency considered. Later in this paper, it is shown that by incorporating the model into the detector formulation, the classification performance is improved dramatically when the relative target/sensor position is uncertain, as it is in field operations.

We consider the fields induced by a highly (but not perfectly) conducting and/or permeable target in free space, due to EMI excitation at kHz frequencies. The problem is solved via a frequency-domain boundary-integral equation formulation. Moreover, to make such an analysis tractable, we specialize to a particular class of targets: those that can be modeled as a body of revolution (BOR) [15] (i.e., targets possessing rotational symmetry). Although here we consider near-field effects for metallic and ferrous targets, the general formulation is very similar to those used previously for far-zone scattering from low-loss dielectric targets [16], [17]. In particular, the problem is formulated in terms of the tangential electric $\mathbf{E}$ and magnetic $\mathbf{H}$ fields on the target surface or equivalently, in terms of electric and magnetic surface currents $\mathbf{J} = \mathbf{n} \times \mathbf{H}$ and $\mathbf{K} = \mathbf{E} \times \mathbf{n}$, respectively, where $\mathbf{n}$ is the outward unit normal. If $\mathbf{E}_1$ and $\mathbf{H}_1$ represent, respectively, the electric and magnetic fields inside the target, and $\mathbf{E}_2$ and $\mathbf{H}_2$ represent the "scattered" fields outside the target, boundary conditions at the interface yield the relationships (enforced at the boundary)

$$\begin{aligned}
&\mathbf{n} \times \left[\mathfrak{L}_1^{EJ}(-\mathbf{J}) + \mathfrak{L}_1^{EK}(-\mathbf{K})\right] \\
&\quad = \mathbf{n} \times \left[\mathfrak{L}_2^{HJ}(-\mathbf{J}) + \mathfrak{L}_2^{EK}(-\mathbf{K})\right] + \mathbf{n} \times \mathbf{E}^i \\
&\mathbf{n} \times \left[\mathfrak{L}_1^{HJ}(-\mathbf{J}) + \mathfrak{L}_1^{HK}(-\mathbf{K})\right] \\
&\quad = \mathbf{n} \times \left[\mathfrak{L}_2^{HJ}(-\mathbf{J}) + \mathfrak{L}_2^{HK}(-\mathbf{K})\right] + \mathbf{n} \times \mathbf{H}^i
\end{aligned} \tag{1}$$

where $\mathbf{E}^i$ and $\mathbf{H}^i$ represent the incident fields. The operators $\mathfrak{L}_n$ involve well-known manipulations of the homogeneous-media Green's function [17]–[19] for medium parameters inside ($n = 1$) and outside ($n = 2$) the body. The problem therefore reduces to solving for $\mathbf{J}$ and $\mathbf{K}$ for particular incident fields $\mathbf{E}^i$ and $\mathbf{H}^i$. In the MoM solution for BOR [15]–[20], $\mathbf{J}$, $\mathbf{K}$, $\mathfrak{L}_n$, $\mathbf{E}^i$, and $\mathbf{H}^i$ are expanded in a Fourier series in the azimuthal variable ϕ, and for each Fourier component, $\mathbf{J}$ and $\mathbf{K}$ are expanded in terms of one-dimensional (1-D) basis functions along the BOR generating arc (see Fig. 1). In this paper, we use subsectional basis functions and testing functions, as in [18]–[20]. The interested reader is referred to [18]–[20] for details concerning implementation of the general algorithm, while here we focus on issues of particular relevance to the EMI problem.

We first consider requirements concerning the subsectional-basis-function discretization of $\mathbf{J}$ and $\mathbf{K}$. For scattering from low-loss targets, it is well known that approximately ten basis functions are required per wavelength [17]. In such problems, this rule is applied to the smallest wavelength of interest in the problem, generally corresponding to the medium inside the target. For the highly conducting targets of interest here, the wavenumber inside the target approximately satisfies $k_1 = (1 - j)/\delta$, where δ is the skin depth. To sample the Green's function phase $\exp(-jk_1R)$ sufficiently, we require $\Delta t/\delta \ll 2\pi$ and $\Delta t/\delta \ll 1$ (for the real and imaginary parts of k_1, respectively), where Δt is the basis-function width. These constraints are usually sufficient to satisfy the outer region sampling requirements (k_2 generally representing the *free-space* wavenumber). Extensive numerical experiments indicate that accurate results are obtained if $\Delta t < \delta/3$.

As discussed above, for EMI applications, we are generally interested in current loop excitation, as distinguished from the plane-wave fields considered for radar problems. While the fields due to a current loop are well known [21], [22], we discuss how such are placed into the BOR framework, as well as appropriate approximations for the EMI problem. In particular, the incident fields are derived from the vector potential [21], [22]

$$A_\phi(\rho, z) \approx \frac{\mu_0 Ia}{4\pi} \int_0^{2\pi} \frac{d\phi' \cos\phi'}{\sqrt{\rho^2 + a^2 + z^2 - 2a\rho\cos\phi'}} \tag{2}$$

where the origin of the local cylindrical coordinate system (ρ, ϕ, z) is situated at the loop center, with axis parallel to z, and I and a are the loop current and radius, respectively. The

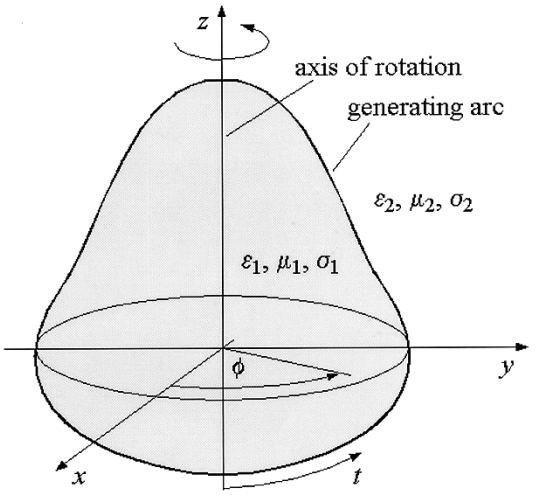


Fig. 1. Equivalent electric and magnetic surface currents for modeling electromagnetic interaction with a highly conducting and/or permeable body of revolution (BOR).

expression in (2) invokes a quasi-static approximation, since, at the wavelengths of interest (in air and soil), the electrical distance between the sensor and target is infinitesimal. Similar approximations can be used (but have not been here) with regard to the Green's function components in the air region ($\mathfrak{L}_2$ in (1)), while the very high conductivity and/or permeability inside the target necessitates a rigorous analysis (i.e., a rigorous formulation of $\mathfrak{L}_1$). The incident electric and magnetic fields are readily computed as

$$H_\rho^i(\rho,z) = -\frac{1}{\mu}\frac{\partial A_\phi}{\partial z},$$
$$H_z^i(\rho,z) = -\frac{1}{\mu\rho}\frac{\partial(\rho A_\phi)}{\partial \rho},$$
$$E_\rho^i(\rho,z) \approx -j\omega A_\phi \tag{3}$$

and these fields are finally expressed in terms of complete elliptical integrals. If the loop axis and the BOR axis are aligned, the fields in (3) can be applied directly to the BOR MoM solution, and only the lowest-order Fourier series mode is excited (reflecting azimuthal symmetry). If the axes are not aligned, a Fourier series representation of the incident tangential fields on the BOR surface is generally required (with the BOR problem solved separately using the incident fields from each such mode [15]–[18]). While the Fourier components for plane-wave incidence can be expressed in closed form [15], we have not found such a simple representation for the loop-induced fields. Therefore, in the general case, we must numerically determine the Fourier coefficients. For example

$$\hat{H}_{mt}(t) = \frac{1}{2\pi}\int_0^{2\pi} H_t^i(t,\phi)\exp(-jm\phi)\,d\phi \tag{4}$$

where H_t^i represents the incident magnetic field along the generating arc (Fig. 1), and (t,ϕ) represents the local BOR coordinate system. Thus, while the space domain fields for the loop can be expressed in closed form, the requisite Fourier components are evaluated numerically. However, the incident fields are generally slowly varying in ϕ and integrals of the type in (4) do not present a significant numerical challenge.

Before proceeding to a comparison of theoretical and measured results, we note that the EMI fields induced by a conducting and/or ferrous target are generally measured in the near zone. Therefore, when calculating the induced fields, we cannot invoke the simplifying far zone approximation generally used for radar-scattering problems [15], [17], [23]. We therefore calculate the EMI "scattered" fields via a rigorous convolution of the calculated currents $\mathbf{J}$ and $\mathbf{K}$ with the free-space Green's function (e.g., with $\mathfrak{L}_2$), performing integrals similar to those used in calculating the components of the MoM impedance matrix. Additionally, we note that an actual EMI sensor does not measure the induced fields, but rather the electromotive force induced on a sensing current loop. To calculate such, we have used appropriate magnetic field components, integrated over the aperture of the sensing loop, to generate a theoretical induced voltage for each target and frequency considered.

Using data collected from a prototype wideband frequency-domain EMI sensor, the GEM-3, the effectiveness of the numeric model is tested. A comparison of the theoretical model and measurements is shown in Section V-B.

The simulation outputs from the model can be used to calibrate the frequency-domain EMI sensor. Let $c(\omega)$ represent the calibration constant for frequency ω, the $K \times 1$ vector $\mathbf{M}$ represent a set of measurements obtained at several (K) positions, and the $K \times 1$ vector $\mathbf{B}$ represent model outputs for the same target and positions. We have the relation that $\mathbf{B}c(\omega) = \mathbf{M}$, and a least-squares method is used to obtain the calibration constants as a function of frequency.

IV. Forward Model-Based Bayesian Classifier Formulation

In a real-world classification scenario, the uncertainty inherent in the sensor output is not only due to additive noise, but also to the fact that the relative position between the sensor and the target is unknown at the point when the measurements are obtained. In this work, we investigate the classification performance of a Bayesian classifier that incorporates modeled wideband EMI signatures as well as position uncertainties and compare its performance to an approach that ignores these uncertainties and assumes the target is at a fixed position corresponding to the mean assumed position.

In this paper, we consider the task of classifying data from one of four known metal objects. It is always true that one of the objects is present, and our goal is to decide which object is present. In real world situations such as landmine detection, it is often the case that a metal object can be located. The task is then to determine whether it is a target or a clutter object. In this case, a library of targets of interest can be established and typical clutter can also be modeled. Alternatively, a statistical model could be imposed for clutter based on localized measurements and the target models can be used as is described here. Thus, this approach can also be applied to an extended set of objects in practice.

In this work, four metal objects are considered (a more detailed description can be found in Section V). Signals used

to measure classification performance are either the modeled wideband frequency-domain EMI responses or measured EMI responses from the GEM-3. Since any sensor is subject to noise, which is usually assumed to follow a Gaussian distribution, the distribution of the sensor outputs (obtained data set of discrete frequencies) while the target/sensor is at a known height and horizontal position, is a Gaussian random vector. The mean of this response is the theoretical response, and the variance is equal to that of the additive noise. Let H_i represent the hypothesis that the ith target is present, where $i = 1, 2, 3, 4$. The received data from the ith target at a known position can be modeled as

$$r_{ij} = A_{ij} + n_j \tag{5}$$

where j corresponds to the discrete frequencies of interest, $j = 1, 2, \ldots, N$, r_{ij} is the received data from the sensor, A_{ij} is the predicted response obtained from the model (as described in detail in Section III, the model can calculate the theoretical frequency-domain EMI responses for a well specified BOR object at a known position) for the ith target at the jth frequency at a known depth and horizontal position relative to the center of the sensor, and n_j is Gaussian noise with zero mean and variance of $\sigma_{n_j}^2$. We assume that n_j's are independent. However, their variance is a function of frequency. Let q_i represent the *a priori* probability that hypothesis H_i is true. We further assume that the cost of a correct decision is zero, and the cost of any wrong decision equals 1. Bayes' solution for this classification problem [24], [25] is to decide that H_i is true if

$$\frac{p(H_i|r)}{p(H_k|r)} = \frac{q_i p(r|H_i)}{q_k p(r|H_k)} > 1 \tag{6}$$

is satisfied for any $k \neq i$. Here $p(H_i|r)$ is the *a posteriori* distribution or discriminant function [26], $p(r|H_i)$ is the probability density or likelihood function of data r given H_i, and r is the received data from the sensor. Assuming the magnitude and the phase of the frequency response are independent, $\mathbf{r}$ is a vector containing both the magnitude and phase information. Therefore, when the sampled data r is received, we decide in favor of hypothesis H_i, where

$$q_i p(\mathbf{r}|H_i) = \max_k \{q_k p(\mathbf{r}|H_k)\} \quad k = 1, 2, 3, 4 \tag{7}$$

Thus, we decide in favor of a hypothesis that has the largest *a posteriori* probability or largest discriminant function at r among all four possible pdf's. Since we usually have no *a priori* knowledge of q_i (in other words, we do not know the probability that a particular target is going to be present), an equal probability assumption for each target is made (i.e., $q_i = 1/4$). Based on the uniform *priori* on q_i, (7) can be further understood as seeking a hypothesis that provides the maximum likelihood among the four possible values. Thus, it can also be referred to as a maximum likelihood (ML) classifier. Since any monotonically increasing function of $p(H_i|\mathbf{r})$ is also a valid discriminant function [26], an alternative discriminant function based on the above assumptions is

$$p(\mathbf{r}|H_i) = (2\pi)^{-N}|\Sigma|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{r}-\mathbf{A}_i)^T\Sigma^{-1}(\mathbf{r}-\mathbf{A}_i)\right] \tag{8}$$

TABLE I
PROBABILITY OF CORRECT CLASSIFICATION OF THE OPTIMAL CLASSIFIER WHEN TARGETS ARE AT A FIXED KNOWN POSITION AS THE NOISE VARIANCE IS INCREASED FROM σ_n^2 TO $2^9\sigma_n^2$

NOISE VARIANCE	PROBABILITY OF CORRECT CLASSIFICATION			
	TARGET 1	TARGET 2	TARGET 3	TARGET 4
σ_n^2	1.0000	1.0000	1.0000	1.0000
$2\sigma_n^2$	1.0000	1.0000	1.0000	1.0000
$2^2\sigma_n^2$	0.9993	1.0000	0.9990	1.0000
$2^3\sigma_n^2$	0.9862	1.0000	0.9877	1.0000
$2^4\sigma_n^2$	0.9365	1.0000	0.9394	1.0000
$2^5\sigma_n^2$	0.8604	1.0000	0.8640	1.0000
$2^6\sigma_n^2$	0.7809	0.9977	0.7747	1.0000
$2^7\sigma_n^2$	0.7134	0.9791	0.6854	1.0000
$2^8\sigma_n^2$	0.6545	0.9254	0.5756	0.9981
$2^9\sigma_n^2$	0.6101	0.8552	0.4554	0.9827

where N is the total number of frequencies used, $\mathbf{r}$ and $\mathbf{A}_i$ are $2N$ by 1 vectors, and Σ is the covariance matrix of $\mathbf{r}$. Given the assumptions on the noise process, Σ is a diagonal matrix with $\sigma_{n_j}^2$ on the jth diagonal, where j corresponds to frequency. Since the coefficient of the exponential term of (8) is the same for all the hypotheses, it can be neglected. After taking the logarithm, the alternative discriminant function simplifies to

$$\log p'(\mathbf{r}|H_i) = -(\mathbf{r}-\mathbf{A}_i)^T\Sigma^{-1}(\mathbf{r}-\mathbf{A}_i) \tag{9}$$

where $(\mathbf{r}-\mathbf{A}_i)^T\Sigma^{-1}(\mathbf{r}-\mathbf{A}_i)$ is often referred to as the Mahalanobis distance from r to $\mathbf{A}_i$ [26]. If Σ is a diagonal matrix with each diagonal element $\sigma_{n_j}^2$, (9) can be expressed as

$$\log p''(\mathbf{r}|H_i) = -\sum_{j=1}^{2N}(r_j - A_{ij})^2 \Big/ \sigma_{n_j}^2 \tag{10}$$

The discriminant function obtained above [(10)] is valid if the height and horizontal position of the object are both known, and the noise is assumed to be independent at each frequency. This solution is optimal only under the assumptions that all the parameters are known, and the sensor is subject only to Gaussian noise. This formulation differs from a bank of matched filters since the noise is not identically distributed and the variance of the noise is a function of frequency, and the signals are not of equal energy. These two facts result in a formulation, which is similar to, but not identical to, the traditional matched filter $r^T A_i$, which is the result of $i.i.d$ additive white Gaussian noise [25], [27].

The performance of the classifier given by (10) is a function of the noise variance and the modeled response. Table I lists the theoretical performance of the classifier as the noise variance is increased from σ_n^2 to $2^9\sigma_n^2$, where σ_n^2 is a vector that contains the noise variance of the magnitude and phase as a function of frequency obtained from experimental data (see Section. V-B). As expected, an increase in the noise variance results in a decrease in the classification performance. This analysis provides

insight into how the classification performance is affected by the sensor noise. Once the sensor is manufactured and well calibrated, the sensor noise cannot be changed artificially. Thus, only simulated data was used to illustrate this effect.

A more realistic assumption for the classification problem is that the height and horizontal position are uncertain, since the exact sensor position where measurements are obtained relative to the underground objects is unknown in practice. In this case, the previously derived processor [as in (10)], which assumed a "fixed" target/sensor orientation, is not the optimal solution. Hence, in order to obtain the optimal discriminant function for the received data, the effect of these random factors must be integrated out, i.e.,

$$p(\mathbf{r}|H_i) = \iiint p(\mathbf{r}|H_i, h, x, y)p(h)p(x, y)\, dh\, dx\, dy \quad (11)$$

where h represents the height of the sensor from the target, x and y represent the horizontal position of the sensor relative to the center of the target, $p(h)$ and $p(x, y)$ are the *a priori* distributions of the position factors, and

$$p(\mathbf{r}|H_i, h, x, y) = \frac{1}{(2\pi)^N |\Sigma|^{1/2}} \exp\Big(-\frac{1}{2}(\mathbf{r} - \mathbf{A}_i(h, x, y))^T \times \Sigma^{-1}(\mathbf{r}\mathbf{A}_i(h, x, y))\Big)$$

where $\mathbf{A}_i(h, x, y)$ is the model prediction (described in detail in Section III) of the ith target response when it is located at the position (h, x, y) relative to the sensor. The model predicts the theoretical frequency-domain EMI response as a function of constitutive parameters, exact dimensions of the object and the horizontal and vertical distance from the center of sensor to that of the object. Monte Carlo integration was implemented in order to calculate the integral in (11).

V. Simulated and Experimental Data

The performance of both the fixed-position processor (10) and the optimal classifier (11) is investigated by using both simulations and measurements for the GEM-3 sensor. In this section, the method used to generate simulated data, the experimental design, and the methods used to take the measurements are described.

Four metal targets are considered for both the simulations and experimental measurements: an aluminum barbell, an aluminum disk, a thick brass disk, and a thin brass disk. The dimensions of these targets are as follows. The diameter of each of the targets is 5.08 cm. The heights of the targets are 2.897 cm, 2.667 cm, 2.34 cm, and 0.3175 cm for the aluminum bar-bell, the aluminum disk, the thick and the thin brass disk, respectively. The response from a target depends on the constitutive parameters, geometry of the target, as well as the horizontal and vertical distance from the center of the sensor to that of the target. In the calculations, 21 linearly spaced frequencies were chosen, ranging from 3990 Hz to 23 970 Hz. These frequencies are within the range that the GEM-3 operates.

A. Simulations

In order to test whether the classification performance is improved by incorporating the model into the classification formulation, several cases were considered. These cases are:

1) fixed position;
2) random height but fixed horizontal position;
3) random horizontal position but fixed height;
4) both height and horizontal position random.

It is assumed that the distribution of the height, h, follows a Gaussian distribution with a mean of 20 cm and variance of 1.53^2 cm^2 and the horizontal position is uniformly distributed in a 20 cm by 20 cm square. To generate the simulated data we specify the constitutive parameters and the dimensions of the target, then generate 10 000 random sets of height (h) and horizontal position (x, y), which follow the distributions described above. Based on these parameters, the wideband EMI response is calculated by the model for each set of h, x, and y. After obtaining the theoretical responses of each target at all the specified positions (which are used in the formulation of the detector), Gaussian noise is added to the theoretical responses to create the simulated data set. Both the processor, which assumes a fixed target/sensor orientation (10), and the optimal classifier (11) are then applied to these data. Results of these classifiers are discussed in Section VI.

B. Measured Data

Using synthetic data to evaluate the performance of the classifier provides useful insight regarding performance bounds, but limiting the analysis to simulated data is not sufficient. Therefore, measurements of the wideband frequency-domain EMI response from the four metal targets were taken using the GEM-3 in order to evaluate the improvement of the classification algorithm in a more realistic scenario. First we consider whether data taken from objects in air is comparable to data obtained when the objects are buried in soil. Figs. 2 and 3 provide the wideband frequency-domain EMI responses measured using the GEM-3 from two metal landmines, a Valmara (an antipersonnel metal landmine) and a VS50 (an antipersonnel metal mine), in air and in North Carolina clay soil (buried 1 in below the surface). These figures indicate that the soil effects can be neglected at least for large metal objects. Therefore, the measured data used to evaluate the performance of various classification techniques was taken in free space.

The experimental set up is as follows. The GEM-3 was mounted on a wooden rack with the sensor head, approximately 1.8 m above the wooden base of the platform. Both rack and platform contained no metal parts. The rack assembly allows placement of a target on a wooden shelf at various distances beneath the sensor head.

First, in order to obtain an estimate of the noise variance associated with the sensor, $\sigma_{n_j}^2$, 100 measurements were taken with the sensor at a fixed position and no target present. We refer to an individual measurement taken without a target present as a background response. This response is subtracted from the responses measured with the target present to estimate the response due to the target alone. Fig. 4 shows a typical plot of the background

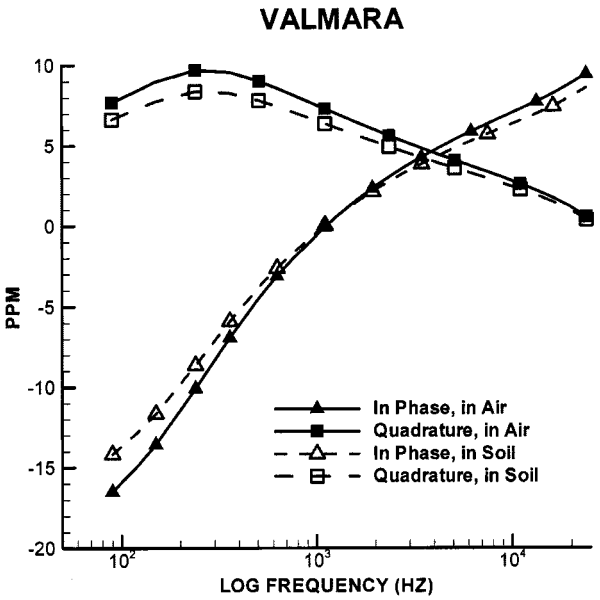


Fig. 2. Wideband frequency-domain response of a Valmara in free space and buried 1 in below the surface of the ground in North Carolina clay soil. The units "ppm" reflect the sensor output multiplied by 10^6.

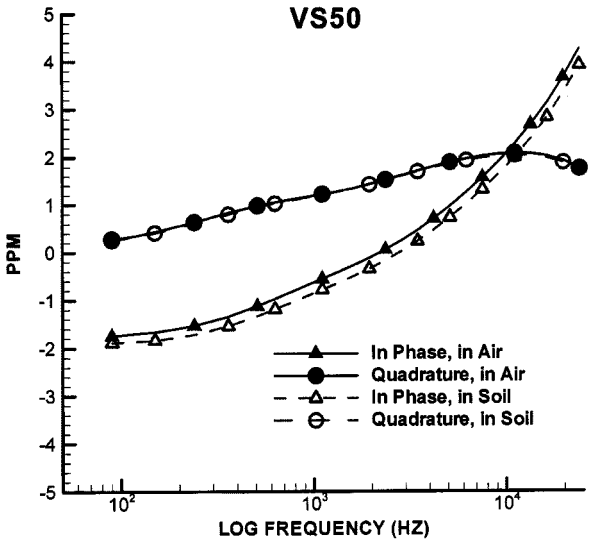


Fig. 3. Wideband frequency-domain response of a VS50 in free space and buried 1 in below the surface of the ground in North Carolina clay soil. The units "ppm" reflect the sensor output multiplied by 10^6.

response. The background response is not the same at all frequencies, and the noise variance is also a function of frequency. Table II lists the variance of the noise for the magnitude and phase of the complex response and the ratio between the mean value of the response and the standard deviation of the noise as a function of frequency, respectively. These estimates were used in the classifier given by (10) and (11).

As described in Section III, to calibrate the sensor, measurements for the four targets were taken so the calibration coefficients could be calculated. Each target was placed beneath the center of the sensor head at distances of 17 cm, 19 cm, 20 cm, 21 cm, and 23 cm. Using these 20 measurements, calibration coefficients were calculated by the least-squares method. Fig. 5

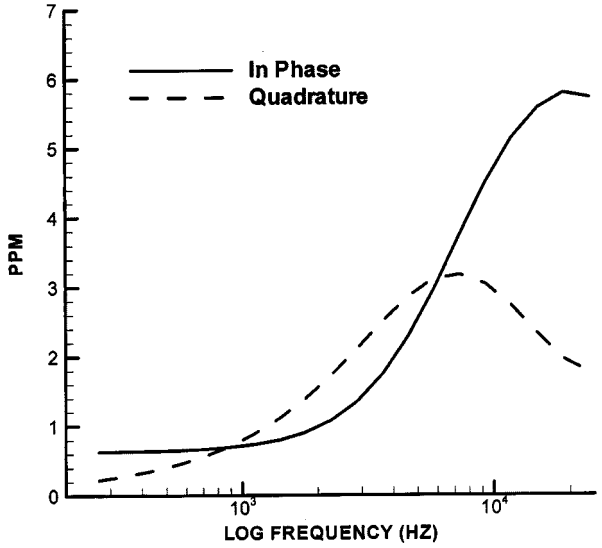


Fig. 4. Wideband frequency-domain response from Earth without the presence of any targets or background response. The units "ppm" reflect the sensor output multiplied by 10^6.

shows the comparison of the theoretical model predictions and the measurements. As has been noted previously, the model predicts the GEM-3 response well [12].

To obtain the data used to evaluate algorithm performance, measurements were taken from each target at seven heights from 17 cm to 23 cm in 1 cm increments. The distribution of height is assumed to be Gaussian with a mean of 20 cm and a variance of 1.53^2 for the simulations. At each height, between 11 and 36 measurements were taken. The exact count was calculated based on the assumed distribution. At each height, the position of each measurement is uniformly distributed within a 20 cm $\times$ 20 cm square. For each target, there were a total of 328 measurements taken. These data were not used to train the algorithm, only to evaluate performance.

VI. Results

We exploit Bayesian decision theory to formulate an optimal classifier to discriminate these targets. In order to show the improvement of the optimal classifier, the performance of a processor that assumes a fixed target/sensor orientation was also evaluated. It was assumed that the sensor is subject to a small amount of additive Gaussian noise. This assumption is verified by the experimental data (see Section V-B). The performance of these classifiers, shown in this section, was evaluated using both synthetic data and experimental measurements.

A. Simulation Results

1) Fixed Height and Horizontal Position: First, the case where all the position parameters are known exactly is considered. The model of each target at the same position and all desired frequencies is calculated. Then, by adding Gaussian random noise with zero mean and variance obtained based on the experimental data (see Table II), 10 000 realizations of simulated data for each target are generated. The decision of which target is present is made based on (7) by using the

TABLE II
VARIANCE OF THE BACKGROUND NOISE AND THE RATIO OF THE MEAN OF THE RESPONSE AND ITS STANDARD DERIVATION AS A FUNCTION OF FREQUENCY FOR THE MAGNITUDE AND PHASE COMPONENTS, RESPECTIVELY

Frequency (Hz)	$\sigma^2_{magnitude}$	$\frac{mean(magnitude)}{\sigma_{magnitude}}$	σ^2_{phase}	$\frac{mean(phase)}{\sigma_{phase}}$
3,990	8.25E-05	3.62E+02	8.86E-02	1.74E+02
5,010	8.99E-05	4.12E+02	5.55E-02	2.00E+02
5,970	9.41E-05	4.52E+02	3.91E-02	2.16E+02
6,990	9.47E-05	4.95E+02	2.83E-02	2.27E+02
8,010	8.96E-05	5.46E+02	2.19E-02	2.31E+02
8,970	8.35E-05	5.96E+02	1.77E-02	2.31E+02
9,990	7.47E-05	6.58E+02	1.44E-02	2.28E+02
11,010	6.57E-05	7.27E+02	1.21E-02	2.22E+02
11,970	5.73E-05	8.00E+02	1.05E-02	2.14E+02
12,990	5.44E-05	8.42E+02	8.99E-03	2.05E+02
14,010	5.70E-05	8.40E+02	8.05E-03	1.92E+02
14,970	6.23E-05	8.18E+02	7.20E-03	1.80E+02
15,990	9.25E-05	6.82E+02	6.23E-03	1.70E+02
16,950	1.11E-04	6.31E+02	5.14E-03	1.64E+02
17,970	1.55E-04	5.42E+02	4.87E-03	1.46E+02
18,990	1.49E-04	5.62E+02	9.35E-03	9.00E+01
19,950	1.63E-04	5.40E+02	1.39E-02	6.25E+01
20,970	7.26E-04	2.60E+02	3.64E-03	1.01E+02
21,950	3.90E-04	3.58E+02	9.92E-03	4.99E+01
22,950	7.81E-04	2.58E+02	9.23E-03	3.99E+01
23,970	1.21E-03	2.10E+02	8.32E-03	3.10E+01

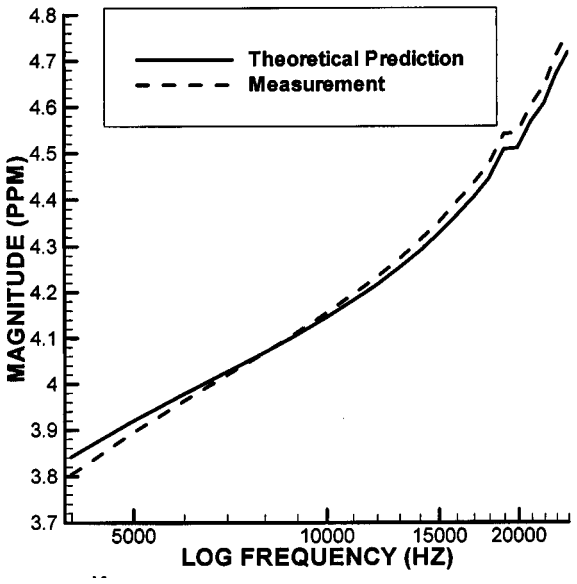


Fig. 5. Comparison of measurements and theoretical predictions for the thin brass disk when the distance from the target to the sensor is 20 cm.

processor expressed in (10), which is optimal for this case. Because of the fact that the wideband EMI signature of these targets is significantly different [12] and the experimentally derived σ_n^2's are low, the performance is perfect.

2) Height Uncertain, Fixed Horizontal Position: Next, the case where only the height of the sensor from the target is unknown and the target is located under the center of the sensor is considered. This situation occurs in a real detection scenario when the sensor operator can accurately center the sensor, but the burial depth of the mine is unknown. The height of the sensor was modeled as a Gaussian distributed random variable with a mean of 20 cm and a variance of 1.53^2 cm^2. Fig. 6 shows the performance of a processor that assumed a fixed target/sensor orientation along with that of the optimal classifier. For the former, it is assumed the target is at the mean height of 20 cm. Clearly, substantial improvements in classification performance are achieved by the optimal classifier over a processor that assumes a fixed target/sensor orientation. This performance is achieved for a relatively small level of uncertainty in the height. The average performance improvement is over 70%.

3) Horizontal Position Uncertain, Fixed Height: Thirdly, we simulate the case where horizontal position is uncertain. It is assumed that the sensor is located at a known, fixed height. Because the exact positions of mines are unknown to the sensor operator during detection, we assumed a uniform distribution in the horizontal plane. Fig. 7 shows the simulation results of the processor that assumes a fixed target/sensor orientation and the optimal classifier when the horizontal positions of targets are uniformly distributed. For the former, it was assumed that the target was at the mean horizontal position and was directly under the sensor. Again, the performance of the optimal classifier is substantially better than that of the processor, which ignores the target/sensor orientation uncertainty. It improves on average by 60%.

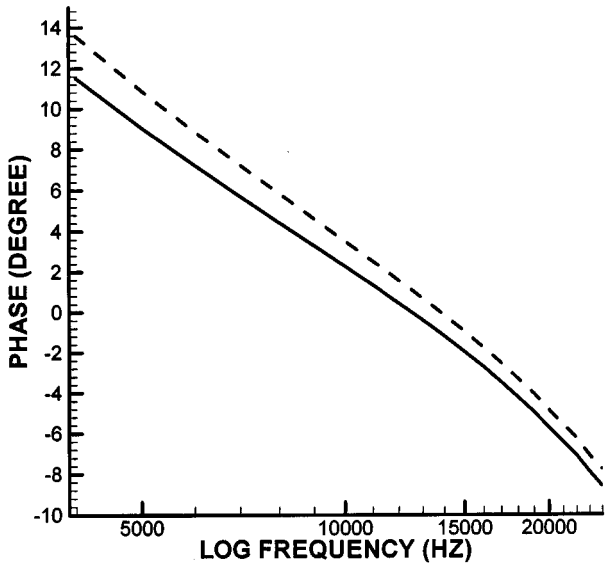


Fig. 6. Comparison of the processor, which ignores target/sensor orientation uncertainty ("fixed position" processor) and the optimal processor under uncertain height, fixed-horizontal position conditions for simulated data.

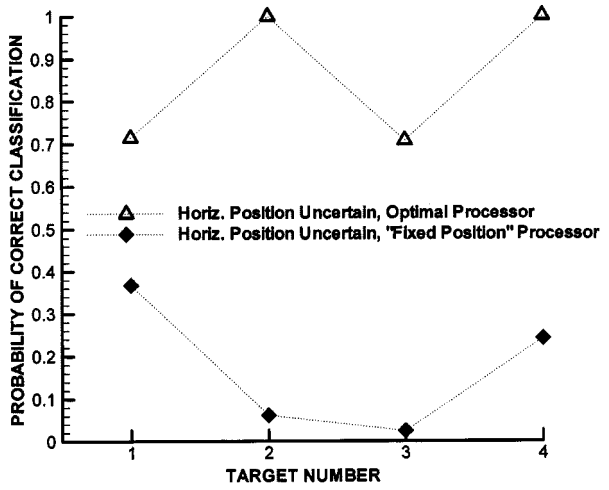


Fig. 7. Comparison of the "fixed position" processor and optimal processor performance under the uncertain horizontal position, but fixed height condition for simulated data.

4) Both Height and Horizontal Position Uncertain: In the final simulation, both height and horizontal position are uncertain. The height is assumed to follow a Gaussian distribution with mean of 20 cm and variance of 1.53^2. The horizontal position follows a uniform distribution (within a 20 cm $\times$ 20 cm square). Fig. 8 illustrates the performance of the two processors. The "fixed orientation" processor assumes that the target is located at the mean height and horizontal position. Performance improves under these conditions by an average of 70% over that of the "fixed position" processor. The results in Figs. 6–8 indicate that for the fixed position processor the performance becomes progressively worse as the position uncertainty increases. Clearly, incorporating the uncertainty of these environmental parameters into the processor affords a significant performance gain over a processor, which ignores this uncertainty.

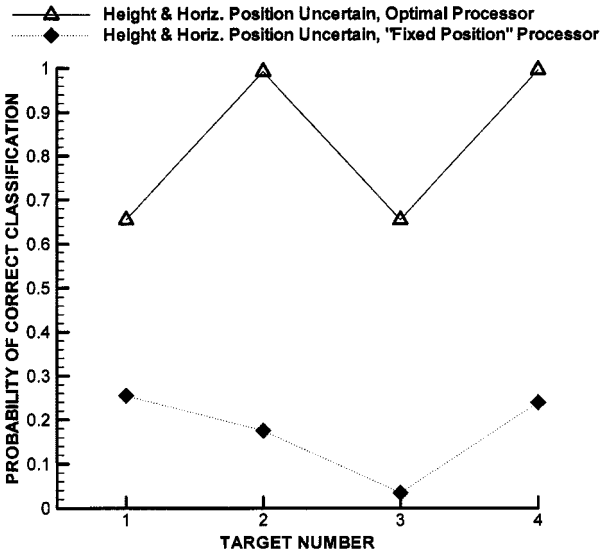


Fig. 8. Comparison of the "fixed position" processor and optimal processor performance when both height and horizontal position are uncertain for simulated data.

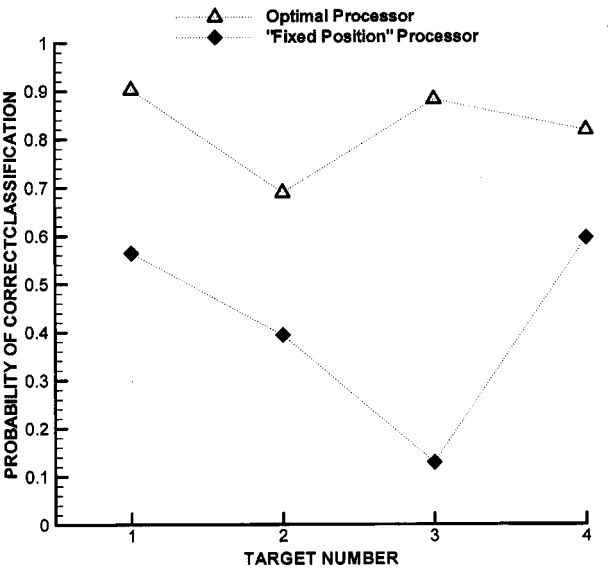


Fig. 9. Comparison of the "fixed position" processor and optimal processor performance under the condition of both height and horizontal position unknown for measured data.

B. Experimental Data

Simulations have shown that significant performance improvements can be achieved when the position uncertainty is incorporated into the classifier. To verify this result, measured data were collected using the GEM-3, as described in Section V-B. In this section, the results of implementing these processors using the measured data are shown.

The same two signal processing algorithms that were applied to simulated data: a fixed position processor, which assumes each target at the mean position, and the optimal classifier, which incorporates the position uncertainty into the processor, were applied to the experimental data. Fig. 9 illustrates the

performance achieved by each of these two algorithms. Clearly, better performance is achieved by the optimal processor. Performance improves on the average by 60%. This improvement, obtained on the measured data, is consistent with that observed in the simulated data set.

VII. Conclusion

In this paper, we utilize a Bayesian decision-theoretic approach to classify metal targets using wideband EMI data. Four manmade metal targets were used. Results from both simulation and measured data, shown in Section V, indicate that incorporating the uncertainty associated with the target/sensor relative position into the processor affords a significant performance gain over a processor that matches to the predicted response at the mean expected target position. It is also noted that, as expected, under conditions of uncertainty the performance of both the fixed orientation processor and the optimal processors drops compared to that of the signal-known-exactly case. Though the optimal classifier can improve performance under uncertain conditions over processors that ignore the uncertainties, it will never achieve the performance obtained when no uncertainty is present.

As expected, simulations have shown that the SNR's of the signal affect the performance of a classifier. Since frequency-domain systems can achieve high SNR's compared to time-domain EMI systems, essentially, it will improve the classification performance.

Our work indicates that we can effectively discriminate different metal targets using wideband EMI signals by incorporating an accurate physical model and models of the uncertainty regarding environmental parameters into the classifier. Performance can be dramatically improved over the standard approach, which ignores environmental uncertainty. In addition, extensive libraries of target signatures do not have to be measured experimentally in order to train the classifier.

This technique can be extended to apply to other applications such as landmine detection and unexploded ordnance (UXO) detection, since in these applications, targets of interests need to be discriminated from metallic clutter in order to reduce false alarm rates. The standard algorithms for these applications, such as matched filters, do not take the uncertainties associated with the target/sensor orientation into account, and only partially exploited the underlying physical nature of the outputs from the sensor. The work shown in this paper provides a promising technique, which integrates both the uncertainties associated with target/sensor orientation and a forward model exploiting the physical signature of wideband frequency-domain EMI response. By developing a model for other signals and sensor modalities, this algorithm can be further applied to other applications that require classification of different targets.

Acknowledgment

The authors would like to thank Dr. D. Keiswetter and Dr. I. J. Won for their assistance with the GEM-3. They would also like to thank R. Weaver for supplying the landmines and for his support of this work through the Joint UXO Coordinate Office (JUXOCO), Fort Belvoir, VA. They would also like to thank the reviewers, Y. Tan for her assistance with the data collection, and Dr. L. Nolte and Dr. S. Tantum for useful discussions regarding this work.

References

[1] L. Collins, P. Gao, and L. Carin, "An improved bayesian decision theoretic approach for land mine detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 811–819, Mar. 1999.

[2] C. E. Baum, "Low Frequency Near-Field Magnetic Scattering from Highly, but Not Perfectly Conducting Bodies," Phillips Laboratory, Interaction Note 499, Nov. 1993.

[3] G. D. Sower and S. P. Cave, "Detection and identification of mines from natural magnetic and electromagnetic resonances," in *Proc. SPIE*, Orlando, FL, 1995.

[4] A. H. Trang, P. V. Czipott, and D. A. Waldron, "Characterization of small metallic objects and non-metallic anti-personnel mines," in *Proc. SPIE*, Orlando, FL, 1997.

[5] S. Vitebskiy and L. Carin, "Short-pulse plane wave scattering from a buried perfectly conducting body of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 112–120, Jan. 1996.

[6] S. Vitebskiy and L. Carin, "Late-time resonant frequencies of buried bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.

[7] C. E. Baum, N. Geng, and L. Carin, "Integral Equations and Polarizability for Magnetic Singularity Identification," Phillips Lab., Interaction Note 524, Mar. 1997.

[8] N. Geng, C. E. Baum, and L. Carin, "On the low-frequency natural response of conducting and permeable targets," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 347–359, Jan. 1999.

[9] I. J. Won, D. A. Keiswetter, and D. R. Hansen, "GEM-3: A monostatic broadband electromagnetic induction sensor," *J. Env. Eng. Geophys.*, vol. 2, pp. 53–64, Aug. 1997.

[10] D. Keiswetter, I. J. Won, B. Barrow, and T. Bell, Object Identification Using Multifrequency EMI Data, UXO Forum, Atlanta, Georgia, May 1999.

[11] P. Gader, H. Frigui, G. Vaillette, B. Nelson, and J. Keller, "New results in fuzzy set based detection of landmines with GPR," in *Detection and Remediation Technologies for Mines and Minelike Targets IV Conf., Int. Symp. Aerospace/Defense Sensing and Controls*, Orlando, FL, Apr. 1999.

[12] N. Geng *et al.*, "Wideband Electromagnetic Induction for Metal-Target Identification: Theory, Measurement and Signal Processing," Duke Univ., Technical Report, Sept. 1997.

[13] I. J. Won, D. Keiswetter, G. R. A. Fields, and L. C. Sutton, "GEM-2: A new multifrequency electromagnetic sensor," *J. Env. Eng. Geophys.*, vol. 1, pp. 129–137, Mar. 1996.

[14] R. F. Harrington, *Field Computation by Moment Methods*. New York: Macmillan, 1968.

[15] J. R. Mautz and R. F. Harrington, "Radiation and scattering from bodies of revolution," *Appl. Sci. Res.*, vol. 20, pp. 405–435, June 1969.

[16] T. Wu and L. L. Tsai, "Scattering from arbitrarily-shaped lossy dielectric bodies of revolution," *Radio Sci.*, vol. 12, pp. 709–718, Sept. 1977.

[17] J. R. Mautz and R. F. Harrington, "Electromagnetic scattering from a homogeneous material body of revolution," *AEÜ*, vol. 33, pp. 71–80, Feb. 1979.

[18] A. W. Glisson and D. R. Wilton, "Simple and efficient numerical methods for problems of electromagnetic radiation and scattering from surfaces," *IEEE Trans. Antennas Propagat.*, vol. 28, pp. 593–603, Sept. 1980.

[19] A. W. Glisson, D. Kajfez, and J. James, "Evaluation of modes in dielectric resonators using a surface integral equation formulation," *IEEE Trans. Microw. Theory Technol.*, vol. 31, pp. 1023–1029, Dec. 1983.

[20] A. W. Glisson, "Integral equation techniques," in *Dielectric Resonators*, D. Kajfez and P. Guillon, Eds. Dedham, MA: Artech House, 1986, pp. 259–325.

[21] C. A. Balanis, *Antenna Theory—Analysis and Design*. New York: Wiley, 1982.

[22] K. Simonyi, *Theoretische Elektrotechnik*. Berlin, Germany: Springer-Verlag, 1980.

[23] C. A. Balanis, *Advanced Engineering Electromagnetics*: Wiley, 1989.

[24] J. C. Hancock and P. A. Wintz, *Signal Detection Theory*. New York: McGraw Hill, 1966, pp. 80–84.

[25] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.

[26] R. J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*. New York: Wiley, 1992, pp. 34–43.
[27] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. San Diego, CA: Academic, 1990, p. 127.



Ping Gao (S'97) received the B.S. and M.E. degrees (honors) in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 1993 and 1996, respectively, and the M.S.E.E. from Duke University, Durham, NC, in 1997, where she is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering.

Since 1997, she has worked as a Research Assistant at Duke University. Her main research interests are in signal processing, detection and estimation theory, statistical signal modeling and analysis, and applied electromagnetics.

Leslie Collins (M'96) [redacted] in Raleigh, NC. She received the B.S.E.E. degree from the University of Kentucky, Lexington, in 1985, and the M.S.E.E. and Ph.D. degrees in electrical engineering, both from the University of Michigan, Ann Arbor, in 1986 and 1995, respectively.

She was a Senior Engineer with the Westinghouse Research and Development Center, Pittsburgh, PA, from 1986 to 1990. In 1995, she became an Assistant Professor in the Electrical and Computer Engineering Department, Duke University, Durham, NC. Her current research interests include incorporating physics-based models into statistical signal processing algorithms, and she is pursuing applications in subsurface sensing as well as enhancing speech understanding by hearing impaired individuals.

Dr. Collins is a member of the Tau Beta Pi, Eta Kappa Nu, and Sigma Xi honor societies.

Philip M. Garber (S'99) [redacted] He received the B.S. degree in electrical engineering from Duke University, Durham, NC, in 1999.

He is currently a Graduate Student with the Radiation Laboratory, University of Michigan, Ann Arbor. His interests are in microwave circuits, communications, and wireless systems.



Norbert Geng (S'91–M'96) [redacted] in Lauchringen, Germany. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Karlsruhe, Germany in 1991 and 1996, respectively.

From 1991 to 1996, he was with the Institute for Microwaves and Electronics, University of Karlsruhe, working on full-wave propagation modeling for radio communication systems. In January 1997, he joined the Department of Electrical and Computer Engineering, Duke University, Durham, NC, as a Visiting Postdoctoral Fellow for 18 months. Since July 1998, he has been with the University of Karlsruhe. His current research interests include computational methods in electromagnetics and wave propagation modeling.

Dr. Geng received the Mannesmann Innovation Award in 1997 for his Ph.D. thesis on full-wave propagation modeling for radio communication systems.

Lawrence Carin (SM'96) [redacted] in Washington, DC, and earned the B.S., M.S., and Ph.D. degrees in electrical engineering at the University of Maryland, College Park, in 1985, 1986, and 1989, respectively.

In 1989, he joined the Electrical Engineering Department, Polytechnic University, New York, as an Assistant Professor, and became an Associate Professor there in 1994. In September 1995, he joined the Electrical Engineering Department, Duke University, Durham, NC, where he is an Associate Professor. His current research interests include short pulse scattering, subsurface sensing, and wave-based signal processing.

Dr. Carin is a member of the Tau Beta Pi and Eta Kappa Nu honor societies. He is currently an Associate Editor of the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION.

Anal. Chem. **1998,** *70,* 2560−2564

Use of Compatible Polymer Blends To Fabricate Arrays of Carbon Black−Polymer Composite Vapor Detectors

Brett J. Doleman, Robert D. Sanner,† Erik J. Severin, Robert H. Grubbs, and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Compatible blends of poly(vinyl acetate) and poly(methyl methacrylate) have been used to produce a series of electrically conducting carbon black composites whose resistance is sensitive to the nature and concentration of an analyte in the vapor phase. The dc electrical resistance response of the composites was found to be a nonlinear function of the mole fraction of poly(vinyl acetate) in the blend. These compatible blend composite detectors provided additional analyte discrimination information relative to a reference detector array that only contained composites formed using the pure polymer phases. The added discrimination power provided by the compatible blend detectors, and thus the added diversity of the enhanced detector array, was quantified through use of a statistical metric to assess the performance of detector arrays in various vapor classification tasks.

Recent work has shown that arrays of chemically sensitive resistors, formed from composites of carbon black with a collection of swellable insulating organic polymers, are broadly responsive to a variety of odors yet allow classification and identification of organic vapors through application of pattern recognition methods.[1] To date, these array elements have been fabricated from a relatively small number of approximately 10−20 organic polymers, with a distinct polymer backbone composition in each detector.[1] Detectors having on the order of 3−10 different polymer backbones have also formed the basis for vapor detector arrays based on bulk conducting polymer films,[2,3] surface acoustic wave devices,[4−6] fiber-optic micromirrors,[7] quartz crystal microbalances,[8] and dye-containing coatings on optical fibers.[9] Although a limited number of polymeric detector compositions might be chosen to perform optimally for a specific application,[4,10] attempts to mimic the sense of olfaction or attempts to address applications in which the sensing task is time dependent or is not defined in advance of the detector array construction will almost certainly require use of extensive, compositionally diverse polymeric detector libraries. Since it is likely to be nontrivial to prepare arrays that contain thousands of chemically distinct polymers using thousands of individual polymer-containing feedstock solutions, combinatorial and/or gradient methods seem appropriate to address this problem. Herein we describe an approach to the construction of detector arrays in which compatible blends of two base polymers are used to create compositionally varying chemically sensitive resistor films. The resolving power of such an array is shown to be superior, in several test tasks, to that of arrays containing an identical number of detectors but composed of only the base polymeric materials.

Nonlinearity of the gas−solid partition coefficient with respect to the mole fraction of the polymers that comprise the compatible blend is crucial to increasing the diversity of a broadly responsive detector array that is fabricated through gradient or combinatorial methods. Otherwise, the response of the blended chemiresistor elements can be predicted precisely from the responses of the base polymeric detector materials. Two polymers, poly(vinyl acetate) (PVA) and poly(methyl methacrylate) (PMMA), were selected to form the compositionally varied chemiresistor composites studied in this work. These materials were chosen because a similar binary polymer blend system, formed from poly(vinyl acetate) and poly(methyl acrylate), has previously been reported to exhibit gas−solid partition coefficients for benzene that are not linearly related to the mole fraction of PVA in the polymer blend.[11]

EXPERIMENTAL SECTION

Eight different PVA/PMMA blend compositions were investigated as carbon black composite chemiresistor vapor detectors. The compatible blend detector fabrication was achieved by combining the two initial base polymer feedstocks to produce solutions of PVA/PMMA having PVA mole fractions (by monomer) of 0.00, 0.11, 0.28, 0.44, 0.64, 0.78, 0.91, and 1.00, respectively. Each stock solution contained 20 mL of tetrahydrofuran, 200 mg of total dissolved polymer, and 86 mg of suspended carbon black.

† Current address: Mail Code L-325, Lawrence Livermore National Laboratory, Livermore, CA 94550.

(1) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298.
(2) Barker, P. S.; Chen, J. R.; Agbor, N. E.; Monkman, A. P.; Mars, P.; Petty, M. C. *Sens. Actuators, B* **1994**, *17*, 143.
(3) Gardner, J. W.; Pearce, T. C.; Friel, S.; Bartlett, P. N.; Blair, N. *Sens. Actuators, B* **1994**, *18*, 240.
(4) Grate, J. W.; Patrash, S. J.; Abraham, M. H. *Anal. Chem.* **1995**, *67*, 2162.
(5) Grate, J. W.; Martin, S. J.; White, R. M. *Anal. Chem.* **1993**, *65*, 5, A987.
(6) Grate, J. W.; Martin, S. J.; White, R. M. *Anal. Chem.* **1993**, *65*, 5, A940.
(7) Hughes, R. C.; Ricco, A. J.; Butler, M. A.; Pfeifer, K. B. *J. Biochem. Biotechnol.* **1993**, *41*, 77.
(8) Chang, S.-M.; Iwasaki, Y.; Suzuki, M.; Tamiya, E.; Karube, I.; Muramatsu, H. *Anal. Chim. Acta* **1991**, *249*, 323.
(9) White, J.; Kauer, J. S.; Dickinson, T. A.; Walt, D. R. *Anal. Chem.* **1996**, *68*, 2191.
(10) Grate, J. W.; Abraham, M. H. *Sens. Actuators, B* **1991**, *3*, 85.
(11) Nunez, E. M.; Myerson, A. S.; Kwei, T. K. *Polym. Eng. Sci.* **1991**, *31*, 1172.

S0003-2700(97)01238-9 CCC: $15.00 © 1998 American Chemical Society
Published on Web 05/13/1998

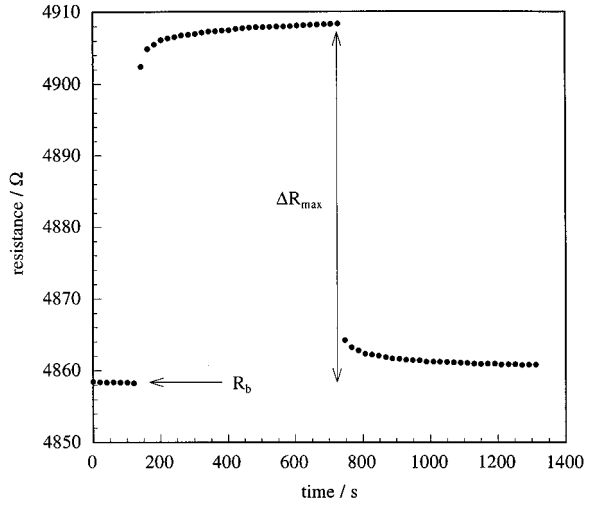


Figure 1. Temporal resistance response of a typical polymer composite chemiresistor detector. This particular carbon black composite detector contained 64% PVA and 36% PMMA by monomer mole fraction. The detector was exposed to 11.3 ppth of methanol in air for 600 s starting at $t = 120$ s. The baseline resistance before the exposure, R_b, and the maximum resistance change during the exposure, ΔR_{max}, were 4858 and 50 Ω, respectively.

Standard glass microscope slides, cut to a size of approximately 2 cm × 2.5 cm, were modified for use as the substrate for each polymer blend detector. Two parallel bands of 20 nm thick chromium (≈2 cm × 1 cm), spaced 0.5 cm apart, were evaporated onto each slide. The chromium bands were then coated with 30 nm of evaporated gold, thus forming robust electrical contacts. Each carbon black−polymer suspension was sonicated for 10 min and was then spin-coated, at 1000 rpm, onto a modified glass slide such that the gap between the slide electrical contacts was spanned by the polymer composite film. The detectors were allowed to dry in ambient air for 12 h before use.

To obtain response data, the detectors were placed into a 1.2 L sampling chamber and electrical leads were attached to the two chromium−gold bands of each detector. The dc resistance of each detector was recorded as a function of time using a Keithley model 7001 channel switcher connected to a Keithley model 2002 multimeter that was interfaced to a personal computer. An automated flow system consisting of LabVIEW software, a personal computer, and electronically controlled solenoid valves and mass flow controllers was used to produce and deliver controlled concentrations of solvent vapors to the detectors in the sampling chamber. The desired vapor concentrations were obtained by passing a stream of carrier gas through a bubbler that had been filled with the solvent of choice and then diluting this flow into a stream of air maintained at a controlled flow rate. The time protocol for each exposure was 120 s of air, followed by 600 s of test vapor in air, ending with another 600 s of air.

Figure 1 displays the resistance response of a typical detector. Upon exposure to a test vapor, the resistance of the composite film increased, and the response then decreased after the vapor exposure was terminated. This behavior has been discussed in detail for a series of pure polymeric compositions that have been used as either carbon black or polypyrrole composites to provide arrays of electrically conductive vapor detectors.[1,12] To assess the performance of the compatible blend composites, all of the detectors were exposed 10 times each, in random order, to five different analytes, with the vapor concentrations chosen to be 11.3 parts per thousand (ppth) of methanol, 5.2 ppth of ethanol, 20.7 ppth of acetone, 8.3 ppth of ethyl acetate, and 8.2 ppth of acetonitrile in air at 21 °C. These concentrations all correspond to 7.1% of the solvent-saturated concentration of each analyte in 21 °C air, under a total atmospheric pressure of 753 Torr. The maximum differential resistance response relative to the baseline resistance ($\Delta R_{max}/R_b$) was used in the analysis of the array performance carried out in this work.

RESULTS AND DISCUSSION

Figure 2 depicts plots of $\Delta R_{max}/R_b$ for the polymer blend chemiresistors upon exposure to ethyl acetate, ethanol, acetonitrile, acetone, and methanol. For each analyte, a statistically significant nonlinearity was observed for the detector response vs the mole fraction of the base polymer feedstocks. Since the nonlinearity is not the same for all solvents, this indicates that useful information is available through use of such compatible blend materials in a detector array for vapor classification.

The ability of a specific detector array to resolve pairs of solvent vapors can be quantified statistically through reference to a generalized resolution factor, rf. This quantity is equivalent to that proposed by Müller[13] and recently used by Gardner and Bartlett[14] and is a multidimensional analogue of the separation factors used in gas chromatography. Resistance responses, $\Delta R_{max}/R_b$, of carbon black−polymer composite detectors, containing ≥20 wt % carbon black, have been shown to vary linearly over at least an order of magnitude in the concentration of the analyte in the vapor phase.[1] Hence, detector arrays which can resolve analytes at one concentration can also be used to resolve analytes at other concentrations. The detector responses were autoscaled to account for the different dynamic ranges of different detectors. The autoscaled response of the jth detector to the ith exposure, A_{ij}, was thus

$$A_{ij} = \frac{(\Delta R_{ij,max}/R_b) - \alpha_j}{\beta_j} \quad (1)$$

where α_j and β_j are the mean and standard deviations, respectively, in the responses of the jth detector to all analytes. The mean response vector, $\vec{x}_a$, of an n-detector array to analyte a is taken as the n-dimensional vector containing the mean autoscaled response of each detector, $\bar{A}_{aj}$, to the ath analyte such that

$$\vec{x}_a = (\bar{A}_{a1}, \bar{A}_{a2}, ..., \bar{A}_{an}) \quad (2)$$

The average separation, $|\vec{d}|$, between two analytes, a and b, in the Euclidean detector response space is then simply the magnitude of the difference between $\vec{x}_a$ and $\vec{x}_b$. The reproducibility of the array responses to the analytes is also important in quantifying the resolving power of the array. A measure of array response reproducibility to analyte a, $\sigma_{a,\vec{d}}$, is obtained by projecting

(12) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2652.
(13) Müller, R. *Sens. Actuators, B* **1991**, *4*, 35.
(14) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators B* **1996**, *33*, 60.

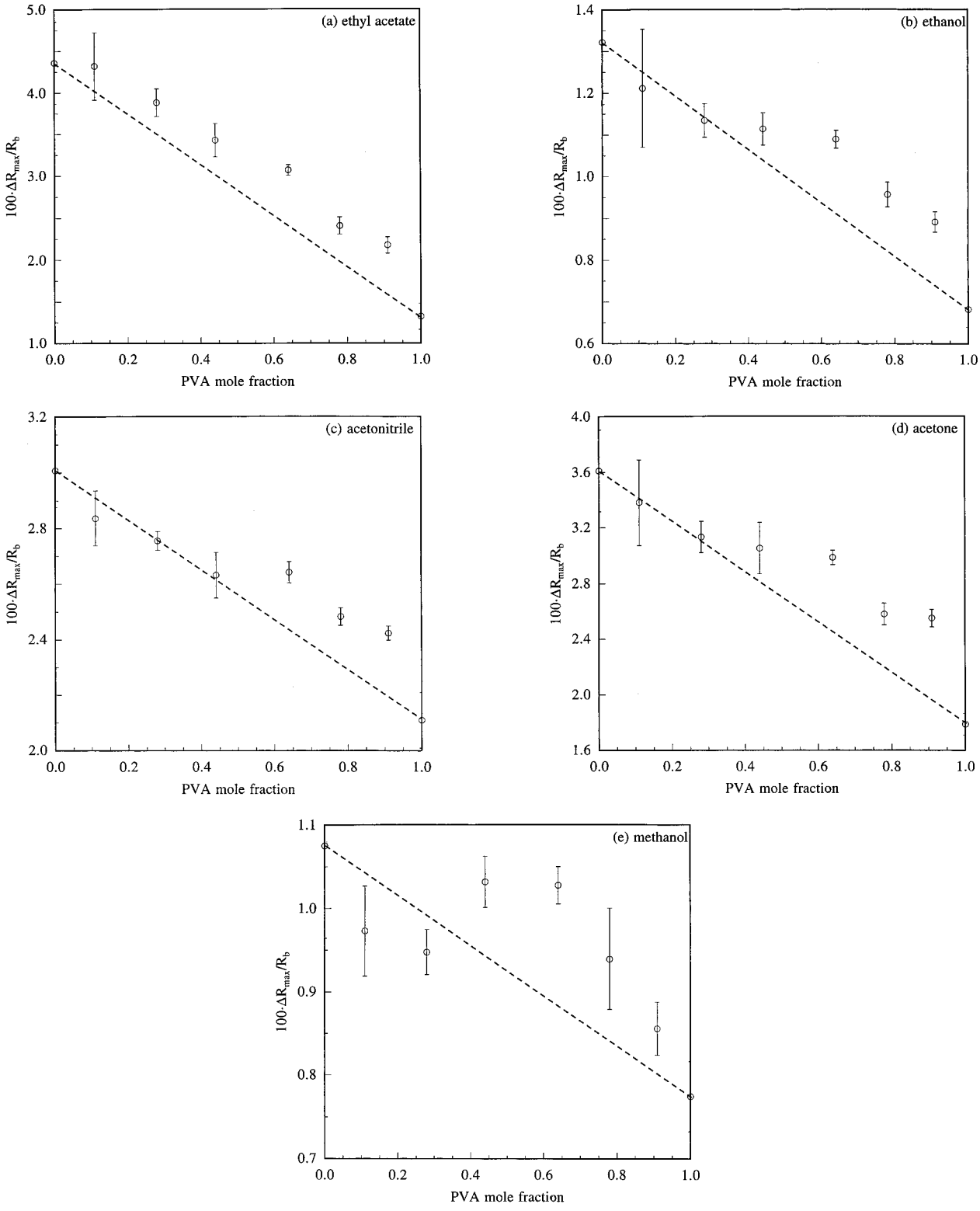


Figure 2. Maximum relative differential resistance response, $\Delta R_{max}/R_b$, of a series of polymer blend−carbon black composite chemiresistors upon exposure to (a) 8.3 ppth of ethyl acetate, (b) 5.2 ppth of ethanol, (c) 8.2 ppth of acetonitrile, (d) 20.7 ppth of acetone, and (e) 11.3 ppth of methanol in air. The plots depict data obtained from 14 detectors of pure PMMA, 10 with 11% (by monomer mole fraction) PVA, 10 with 28% PVA, 15 with 44% PVA, 10 with 64% PVA, 15 with 78% PVA, 10 with 91% PVA, and 15 of pure PVA. The responses plotted for each mole fraction are the mean $\Delta R_{max}/R_b$ values for 10 exposures to each set of detectors containing the specified mole fraction of PVA, while the error bars represent one standard deviation unit. Dashed lines were drawn, joining the end points, as a guide to the eye indicating a linear response relationship.

the array response vectors for each exposure to analyte *a* onto the vector $\vec{d}$, and calculating the standard deviation in these scalar projections about the projection of the mean response vector, $\vec{x}_a$ onto $\vec{d}$. The same procedure is repeated for analyte *b* of the *a,b*

Table 1. Resolution Factors Quantifying the Average Ability of 8-Detector Arrays, Composed from Four Different Detector Sets, To Resolve Pairwise the Collection of Test Analytes[a]

sensors used	overall av rf	methanol vs ethanol	methanol vs ethyl acetate	methanol vs acetonitrile	methanol vs acetone	ethanol vs ethyl acetate	ethanol vs acetonitrile	ethanol vs acetone	ethyl acetate vs acetonitrile	ethyl acetate vs acetone	acetonitrile vs acetone
set A	52	25	61	90	44	58	93	42	58	20	27
set B	60	19	67	104	81	67	110	81	31	17	26
set C	81	4.6	122	102	181	103	93	148	17	31	8.7
set D	60	23	84	93	60	82	96	58	55	22	26

[a] The overall average rf represents the average resolution factor across all analyte pairs for random combinations of detectors from a given detector set. The results for set A, set B, and set D were obtained by averaging over 500 randomly selected 8-detector arrays composed of only the detectors within each respective set. The results for set C were obtained by averaging over all 45 possible 8-detector combinations of the detectors within the set. Set A contained all 14 of the 0% PVA detectors and all 15 of the 100% PVA detectors (i.e., all the base polymer detectors). Set B contained all 99 of the prepared detectors ranging from 0% to 100% PVA Content. Set C contained only the 10 detectors with 91% PVA. Set D contained all 14 of the 0% PVA detectors, all 10 of the 91% PVA detectors, and all 15 of the 100% PVA detectors.

analyte pair, allowing a pairwise resolution factor to be defined as

$$\mathrm{rf} = \frac{|\vec{d}|}{\sqrt{\sigma_{a,\vec{d}}^{\,2} + \sigma_{b,\vec{d}}^{\,2}}} \qquad (3)$$

This metric allows quantification of the ability to resolve pairwise the vapors of concern in the test analyte set based on the response patterns that they produce on the detector array. Because the functional form of the response of the various polymer composite chemiresistors was very similar, this procedure can be used to provide an objective measure of array performance, as opposed to performing a subjective assessment of the performance of task-specific neural network classifiers on functionally dissimilar responses of various array elements.[15] It is important to realize, however, that the results are nevertheless coupled to the metric used to evaluate the responses and that different algorithms, such as, for example, Fisher linear discriminants, which are linear data analysis methods that are not confined to pass through the mean response values of the analytes of concern,[16] may well yield different conclusions from the same response data.

The responses produced by a set of 99 detectors, 14 detectors with pure PMMA, 10 with 11% PVA, 10 with 28% PVA, 15 with 44% PVA, 10 with 64% PVA, 15 with 78% PVA, 10 with 91% PVA, and 15 with pure PVA, were investigated using this approach. The performances of 8-detector combinations from different sets of detectors were evaluated to determine if arrays containing some of the compatible blend polymer detectors would perform better than arrays containing only detectors made from the base polymers, for certain test tasks. The performance of each studied array was measured by its ability to resolve the solvents pairwise, as given by the calculated rf values obtained using the simple linear data analysis method described above.

Results are presented for four sets of detectors. Set A contained all 14 detectors with 0% PVA and all 15 detectors with 100% PVA (i.e., all the base polymer detectors). Set B contained all 99 of the prepared detectors ranging from 0% through to 100% PVA content. Set C contained only the 10 detectors with 91% PVA. Set D contained all 14 of the 0% PVA detectors, all 10 of the 91% PVA detectors, and all 15 of the 100% PVA detectors. Since there are extremely large numbers of possible 8-detector combinations from within sets A, B, and D ($\approx 10^{11}$ unique 8-detector combinations out of 99 set B detectors), 500-member subsets of the total number of 8-detector array combinations were selected randomly and their corresponding rf values were calculated. For set C, rf values for all 45 possible 8-detector combinations out of 10 detectors were calculated. The results of the calculated resolution factors for arrays of 8-detectors within each set were averaged and are presented in Table 1.

Clearly, the inclusion of compatible blend detectors produced a statistically significant improvement in maximizing the overall average rf, which is the average ability of all calculated 8-detector array combinations within a set of detectors to resolve all analyte pairs using the metric defined above. For example, sets B, C, and D, which contained compatible blend detectors, had overall average rf's of 60, 81, and 60, respectively, whereas the base polymer detector arrays (set A) had an overall average rf of 52. The array performance in separating the pair of solvents, ethyl acetate vs acetone, that was worst resolved by set A (base polymer detectors) could also be improved by using 8-detector arrays containing only 91% PVA detectors (set C) or by including these detectors in arrays that contained the base polymer detectors (set D). Set D arrays, containing blended polymers, exhibit a larger overall average rf, a larger rf for the worst resolved analyte pair, and resolved 7 of the 10 analyte pairs better than did the base polymer arrays of set A.

The results presented in Table 1 show that each of the four detector sets produces arrays with at least one advantage over the others, such as the ability to resolve at least one analyte pair better than the other selected arrays. This reinforces the concept that more detector diversity is desirable, since it provides a larger basis of chemically unique detectors from which to tailor arrays for specific tasks.

Another significant conclusion arising from the data is that the classification of these various vapors, at fixed concentrations, is statistically robust from the array response even though the individual detectors themselves were not designed to possess high selectivity toward a specific analyte. For example, a pairwise resolution factor of 4.5 implies that, in a single presentation of the challenge vapors to the detector array, a given vapor can be

(15) Zupan, J.; Gasteiger, J. *Neural Networks for Chemists*; VCH: New York, 1993; p 305.

(16) Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; John Wiley & Sons: New York, 1973; p 482.

distinguished from the other member of the test pair with >99.9% confidence level using the simple statistical pairwise method of data analysis described herein. This level of performance was met or exceeded by all of the eight-element detector arrays of Table 1 for all of the test vapor pairs used in this work, even though the array elements were not chosen in advance specifically to perform any particular set of vapor classification tasks.

Utilization of a nonlinear response of binary, tertiary, and quarternary blend composite chemiresistors to various solvent vapors should offer the opportunity to increase significantly the diversity of a polymer composite detector array and therefore to increase its classification performance relative to an array that contains chemiresistors fabricated from the pure polymeric phases alone. The binary polymer blend advantages reported herein are in agreement with those recently published using a different detector modality, polymer−dye optical detectors.[17] The exact performance gain of any specific array will likely be task dependent and must be evaluated for each application of concern. We note, however, that the olfactory bulb of canines has approximately 100 million receptor cells and that humans have over 1000 different olfactory receptor proteins.[18] Thus, attempts to mimic functionally the olfactory sense are unlikely to be realizable without exploitation of methodologies capable of incorporating extensive diversity into a polymer-based vapor-sensing array.[18] Extension of the approach described herein to other blends and a comparison of the detector diversity that can be achieved through the use of block and random copolymers as a complement to the use of compatible blends in detector arrays will be reported separately.

ACKNOWLEDGMENT

We thank NASA, the Army Research Office (Grants DAAH04-96-1-0048 and DAAG55-97-1-0187), and DARPA for support of this work. B.J.D. acknowledges the NSERC Canada for a graduate fellowship, and R.D.S. acknowledges the University of California Lawrence Livermore National Laboratory for support from its Professional Research and Teaching Leave program while at Caltech.

Received for review November 12, 1997. Accepted March 19, 1998.

AC971238H

(17) Dickinson, T. A.; Walt, D. R.; White, J.; Kauer, J. S. *Anal. Chem.* **1997**, *69*, 3413.

(18) Axel, R. *Sci. Am.* **1995**, 154.

ELSEVIER

Sensors and Actuators B 72 (2001) 41–50

SENSORS AND ACTUATORS B CHEMICAL

www.elsevier.nl/locate/sensorb

Comparison of odor detection thresholds and odor discriminablities of a conducting polymer composite electronic nose versus mammalian olfaction

Brett J. Doleman, Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, Noyes Laboratory, 127-72 Pasadena, CA 91125, USA

Received 2 May 2000; received in revised form 20 July 2000; accepted 27 July 2000

Abstract

Response data from an array of conducting polymer composite vapor detectors that form an electronic nose were collected for the purpose of comparing selected, quantitatively measurable, phenomena in odor detection and classification to the olfactory characteristics of monkeys and humans. Odor detection thresholds and discriminability between structurally similar pairs of odorants were the two primary quantities evaluated for this comparison. Comparisons were only made for volatile organic vapors as opposed to aroma active odorant vapors. Electronic nose detection thresholds for a homologous series of *n*-alkane and 1-alcohol odorants were determined and the results were compared to literature values for the mean olfactory detection thresholds observed in psychophysical experiments on humans exposed to these same vapors. The trends in odor detection thresholds of the electronic nose towards the tested analytes were very similar to those exhibited by humans. The discrimination performance of the electronic nose for distinguishing between pairs of odorants within incrementally varying series of esters, carboxylic acids and alcohols were also compared to the published data of Laska and co-workers on the psychophysical performance of humans and monkeys for these same odorant pairs. Similar trends were generally observed between the humans, monkeys, and the electronic nose in that discrimination performance increased as the compounds of an odorant pair became more structurally dissimilar. With use of the Fisher linear discriminant algorithm for classification of these test pairs of odorants, the electronic nose exhibited significantly better discriminability than humans or monkeys for the odorant pairs evaluated in this work under the test conditions for which the discriminability was evaluated. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Electronic nose; Conducting polymer composites; Odor detection threshold; Odor discriminabilities

1. Introduction

The human olfactory epithelium contains $\approx 10^6$–10^7 total olfactory receptor neurons, which each are thought to contain only one of $\approx 10^3$ different types of olfactory receptor proteins [1–8]. It has been recently shown that one odorant receptor recognizes multiple odorants and that one odorant is recognized by multiple odorant receptors, but that different odorants are recognized by different combinations of odorant receptors [9]. These observations support some previous hypotheses of a distributed olfactory response to odorants as a possible mechanism for detection of odorants by the olfactory system [3–6,10–13]. Broad responsiveness to odorants has also been experimentally observed from electrophysiological recordings in the mammalian olfactory epithelium and in mitral/tufted cells [14].

Several artificial/electronic olfactory systems are being developed based on the notion of using broadly cross-reactive arrays of chemical sensors [15–21]. Such array modalities can include conducting organic polymers, polymer-coated surface acoustic wave resonators, polymer-coated quartz crystal microbalances, conducting polymer composites, dye-impregnated polymers coated on optical fibers, and electrochemical gas sensors [1,15–22]. These systems share the common trait of ''cross-reactivity'', in that a given detector responds to a variety of odorants and that a given odorant elicits a response from a variety of detectors. Pattern recognition algorithms are then used to classify and quantify odorants based on the data stream produced by the detector array. The gross behavior of such systems therefore resembles mammalian olfaction in that the system need not be designed in advance to detect a particular odorant. Instead, the systems can

* Corresponding author. Tel.: +1-626-395-6335; fax: +1-626-795-7487.
E-mail address: nslewis@caltech.edu (N.S. Lewis).

0925-4005/01/$ – see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0925-4005(00)00635-3

learn new patterns and can associate these patterns with new odorants through use of training and data storage functions.

The particular electronic nose implementation evaluated in this work is comprised of an array of conducting polymer composite vapor detectors [23,24]. Each detector is a chemically-sensitive thin film that undergoes a change in electrical resistance upon exposure to a vapor. The resistors are made from composites having regions of an electrical conductor (typically carbon black, although other conductors can also be used) interspersed into regions of an insulating organic material, typically an organic polymer. Sorption of an odorant into the polymer causes a swelling of the film, which decreases the connectivity between the regions that comprise the conducting phase of the composite. This change in connectivity modifies the electrical resistance of the detector. Removal of the stimulus leads to desorption of the odorant and a decrease in resistance back to the original baseline response value. Different polymers display different gas/polymer sorption constants so that the signals obtained from an array of detectors, each made of a compositionally different conductor/insulator composite, provide a fingerprint that can be used for identification, classification, and quantification of odorants. This particular electronic nose array implementation is advantageous because the polymers that can be used in the detectors can vary widely in their chemical and physical properties. A wide range of odorants can therefore be discriminated based on the response patterns that the odorants produce on the sensor array [23,24]. In addition, electrical resistance is an easily measured experimental quantity, and the steady state signal response signals for odorants tested to date are linearly dependent on the concentration of analyte in the gas phase above the detector [23,25]. This combination of features allows for a very simple algorithmic implementation to obtain robust odor classification and quantification under a variety of ambient conditions.

Some of the questions of interest in these artificial systems relate to how their performance, on a system-level basis, compares to that of mammalian olfaction. To make this comparison, in this work we specifically addressed how well the electronic nose, as compared to humans and monkeys, can discriminate between selected pairs of odorants. In addition, we have determined odor intensity thresholds for the current (highly unoptimized) version of the conducting polymer composite electronic nose, in order to evaluate trends in detection limits of humans relative to the detection capabilities of the electronic nose. Comparisons were only made for volatile organic vapors as opposed to aroma active odorants. These discrimination and odor sensitivity data are interesting in their own right but additionally are of importance in benchmarking the status of the electronic nose towards the long-range goal of constructing an artificial/electronic system that functionally mimics the human sense of olfaction.

2. Experimental

An automated system consisting of LabVIEW software, a Pentium computer, a Keithley channel switcher, a Keithley multimeter, and electronically-controlled solenoid valves and mass flow controllers, was used to deliver selected concentrations of solvent vapors to the detectors in the electronic nose array. The system was also used to monitor the resistance of the detectors. The apparatus has been described in detail previously [23]. The analytes were purchased from Aldrich and Pfaltz & Bauer.

To obtain the desired analyte concentration, a stream of carrier gas was passed through a bubbler that had been filled with the solvent of choice. The carrier gas for all experiments was oil-free air, obtained from the general compressed air laboratory source, containing 1.10 ± 0.15 ppth (parts per thousand) of water vapor. The air was filtered to remove particulates, but deliberately was not dehumidified nor otherwise purified. Fluctuations in laboratory temperature, $21.5 \pm 1.5^{\circ}$C, could cause a $\sim$10% error in setting and controlling the vapor concentrations between nominally identical exposures over the course of the data collection analyzed in this work. No temperature control of the apparatus or of the carbon black-polymer composite detectors was performed. Saturation of the carrier gas with the solvent vapor was verified through measurement of the rate of mass loss of the solvent in the bubbler. The mass flow controllers were verified independently to be within specification prior to and after an experimental run. In addition, calibrations of the flow system using a flame ionization detector (FID) (Model 300 HFID, California Analytical Instruments, Inc.) verified that the analyte concentrations actually delivered to the sensors were those expected from the settings of the mass flow controllers. The flame ionization detector was calibrated using toluene/air standard calibration gas mixtures sold by Scott Specialty Gases, Inc. Corrections were made for the sensitivity changes of the FID unit for toluene versus other gases when determining the concentrations of the other analytes of interest in this work.

Detector fabrication methods have been described in detail in the literature. [23,24]. All detectors evaluated in this work were made from composites of carbon black with insulating organic polymers. The detector array used in determining the electronic nose detection thresholds for the homologous series of *n*-alkane (*n*-pentane, *n*-hexane, *n*-heptane, *n*-octane and *n*-nonane) and 1-alcohol (methanol, ethanol, 1-propanol, 1-butanol, 1-pentanol) odorants consisted of 20 conducting polymer composite detectors. Two detector copies were made from each of the 10 polymers listed in Table 1 (purchased from Aldrich and Polysciences). The polymers were chosen to encompass a broad range of chemical properties, but were selected with no specific knowledge of how well the detectors would perform for the odor discrimination tasks of concern in this study. Each detector was fabricated by spin-coating mixtures containing a dissolved polymer and suspended carbon

Table 1
Polymers contained in the detectors of the electronic nose arrays

Detector	Polymer
1	Poly(4-vinyl phenol)
2	Poly(*N*-vinyl pyrrolidone)
3	Poly(sulfone)
4	Poly(methyl methacrylate)
5	Poly(caprolactone)
6	Poly(ethylene-*co*-vinyl acetate), 82% ethylene
7	Poly(ethylene oxide)
8	Poly(ethylene)
9	Poly(vinylidene fluoride)
10	Poly(ethylene glycol)

black particles onto a glass slide, as described previously [23,24].

A second detector array, made from the same polymers and using the same fabrication techniques, was used to test the ability of the electronic nose to discriminate pairwise between various odorants within series of esters (isopentyl acetate, isopentyl propionate, isopentyl butanoate, isopentyl pentanoate, isopentyl hexanoate, ethyl acetate, *n*-propyl acetate, *n*-butyl acetate, *n*-pentyl acetate, *n*-hexyl acetate, *n*-octyl acetate, *n*-decyl acetate, isopropyl acetate and isobutyl acetate), alcohols (ethanol, 1-propanol, 1-butanol, 1-pentanol, 1-hexanol, 1-heptanol, 1-octanol, 2-pentanol and 3-pentanol) and carboxylic acids (*n*-propanoic acid, *n*-butanoic acid, *n*-pentanoic acid, *n*-hexanoic acid, *n*-heptanoic acid, isobutanoic acid, isopentanoic acid and isohexanoic acid).

For detection threshold measurements, the the detector response was taken as the difference between the baseline resistance and the maximum resistance observed during the 10 min exposure to the odorant-containing air flow. This 10 min odorant exposure was always preceded by, and followed by, a 2 min exposure to a flow of air that was obtained from the general laboratory source. The resistance values were measured with a Keithley Model 2002 multimeter by integrating over a measurement aperture of 167 ms. Detection thresholds were defined as the lowest concentration at which any detector in the electronic nose array had a response with a signal to noise (s/n) ratio of 3. Thus, at the detection threshold an odorant could be detected but would generally not be identifiable without use of a sophisticated signal processing algorithm, due to the lack of significant responses from multiple detectors at the S/N level of 3:1. This definition of the detection threshold was employed previously in a study of surface acoustic wave-based vapor detectors [26]. The noise of a given detector was defined as the standard deviation in the residuals about a nine-point moving average of the baseline resistance values, spanning approximately 35 s. Using conventional procedures as described in analytical chemistry textbooks [27], the expected detection thresholds were then estimated based on the noise levels typical of the better detectors (i.e. slightly less than 10 ppm for the detectors containing poly(ethylene oxide), poly(ethylene) and poly(ethylene glycol), which had baseline resistances of approximately 1.3×10^4, 1.2×10^4 and $4.0 \times 10^3\ \Omega$, respectively over the course of the experiment) in combination with the knowledge of the signal amplitude at a known, relatively high concentration of odorant, and prior reports that the steady state detector response values generally vary linearly as a function of odorant concentration [23,25,28]. To verify these detection level estimates, each of the 10 odorants was then delivered, at three separate concentrations which generall spanned approximately an order of magnitude near the expected detection threshold. For each odorant, the largest signal to noise ratio of any of the detectors, at the lowest concentration where a signal was detected, was used to determine the detection threshold by linearly extrapolating down to a signal to noise ratio of 3. The extrapolations were never larger than a factor of 5 and were typically less than a factor of 3.

For quantifying odor discrimination, the electronic nose was exposed 10 times to each of the odorants to be discriminated. The order was randomized within subsets of a maximum of eight odorants, because the automated odorant delivery apparatus could handle only eight odorants at a time. The experimental protocol for each exposure was 5 min of clean air flow, followed by 5 min of air flow containing the odorant at a partial pressure corresponding to 1% of its vapor pressure, followed by another 5 min of clean air flow. Only the steady state signal response of the detectors was used in evaluating the discrimination capability of the electronic nose. Use of these values reflects only the relative partition coefficients of the analytes into the various polymer films of the detectors, and does not utilize kinetic sorption information in classifying odorants. In addition, use of the steady state response value largely eliminates variation in response properties due to fabrication-related parameters such as inhomogenieties in detector film thickness and/or variability in the loading of the carbon black particles in the polymeric detector films [28]. The time required to reach this steady state response value depended on the type of polymer and type of analyte, and on the thickness of each polymer in the detector array. In addition, in the most rapidly responding cases the response time depended on the mass transport properties of analyte from the bubbler source through the delivery system to the vicinity of the detectors. Response times for many detectors were below 5 s, although some polymer/analyte combinations required longer times to reach their steady state values. The steady state response data were processed to extract the relative differential resistance response signals, as described previously [23–25] upon exposure to the test vapors of interest.

The pairwise odor discrimination performance of the electronic nose was evaluated using the Fisher linear discriminant algorithm [23–25]. A resolution factor (rf) for any solvent pair can be obtained along any vector, $\vec{w}$, from the vector projection onto $\vec{w}$ of the distance between the cluster

centroids, $d_{\vec{w}}$, divided by the sum of the projected standard deviations, $\sigma_{a,\vec{w}}$, and $\sigma_{b,\vec{w}}$, for data arising from repeated exposures for two different analytes, *a* and *b*. The resulting numerical resolution factor along $\vec{w}$ is defined as

$$\mathrm{rf} = \frac{d_{\vec{w}}}{\sqrt{\sigma^2_{a,\vec{w}} + \sigma^2_{b,\vec{w}}}}$$

The Fisher linear discriminant operates by searching for the vector, $\vec{w}$, such that the rf value is maximized along this optimal discriminant vector. Assuming a Gaussian distribution relative to the mean value of the data points in a given cluster, the probabilities of correctly identifying an analyte as *a* or *b* are approximately 72, 92 and 98% from a single presentation when analytes *a* and *b* are separated with resolution factors of 1.0, 2.0, or 3.0, respectively. Data extracted from multiple exposures of an analyte estimate the statistical distributions of the clustered data.

Because the 20-detector array contained two copies of nominally the same 10-detector array, random combinations of 10-detectors (constrained to each contain exactly one of each of the 10 polymer types) were evaluated to obtain a measure of variability in the ability of a 10-detector electronic nose to discriminate odorants. For each randomly selected 10-detector array, the response signals were normalized and then the Fisher linear discriminant method was used to quantify the ability of the electronic nose to distinguish the odorants [23]. The rf for resolving a given odorant pair was then taken as the mean rf over the results from 10 randomly selected 10-detector arrays.

3. Results and discussion

3.1. Determination of detection thresholds for the electronic nose and comparison with human olfactory threshold levels

3.1.1. Determination of detection thresholds for the electronic nose

The detection thresholds obtained for the electronic nose upon exposure to the homologous series of alkanes (*n*-pentane to *n*-nonane) and alcohols (methanol to 1-pentanol) are shown in Fig. 1a and b. Within each homologous series of odorants, the electronic nose detection threshold values decreased with decreasing vapor pressure of the odorant. This behavior is in accord with prior work, which showed that lower vapor pressure odorants have larger mean partition coefficients for sorption into polymer composite detector films [29]. In the case of the homologous series of *n*-alkanes, detectors containing poly(ethylene oxide) and poly(ethylene) exhibited similar signal to noise ratios and were the detectors that defined the detection threshold of the array towards these odorants. In contrast, a poly(ethylene glycol)-based detector exhibited the largest signal to noise ratios for methanol, ethanol and 1-propanol while a

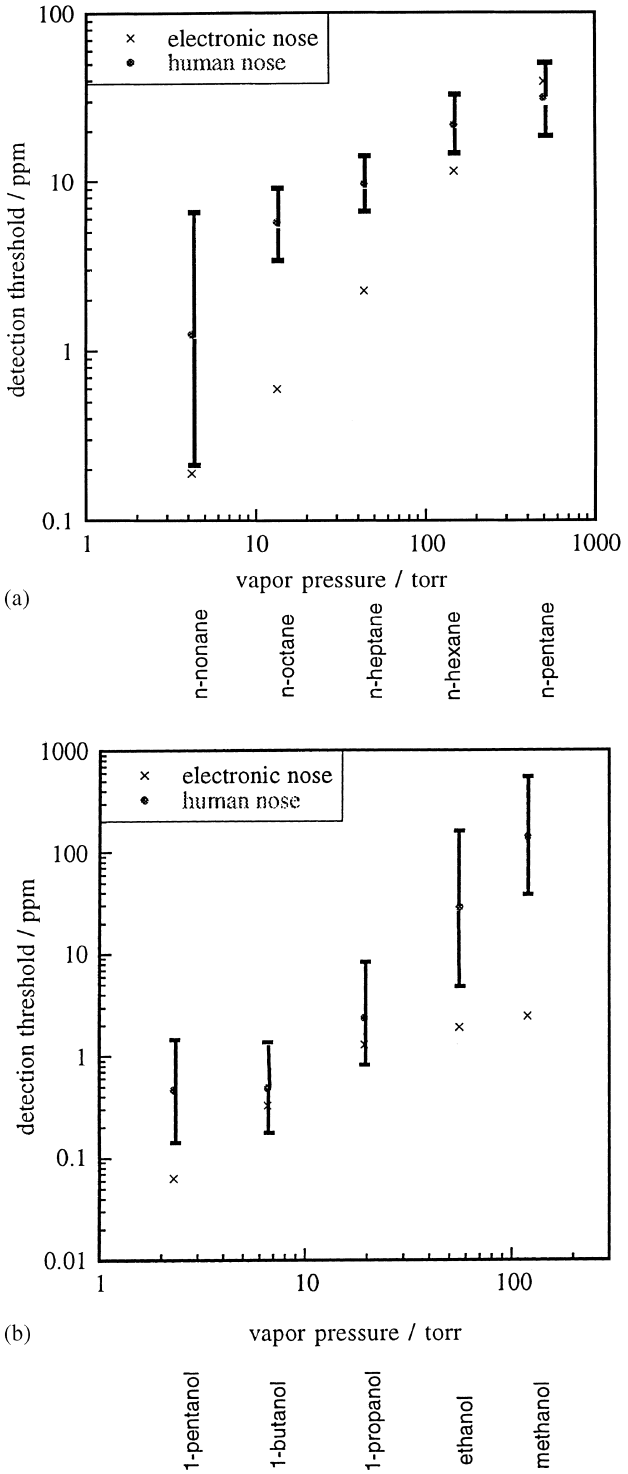


Fig. 1. Plots of the detection threshold in parts per million (ppm) of odorant in air, for the average human [30] and the electronic nose, vs. odorant vapor pressure for homologous series of (a) alkanes from *n*-pentane through *n*-nonane and (b) alcohols from methanol through 1-pentanol. The detection threshold data for humans are mean values for many humans and also are averaged over the results of many investigators as reported in the literature. Estimates of the standard deviations of the reported mean detection values for a given odorant are shown. The standard deviations for a particular set of electronic nose detectors upon repeated exposures to a given concentration of analyte are approximately the height of the "X" symbols depicted in the figures.

poly(ethylene oxide)-based detector exhibited the largest signal to noise ratios for 1-butanol and 1-pentanol. The specific detector that allowed the array to detect the lowest concentration of a given odorant, defining the array detection threshold, thus varied somewhat depending on the affinity of the detectors for the odorant class of concern.

3.2. Comparison of detection thresholds for human and electronic noses

Data for human olfactory threshold values in detecting members of the homologous series of *n*-alkanes and 1-alcohols were obtained from the literature [30]. The values plotted in Fig. 1 represent mean detection thresholds as averaged over the work of many authors in testing many humans [30]. The averaging over many studies was performed in an attempt to obtain a mean detection threshold value without bias towards, or away from, the methodology of any particular investigator or with respect to any particular subgroup of humans that had been exposed to the odorants of interest during the course of a single study. The human detection threshold is defined as the lowest concentration of an odorant that an average human is able to detect relative to background air, without necessarily classifying the odorant type. This measure of odor detection sensitivity is consistent with the definition employed for the electronic nose. In fact, it can be postulated that in the case of humans, odorant detection at threshold likely involves the brain receiving a signal from only the olfactory receptor or receptors most strongly responsive to a given odorant. This is also consistent with the electronic nose detection threshold, as defined above. Comparisons were only made for volatile organic vapors as opposed to aroma active odorant vapors.

The data of Fig. 1 indicate that for the odorants studied in this work, the electronic nose detection thresholds are typically lower than the mean human detection thresholds, with the sole exception being *n*-pentane, for which the mean human detection threshold is slightly lower detection threshold than the electronic nose. Typically, the electronic nose detection thresholds are lower than mean human detection thresholds by a factor of 2–10, but of course the variance between human subjects and between human trials is such that the best performing humans in certain trials might well outperform the particular electronic nose implementation utilized in this work under the test conditions and detector configurations employed herein. In terms of thermodynamic activities, the fraction of a given odorant's vapor pressure (i.e. the partial pressure of the odorant, P, relative to its vapor pressure, P^0) at which the detection threshold lies for the electronic nose is in the range of 2.0×10^{-5}–6.0×10^{-5} for the odorants studied herein. For comparison, for the same set of test odorants the minimum fraction of an odorant's vapor pressure that is detectable by the average human is in the range of 4.9×10^{-5}–9.0×10^{-4}.

The underlying physicochemical principles for the odor detection threshold trend have recently been elucidated for certain detector compositions of the polymer composite electronic nose [31,32], but the analogous principles are not fully defined for the human olfactory system. Carbon black/organic-polymer composite electronic nose detectors exhibit similar steady state relative differential resistance response levels when exposed to most odorants at a constant value of P/P^0, due to the thermodynamic relationships between the chemical potential of the analyte in the gas phase and the mole fraction of analyte that will sorb into the polymeric phase under equilibrium conditions [29]. Thus, the electronic nose would be expected to exhibit similar detection thresholds for the tested odorants in terms of the minimum detectable fraction of vapor pressure (i.e. $P/P^0 \sim 4.0 \times 10^{-5}$), and this behavior was in fact observed experimentally herein. Humans also exhibit a trend in mean detection threshold of these odorants that tracks with the vapor pressure of the odorant, presumably reflecting at least in part the tendency of the odorant to sorb into the mucous. For the compounds evaluated in this study, the mean human odor detection threshold value is somewhat higher than the detection threshold of the current conducting polymer composite electronic nose. Given the standard deviations of the detection thresholds over various trials and groups of humans, there is some error in the absolute comparison of these two types of data sets, but clearly the data indicate that the detection limits of the human and electronic nose systems are comparable for the odorants investigated in this work.

We note, however, two significant classes of exceptions, notably the cases of thiols and biogenic amines, for which human odor detection thresholds are several orders of magnitude lower than those of the electronic nose. This suggests the presence of either highly specific receptor types and/or highly specific processing algorithms in the human olfactory system towards these odorants. In addition, the human nose is highly sensitive to certain pyrazines, thiazoles, to some aldehydes, and to many other compounds that have aroma activity as opposed to merely being volatile organic compounds. Although we have not investigated the detection threshold behavior of the electronic nose to these compounds as part of the present study, we expect that the current implementation of the conducting polymer composite electronic nose will exhibit detection thresholds for these odorants that are significantly higher than the mean human olfactory detection thresholds for these classes of analytes.

3.3. Pairwise odorant discrimination abilities of the electronic nose versus pairwise odor discrimination abilities of humans and monkeys

3.3.1. Odorant discrimination ability of the electronic nose

Assessing the discrimination ability of the electronic nose requires a choice of a data analysis algorithm. We have

chosen to use a statistical approach, in which a hyperplanar decision boundary is drawn in the *n*-dimensional space that encompasses the steady state resistance response data from the array of *n*-detectors responding to the presence of an odorant. The electronic nose containing 10-detector types was able to discriminate pairwise between all tested odorants with a resolution factor of at least 3.7. The minimum value was observed for the case of *n*-hexanoic acid versus isohexanoic acid, while the median rf across all tested odorant pairs was 29. Assuming that the statistical distributions of the collected data samples are representative of the actual statistical distributions, an rf of ∼3.0 corresponds to a probability of ∼98% of correctly identifying an odorant as *a* or *b* as a result of a single presentation. Hence, the electronic nose can easily discriminate between all of the test odorant pairs evaluated in this study under the experimental conditions employed in our test protocol.

3.3.2. Comparison of the trends in discrimination abilities between the conducting polymer composite electronic nose and mammalian olfaction

Data assessing the abilities of monkeys (*Saimiri sciureus*) and humans to discriminate pairwise between various ester, alcohol and carboxylic acid odorants were obtained from Laska et al. [33–35]. In those experiments, an average probability of correctly distinguishing between a given pair of odorants was determined by averaging across multiple trials and multiple test subjects. The odorants were presented at a perceived intensity-matched vapor concentration that was arrived at via 1:100 dilution of the pure odorants with diethyl phthalate.

Because the electronic nose exhibited >99% statistical probability of correctly distinguishing between members of every tested analyte pair in a single presentation of analyte to the detector array, depicting the discrimination performance of humans and/or primates versus the discrimination performance of the electronic nose on a common, linear axis would yield essentially no information concerning the trends in electronic nose performance as a function of chemical differences between pairs of odorants. Thus, the measure of distinguishing ability for the electronic nose was taken as the resolution factor, which scales with the separation of the clustered odorant responses relative to the widths of the clusters in detector space, and which is a statistically meaningful quantity based on the variation of the data for repeated exposures of the detector array to analyte pairs of interest. The resolution factor values can be converted into percentage of correct decision rates using conventional statistical properties of Gaussian distributions of data relative to the separation of the mean values of the data sets. The measure of distinguishing ability for the humans and monkeys was taken as the percentage of correct decisions in the task of distinguishing between members of an odorant pair. The experiments with human subjects utilized a forced-choice triangular test, so that the probability of a correct decision by random chance was 33%, whereas random chance for the monkey experiments would produce a 50% correct decision rate. Figs. 2–4 display the data for odorant discrimination of isopentyl acetate from a series of other esters (Fig. 2), *n*-pentanoic acid from a series of other carboxylic acids (Fig. 3), and *n*-pentanol from a series of other alcohols (Fig. 4).

Several correlations are apparent between trends in the discrimination properties of the electronic nose, the monkey olfactory system, and the human olfactory system. In most cases, as the odorants become more chemically dissimilar, the members of an odorant pair become easier to discriminate for the electronic nose, monkeys and humans. For example, in Fig. 2a, the task of discriminating between isopentyl acetate versus isopentyl propionate, isopentyl butanoate, isopentyl pentanoate, and finally isopentyl hexanoate, becomes progressively easier for all three olfactory systems. This makes sense because the odorant molecules in a pair are becoming progressively more dissimilar structurally. Similarly, in Fig. 3b, *n*-pentanoic acid is more difficult for each of the three olfactory systems to discriminate from isopentanoic acid than from either of isobutanoic acid or isohexanoic acid.

There are also subtle differences in the discrimination trends. For example, in discriminating *n*-pentanoic acid from each of *n*-propanoic acid, *n*-butanoic acid, *n*-hexanoic acid and *n*-heptanoic acid (Fig. 3a), the electronic nose has more difficulty in discriminating *n*-pentanoic acid from the longer chain acids relative to the shorter chain acids. Conversely, both mammals can more easily discriminate *n*-pentanoic acid from *n*-hexanoic acid than from *n*-butanoic acid. A similar observation can be made from Fig. 4, which indicates that monkeys more easily discriminate 1-pentanol from 1-heptanol and 1-octanol than from 1-propanol and ethanol, while the opposite is true for the electronic nose. Typically, the chemical difference between consecutive molecules in a homologous series decreases with the addition of each additional carbon atom. Thus, the relative difficulty of the electronic nose in discriminating between longer chain molecules versus shorter chain molecules, with molecules in each task differing by the same number of carbon atoms, is readily understandable. The opposite observation in specific instances for mammals could possibly indicate odorant receptor proteins having geometrically defined binding sites that more effectively differentiate between certain odorant geometries than do the polymeric detectors used in the current implementation of the electronic nose. Observations confirming the differential responses of an olfactory receptor neuron in vitro as a function of odorant chain length have recently been published [9].

It is also interesting to note that for the pairs of odorants evaluated in this work, the quantitative discrimination performance of the conducting polymer composite electronic nose is significantly higher than that of the monkey or human olfactory system (cf. Figs. 2–4). Interestingly, the data of Buck and co-workers, who determined Ca^{2+}(aq) production rates for a collection of olfactory receptors in response to a

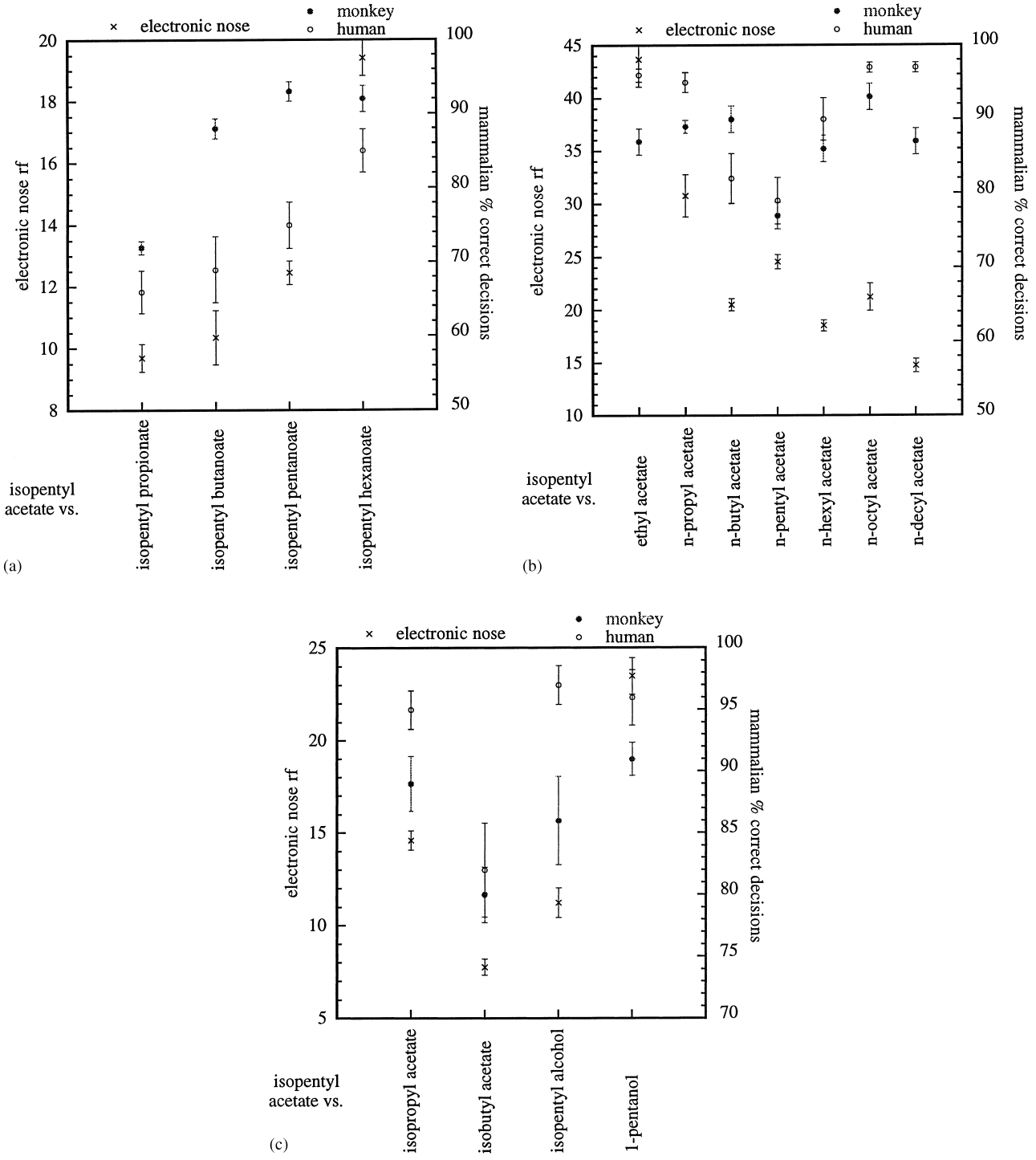
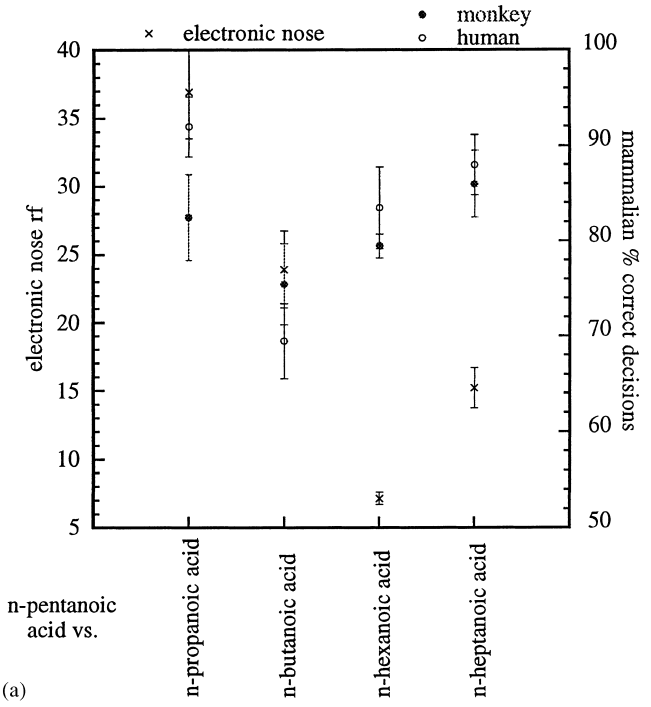
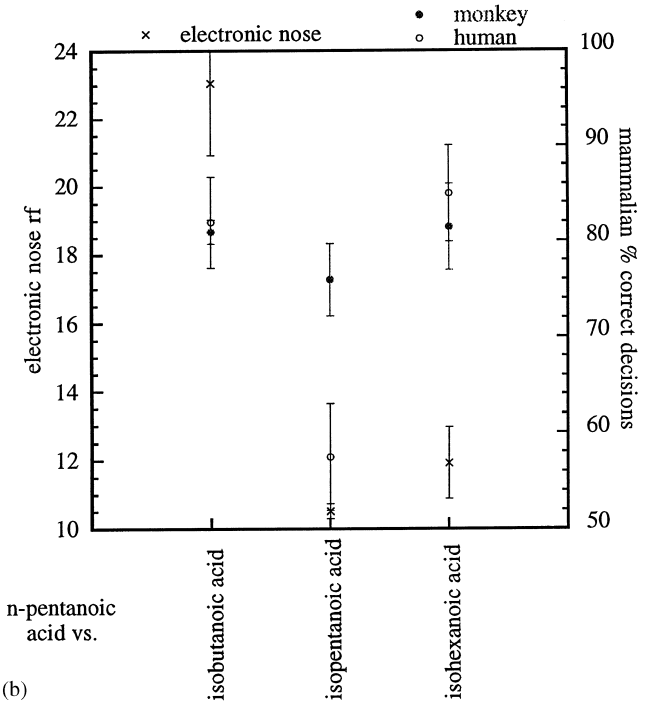


Fig. 2. Plots of trends in the abilities of the electronic nose, monkeys and humans to discriminate isopentyl acetate from other ester odorants (a,b) and to discriminate isopentyl acetate from isopentyl alcohol and 1-pentanol (c). The electronic nose data are plotted with reference to the left scale and the data points represent a mean resolution factor (rf). The monkey and human data are plotted with reference to the ordinate scale on the right of the figure and represent the mean probability of correctly discriminating between each specific pair of odorants averaged over 20 humans and over 5 monkeys, as experimentally determined by Laska [35]. The error bars represent 1 standard deviation unit of confidence in the mean values.

series of alcohols and carboxylic acids, show distinctly different receptor cell Ca^{2+} response patterns even for closely structurally-related compounds such as 1-octanol and 1-nonanol [9]. Reference to Fig. 4 indicates that humans and monkeys cannot effectively discriminate 1-pentanol and 1-hexanol nearly as well as their discrimination performance on more chemically dissimilar odorants. More chemically dissimilar receptor firing patterns might be expected for 1-pentanol relative to 1-hexanol than for 1-octanol relative to 1-nonanol because 1-pentanol is more chemically dissimilar



(a)



(b)

Fig. 3. Plots of trends in the abilities of the electronic nose, monkeys and humans to discriminate between *n*-pentanoic acid and two series of other carboxylic acid odorants. The electronic nose data are plotted with reference to the ordinate scale on the left of the figure, and the data points represent a mean resolution factor (rf). The monkey and human data are plotted with reference to the ordinate scale on the right of the figure and represent the mean probability of correctly discriminating between each specific pair of odorants averaged over 10 humans and over 4 monkeys, as experimentally determined by Laska [34]. The error bars represent 1 standard deviation unit of confidence in the mean values.

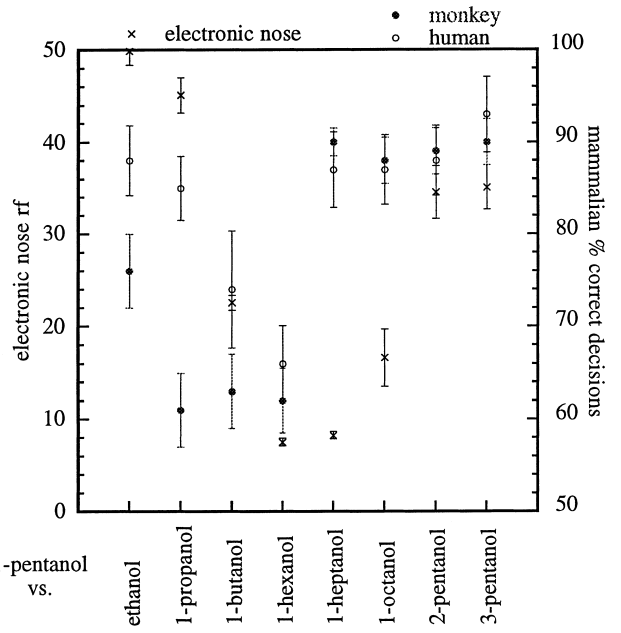


Fig. 4. Plots of trends in the abilities of the electronic nose, monkeys and humans to discriminate between 1-pentanol and a series of other alcohol odorants. The electronic nose data are plotted with reference to the ordinate scale on the left of the figure, and the data points represent a mean resolution factor (rf). The monkey and human data are plotted with reference to the ordinate scale on the right of the figure, and represent the mean probability of correctly discriminating between each specific pair of odorants averaged over 10 humans and over 4 monkeys, as experimentally determined by Laska et al. [33]. The error bars represent 1 standard deviation unit of confidence in the mean values.

to 1–hexanol than is 1-octanol relative to 1-nonanol. If this expectation is found to be valid experimentally, it would imply that the processing in the olfactory system of humans and monkeys results in a reduced ability to discriminate odors relative to the inherent differences in data content arising from the receptor response pattern produced by the outputs of the receptors in the olfactory bulb.

Several other factors could potentially confound the comparisons made above between the discrimination characteristics of the biological and electronic olfaction system for the compounds of interest in this work. First, the electronic nose measurements were collected for analytes in a background of laboratory air, while the human and primate discrimination data were collected for analytes diluted in diethyl phthlate. The human and primate data were also collected using higher background humidity levels than those in the air background used in the electronic nose discrimination measurements. The steady state relative differential resistance signals from carbon black-polymer composite detectors exposed to a wide series of test odorants have been shown to be essentially independent of whether the background ambient is pure air or is instead air with other analytes present [25]. Thus, introduction of diethyl phthlate as a background ambient component into the electronic nose measurements, and/or an increase in humid-

ity in the background carrier gas, would not be expected to produce a significant difference in the electronic nose discrimination performance relative to the data reported herein, and therefore, is not expected to significantly affect the comparisons in discrimination abilities discussed above. In addition, although the steady state response time for some of the electronic nose detectors was longer than the detection times used to evaluate the performance of the human system, the response time of the electronic nose detectors was not used as a classifying feature of the data, and only steady state values were utilized in the data analysis. The response times of the electronic nose detectors are a strong function of the system design, the delivery system characteristics, the polymer film thickness in the detectors, the detector readout electronics, and other properties which could be readily optimized if one wanted to reach steady state values for all detectors in a 1–5 s timescale. We have therefore used only the more fundamental, steady state response values of the electronic nose (which are ultimately not system configuration dependent) in our comparisons with the discrimination characteristics exhibited by the mammalian olfactory system.

An additional factor to be considered in comparing biologically-based olfaction to that exhibited by artificial devices is that the human olfactory system is non-linear in its response with respect to the concentration of analyte and/or analyte mixtures [36] whereas, the electronic nose signals from an individual detector are a linear function of the concentration of analyte in the gas phase, at least for partial pressures of odorants up to 5% of an analyte's vapor pressure at room temperature [25]. The discrimination metric for the electronic nose between an analyte pair is thus not a strong function of the analyte concentration (except as the variance between the classes of data is increased when the noise floor is closely approached). This behavior is not true for the biological system, however, in that discriminability can be a strong function of odorant concentration. Thus, for odorants whose perceived olfactory odor quality changes as the odor intensity changes, precise comparisons between discriminability of odorant pairs of the human and electronic nose will be dependent on the precise conditions under which the comparisons are made. In this work, we have chosen to make the comparison for odors from literature data that have been collected at a commonly perceived odorant intensity level for humans and monkeys, and for which electronic nose data have been obtained at a common thermodynamic activity level for each analyte. Furthermore, the activities of analyte in the gas phase were comparable between the human and primate studies and those used in the electronic nose performance evaluation. Information on the dependence of the mammalian olfactory discriminability of odorant pairs as a function of odorant concentration will be very valuable in making more robust comparisons between biological and artificial systems and ultimately in constructing an artificial device that mimics functionally certain aspects of human olfaction.

4. Conclusions

The results presented herein show some interesting similarities between the system-level performance features of the electronic nose and those of mammalian olfaction. Specifically, for the volatile organic, non-aroma active odorants evaluated in this work, the absolute detection thresholds and trends in the detection thresholds of the conducting polymer composite electronic nose are similar to trends in mean detection thresholds of humans. The electronic nose can typically detect a minimum odorant partial pressure of approximately 4×10^{-5} of the odorant's vapor pressure, which is comparable to the mean human detection thresholds reported for the particular odorants evaluated in this study. Aroma active compounds would be detectable at lower concentration levels by the human olfactory system than by the existing electronic nose implementation. Some similarities were also found to exist in odorant discrimination abilities between the systems, in that odorants typically become easier to discriminate pairwise for primates, humans, and the electronic nose as the members of an odorant pair become more chemically dissimilar structurally. The quantitative, statistically based discrimination performance of the conducting polymer composite electronic nose was significantly higher than that of the monkey or human olfactory system for the particular pairs of odorants that were evaluated in this study. The quantitative comparisons between the electronic nose and mammalian olfaction provided in this study are an initial contribution towards the ambitious goal of designing an electronic analogue to the mammalian olfactory sense.

Acknowledgements

We sincerely thank Dr. Matthias Laska and his co-workers from the Department of Medical Psychology at the University of Munich Medical School for providing us with their data on the odorant discriminating ability of humans and monkeys. We thank NASA, the Army Research Office and DARPA for their support of this work, with primary support under a MURI grant from the Army Research Office. B.J.D. acknowledges the Government of Canada for an NSERC 1967 Centennial Graduate Fellowship.

References

[1] R. Axel, The molecular logic of smell, Sci. Am. 273 (1995) 154–159.
[2] H. Breer, I. Wanner, J. Strotmann, Molecular genetics of mammalian olfaction, Behav. Genet. 26 (1996) 209–219.
[3] D. Lancet, Vertebrate olfactory reception, Ann. Rev. Neurosci. 9 (1987) 329–355.
[4] D. Lancet, U. Pace, The molecular basis of odor recognition, Trends Biochem. Sci. 12 (1987) 63–66.
[5] D. Lancet, N. Ben-Arie, Olfactory receptors, Curr. Biol. 3 (1993) 668–674.

[6] K. Mori, Y. Yoshihara, Molecular recognition and olfactory processing in the mammalian olfactory system, Prog. Neurobiol. 45 (1995) 585–619.
[7] K. Mori, H. Nagao, Y. Yoshihara, The olfactory bulb: coding and processing of odor molecule information, Science 286 (1999) 711–715.
[8] R. Reed, Signaling pathways in odorant detection, Neuron 8 (1992) 205–209.
[9] B. Malnic, J. Hirono, T. Sato, L.B. Buck, Combinatorial receptor codes for odors, Cell 96 (1999) 713–723.
[10] K.A. Hamilton, J.S. Kauer, Intracellular potentials of salamander mitral/tufted neurons in response to odor stimulation, Brain Res. 338 (1985) 181–185.
[11] J.G. Hildebrand, G.M. Shepherd, Mechanisms of olfactory discrimination: converging evidence for common principles across phyla, Annu. Rev. Neurosci. 20 (1997) 595–631.
[12] J. Kauer, Contributions of topography and parallel processing to odor coding in the vertebrate olfactory pathway, Trends Neurosci. 14 (1991) 79–85.
[13] J. White, K.A. Hamilton, S.R. Neff, J.S. Kauer, Emergent properties of odor information coding in a representational model of the salamander olfactory bulb, J. Neurosci. 12 (1992) 1772–1780.
[14] K. Mori, N. Mataga, K. Imamura, Differential specificities of single mitral cells in rabbit olfactory bulb for a homologous series of fatty-acid odor molecules, J. Neurophys. 67 (1992) 786–789.
[15] K.J. Albert, N.S. Lewis, C.L. Schauer, G.A. Sotzing, S.E. Stitzel, T.P. Vaid, D.R. Walt, Cross-reactive chemical sensor arrays, Chem. Rev. 100 (2000) 2595–2626.
[16] J.W. Gardner, P.N. Bartlett, A brief-history of electronic noses, Sens. Actuators B 18 (1994) 211–220.
[17] J.W. Gardner, P.N. Bartlett, Electronic Noses: Principles and Applications, Oxford University Press, Oxford, 1999, pp. 264.
[18] J.W. Gardner, H.V. Shurmer, P. Corcoran, Integrated tin oxide odor sensors, Sens. Actuators B 4 (1991) 117–121.
[19] K. Persaud, G. Dodd, Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose, Nature 299 (1982) 352–355.
[20] H.V. Shurmer, An electronic nose — a sensitive and discriminating substitute for a mammalian olfactory system, IEEE Proc. — G Circ. Dev. Syst. 137 (1990) 197–204.
[21] J. White, T.A. Dickinson, D.R. Walt, J.S. Kauer, An olfactory neuronal network for vapor recognition in an artificial nose, Biol. Cybernetics 78 (1998) 245–251.
[22] D.R. Walt, T. Dickinson, J. White, J. Kauer, S. Johnson, H. Engelhardt, J. Sutter, P. Jurs, Optical sensor arrays for odor recognition, Biosens. Bioelectron. 13 (1998) 697–699.
[23] B.J. Doleman, M.C. Lonergan, E.J. Severin, T.P. Vaid, N.S. Lewis, Quantitative study of the resolving power of arrays of carbon black-polymer composites in various vapor sensing tasks, Anal. Chem. 70 (1998) 4177–4190.
[24] B.J. Doleman, R.D. Sanner, E.J. Severin, R.H. Grubbs, N.S. Lewis, Use of compatible polymer blends to fabricate arrays of carbon black-polymer composite vapor detectors, Anal. Chem. 70 (1998) 2560–2654.
[25] E.J. Severin, B.J. Doleman, N.S. Lewis, An investigation of the concentration dependence and response to analyte mixtures of carbon black-insulating organic polymer composite vapor detectors, Anal. Chem. 72 (2000) 658–668.
[26] E.T. Zellers, J. Park, T. Hsu, W.A. Groves, Establishing a limit of recognition for a vapor sensor array, Anal. Chem. 70 (1998) 4191–4201.
[27] D.A. Skoog, D.M. West, Fundamentals of Analytical Chemistry, CBS College Publishing, New York, 1982, p. 67.
[28] M.C. Lonergan, E.J. Severin, B.J. Doleman, S.A. Beaber, R.H. Grubbs, N.S. Lewis, Array-based vapor-sensing using chemically-sensitive, carbon black-polymer resistors, Chem. Mater. 8 (1996) 2298–2312.
[29] B.J. Doleman, E.J. Severin, N.S. Lewis, Trends in odor intensity for humans and electronic noses: relative roles of odorant vapor pressure versus molecularly specific odorant binding, Proc. Natl. Acad. Sci. U. S. A 95 (1998) 5442–5447.
[30] M. Devos, F. Patte, J. Rouault, P. Laffort, L.J. Van Gemert, Standardized Human Olfactory Thresholds, Oxford University Press, New York, 1990.
[31] M.J. Swann, A. Glidle, L. Cui, J.R. Barker, J.M. Cooper, The determination of gaseous molecular density using a hybrid vapour sensor, Chem. Commun. (1998) 2753–2754.
[32] E.J. Severin, N.S. Lewis, Relationships among resonant frequency changes on a coated quartz crystal microbalance, thickness changes, and resistance responses of polymer-carbon black composite chemiresistors, Anal. Chem. 72 (2000) 2008–2015.
[33] M. Laska, P. Teubner, Olfactory discrimination ability for homologous series of aliphatic alcohols and aldehydes, Chem. Senses 24 (1999) 263–270.
[34] M. Laska, P. Teubner, Odor structure-activity relationships of carboxylic acids correspond between squirrel monkeys and humans, Am. J. Physiol.-Reg. I. 43 (1998) R1639–R1645.
[35] M. Laska, D. Freyer, Olfactory discrimination ability for aliphatic esters in squirrel monkeys and humans, Chem. Senses 22 (1997) 457–465.
[36] F.T. Schiet, W.S. Cain, Odor intensity of mixed and unmixed stimuli under environmentally realistic conditions, Perception 19 (1990) 123–132.

GEOPHYSICS, VOL. 66, NO. 4 (JULY-AUGUST 2001); P. 1258–1266, 7 FIGS., 2 TABLES.

The application of the perfectly matched layer in numerical modeling of wave propagation in poroelastic media

Y. Q. Zeng, J. Q. He, and Q. H. Liu*

ABSTRACT

The perfectly matched layer (PML) was first introduced by Berenger as a material absorbing boundary condition (ABC) for electromagnetic waves. In this paper, a method is developed to extend the perfectly matched layer to simulating seismic wave propagation in poroelastic media. This nonphysical material is used at the computational edge of a finite-difference algorithm as an ABC to truncate unbounded media. The incorporation of PML in Biot's equations is different from other PML applications in that an additional term involving convolution between displacement and a loss coefficient in the PML region is required. Numerical results show that the PML ABC attenuates the outgoing waves effectively.

INTRODUCTION

Elastic wave propagation in a poroelastic medium is governed by Biot's theory (Biot, 1956a, b, 1962a, b). In a homogeneous medium, analytical solutions for Biot's equations can be obtained easily, but such solutions are in general impossible for an arbitrary heterogeneous medium. Elastic wave propagation in complex poroelastic media is of great interest in geophysics and other branches of applied sciences, such as petroleum engineering, structure mechanics, and seismology. Numerical methods are usually employed to solve these problems. Examples of numerical methods include the finite-difference, finite element, and pseudospectral methods. In this paper, a finite-difference method will be used.

In numerical simulation of wave propagation in an unbounded medium, the imposition of artificial boundaries introduces spurious reflections which will affect the accuracy of numerical solutions. Although the problem can be overcome by increasing the size of the model, it is not always feasible because of the large amount of computer memory required for long-time solutions. It is thus highly desirable to eliminate these reflections by using absorbing boundary conditions (ABCs) when simulating an unbounded medium. Developing an effective and stable ABC is always one of the most important tasks in numerical modeling of wave propagation. Smith (1974) proposed an ABC for finite-difference and finite element methods. In his method, the Dirichlet and Neumann conditions are used alternatively, and the solutions from these two conditions are superimposed. Although easy to implement, this method greatly increases the computation time. A widely used ABC in seismic modeling proposed by Clayton and Engquist (1977) is the one-way wave equation based on the paraxial approximations of the acoustic or elastic equations. Similar approaches were proposed by several authors, including Reynolds (1978). Although effective for small incidence angles, these absorbing boundary conditions degrade for large angles of incidence. Furthermore, they are known to have instability problem when the Poisson's ratio is greater than 2.0 (Mahrer, 1986; Stacey, 1988). Another approach is to add a spatial filter or damping taper to the boundaries (Cerjan et al., 1985; Kosloff and Kosloff, 1986). In this so-called sponge absorber method, the transition zone from the inner region to the outer boundary should be thick and smooth. Liao et al. (1984) developed an ABC based on the principle of plane wave and interpolation. Although effective even for perpendicular inhomogeneous boundary intersecting the absorbing boundary, Liao's ABC requires double precision for stability.

In 1994, Berenger proposed a highly effective perfectly matched layer (PML) as an absorbing boundary condition for electromagnetic waves. It has since been widely used for finite-difference and finite element methods (e.g., Chew and Weedon, 1994; Liu, 1997). Chew and Liu (1996) first proved that such a perfectly matched layer also exists for elastic waves in spite of the coupling of *S*- and *P*-waves at an elastic interface. In the continuous limit, the PML has zero reflection to the regular elastic medium, although a small reflection can result from discretization in the PML scheme. Hastings et al. (1996) independently implemented the PML ABC for two-dimensional problems using potentials. The PML ABC has also been extended to model acoustic waves and electromagnetic waves in

Manuscript received by the Editor August 6, 1999; revised manuscript received November 30, 2000.
*Duke University, Department of Electrical Engineering, 130 Hudson Hall, Box 90291, Durham, North Carolina 27708-0291. E-mail: yz3@ee.duke.edu; Qing.Liu@ee.duke.edu; jhe@ee.duke.edu.
© 2001 Society of Exploration Geophysicists. All rights reserved.

lossy media (Liu, 1997; Liu and Tao, 1997) as well as electromagnetic and elastic waves in cylindrical and spherical coordinates (Liu, 1999; He and Liu, 1999).

These previous ABCs, however, have been developed for electromagnetic, acoustic, and elastic waves in solids. Little attention has been paid to elastic waves in porous media. In this work, we extend the PML to truncate unbounded poroelastic media for numerical solutions using a finite-difference method. We adopt the method of complex coordinates (Chew and Weedon, 1994; Liu, 1997; Liu and Tao, 1997) to formulate the PML for poroelastic media. Two-dimensional finite-difference results show the efficiency of the PML ABC.

PROBLEM FORMULATION

Biot's theory for elastic waves in porous media was established on a macroscopic level (Biot, 1956a, b, 1962a, b). The anelastic effects arise from viscous interactions between the fluid and the solid matrix. The following assumptions are used in the theory: (1) seismic wavelength is large in comparison to the pore size, (2) the deformations are small, (3) the liquid phase is continuous, such that pores are connected and the disconnected pores are part of the matrix, (4) the solid matrix is elastic, (5) the medium is statistically isotropic, and (6) gravity forces are neglected. At low frequency, wave propagation in a heterogeneous, porous medium is described by following equations:

$$2\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) + \frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) = \frac{\partial^2}{\partial t^2}(\rho u_i + \rho_f w_i), \tag{1}$$

$$\frac{\partial}{\partial x_i}(\alpha M e - M\xi) = \frac{\partial^2}{\partial t^2}(\rho_f u_i + m w_i) + \frac{\eta}{\kappa}\frac{\partial w_i}{\partial t}, \tag{2}$$

where $m = a\rho_f/\phi$ and

$$M = \left(\frac{\phi}{K_f} + \frac{\alpha-\phi}{K_s}\right)^{-1},$$
$$\alpha = 1 - \frac{K_b}{K_s}.$$

In these equations, the parameters describing the physical properties of the medium are as follows: μ = shear modulus of dry porous matrix, λ_c = Lame constant of saturated matrix, ϕ = porosity, η = viscosity, κ = permeability, ρ = the overall density of the saturated medium determined by $\rho_f\phi + (1-\phi)\rho_s$, ρ_s = density of solid material, ρ_f = density of fluid, a = tortuosity of the matrix, K_s = bulk modulus of the solid, K_f = bulk modulus of the fluid, and K_b = bulk modulus of the dry porous frame.

In equations (1) and (2), $i,\ j = 1, 2, 3$; u_i is the ith component of the displacement vector of the solid material and $w_i = \phi(U_i - u_i)$ is the ith component of the displacement vector of the pore fluid relative to that of the solid, U_i is the displacement vector of the pore fluid, $e = \nabla\cdot\mathbf{u}$ is the dilatation for the solid motion, $\xi = -\nabla\cdot\mathbf{w}$ is the dilatation for the relative motion between the fluid and the solid, and $e_{ij} = [(\partial u_j/\partial x_i + \partial u_i/\partial x_j)/2]$, is the strain tensor in the porous medium. In Cartesian coordinates, x_1, x_2 and x_3 are equivalent to x, y and z.

In a two-dimensional XZ plane, equations (1) and (2) can be reorganized as

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial^2 u_1}{\partial t^2} &= 2m\frac{\partial}{\partial x_1}\left(\mu\frac{\partial u_1}{\partial x_1}\right) + m\frac{\partial}{\partial x_2}\left(\mu\frac{\partial u_1}{\partial x_2}\right)\\ &+ m\frac{\partial}{\partial x_2}\left(\mu\frac{\partial u_2}{\partial x_1}\right) + m\frac{\partial}{\partial x_1}\left(\lambda_c\frac{\partial u_1}{\partial x_1}\right)\\ &+ m\frac{\partial}{\partial x_1}\left(\lambda_c\frac{\partial u_2}{\partial x_2}\right) + m\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial w_1}{\partial x_1}\right)\\ &+ m\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial w_2}{\partial x_2}\right) - \rho_f\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial u_1}{\partial x_1}\right)\\ &- \rho_f\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial u_2}{\partial x_2}\right) - \rho_f\frac{\partial}{\partial x_1}\left(M\frac{\partial w_1}{\partial x_1}\right)\\ &- \rho_f\frac{\partial}{\partial x_1}\left(M\frac{\partial w_2}{\partial x_2}\right) + \rho_f\frac{\eta}{\kappa}\frac{\partial w_1}{\partial t}, \end{aligned}\tag{3}$$

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial^2 w_1}{\partial t^2} &= \rho\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial u_1}{\partial x_1}\right) + \rho\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial u_2}{\partial x_2}\right)\\ &+ \rho\frac{\partial}{\partial x_1}\left(M\frac{\partial w_1}{\partial x_1}\right) + \rho\frac{\partial}{\partial x_1}\left(M\frac{\partial w_2}{\partial x_2}\right)\\ &+ 2\rho_f\frac{\partial}{\partial x_1}\left(\mu\frac{\partial u_1}{\partial x_1}\right) - \rho_f\frac{\partial}{\partial x_2}\left(\mu\frac{\partial u_1}{\partial x_2}\right)\\ &- \rho_f\frac{\partial}{\partial x_2}\left(\mu\frac{\partial u_2}{\partial x_1}\right) - \rho_f\frac{\partial}{\partial x_1}\left(\lambda_c\frac{\partial u_1}{\partial x_1}\right)\\ &- \rho_f\frac{\partial}{\partial x_1}\left(\lambda_c\frac{\partial u_2}{\partial x_2}\right) - \rho_f\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial w_1}{\partial x_1}\right)\\ &- \rho_f\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial w_2}{\partial x_2}\right) - \rho\frac{\eta}{\kappa}\frac{\partial w_1}{\partial t}, \end{aligned}\tag{4}$$

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial^2 u_2}{\partial t^2} &= m\frac{\partial}{\partial x_1}\left(\mu\frac{\partial u_2}{\partial x_1}\right) + m\frac{\partial}{\partial x_1}\left(\mu\frac{\partial u_1}{\partial x_2}\right)\\ &+ 2m\frac{\partial}{\partial x_2}\left(\mu\frac{\partial u_2}{\partial x_2}\right) + m\frac{\partial}{\partial x_2}\left(\lambda_c\frac{\partial u_1}{\partial x_1}\right)\\ &+ m\frac{\partial}{\partial x_2}\left(\lambda_c\frac{\partial u_2}{\partial x_2}\right) + m\frac{\partial}{\partial x_2}\left(\alpha M\frac{\partial w_1}{\partial x_1}\right)\\ &+ m\frac{\partial}{\partial x_2}\left(\alpha M\frac{\partial w_2}{\partial x_2}\right) - \rho_f\frac{\partial}{\partial x_2}\left(\alpha M\frac{\partial u_1}{\partial x_1}\right)\\ &- \rho_f\frac{\partial}{\partial x_2}\left(\alpha M\frac{\partial u_2}{\partial x_2}\right) - \rho_f\frac{\partial}{\partial x_2}\left(M\frac{\partial w_1}{\partial x_1}\right)\\ &- \rho_f\frac{\partial}{\partial x_2}\left(M\frac{\partial w_2}{\partial x_2}\right) + \rho_f\frac{\eta}{\kappa}\frac{\partial w_2}{\partial t}, \end{aligned}\tag{5}$$

and

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial^2 w_2}{\partial t^2} &= \rho\frac{\partial}{\partial x_2}\left(\alpha M\frac{\partial u_1}{\partial x_1}\right) + \rho\frac{\partial}{\partial x_2}\left(\alpha M\frac{\partial u_2}{\partial x_2}\right)\\ &+ \rho\frac{\partial}{\partial x_2}\left(M\frac{\partial w_1}{\partial x_1}\right) + \rho\frac{\partial}{\partial x_2}\left(M\frac{\partial w_2}{\partial x_2}\right)\\ &- \rho_f\frac{\partial}{\partial x_1}\left(\mu\frac{\partial u_2}{\partial x_1}\right) - \rho_f\frac{\partial}{\partial x_1}\left(\mu\frac{\partial u_1}{\partial x_2}\right)\\ &- 2\rho_f\frac{\partial}{\partial x_2}\left(\mu\frac{\partial u_2}{\partial x_2}\right) - \rho_f\frac{\partial}{\partial x_2}\left(\lambda_c\frac{\partial u_1}{\partial x_1}\right)\\ &- \rho_f\frac{\partial}{\partial x_2}\left(\lambda_c\frac{\partial u_2}{\partial x_2}\right) - \rho_f\frac{\partial}{\partial x_2}\left(\alpha M\frac{\partial w_1}{\partial x_1}\right)\\ &- \rho_f\frac{\partial}{\partial x_2}\left(\alpha M\frac{\partial w_2}{\partial x_2}\right) - \rho\frac{\eta}{\kappa}\frac{\partial w_2}{\partial t}. \end{aligned}\tag{6}$$

We aim to solve these equations numerically for the displacement fields in an unbounded medium. To this end, we first formulate the PML as the absorbing boundary condition.

Complex coordinates for absorptive media

In order to introduce the PML for seismic waves in absorptive media, equations (1) and (2) will be modified using the complex stretched coordinates. In the frequency domain, the complex coordinate-stretching variable is chosen as

$$\tilde{x}_j = \int_0^{x_j} e_{x_j}(x_j)\,dx_j, \quad e_{x_j} = a_{x_j} + i\frac{\omega_{x_j}}{\omega} \quad (j=1,2,3), \tag{7}$$

where $i=\sqrt{-1}$, $a_{x_j}\geq 1$ and $\omega_{x_j}\geq 0$.

In the PML formulation, the regular coordinate variable x_j is replaced by the complex coordinate variable $\tilde{x}_j$. The derivative $\partial/\partial\tilde{x}_j$ can be expressed in terms of the complex coordinate-stretching variables

$$\frac{\partial}{\partial\tilde{x}_j} = \frac{1}{e_{x_j}}\frac{\partial}{\partial x_j} \quad (j=1,2,3), \tag{8}$$

where a time dependence of $e^{-i\omega t}$ is implied. In equation (7), for a PML region, the real part a_{x_j} is a scaling factor, and the imaginary part ω_{x_j} represents a loss in the PML. In a regular non-PML region, $a_{x_j}=1$ and $\omega_{x_j}=0$. Replacing the spatial derivatives in equations (3)–(6) with those in terms of complex coordinates, we arrive at the frequency domain wave equations for porous medium. In particular, equations (1) and (2) become in complex coordinates

$$2\sum_j \frac{\partial}{\partial\tilde{x}_j}(\mu e_{ij}) + \frac{\partial}{\partial\tilde{x}_i}(\lambda_c e - \alpha M\xi) = \frac{\partial^2}{\partial t^2}(\rho u_i + \rho_f w_i), \tag{9}$$

$$\frac{\partial}{\partial\tilde{x}_i}(\alpha M e - M\xi) = \frac{\partial^2}{\partial t^2}(\rho_f u_i + m w_i) + \frac{\eta}{\kappa}\frac{\partial w_i}{\partial t}. \tag{10}$$

The zero reflection of the PML has been proved by Chew and Liu (1996) for elastic media. Their proof was based on the reflection and transmission coefficients at a planar interface between the PML medium and a regular medium. It was later found that this lengthy proof is actually unnecessary (Teixeira and Chew, 1997; Liu, 1999). This is also true for poroelastic media. Indeed, the fact that the PML has zero reflection to the regular poroelastic medium follows the simple observation that equations (9) and (10) have exactly the same form as the original equations (1) and (2). Thus, the same solutions obtained for the regular media can be mapped to the PML media through a simple analytic continuation of the spatial variables to a complex space. This simple method for the construction of PML equations is especially useful for non-Cartesian coordinates (Teixeira and Chew, 1997; He and Liu, 1999; Liu, 1999).

Splitting of equations in time domain

Before incorporating the stretching variables, each of equations (3)–(6) is split into three equations, corresponding to three displacement variables. For example, for the horizontal displacement u_1 of solid, we have

$$u_1 = u_1^a + u_1^b + u_1^c,$$

where

$$u_1^a \sim \frac{\partial}{\partial x_1}\left(g\frac{\partial u_1}{\partial x_1}\right),$$
$$u_1^b \sim \frac{\partial}{\partial x_1}\left(g\frac{\partial u_1}{\partial x_2}\right),$$
$$or\ \frac{\partial}{\partial x_2}\left(g\frac{\partial u_1}{\partial x_1}\right),$$
$$u_1^c \sim \frac{\partial}{\partial x_2}\left(g\frac{\partial u_1}{\partial x_2}\right).$$

Here g represents a function of space. We then transform all equations to the frequency domain with complex coordinates.

Applying these schemes to equations (3)–(6), we obtain a total of 12 equations with respect to u_1^a, u_1^b, u_1^c, u_2^a, u_2^b, u_2^c, w_1^a, w_1^b, w_1^c, w_2^a, w_2^b, and w_2^c. For example, for u_1^a, the equation is

$$\begin{aligned}
\rho_f\frac{\eta}{\kappa}(i\omega)w_1^a + (m\rho-\rho_f^2)(-\omega^2)u_1^a = {} & \frac{2m}{e_{x_1}^2}\left[-\frac{e'_{x_1}}{e_{x_1}}\left(\mu\frac{\partial u_1}{\partial x_1}\right)\right. \\
& \left.+\frac{\partial}{\partial x_1}\left(\mu\frac{\partial u_1}{\partial x_1}\right)\right] + \frac{m}{e_{x_1}^2}\left[-\frac{e'_{x_1}}{e_{x_1}}\left(\lambda_c\frac{\partial u_1}{\partial x_1}\right)+\frac{\partial}{\partial x_1}\left(\lambda_c\frac{\partial u_1}{\partial x_1}\right)\right] \\
& + \frac{m}{e_{x_1}^2}\left[-\frac{e'_{x_1}}{e_{x_1}}\left(\alpha M\frac{\partial w_1}{\partial x_1}\right)+\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial w_1}{\partial x_1}\right)\right] \\
& - \frac{\rho_f}{e_{x_1}^2}\left[-\frac{e'_{x_1}}{e_{x_1}}\left(\alpha M\frac{\partial u_1}{\partial x_1}\right)+\frac{\partial}{\partial x_1}\left(\alpha M\frac{\partial u_1}{\partial x_1}\right)\right] \\
& - \frac{\rho_f}{e_{x_1}^2}\left[-\frac{e'_{x_1}}{e_{x_1}}\left(M\frac{\partial w_1}{\partial x_1}\right)+\frac{\partial}{\partial x_1}\left(M\frac{\partial w_1}{\partial x_1}\right)\right],
\end{aligned} \tag{11}$$

where $e'_{x_1} = i(d\omega_{x_1}/dx_1)/\omega$. Then we transform all these 12 equations back to the time domain. For example, in the time domain, equation (11) becomes

$$\begin{aligned}
-\rho_f\frac{\eta}{\kappa}\left[\frac{\partial w_1^a}{\partial t} + 2\omega_{x_1}w_1^a + \omega_{x_1}^2\int_0^t w_1^a\,d\tau\right] - (m\rho-\rho_f^2) \\
\times\left[-\frac{\partial^2 u_1^a}{\partial t^2} - 2\omega_{x_1}\frac{\partial u_1^a}{\partial t} - \omega_{x_1}^2 u_1^a\right] = R,
\end{aligned} \tag{12}$$

where R is the inverse transformation of the right side of equation (11) times $e_{x_1}^2$. The multiplication of e'_{x_1}/e_{x_1} and $\mu(\partial u_1/\partial x_1)$ in the frequency domain leads to the convolution of these two terms in time domain. Since e'_{x_1}/e_{x_1} is an exponential function in the time domain, this convolution can be obtained efficiently as follows. Let $f(t)$ represent the inverse Fourier transform of $(e'_{x_1}/e_{x_1})\,\mu(\partial u_1/\partial x_1)$. It can be computed via

$$\begin{aligned}
f(t) = e^{-\omega_{x_1}(x_1)\triangle t}\left[f(t-\triangle t) + \mu\omega'_{x_1}(x_1)\right. \\
\left.\int_{t-\triangle t}^{t} e^{-\omega_{x_1}(x)(t-\tau)}\frac{\partial u_1(\tau)}{\partial x_1}\,d\tau\right].
\end{aligned}$$

This requires the storage of u_1 at $(n-1)$th and nth time steps only. Similar equations can be obtained for u_1^b, u_1^c, and other field components.

It is worthwhile to note that these splittings are only necessary inside the PML region. The number of cells in the PML region is about $2m(N_x+N_z)$, where m is the thickness of the

PML region, and N_x and N_z are the total number of nodes in the x and z directions, respectively. For a large-scale problem, the ratio between this number and the total number of nodes $2m(N_x + N_z)/N_x N_z$ is negligible.

Finite-difference implementation

In the last section, we obtained 12 time-domain equations for 12 split field variables u_1^a, u_1^b, u_1^c, u_2^a, u_2^b, u_2^c, w_1^a, w_1^b, w_1^c, w_2^a, w_2^b, and w_2^c, which already incorporate the PML boundary condition.In numerical implementation, the time and space are discretized by time step $\triangle t$ and node spacing $\triangle x$ and $\triangle z$. The unbounded medium is truncated into a finite computational domain with a total of $N_x \times N_z$ grid nodes. Then, the central finite-difference method is applied to each equation to obtain the updated displacement at next time step. In this FD scheme, both spatial and temporal derivatives are obtained by central differencing. Thus, it is an explicit scheme with second-order accuracy both in spatial and temporal domains.

With the PML boundary condition, the computational domain is partitioned into a regular interior region and the boundary PML region. Unfortunately, the stability condition of the finite-difference method for the PML material does not seem to have a closed-form solution. However, numerical experiments suggest that the stability condition is the same as the regular medium when the damping coefficient $b = \eta/\kappa$ is negligible.

The absorption of outgoing wave is achieved by the PML region, which consists of several cells of PML materials with a quadratically or linearly tapered profile to increase the attenuation toward the outer boundary. In this paper, 20 cells of PML with a quadratic profile for ω_{x_j} are used. So the ω_x and ω_z have the following forms

$$\omega_{x_i} = \begin{cases} 2\pi a_0 f_0 (\ell_{x_i}/L_{PML})^2, & \text{inside PML,} \\ 0, & \text{outside PML,} \end{cases}$$

where f_0 is the dominant frequency of the source described in the next section, L_{PML} is the thickness of the PML cells, and ℓ_{x_i} is the distance from the interface between the interior region and PML region. After testing several values for a_0, we found 1.79 has the best absorption of outgoing waves. The outer boundary conditions for the PML are those for a hard boundary (or alternatively, a soft boundary).

Discussions

The above formulation for the perfectly matched layer is based on the complex coordinates and is a natural extension of our earlier work on PML for acoustic and elastic waves (Chew and Liu, 1996; Liu and Tao, 1997, Liu, 1998, 1999). The novelty of the PML formulation using the complex coordinates is that the governing equations and the solution forms remain identical to those in the regular coordinates. Therefore, in the continuum, the boundary between a PML material and a regular material is reflection free. (The small reflections in the numerical results are solely due to the discretization and the truncation.)

The fact that the PML is reflection free is fundamentally different from the classical sponge absorbing-boundary method (e.g., Kosloff and Kosloff, 1986). Although the sponge absorbing-boundary method has been used widely and successfully in many applications, the boundary between a regular material and a sponge absorber is not reflection free even in the continuum limit. Indeed, consider a simple 1-D acoustic case with a constant density ρ and constant acoustic velocity c. At the boundary between such a medium and a sponge absorber half-space with the same density and acoustic velocity, and a constant absorbing parameter γ defined by Kosloff and Kosloff (1986), it is straightforward to derive the reflection coefficient

$$R = \frac{1 - \sqrt{1 + i\dfrac{2\gamma}{\omega} - \dfrac{\gamma^2}{\omega^2}}}{1 + \sqrt{1 + i\dfrac{2\gamma}{\omega} - \dfrac{\gamma^2}{\omega^2}}},$$

which is nonzero for any nonzero real values of γ. This reflection coefficient decreases as γ decreases, which is the basis for the tapered sponge absorber in most applications. It is worthwhile to note that this particular formulation of the sponge method can be improved for the 1-D case. Indeed, as the perfectly matched condition for the 1-D case requires only the match in impedance, it is quite straightforward to design a sponge layer so that there is no reflection for the 1-D case. This improved sponge method formulation uses partial differential equations:

$$\frac{\partial p}{\partial t} + \gamma p = -\rho c^2 \frac{\partial v}{\partial x},$$
$$\frac{\partial v}{\partial t} + \gamma v = -\rho^{-1} \frac{\partial p}{\partial x},$$

where p is the pressure field, and v is the particle velocity field. An example of velocity-stress formulation for poroviscoelastic media is given in Carcione (1998). It is straightforward to show that the reflection coefficient of this sponge layer is zero because of the match in impedance between regular medium and the sponge. However, such reflection-free sponge method formulations are not possible in the multidimensional case, as is well known in seismic modeling. This the major difference between the sponge method and the PML method can also be explained by the difference in the partial differential equations they solve. In particular in the frequency domain, for the special 2-D acoustic case, the sponge method solves (Kosloff and Kosloff, 1986)

$$\rho c^2 \left[\frac{\partial}{\partial x} \frac{1}{\rho} \frac{\partial p}{\partial x} + \frac{\partial}{\partial y} \frac{1}{\rho} \frac{\partial p}{\partial y} \right] = (-\omega^2 + \gamma^2 - i2\omega\gamma) p,$$

whereas the PML method solves

$$\rho c^2 \left[\frac{1}{e_x} \frac{\partial}{\partial x} \frac{1}{\rho e_x} \frac{\partial p}{\partial x} + \frac{1}{e_y} \frac{\partial}{\partial y} \frac{1}{\rho e_y} \frac{\partial p}{\partial y} \right] = -\omega^2 p,$$

where e_x and e_y are given by equation (7). Obviously, the anisotropic nature of the attenuation factors (i.e., different ω_x and ω_y) is present in the PML method, but not in the sponge method. For the sponge method, the reader is also refered to Carcione et al. (1988).

We argue that the zero reflection coefficient of the PML interface allows a larger attenuation factor in the PML absorber, and thus a more effective absorbing boundary than the classical sponge absorber. Along with this advantage comes the additional cost in terms of the computer memory requirement in the PML formulation. Specifically, within this PML formulation there are 12 split field variables instead of the original

four field variables. Fortunately, this increase in the memory is necessary only within the PML cells surrounding the interested domain and is quite negligible for large-scale problems.

NUMERICAL EXAMPLES

We have implemented the finite-difference algorithm with the perfectly matched layers as an absorbing boundary condition for Biot's equations in two-dimensional porous media. Unlike the continuous case, a small reflection will occur at the PML interface due to the discretization and truncation. This reflection is proportional to the contrast in the coordinate-stretching variables. Therefore, to minimize the reflection from the PML region, we use a quadratic profile with 20 cells of perfectly matched layers at the computational edge.

In the following examples, a line, pure compressional source is used to excite the seismic wave field. The source time function is the first derivative of the Gaussian function:

$$S(t) = (t - t_0)e^{-[\pi f_0(t-t_0)]^2},$$

where f_0 is the dominant frequency, and t_0 is the central time of the wavelet.

A bulk source is used in this paper. The source energy is partitioned linearly between the solid and the fluid phases:

$$W_f = \phi,$$
$$W_s = (1 - \phi),$$

where W_f is the weighting factor for the fluid motion, and W_s is for the solid motion.

A homogeneous model

A homogeneous model is used to test the effectiveness of the PML ABC. The material properties of the model are listed in Table 1. The size of the model is $N_x \times N_z = 200 \times 200$ nodes with 20 cells of PML on each side of the computational domain.

Table 1. Material and Biot's parameters for a homogeneous porous medium (from Özdenvar and McMechan, 1997).

Solid	
Bulk modulus K_s (Pa)	3.5×10^{10}
Density ρ_s (kg/m^3)	2650.0
Matrix	
Bulk modulus K_b (Pa)	4.17×10^9
Shear modulus μ (Pa)	1.855×10^9
Density (with fluid) ρ (kg/m^3)	2167.0
Porosity ϕ	0.3
Permeability κ (m^2)	1.0×10^{-12}
Tortuosity a	2.0
Fluid	
Bulk modulus K_f (Pa)	2.4×10^9
Density ρ_f (kg/m^3)	1040.0
Viscosity η (Pa s)	1.0×10^{-3}
Biot's parameters	
P (N/m^2)	9.02×10^9
Q (N/m^2)	1.23×10^9
R (N/m^2)	6.35×10^8
Seismic characteristics	
Velocity of fast P-wave (m/s)	2365
Velocity of slow P-wave (m/s)	775
Velocity of S-wave (m/s)	960

The spatial and temporal steps are $\triangle x = \triangle z = 1.5$ m and $\triangle t = 0.0001$ s. A line monopole source with a dominant frequency of 40 Hz is located at the center of the model. To inspect the reflection from the PML interface, we put one array of receivers along a vertical line running through the source.

The waveforms of vertical displacement in solid at nine locations are shown in Figure 1. In this figure, a reference result without reflection is also shown. This reference result is obtained from a much larger model with Dirichlet boundary in which reflections have not arrived within the time window of interest. Note that the results are normalized with respect to the peak value of the field at the first receiver. Because the fluid is viscous, the slow P-wave is rapidly attenuated as it propagates. We observe no obvious reflections in these waveforms.

Figure 2 shows snapshots of the vertical components of the displacement of the solid at 0.04, 0.06, 0.1, and 0.13 s when the damping coefficient b has the realistic value of 10^8N s/m^4. In this case the slow wave is diffusive and appears as a static mode at the source location (Zhang, 1999).

From the above results, it is seen that the PML method provides an effective attenuation to outgoing waves. Numerical experiments also demonstrate that this method is stable as long as $\triangle t \leq \min(\triangle x, \triangle z)/\sqrt{2}V_{\max}$ when the damping coefficient b is small. It is also noted that in this example the ratio of V_f/V_s is 2.46, which is much greater than 2 allowed in other conventional absorbing boundary conditions. We calculated the wave field with Clayton-Engquist (CE) absorbing boundary condition on the same model, and found the method becomes unstable after about 2000 time steps.

A two-layer model

In order to demonstrate the behavior of this algorithm for inhomogeneous media, we test the PML absorbing boundary condition on a two-layer model. In contrast to the homogeneous model, multiwave phases (such as reflected, transmitted, and converted phases) will be present. Some of them are relatively weak compared to the direct wave phases and are very vulnerable to the artificial reflections from the boundaries. Thus, whether these weak wave arrivals can be distinguished

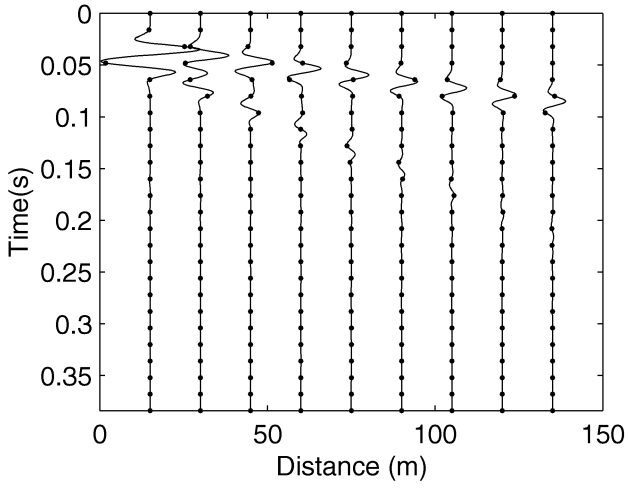


FIG. 1. Waveforms of vertical component of the solid displacement at nine vertical locations. Solid line: results from the model with PML. Dots: results with reflection.

determines the effectiveness of the absorbing boundary condition. In this example, we compare the numerical results using PML ABC with those of CE ABC. The geometry of the model is shown in Figure 3. A line monopole source is located at the center of the model which is in the upper layer. The model parameters are shown in Table 2.

The snapshot in Figure 4a at $t = 0.055$ s shows that the solid and fluid particle displacements are in phase for fast P-wave and out of phase for slow P-wave, and the fast P-wave begins to hit the interface. At $t = 0.085$ s (Figure 4b), transmitted and reflected wave phases are produced from the fast P-wave. At $t = 0.12$ s (Figure 4c), the slow P-wave reaches the interface and produces reflected and transmitted wave phases. Figure 4d shows the wave field at $t = 0.18$ s. In this figure, a very small reflection occurs from the edge.

Table 2. Material parameters for a two-layer model of porous media.

Material properties	Layer 1	Layer 2
Solid		
Bulk modulus K_s (Pa)	3.9×10^{10}	5.2×10^9
Shear modulus μ (Pa)	5.25×10^9	2.4×10^9
Density ρ_s (kg/m^3)	2588.0	2250.0
Matrix		
Bulk modulus K_b (Pa)	4.12×10^9	2.2×10^9
Density (with fluid) ρ (kg/m^3)	2167.0	2167.0
Porosity ϕ	0.25	0.1
Tortuosity a	2.49	2.42
Fluid		
Bulk modulus K_f (Pa)	2.21×10^9	2.25×10^9
Density ρ_f (kg/m^3)	952.4	1040.0
η/κ		
Damping coefficient b (N s/m^4)	3.38×10^5	3.33×10^6

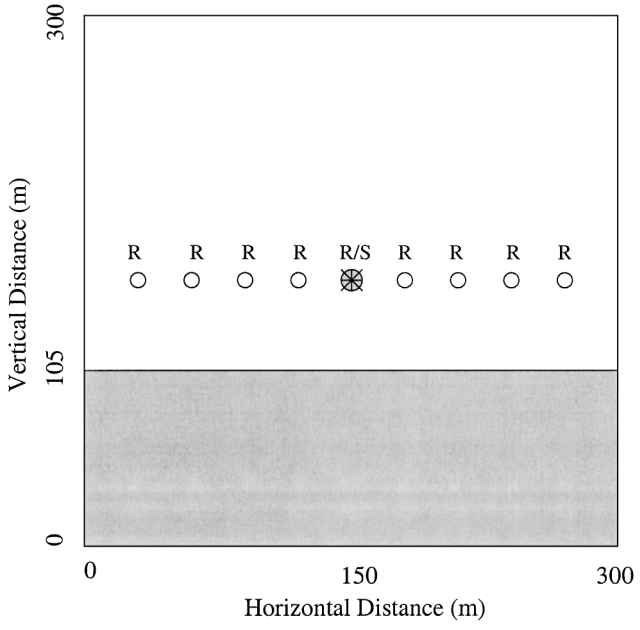


FIG. 3. Configuration of the two-layer model. R: receiver, S: source.

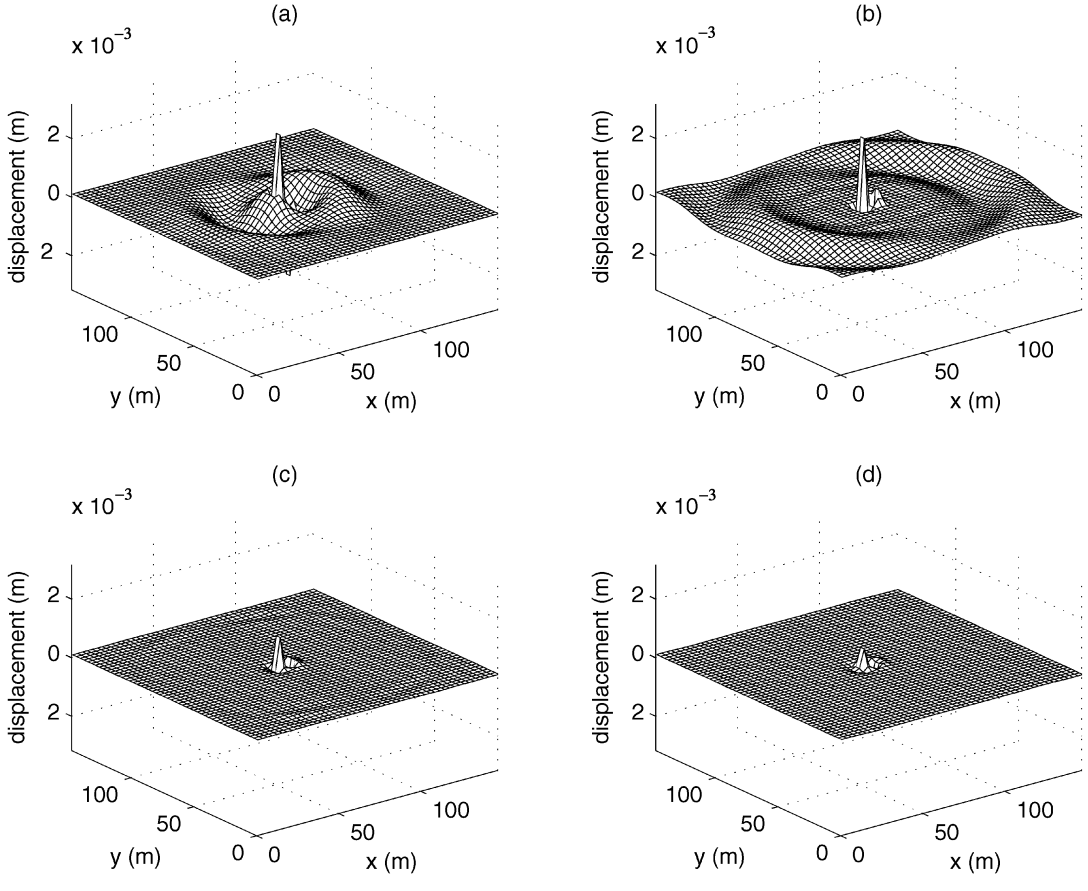


FIG. 2. Snapshots of the vertical components of the solid displacement at 0.04 s (a), 0.06 s (b), 0.1 s (c), and 0.13 s (d). The model parameters are shown in Table 1 except $b = 10^8$N s/m. A line monopole source is used.

In Figures 5 and 6, we present the synthetic seismograms for this model. The receivers are located along the line running through the source and parallel to the interface. The results from the algorithm with PML ABC are shown in Figure 5, those with CE ABC are shown in Figure 6. In order to compare the PML ABC with CE ABC, two snapshots from the method with CE ABC and two from the method with PML ABC are shown together in Figure 7. Clearly, the PML ABC has a far better performance.

CONCLUSIONS

A finite-difference method combined with the perfectly matched layer (PML) absorbing boundary condition is developed for modeling seismic wave propagation in poroelastic media. Within the boundary region of the computational domain, perfectly matched layers are used to attenuate outgoing seismic waves. The numerical results show that the outgoing waves are effectively absorbed and the reflection is very small. In contrast to some other existing absorbing boundary conditions, this new absorbing boundary condition is stable even when the ratio of the *P*-wave velocity to *S*-wave velocity is much greater than 2. Furthermore, since the PML boundary condition is directly incorporated into the wave equations, it can be used in the pseudospectral method.

ACKNOWLEDGMENT

This work is supported in parts by the Army Research Office under Grant DAAH04-96-1-0448 (Demining MURI), by the U.S. EPA under a PECASE grant CR-825-225-010, by the

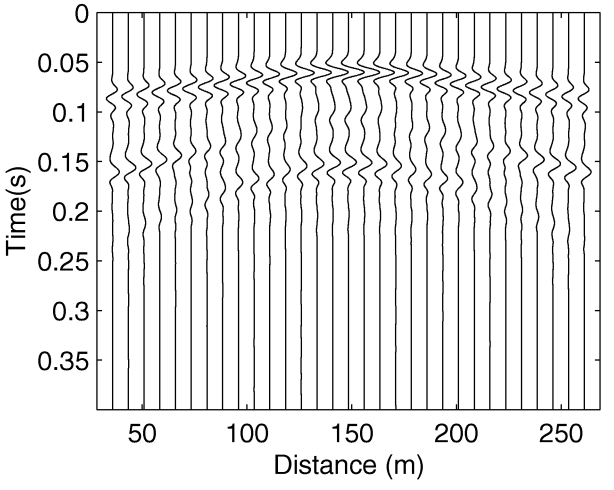


FIG. 5. Waveforms of vertical component of the solid displacement from the two-layer model with PML boundary conditions.

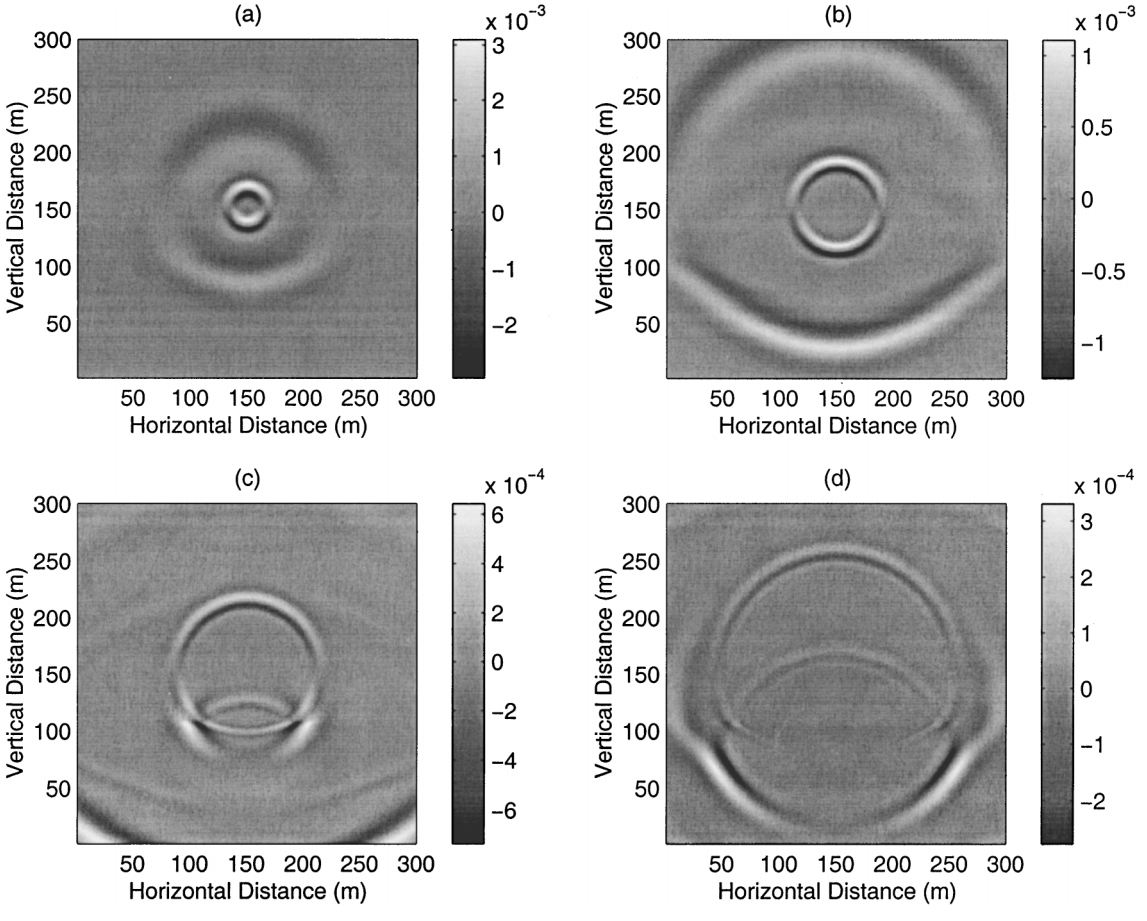


FIG. 4. Snapshots of the vertical component of the solid displacement at four different times 0.055 s (a), 0.085 s (b), 0.12 s (c), and 0.18 s (d) from the finite-difference solutions with PML. The model parameters are shown in Table 2. A monopole line source is used.

NSF under a CAREER grant ECS-9702195. The authors thank the Associate Editor and two anonymous reviewers for their suggestions to improve the manuscript.

REFERENCES

Berenger, J. P., 1994, A perfectly matched layer for absorption of electromagnetic waves: J. Comput. Phys., **114**, 185–200.

Biot, M. A., 1956a, Theory of propagation of elastic waves in a fluid-saturated porous solid. 1. Low-frequency range: J. Acoust. Soc. Am., **28**, 168–178.

———1956b, Theory of propagation of elastic waves in a fluid-saturated porous solid. 1. Higher-frequency range: J. Acoust. Soc. Am., **28**, 179–191.

———1962a, Mechanics deformation and acoustic propagation in porous media: J. Appl. Phys., **33**, 1482–1498.

———1962b, Generalized theory of acoustic propagation in porous dissipative media: J. Acoust. Soc. Am., **34**, 1254–1264.

Carcione, J. M., 1998, Viscoelastic effective rheologies for modelling wave propagation in porous media: Geophys. Prosp., **46**, 249–270.

Carcione, J. M., Kosloff, D., and Kosloff, R., 1988, Viscoacoustic wave propagation simulation in the earth: Geophysics, **53**, 769–777.

Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M., 1985, A nonreflecting boundary condition for discrete acoustic and elastic wave equations: Geophysics, **50**, 705–708.

Chew, W. C., and Liu, Q. H., 1996, Perfectly matched layers for elastodynamics: A new absorbing boundary condition: J. Comp. Acoust., **4**, no. 4, 72–79.

Chew, W. C., and Weedon, W. H., 1994, A 3-D perfectly matched medium from modified Maxwell's equations with stretched coordinates: Microw. Opt. Technol. Lett., **7**, 599–604.

Clayton, R., and Engquist, B., 1977, Absorbing boundary conditions for acoustic and elastic wave equations: Bull. Seism. Soc. Am., **67**, 1529–1540.

Hastings, F. D., Schneider, J. B., and Broschat, S. L., 1996, Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation: J. Acoust. Soc. Am., **100**, 3061–3069.

He, J., and Liu, Q. H., 1999, A nonuniform cylindrical FDTD algorithm with improved PML and quasi-PML absorbing boundary conditions: IEEE Trans. Geosci. Remote Sensing, **37**, 1066–1072.

Kosloff, R., and Kosloff, D., 1986, Absorbing boundary for wave propagation problems: J. Comp. Phys., **63**, 363–376.

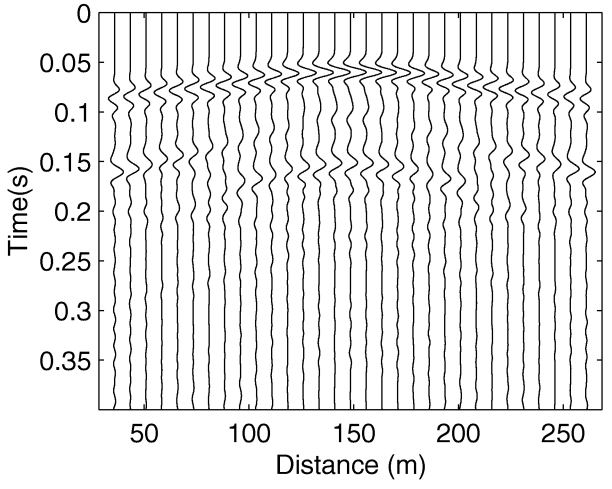


FIG. 6. Waveforms of vertical component of the solid displacement from the two-layer model with CE boundary conditions.

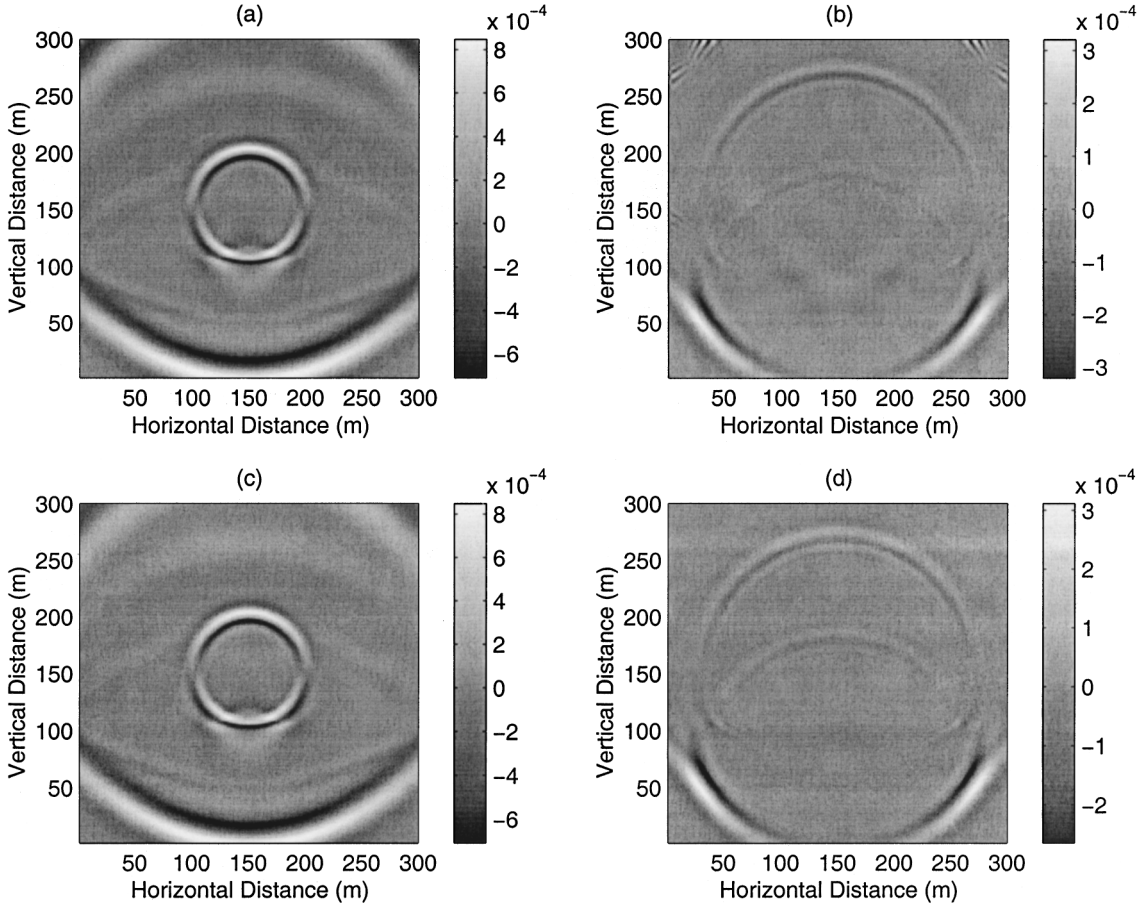


FIG. 7. Snapshots at times 0.1 s (a) and 0.2 s (b) for a two-layer model from the method with CE ABC. Snapshots at times 0.1 s (c) and 0.2 s (d) for the same model from the method with PML ABC. The model parameters are shown in Table 2 . A monopole line source is used.

Liao, Z. P., Wong, H. L., Yang, B. P., and Yuan, Y. F., 1984, A transmitting boundary for transient wave analysis: Sci. Sinica A, **27**, 1063–1076.

Liu, Q. H., 1997, An FDTD algorithm with perfectly matched layers for conductive media: Microw. Opt. Technol. Lett., **14**, 134–137.

———1998, The PSTD algorithm for acoustic waves in inhomogeneous, absorptive media: IEEE Trans. Ultrason., Ferroelect., Freq. Contr., **45**, 1044–1055.

———1999, Perfectly matched layers for elastic waves in cylindrical and spherical coordinates: J. Acoust. Soc. Am., **105**, 2075–2084.

Liu, Q. H. and Tao J., 1997, The perfectly matched layer for acoustic waves in absorptive media: J. Acoust. Soc. Am., **102**, 2072–2082.

Mahrer, K. D., 1986, An empirical study of instability and improvement of absorbing boundary conditions for elastic wave equation: Geophysics, **51**, 1499–1501.

Özdenvar, T., and McMechan, G. A., 1997, Algorithm for staggered-grid computations for poroelastic, elastic, acoustic, and scalar wave equations: Geophys. Prosp., **45**, 403–420.

Reynolds, A. C., 1978, Boundary conditions for the numerical solution of wave propagation problems: Geophysics, **43**, 1099–1110.

Smith, W. D., 1974, A nonreflecting plane boundary for wave propagation problems: J. Comp. Phys., **15**, 492–503.

Stacey, R., 1988, Improved transparent boundary formulations for the elastic-wave equation: Bull. Seism. Soc. Am., **78**, 2089–2097.

Teixeira, F. L., and Chew, W. C., 1997, PML-FDTD in cylindrical and spherical coordinates: IEEE Microwave Guided Wave Lett., **7**, 285–287.

Zhang, J., 1999, Quadrangle-grid velocity-stress finite-difference method for poroelastic wave equations: Geophys. J. Internat., **139**, 171–182.

Acoustic Detection of Buried Objects in 3-D Fluid Saturated Porous Media: Numerical Modeling

Yan Qing Zeng, *Student Member, IEEE*, and Qing Huo Liu, *Senior Member, IEEE*

***Abstract*—Acoustic waves can be a viable tool for the detection and identification of land mines, unexplored ordnance (UXO), and other buried objects. Design of acoustic instruments and interpretation and processing of acoustic measurements call for accurate numerical models to simulate acoustic wave propagation in a heterogeneous soil with buried objects. Compared with the traditional seismic exploration, high attenuation is unfortunately ubiquitous for shallow surface acoustic measurements because of the loose soil and the fluid in its pore space. To adequately model such acoustic attenuation, we propose a comprehensive multidimensional finite-difference time-domain (FDTD) model to simulate the acoustic wave interactions with land mines and soils based on the Biot theory for poroelastic media. For the truncation of the computational domain, we use the perfectly matched layer (PML). The method is validated by comparison with analytical solutions. Unlike the pure elastic wave model, this efficient PML-FDTD model for poroelastic media incorporates the interactions of waves and the fluid-saturated pore space. Several typical land mine detection measurements are simulated to illustrate the application.**

***Index Terms*—Acoustic waves, finite-difference time-domain (FDTD), landmine detection, perfectly matched layer, poroelasticity.**

I. Introduction

RECENTLY, acoustic methods have received considerable attention in the characterization of shallow objects. It has been shown that acoustic waves can be a viable tool for the detection and identification of land mines and unexplored ordnance (UXO) [1], [2] (also see some experimental papers in this special issue). The critical issue of acoustic wave coupling into the soil has been addressed by using low-frequency loudspeakers, laser, off-set mechanical source, and electric spark sources [2], [1]. These recent experimental efforts have shown promising results for acoustic landmine detection.

Design of acoustic instruments, interpretation and processing of acoustic measurements call for accurate numerical models to simulate acoustic wave propagation in a heterogeneous soil with buried objects. Conventional methods for acoustic modeling include the popular finite-difference time-domain (FDTD) solution of pure elastic media [3], [4]. Unfortunately, the pure elastic model cannot adequately incorporate the attenuation mechanism in its governing equations, although some approximate models are possible [5].

Manuscript received July 5, 2000; revised December 28, 2000. This work was supported by the Army Research Office under Grant DAAH04-96-1-0448 (Demining MURI).

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA (qhliu@ee.duke.edu).

Publisher Item Identifier S 0196-2892(01)04835-5.

Compared with the traditional seismic exploration, high attenuation is unfortunately ubiquitous for shallow surface acoustic measurements because of the loose soil and the fluid in its pore space. To adequately model such acoustic attenuation, we propose a comprehensive model to simulate the acoustic wave interactions with land mines and soils based on the Biot theory [6]–[8] for poroelastic media. The FDTD method is then used to solve the Biot equations [8]. Similar to the velocity-stress finite-domain (FD) method [9], a velocity-strain, FD method is developed in a staggered grid for heterogeneous poroelastic media. In this method, Biot equations [8] are reformulated into first-order equations to arrive at a leap-frog system in a staggered grid both in time and space domains.

In order to simulate an unbounded medium, an absorbing boundary condition (ABC) must be implemented to truncate the computational domain in numerical algorithms. Several ABCs have been developed for numerical simulation of wave propagation. Cerjan *et al.* [10] introduce a simple damping taper to the boundaries that attenuates outgoing waves. Since this lossy layer is not perfectly matched to the interior region, however, it requires a substantial number of layers for the taper to be effective. Clayton and Engquist [11] (CE) use the paraxial approximation to the wave equation to make the boundary transparent to outgoing waves. The CE ABC can lead to instability when the Poisson's ratio is greater than two [12]. Since Berenger [13] proposed the highly effective perfectly matched layer (PML) as an absorbing boundary condition for electromagnetic waves, it has been widely used for FD and finite-element methods. Chew and Liu [14] first proposed the PML for elastic waves in solids, and proved the zero reflections from PML to the regular elastic medium. Hastings *et al.* [15] have independently implemented the PML ABC for two-dimensional problems by using potentials. The PML has also been extended to model acoustic waves and electromagnetic waves in lossy media [16] and to cylindrical and spherical coordinates [17]. Recently, PML has been applied to the second-order Biot's equations for two-dimensional (2-D) fluid-saturated poroelastic media [18], which requires a complicated convolution. In this paper, we apply the PML to three-dimensional (3-D) using the first-order partial differential equation (PDE) system.

The difference between elastic model and poroelastic model is investigated by studying surface wave amplitude variation with offset (AVO) in three different types of soil: dry sand, fully water saturated sand and partly water saturated sand. The interaction of elastic wave with a plastic mine buried in dry sand is simulated. We also simulate two typical configurations for land mine detection, namely the excitation in soil [1] and the excitation in air [2]. The results show that the wave responses

0196–2892/01$10.00 © 2001 IEEE

are significantly affected by the existence of a mine-like object through surface waves. After processing, the target can be detected by using surface acoustic measurements. The 2-D and 3-D poroelastic codes provide a useful tool for design and detection issues in acoustic characterization of buried objects.

II. Formulation

A. Governing Equations

The propagation of acoustic waves in fluid-saturated porous media is different from that in single phase elastic media. In addition to the regular P (compressional) waves and S (shear) waves in solid elastic media, a slow P wave resulting from the relative motion between the solid frame and pore fluid may be present in porous media. Thus, energy dissipation in porous media is different from that in solid elastic media. Based on continuum mechanics and macroscopic constitutive relationship, Biot [6]–[8] developed a theory of wave motion in a poroelastic solid saturated with a viscous compressible fluid. Biot's theory was confirmed by Burridge and Keller [19] based on the dynamic equations governing the behavior of medium on a microscopic scale. Plona [20] also confirmed Biot's theory through experiments.

In an isotropic, heterogeneous porous elastic medium, the parameters describing the physical properties of the medium are as follows:

μ	shear modulus of dry porous matrix;
λ_c	Láme constant of saturated matrix;
ϕ	porosity;
η	viscosity;
κ	permeability;
ρ	overall density of the saturated medium determined by $\rho_f\phi + (1-\phi)\rho_s$;
ρ_s	density of solid material;
ρ_f	density of fluid;
a	tortuosity of the matrix;
K_s	bulk modulus of the solid;
K_f	bulk modulus of the fluid;
K_b	bulk modulus of the dry porous frame.

The macroscopic displacements and strains are defined as

u_i	ith component of displacements of solid particle;
U_i	ith component of displacements of fluid particle;
w_i	ith component of relative displacement, $w_i = \phi(U_i - u_i)$;
e_{ij}	ij component of strain tensor in porous medium;

$$e_{ij} = \left[\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right) \Big/ 2\right];$$

e	dilatation for the solid motion $e = \sum_{i=1,2,3} e_{ii}$;
ξ	dilatation for the relative motion $\xi = -\sum_{i=1,2,3} (\partial w_i/\partial x_i)$.

For a 3-D isotropic, heterogeneous and porous elastic medium, wave propagation is governed by Biot's equations [8]

$$2\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) + \frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) = \frac{\partial^2}{\partial t^2}(\rho u_i + \rho_f w_i) \tag{1}$$

$$\frac{\partial}{\partial x_i}(\alpha M e - M\xi) = \frac{\partial^2}{\partial t^2}(\rho_f u_i + m w_i) + \frac{\eta}{\kappa}\frac{\partial w_i}{\partial t} \tag{2}$$

where $m = a\rho_f/\phi$ and

$$M = \frac{1}{\dfrac{\phi}{K_f} + \dfrac{(\alpha-\phi)}{K_s}}$$

$$\alpha = 1 - \frac{K_b}{K_s}.$$

Let $\mathbf{v^s} = \partial\mathbf{u}/\partial t$ be the velocity of the solid particle and $\mathbf{v^f} = \partial\mathbf{w}/\partial t$ be the velocity of the pore fluid relative to the solid frame. Then the second-order equations (1) and (2) can rearranged as the first-order equations

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial v_i^s}{\partial t} &= 2m\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) + m\frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) \\ &\quad - \rho_f\frac{\partial}{\partial x_i}(\alpha M e - M\xi) + \rho_f\frac{\eta}{\kappa}v_i^f \end{aligned} \tag{3}$$

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial v_i^f}{\partial t} &= \rho\frac{\partial}{\partial x_i}(\alpha M e - M\xi) - \rho\frac{\eta}{\kappa}v_i^f - 2\rho_f\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) \\ &\quad - \rho_f\frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi). \end{aligned} \tag{4}$$

The time derivatives of strains e_{ij} and ξ can be expressed as

$$\frac{\partial e_{ij}}{\partial t} = \left[(\partial v_j^s/\partial x_i + \partial v_i^s/\partial x_j)/2\right] \tag{5}$$

$$\frac{\partial \xi}{\partial t} = -\nabla\cdot\mathbf{v^f}. \tag{6}$$

In this work, an explicit second-order FD scheme is applied to convert (3)–(6) into a leap-frog system for the strain field e_{ij} ξ and velocity field $\mathbf{v^s}$ and $\mathbf{v^f}$. With proper absorbing boundary conditions, these equations can be solved numerically for the wave field in an unbounded medium.

Equations (3)–(6) predict the existence of three different waves in fluid-saturated poroelastic media: a shear wave and two compressional waves with a faster and a slower propagation velocities [6]–[8].

B. Note for Pure Fluid and Solid

Biot's theory is more general than the elastodynamic equations, and includes the latter as a special case. Therefore, equations (3) and (4) may also be used to describe the response of single-phase elastic media as well as a single-phase fluid. For a single-phase solid, we simply let porosity $\phi \to 0$ and the initial condition for the relative motion equal to zero. On the other hand, a single-phase fluid can be modeled by setting $K_f = K_s$ and $\mu = 0$ explicitly in these equations. Given the zero initial relative motion, equation (4) will guarantee there is no relative motion for all time. Therefore, there is no need to introduce special treatment even if some regions of the model have pure fluid and solid.

C. Equations for the PML Absorbing Boundary Condition

In this paper, the PML is used to truncate the unbounded medium to absorb all outgoing waves. This artificial absorptive medium is introduced in the regular medium by modifying equations (3)–(6) with complex coordinates [14], [17]. In the frequency domain with $e^{-i\omega t}$ convention, a complex coordinate variable is chosen as

$$\widetilde{x_j} = \int_0^{x_j} e_j(x_j')\, dx_j' \tag{7}$$

$$e_j = a_j + i\,\frac{\omega_j}{\omega} \qquad (j = 1, 2, 3) \tag{8}$$

where $a_j \geq 1$ is a scaling factor, and $\omega_{x_j} \geq 0$ is an attenuation factor. The operator $\partial/\partial x_j$ can be expressed in terms of the regular coordinate

$$\frac{\partial}{\partial \widetilde{x_j}} = \frac{1}{e_j}\,\frac{\partial}{\partial x_j}. \tag{9}$$

The PML formulation is to replace x_j in (4)–(6) by the corresponding complex coordinate $\widetilde{x_j}$. In a regular non-PML region, $a_j = 1$ and $\omega_j = 0$. In order to simplify PML equations, the field variables are split as the following:

$$v_j^s = \sum_{k=1}^{3} v_j^{s(k)}$$

$$v_j^f = \sum_{k=1}^{3} v_j^{f(k)}$$

where $v_i^{s(k)}$ and $v_i^{f(k)}$ represent the split field variables containing space derivative $\partial/\partial x_k$ only . For example, (3) for $i = 1$ can be split into the following three equations:

$$(m\rho - \rho_f^2)\,\frac{\partial v_1^{s(1)}}{\partial t} = 2m\,\frac{\partial}{\partial x_1}(\mu e_{11}) + m\,\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M\xi) - \rho_f\,\frac{\partial}{\partial x_1}(\alpha M e - M\xi)$$

$$(m\rho - \rho_f^2)\,\frac{\partial v_1^{s(2)}}{\partial t} = 2m\,\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f\,\frac{\eta}{\kappa}\,v_1^f$$

$$(m\rho - \rho_f^2)\,\frac{\partial v_1^{s(3)}}{\partial t} = 2m\,\frac{\partial}{\partial x_3}(\mu e_{13}).$$

The diagonal strain components e_{jj} need not be split. However, other strain components have to be split as $e_{jl} = \sum_{k=1}^{3} e_{jl}^{(k)}$ for $j \neq l$, and $\xi = \sum_{k=1}^{3} \xi^{(k)}$, for example

$$\frac{\partial e_{jj}}{\partial t} = \frac{\partial v_j^s}{\partial x_j}$$

$$e_{12} = e_{12}^{(1)} + e_{12}^{(2)}$$

$$\frac{\partial e_{12}^{(1)}}{\partial t} = \frac{1}{2}\,\frac{\partial v_2^s}{\partial x_1}$$

$$\frac{\partial e_{12}^{(2)}}{\partial t} = \frac{1}{2}\,\frac{\partial v_1^s}{\partial x_2}$$

$$\xi = \xi^{(1)} + \xi^{(2)} + \xi^{(3)}$$

$$\frac{\partial \xi^{(1)}}{\partial t} = -\frac{\partial v_1^f}{\partial x_1}$$

$$\frac{\partial \xi^{(2)}}{\partial t} = -\frac{\partial v_2^f}{\partial x_2}$$

$$\frac{\partial \xi^{(3)}}{\partial t} = -\frac{\partial v_3^f}{\partial x_3}.$$

In frequency domain PML formulation, after x_j is replaced by $\widetilde{x_j}$, equations for v_1^s, for example, can be rewritten as

$$\begin{aligned}
&(m\rho - \rho_f^2)\,(-i\omega)\left(1 + i\,\frac{\omega_1}{\omega}\right) v_1^{s(1)} \\
&\quad = 2m\,\frac{\partial}{\partial x_1}(\mu e_{11}) + m\,\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M\xi) \\
&\qquad - \rho_f\,\frac{\partial}{\partial x_1}(\alpha M e - M\xi) \\
&(m\rho - \rho_f^2)\,(-i\omega)\left(1 + i\,\frac{\omega_2}{\omega}\right) v_1^{s(2)} \\
&\quad = 2m\,\frac{\partial}{\partial x_2}(\mu e_{12}) + \left(1 + i\,\frac{\omega_2}{\omega}\right)\rho_f\,\frac{\eta}{\kappa}\,v_1^f \\
&(m\rho - \rho_f^2)\,(-i\omega)\left(1 + i\,\frac{\omega_3}{\omega}\right) v_1^{s(3)} \\
&\quad = 2m\,\frac{\partial}{\partial x_3}(\mu e_{13}).
\end{aligned}$$

By taking an inverse Fourier transform (FT), the above equations yield the time-domain PML equations

$$\begin{aligned}
&(m\rho - \rho_f^2)\left(\frac{\partial v_1^{s(1)}}{\partial t} + \omega_1 v_1^{s(1)}\right) \\
&\quad = 2m\,\frac{\partial}{\partial x_1}(\mu e_{11}) + m\,\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M\xi) \\
&\qquad - \rho_f\,\frac{\partial}{\partial x_1}(\alpha M e - M\xi)
\end{aligned} \tag{10}$$

$$\begin{aligned}
&(m\rho - \rho_f^2)\left(\frac{\partial v_1^{s(2)}}{\partial t} + \omega_2 v_1^{s(2)}\right) \\
&\quad = 2m\,\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f\,\frac{\eta}{\kappa}\left(v_1^f + \omega_2 \int_{-\infty}^{t} v_1^f\, d\tau\right)
\end{aligned} \tag{11}$$

$$\begin{aligned}
&(m\rho - \rho_f^2)\left(\frac{\partial v_1^{s(3)}}{\partial t} + \omega_3 v_1^{s(3)}\right) \\
&\quad = 2m\,\frac{\partial}{\partial x_3}(\mu e_{13})
\end{aligned} \tag{12}$$

$$\frac{\partial e_{11}}{\partial t} = \frac{\partial v_1}{\partial x_1} - \omega_1 e_{11} \tag{13}$$

$$\frac{\partial e_{12}^{(1)}}{\partial t} = \frac{1}{2}\,\frac{\partial v_2^s}{\partial x_1} - \omega_1 e_{12}^{(1)} \tag{14}$$

$$\frac{\partial e_{12}^{(2)}}{\partial t} = \frac{1}{2}\,\frac{\partial v_1^s}{\partial x_2} - \omega_2 e_{12}^{(2)} \tag{15}$$

$$\frac{\partial \xi^{(1)}}{\partial t} = -\frac{\partial v_1^f}{\partial x_1} - \omega_1 \xi^{(1)} \tag{16}$$

$$\frac{\partial \xi^{(2)}}{\partial t} = -\frac{\partial v_2^f}{\partial x_2} - \omega_2 \xi^{(2)} \tag{17}$$

$$\frac{\partial \xi^{(3)}}{\partial t} = -\frac{\partial v_3^f}{\partial x_3} - \omega_3 \xi^{(3)}. \tag{18}$$

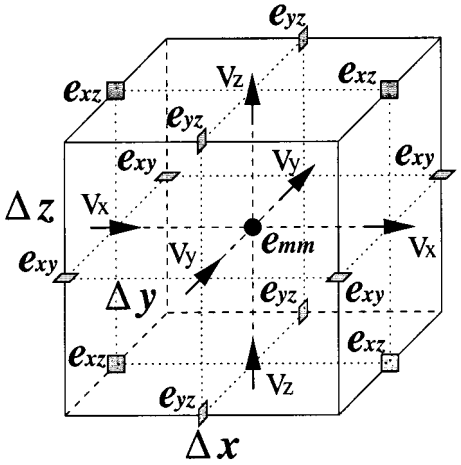


Fig. 1. Relative locations of field components in a unit cell of staggered grid.

Similar equations can be obtained in the same way for other components v_2^s, v_3^s, v_1^f, v_2^f, v_3^f, and e_{13}, e_{23}. Within the PML region, (3) and (4) are split into 18 equations for 3-D. Equation (5) is split into nine equations, and (6) is split into three equations. So the total number of the equations is 30 for the PML region, compared to 13 for the regular interior region. Thus, the memory requirement within the PML region is about two and half times that required by a regular medium for 3-D problems. This extra memory requirement in PML region is offset by the effectiveness of PML in absorbing the outgoing waves.

D. FD Implementation

The governing equations for the PML absorbing boundary condition, such as (10)–(18), are first-order partial differential equations for particle velocity and strain. They can be solved with different numerical methods. For the evaluation of seismic and acoustic responses, the FD method is widely used because of its flexibility and simplicity. Here, we use the explicit second-order FD method with a staggered grid in both spatial and temporal domains.

To implement a 3-D FD method to the PML equations, the material parameters and unknown field components are discretized on a regular 3-D grid at the intervals Δx_1, Δx_2 and Δx_3. The time domain is also discretized with time step Δt. For the FD implementation of Biot's equations on a staggered grid in Fig. 1, the velocity field components are located at the cell's face centers, while material parameters and normal strains are located at the center of the cell and shear strains are located at the six edge centers. The strain field is computed at $n\Delta t$ and velocity field is computed at $(n+1/2)\Delta t$. This staggered grid is similar to that for elastic waves in a solid [4], [22].

With this discretization, a leap-frog time-stepping system can be obtained. In order to simplify the layout of the formulas, the governing equations with PML boundary condition can be generalized as the first-order differential equation. For example, (11), (13), and (16) can be rewritten as

$$\frac{\partial v_1^{s(2)}}{\partial t} + c_0 v_1^{s(2)} = c_1 \int_{-\infty}^{t} v_1^f \, d\tau + c_2 \left[2m \frac{\partial}{\partial x_2} (\mu e_{12}) + \rho_f \frac{\eta}{\kappa} (v_1^f) \right] \quad (19)$$

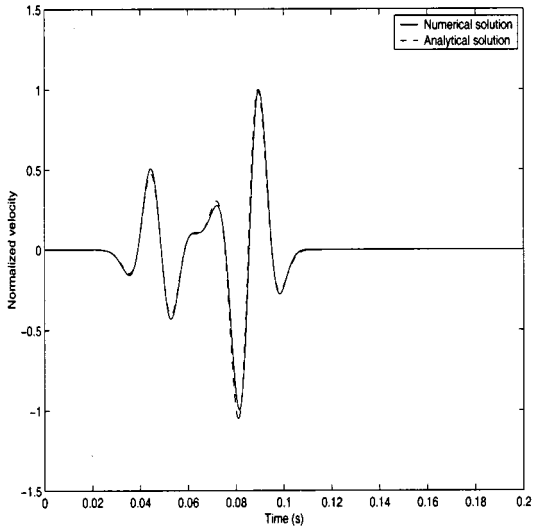


Fig. 2. Comparison of analytical and numerical solutions for the vertical velocity component in a homogeneous solid.

$$\frac{\partial e_{11}}{\partial t} + \omega_1 e_{11} = \frac{\partial v_2^s}{\partial x_1} \quad (20)$$

$$\frac{\partial \xi^1}{\partial t} + \omega_1 \xi^1 = -\frac{\partial v_1{}^f}{\partial x_1} \quad (21)$$

where c_0, c_1, and c_2 space-dependent coefficients. Then the time-stepping equations can be written as

$$v_1^{s(2)}\left[j_1, j_2, j_3, \left(n+\tfrac{1}{2}\right)\right] = f_1 v_1^{s(2)}\left[j_1, j_2, j_3, \left(n-\tfrac{1}{2}\right)\right] + f_2 R_1 \quad (22)$$

$$e_{11}[j_1, j_2, j_3, (n+1)] = g_1 e_{11}[j_1, j_2, j_3, n] + g_2 R_2 \quad (23)$$

$$\xi^1[j_1, j_2, j_3, (n+1)] = g_1 \xi^1[j_1, j_2, j_3, n] - g_2 R_3 \quad (24)$$

$$f_1 = \frac{c_0/2 - 1/\Delta t}{1/\Delta t + c_0/2}$$

$$f_2 = \frac{1}{1/\Delta t + c_0/2}$$

$$g_1 = \frac{\omega_1/2 - 1/\Delta t}{1/\Delta t + \omega_1/2}$$

$$g_2 = \frac{1}{1/\Delta t + \omega_1/2}$$

where R_1, R_2, and R_3 are right-hand sides of equations (19), (20), and (21), respectively. It should be noted that the material parameters in the above equations must be properly averaged in order to arrive at a higher accuracy [4]. In order to save computer storage, the computational domain is divided into a PML region and an interior region. The absorption of outgoing waves is achieved by the PML region, which consists of several cells of PML materials with a quadratically tapered ω_j profile to in-

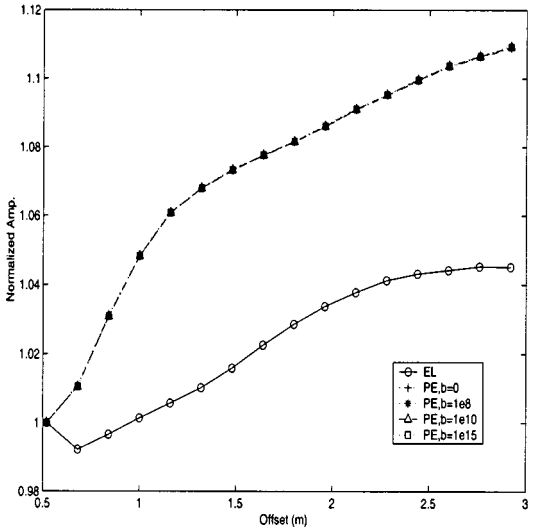


Fig. 3. Comparison between the elastic and poroelastic model for dry sand.

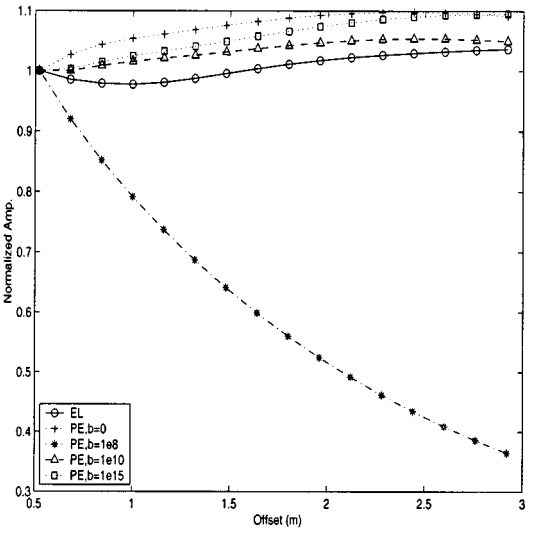


Fig. 4. Comparison between the elastic and poroelastic model for fully water-saturated sand.

crease the attenuation toward the outer boundary. In this paper, ω_j of the PML region is chosen as

$$\omega_j(j) = \frac{(M-1/2-j)^2}{(M-1/2)^2}\,\omega_{j,\max} \tag{25}$$

where $\omega_{j,\max}$ is the value at the center of the cell at the outermost boundary. At the outer boundary, the velocity components are forced to be zero. For convenience, $\omega_{j,\max}$ can be expressed in terms of dominant frequency and normalized coefficient. Then (25) becomes

$$\omega_j = \begin{cases} 2\pi a_0 f_0 \left(l_{x_j}/L_{\mathrm{PML}}\right)^2, & \text{inside PML,} \\ 0, & \text{outside PML} \end{cases} \tag{26}$$

where

f_0 dominant frequency of the source;
L_{PML} thickness of the PML region;
l_{x_j} distance from the interface between the interior region and PML region.

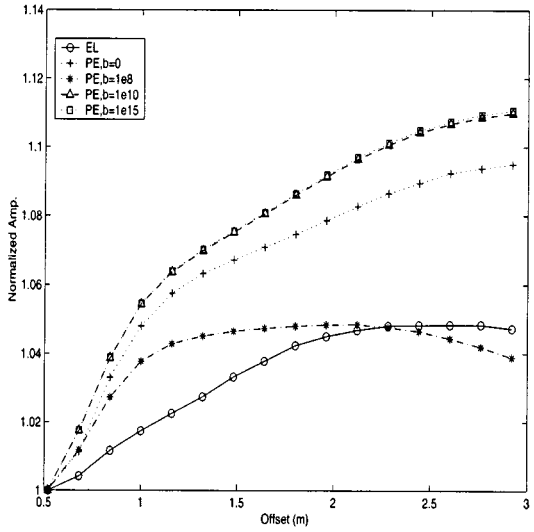


Fig. 5. Comparison between the elastic and poroelastic model for partly water-saturated sand.

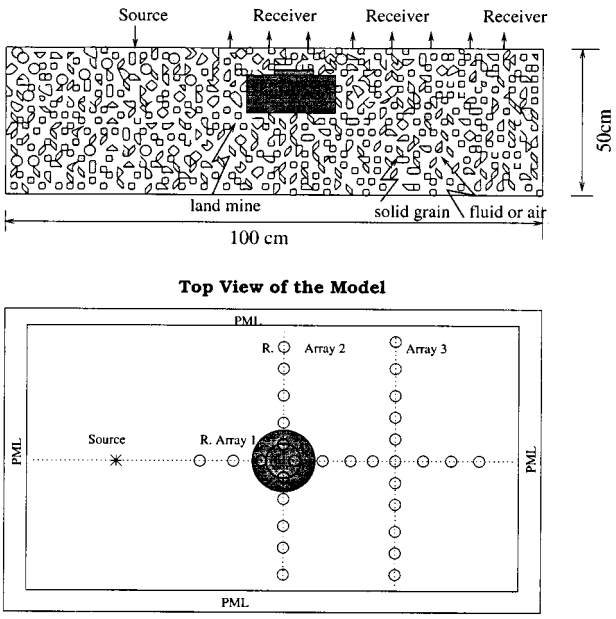


Fig. 6. Geometry of a buried land mine model.

III. Numerical Results

In the following numerical simulations, the source energy is partitioned linearly between the solid and the fluid phases with factors

$$W_f = \phi$$
$$W_s = (1-\phi)$$
$$W_r = \phi|W_f - W_s|$$

where

W_f weighting factor for the fluid motion;
W_s solid motion;
W_r relative motion between solid frame and pore fluid.

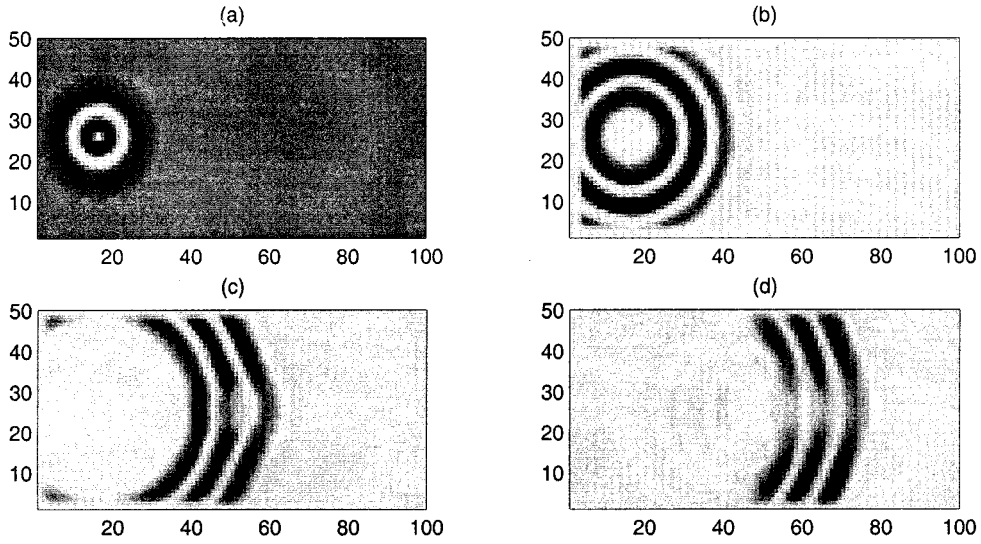


Fig. 7. Snapshots at (a) 2.3 ms, (b) 4.0 ms, (c) 6.0 ms, and (d) 7.8 ms. A dipole source is on the free surface.

A. Validation of Numerical Results

An analytical solution for the particle velocity field in a homogeneous, fluid-saturated poroelastic medium subject to a point source in 3-D space or a line source in 2-D space can be obtained in a closed form via potential functions [9]. In a poroelastic medium with an ideal nonviscous fluid, a purely dilatational source will only excite P waves given by scalar potentials

$$\psi_s(r,\,t) = \frac{\alpha s\left(t - \frac{r}{V_f}\right) + \beta s\left(t - \frac{r}{V_s}\right)}{4\pi r}$$

$$\psi_f(r,\,t) = \frac{\alpha A_f s\left(t - \frac{r}{V_f}\right) + \beta A_s s\left(t - \frac{r}{V_s}\right)}{4\pi r}$$

where

r	distance;
$s(t)$	source time function;
A_f and A_s	ratios between the solid and fluid motion for the fast P-wave and the slow P-wave, and the coefficients;
α and β	determined by the regularity conditions;
V_f and V_s	velocities of the fast P-wave and the slow P-wave, respectively.

The displacement field in the solid and fluid is then written respectively as $\mathbf{u} = \nabla\psi_s$ and $\mathbf{U} = \nabla\psi_f$.

The validation of the numerical method can be done by comparing the numerical results with the above analytical solution. A homogeneous model of sand saturated with water is considered. A P-wave point source of the first derivative Gaussian time function with $f_0 = 40$ Hz is located at (0,0,0). Then the solution at (30 m, 30 m, 5 m) is calculated. The numerical result and analytical solution in Fig. 2 have an excellent agreement.

B. Comparison Between Elastic and Poroelastic Models

In the conventional elastic model, a single-phase medium is considered. There is no energy dissipation in such an elastic medium. In poroelastic model, however, a multiphase medium (usually solid frame and pore fluid) is considered. The in-phase motion between solid and fluid leads to the regular P-wave and S-wave, and the out-phase motion between solid and fluid leads to the slow P wave. Because the fluid is viscous, there is energy dissipation between the solid frame and the pore fluid. In Biot's theory, the ratio of viscosity and permeability ($b = \eta/\kappa$)

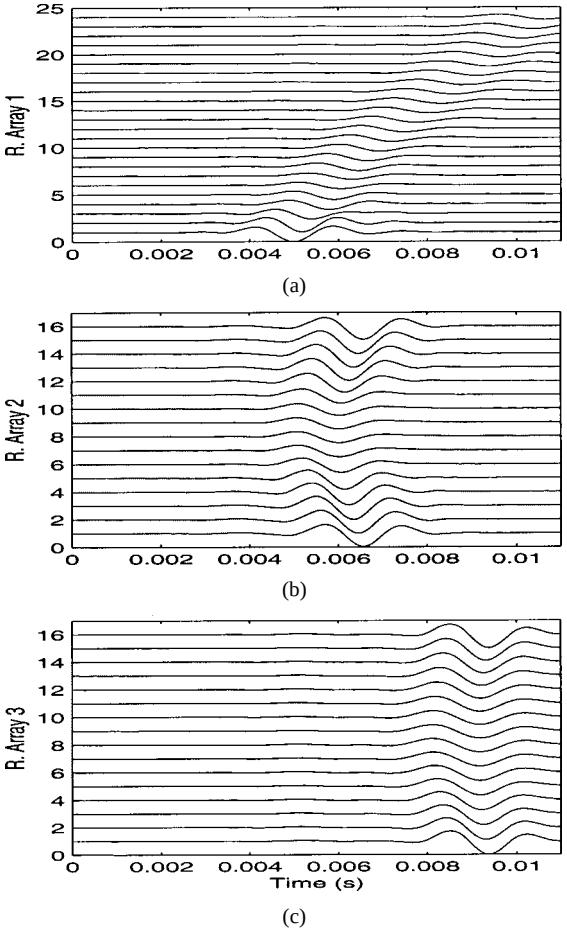


Fig. 8. Waveforms of vertical velocity of solid frame at receiver (a) array 1, (b) array 2, (c) array 3. A dipole source is on the free surface.

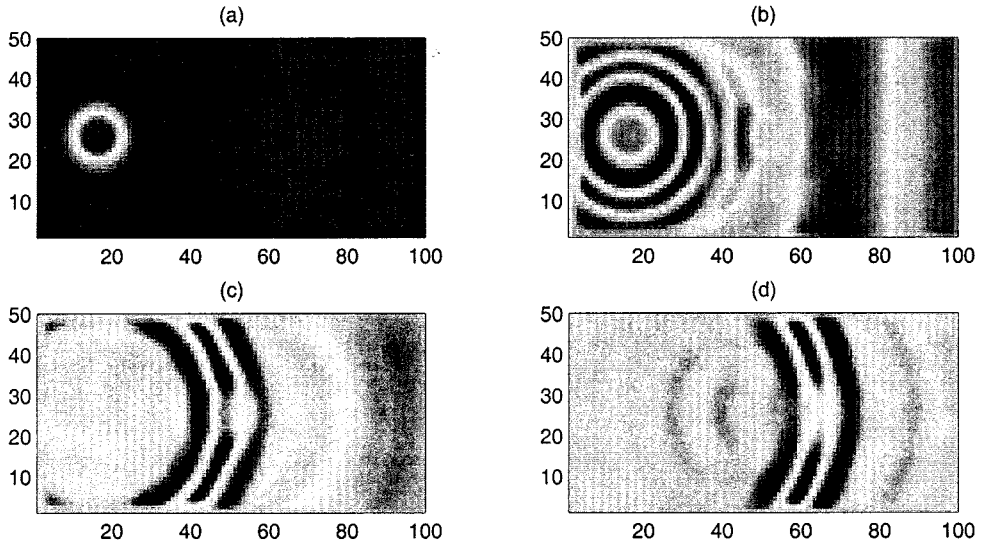


Fig. 9. Snapshots at (a) 2.3 ms, (b) 4.0 ms, (c) 6.0 ms, and (d) 7.8 ms. A monopole source is in the air.

determines this dissipation, as shown in (3) and (4). The difference between an elastic model and a poroelastic model is investigated here by studying the effect of b on AVO response in three different ground media: dry sand, fully water saturated sand, and partly water saturated sand (70% air and 30% water). For a half-space problem, since the source and receivers are usually on the free surface and the surface wave is the dominant signal at low frequencies, only the surface wave is investigated here. In this case, we model the source as a vertical dipole right on the ground surface and displaced horizontally away from the buried object, a scenario similar to the source in [1].

Figs. 3–5 show the effects of b on the surface wave AVO for dry sand, fully water saturated sand, and partly water-saturated sand models, respectively. For the dry sand, the difference between elastic and poroelastic models is obvious, but b has little effect on AVO. For the fully water saturated sand model, b has a significant effect on AVO. When b is very large, the AVO response is very close to that of the elastic model. For the partly water saturated sand model, the difference between the elastic and poroelastic model is obvious, and b also has a significant effect on AVO.

C. Acoustic Land Mine Detection Systems

This method is used to model a plastic antipersonnel mine that is difficult to detect with the more conventional electromagnetic induction sensors (i.e., metal detectors). The geometry of the model of a buried mine is shown in Fig. 6. The source can be either on the surface or in the air. There are three arrays of receivers on the surface to pick up the signal of vertical velocity.

The parameters for the plastic mine are chosen as $\rho = 1200$ kg/m^3, S-wave velocity $v_{sh} = 1100$ m/s and P-wave velocity $v_p = 2700$ m/s. The material in the ground may be considered as a two-phase composite material consisting of granular solid and pore fluid. The nature of this composite varies with environment, geographic location, and with depth below the surface of ground. In this model, the soil is chosen as dry sand and the porosity is 0.35. The fast P wave velocity is 250 m/s and S wave velocity is 87 m/s.

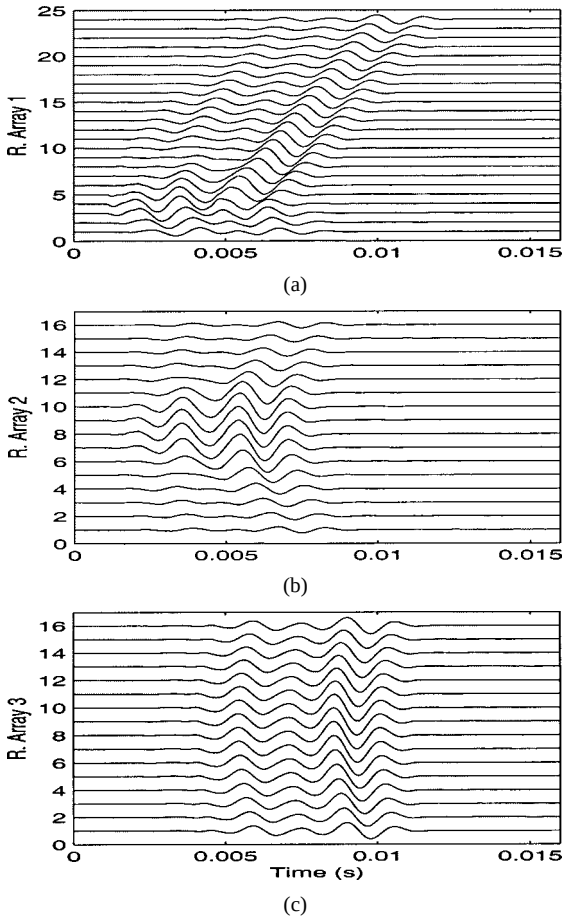


Fig. 10. Waveforms of scattered field of solid frame at receiver (a) array 1, (b) array 2, (c) array 3. A monopole source is in the air.

1) Source on the Ground Surface: To model the land-mine detection system in [1], a half space with a dipole source on the ground surface is simulated. The source time function is the Blackman–Harris window function with the center frequency 450 Hz. Fig. 7 shows the four snapshots of the total field on the free ground surface at 2.3 ms, 4.0 ms, 6.0 ms and 7.8 ms. The

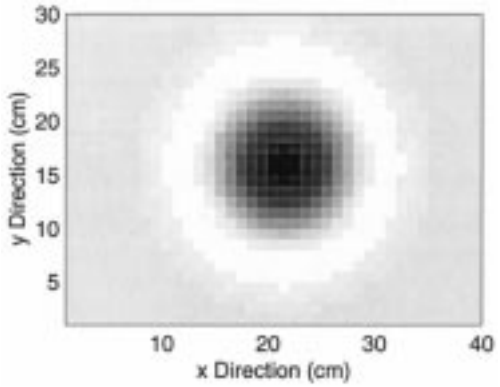


Fig. 11. Energy distribution of the scattered field when the source is in the air above the mine.

primary wave on these snapshots are surface waves. Compared to the surface waves, the P wave is very weak. The obvious surface wave reflections by the mine are shown on the snapshots at 6.0 ms and 7.8 ms. The influence of the buried mine on the right-propagating surface wave can be seen clearly on the snapshots. There is a decrease on amplitude when the surface wave passes over the buried mine. Fig. 8(a) shows the waveforms of receiver array 1. The reflected surface waves are presented on the first five traces. Fig. 8(b) and (c) show the waveforms of receiver arrays 2 and 3 respectively. The decrease of the amplitude can be easily seen on the time traces of receiver 6 to 11.

2) Source in the Air: Another land-mine detection prototype uses airborne acoustic waves to couple into the ground surface to detect minelike targets [2]. In our simulation, the model consists of two parts: air (pure fluid) and soil (poroelastic medium) with a plastic land mine. A monopole source is in the air 2 cm above the ground surface. The locations of the receivers are the same as those in the previous model. Fig. 9 shows the four snapshots of vertical velocity field on the ground surface at 2.3 ms, 4.0 ms, 6.0 ms and 7.8 ms. We can see that if a minelike target is present below the surface, the ground vibrational velocity will show distinct changes due to reflection and scattering of surface waves. Fig. 10(a)–(c) show time traces of the scattered field at receiver arrays 1, 2, and 3 respectively. The signature of the buried land mine is clearly present in these figures. It is interesting to note that both direct arrival and the surface waves are scattered by the object into surface waves.

The above results are shown in time domain. However, the prototype in [2] actually displays the acoustic energy in the frequency domain. This display can also be easily done by our time-domain solution. Displayed in Fig. 11 is the energy distribution of the scattered field for a monopole source directly above the mine, showing the presence of the buried object.

IV. Conclusion

A particle velocity-strain, FD method combined with the perfectly matched layer (PML) has been developed for the simulation of acoustic waves propagating in 3-D poroelastic media. The results show the scheme is stable even if the ratio of the fast P wave velocity to shear wave velocity is greater than 2.

The difference between elastic model and poroelastic model and moisture effect are investigated by studying surface wave amplitude variation with offset (AVO) in three different ground media: dry sand, fully water saturated sand and partly water saturated sand. The difference between these models is significant. For air saturated sand, the attenuation factor has little influence on AVO. For fully or partly water saturated sand, the attenuation factor greatly affects the AVO response.

This numerical method has been used to investigate the interaction of acoustic waves with a buried mine-like object. Two land-mine detection prototypes are simulated: a half space model with a dipole source on the free ground surface and a half space model with a monopole source in the air. The results show that the wave responses are substantially changed by the presence of the mine-like object, and surface waves play as a primary mechanism of detection at the low frequencies.

References

[1] W. R. Scott Jr., C. T. Schröder, and J. S. Martin, "An hybrid acousto/electromagnetic technique for locating land mines," in *Proc. Int. Geoscience and Remote Sensing Symp.*, Seattle, WA, 1998, pp. 216–218.

[2] J. M. Sabatier and N. Xiang, "Systematic investigation on acoustic-to-seismic responses of landmines buried in soil," *J. Acoust. Soc. Amer.*, vol. 107, p. 2896, 1999.

[3] J. Virieux, "SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method," *Geophysics*, vol. 49, pp. 1933–1957, 1986.

[4] Q. H. Liu, F. Daube, C. Randall, E. Schoen, H. Liu, and P. Lee, "A 3D finite difference simulation of sonic logging," *J. Acoust. Soc. Amer.*, vol. 100, pp. 72–79, 1996.

[5] Q. H. Liu and C. Chang, "Compressional head waves in attenuative formations: Forward modeling and inversion," *Geophysics*, vol. 61, pp. 1908–1920, 1996.

[6] M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. 1. Low-frequency range," *J. Acoust. Soc. Amer.*, vol. 28, pp. 168–178, 1956.

[7] ——, "Theory of propagation of elastic waves in a fluid-saturated porous solid. 1. High-frequency range," *J. Acoust. Soc. Amer.*, vol. 28, pp. 179–191, 1956.

[8] ——, "Mechanics deformation and acoustic propagation in porous media," *J. Appl. Phys.*, vol. 33, pp. 1482–1498, 1962.

[9] N. Dai, A. Vafidis, and E. R. Kanasewich, "Wave propagation in heterogeneous, porous media: A velocity-stress, finite-difference method," *Geophysics*, vol. 60, pp. 327–340, 1995.

[10] C. Cerjan, D. Kosloff, R. Kosloff, and M. Reshef, "A nonreflecting boundary condition for discrete acoustic and elastic wave equations," *Geophysics*, vol. 50, pp. 705–708, 1985.

[11] R. Clayton and B. Engquist, "Absorbing boundary conditions for acoustic and elastic wave equations," *Bull. Seismol. Soc. Amer.*, vol. 67, pp. 1529–1540, 1977.

[12] K. D. Mahrer, "An empirical study of instability and improvement of absorbing boundary conditions for elastic wave equation," *Geophysics*, vol. 51, pp. 1499–1501, 1986.

[13] J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, vol. 114, pp. 185–200, 1994.

[14] W. C. Chew and Q.H. Liu, "Perfectly matched layers for elastodynamics: A new absorbing boundary condition," *J. Comput. Acoust.*, vol. 4, pp. 72–79, 1996.

[15] F. D. Hastings, J. B. Schneider, and S. L. Broschat, "Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation," *J. Acoust. Soc. Amer.*, vol. 100, pp. 3061–3069, 1996.

[16] Q. H. Liu and J. Tao, "The perfectly matched layer for acoustic waves in absorptive media," *J. Acoust. Soc. Amer.*, vol. 102, pp. 2072–2082, 1997.

[17] Q. H. Liu, "Perfectly matched layers for elastic waves in cylindrical and spherical coordinates," *J. Acoust. Soc. Amer.*, vol. 105, pp. 2075–2084, 1999.

[18] Y. Q. Zeng, J. He, and Q. H. Liu, "The application of the perfectly matched layer in numerical modeling of wave propagation in poroelastic media," Res. Rep., Duke Univ., Durham, NC, 1999.

[19] R. Burridge and J. B. Keller, "Poroelasticity equations derived from microstructure," *J. Acoust. Soc. Amer.*, vol. 70, pp. 1140–1146, 1981.

[20] J. Plona, "Observation of the second bulk compressional wave in a porous medium at ultrasonic frequencies," *Appl. Phys. Lett.*, vol. 36, pp. 259–261, 1980.
[21] G. Kneib and C. Kerner, "Accurate and efficient seismic modeling in random media," *Geophysics*, vol. 58, pp. 576–588, 1993.
[22] Q. H. Liu, "Some current trends in numerical methods for transient acoustic and elastic waves in multidimensional inhomogeneous media," *Current Topics Acoust. Res.*, vol. 2, pp. 31–42, 1998.



Yan Qing Zeng (S'00) received the B.S. degree in 1987, the M.S. degree in geophysics in 1990 from Peking University, Beijing, China, and the M.S. degree in civil and environmental engineering from Duke University, Durham, NC, in 1999, where he is currently pursuing the Ph.D. degree in electrical and computer engineering.

He was a Faculty Member, Department of Geophysics, Peking University, from 1990 to 1995. His interests include computational acoustics and electromagnetics, and numerical modeling of wave propagation in poroelastic media.

Qing Huo Liu (S'88–M'89–SM'94) received the Ph.D. degree in electrical engineering from the University of Illinois, Urbana, in 1989.

He was with the Electromagnetics Laboratory, University of Illinois, as a Research Assistant from September 1986 to December 1988, and as a Postdoctoral Research Associate from January 1989 to February 1990. He was a Research Scientist and Program Leader with Schlumberger-Doll Research, Ridgefield, CT from 1990 to 1995, and from October 1995 to May 1999, he was with New Mexico State University, Las Cruces. Since June 1999, he has been an Associate Professor of Electrical Engineering with Duke University, Durham, NC. He has published more than 170 papers in refereed journals and conference proceedings. His research interests include computational electromagnetics and acoustics, wave propagation in inhomogeneous media, geophysical subsurface sensing, and inverse problems.

Dr. Liu is a member of Phi Kappa Phi, Tau Beta Pi, SEC, and a full member of the U.S. National Committee of URSI Commissions B and F. Currently, he serves as an Associate Editor for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, for which he also served as a Guest Editor for a special issue on computational methods. He received the Presidential Early Career Award for Scientists and Engineers (PECASE) from the National Science and Technology Council, the Early Career Research Award from the Environmental Protection Agency in 1996, and a CAREER Award from the National Science Foundation in 1997.

Anal. Chem. **2001,** *73,* 884−892

Detection and Classification Characteristics of Arrays of Carbon Black/Organic Polymer Composite Chemiresistive Vapor Detectors for the Nerve Agent Simulants Dimethylmethylphosphonate and Diisopropylmethylphosponate

Alan R. Hopkins and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Arrays of conducting polymer composite vapor detectors have been evaluated for performance in the presence of the nerve agent simulants dimethylmethylphosphonate (DMMP) and diisopropylmethylphosponate (DIMP). Limits of detection for DMMP on unoptimized carbon black/organic polymer composite vapor detectors in laboratory air were estimated to be 0.047−0.24 mg m^{-3}. These values are lower than the EC_{50} value (where EC_{50} is the airborne concentration sufficient to induce severe effects in 50% of those exposed for 30 min) for the nerve agents sarin (methylphosphonofluoridic acid, 1-methylethyl ester) and soman (methylphosphonofluoridic acid, 1,2,2-trimethylpropyl ester), which has been established as ∼0.8 mg m^{-3}. Arrays of these vapor detectors were easily able to resolve signatures due to exposures to DMMP from those due to DIMP or due to a variety of other test analytes (including water, methanol, benzene, toluene, diesel fuel, lighter fluid, vinegar, and tetrahydrofuran) in a laboratory air background. In addition, DMMP at 27 mg m^{-3} could be detected and differentiated from the signatures of the other test analytes in the presence of backgrounds of potential interferences, including water, methanol, benzene, toluene, diesel fuel, lighter fluid, vinegar, and tetrahydrofuran, even when these interferents were present in much higher concentrations than that of the DMMP or DIMP being detected.

Arrays of chemically sensitive resistors fabricated from composites of carbon black with insulating organic polymers have received significant attention recently for use in detecting, quantifying, and discriminating among various organic vapors.[1−4] One of the potential advantages of this approach to vapor sensing[5−20] is that it affords the possibility of fabricating very small size, low-power, and lightweight detector arrays that have hundreds of compositionally different detector pixels. For example, hundreds of such detectors have been deposited and read out in our laboratory on a Si chip having an area of 0.04 cm^2;[21,22] thus, one could envision developing wearable badge-type detectors for personal environmental monitoring devices based on this technology.

In this work, we have evaluated the detection and discrimination capabilities of an array of carbon black/organic polymer composite chemiresistors toward the nerve agent simulants dimethylmethylphosphonate (DMMP)[23,24] and diisopropylmeth-

(1) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298−2312.
(2) Doleman, B. J.; Sanner, R. D.; Severin, E. J.; Grubbs, R. H.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 2560−2564.
(3) Swann, M. J.; Glidle, A.; Cui, L.; Barker, J. R.; Cooper, J. M. *Chem Commun.* **1998**, *24*, 2753−2754.
(4) Patel, S. V.; Jenkins, M. W.; Hughes, R. C.; Yelton, W. G.; Ricco, A. J. *Anal. Chem.* **2000**, *72*, 1532−1542.
(5) Gardner, J. W.; Pearce, T. C.; Friel, S.; Bartlett, P. N.; Blair, N. *Sens. Actuators, B* **1994**, *18*, 240−243.
(6) Pearce, T.; Gardner, J. *Analyst* **1998**, *123*, 2047−2055.
(7) Gardner, J. W.; Gardner, W. E. *Insight* **1997**, *39*, 865−869.
(8) Pearce, T. C.; Gardner, J. W.; Friel, S.; Bartlett, P. N.; Blair, N. *Analyst* **1993**, *118*, 371−377.
(9) Baldacci, S.; Matsuno, T.; Toko, K.; Stella, R.; De Rossi, D. *Sens. Mater.* **1998**, *10*, 185−200.
(10) Hodgins, D.; Simmonds, D. *J. Autom. Chem.* **1995**, *17*, 179−185.
(11) Hirschfelder, M.; Ulrich, D.; Hoberg, E.; Hanrieder, D. *Gartenbauwissenschaft* **1998**, *63*, 185−190.
(12) Simon, J. E.; Hetzroni, A.; Bordelon, B.; Miles, G. E.; Charles, D. J. *J. Food Sci.* **1996**, *61*, 967.
(13) Broda, S.; Schnitzler, W. *Dtsch. Lebensm.-Rundsch.* **1998**, *94*, 13−16.
(14) Stuetz, R. M.; White, M.; Fenner, R. A. *J. Water Serv. Res. Technol.-Aqua* **1998**, *47*, 223−228.
(15) Stuetz, R. M.; Engin, G.; Fenner, R. A. *Water Sci. Technol.* **1998**, *38*, 331−335.
(16) Stuetz, R. M.; Fenner, R. A.; Engin, G. *Water Res.* **1999**, *33*, 453−461.
(17) Stuetz, R. M. Fenner, R. A., Engin, G. *Water Res.* **1999**, *33*, 442−452.
(18) Chandiok, S.; Crawley, B.; Oppenheim, B.; Chadwick, P.; Higgins, S.; Persaud, K. *J. Clin. Pathol.* **1997**, *50*, 790−791.
(19) Gardner, J. W.; Craven, M.; Dow, C.; Hines, E. L. *Meas. Sci. Technol.* **1998**, *9*, 120−127.
(20) Persaud, K.; Pisanelli, A.; Szyszko, S.; Reichl, M. J.; Horner, G.; Rakow, W.; Keding, H.; Wessels, H. *Sens Actuators, B* **1994**, *18*, 211−220.
(21) Dickson, J. A.; Freund, M. S.; Lewis, N. S.; Goodman, R. M. *Solid State Sensor and Actuator Workshop*, Hilton Head Island, SC, June 2−8, 2000.
(22) Briglin, S. M.; Burl, M. C.; Freund, M. S.; Lewis, N. S.; Matzger, A.; Ortiz, D. N.; Tokumaru, P. Proc. SPIE 14th Annual International Symposium on Aerospace/Defence Sensing, Simulation, and Controls, 2000; Vol. 4038.
(23) Grate, J. W.; Rose-Pehrsson, S. L.; Venezky, D. L.; Klusty, M.; Wohltjen, H. *Anal. Chem.* **1993**, *65*, 1868−1881.

10.1021/ac0008439 CCC: $20.00 © 2001 American Chemical Society
Published on Web 02/28/2001

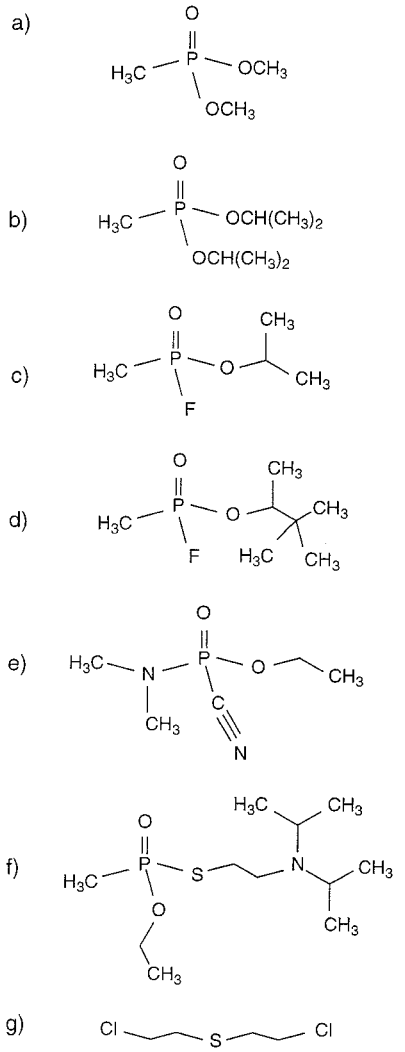


Figure 1. Structures of the chemical warfare simulants (a) dimethylmethylphosphonate (DMMP) and (b) diisopropylmethylphosphonate (DIMP) and of the nerve agents (c) sarin (methylphosphonofluoridic acid, 1-methylethyl ester), (d) soman (methylphosphonofluoridic acid, 1,2,2-trimethylpropyl ester), (e) tabun (dimethylphosphoramidocyanidic acid, ethyl ester), (f) VX (methylphosphonothioic acid, *S*-[2-[bis-(1-methylethyl)amino]ethyl]-*O*-ethyl ester), and (g) sulfur mustard (1,1′-thiobis(2-chloroethane)).

ylphosphonate (DIMP) (Figure 1). Table 1 indicates some chemical warfare agents and concentrations of interest.[25] DMMP is often regarded as a simulant for sarin (methylphosphonofluoridic acid, 1-methylethyl ester) and DIMP as a simulant for soman (methylphosphonofluoridic acid, 1,2,2-trimethylpropyl ester), so DMMP and DIMP have been the focus of our initial efforts in this area. In this work, we report the detection limits of these detector arrays, under controlled laboratory conditions, for DMMP or DIMP in background air, as well as for DMMP or DIMP in the presence of a variety of different background analytes including water vapor, diesel fuel, organic solvent vapors, and other selected possible interferences. Additionally, we have evaluated the ability

(24) Ballantine, D. S.; Rose, S. L.; Grate, J. W.; Wohltjen, H. *Anal. Chem*. **1986**, *58*, 3058−3066.

(25) *Chemical and Biological Terrorism Research and Development to Improve Civilian Medical Response*, Institute of Medicine and National Research Council, National Academy Press: Washington, DC, 1999.

Table 1. Human Exposure Guidelines for Selected Chemical Warfare Agents (mg m^{-3})

agent	AEL[a]	EC_{50}[b]	ED_{50}[c]
sarin	0.0001	<0.8	1000
soman	0.00003	<0.8	200
tabun	0.0001	<1.7	<880
VX	0.00001	<0.3	<2.5
sulfur mustard	0.003	3.33	600

[a] AEL, the maximum airborne exposure concentration for an 8-h workday. [b] EC_{50}, the airborne concentration sufficient to induce severe effects in 50% of those exposed for 30 min. [c] ED_{50}, the amount of liquid agent on the skin sufficient to produce severe effects in 50% of the exposed population.

Table 2. Polymers Used in the Carbon Black-Polymer Composite Detector Array

detector no.	polymer
1	poly(ethylene oxide), PEO
2	poly(ethylene-*co*-vinyl acetate), 45% vinyl acetate, PEVA
3	poly(butadiene-*co*-styrene), 72% butadiene, PBS
4	poly(vinylcarbazole), PVC
5	poly(vinyl acetate), PVA
6	poly(caprolactone), PCL
7	polysulfone
8	poly(vinylpyrrolidone), PVP
9	poly(4-vinylphenol), PVPH
10	poly(methyloctadecylsiloxane), PMODS

of these detectors to discriminate among DMMP, DIMP, and a collection of other selected organic vapors under the measured test conditions.

EXPERIMENTAL SECTION

The carbon black used in the composites was Black Pearls 2000 (BP2000), a furnace black material that was generously donated by Cabot Co. (Billerica, MA). The polymer used in the composites (Table 2) was (detector number, polymer, manufacturer): 1, poly(ethylene oxide), Polysciences; 2, poly(55% ethylene-*co*-45%vinyl acetate), Polysciences; 3, poly(72% butadiene-*co*-28% styrene), Scientific Polymer Products; 4, poly(vinylcarbazole), Polysciences, 5, poly(vinyl acetate), Scientific Polymer Products; 6, poly(caprolactone), Polysciences; 7, polysulfone, Polysciences; 8 poly(vinylpyrrolidone), Scientific Polymer Products; 9, poly(4-vinylphenol), Polysciences; 10, poly(methyloctadecylsiloxane), Polysciences. To form contacts on the substrate, two parallel bands of gold, 50−100 nm thick and separated by 1 mm, were deposited onto conventional 7.5 cm × 2.5 cm glass slides (Corning Inc.). The slides were then cut into strips to produce 0.7 cm × 2.5 cm pieces of glass. The detector films were made from a solution of the polymer into which carbon black had been suspended. A 160-mg sample of one of the insulating polymers was dissolved in 20 mL of solvent and 40 mg of carbon black was then suspended in this solution, to produce a composition of 80% polymer and 20% carbon black by weight of solids. The solvent was either toluene or tetrahydrofuran (THF), depending on the solubility of the polymer. The solutions were sonicated for 5 min to suspend the carbon black. A single solution that contained the polymer and the carbon black was used to prepare all the detectors of a given composition that were used in this work. An aliquot of the

suspension was spin coated, at 1500 rpm, onto a glass substrate using a Headway (Garland, TX) spin coater, and the resulting film was allowed to dry in air for 24 h. Multiple coatings of the suspension were applied to each substrate to yield detectors having resistance values of approximately a few hundred kiloohms.

The solvents used in this study were THF, benzene, methanol, toluene, DMMP, and DIMP. These solvents were purchased from Aldrich and were used without further purification. The vinegar (Lucky Brand), lighter fluid (butane and other low molecular weight hydrocarbons), and diesel fuel (Unocal 76) were purchased from commercial consumer sources.

The apparatus used to generate known concentrations of organic vapors and to acquire resistance versus time data from the detectors has been described previously.[1,2,26] Briefly, solvents were saturated with a background of laboratory air and the resultant flowing gas stream was mixed with an independently controlled flow of air to produce the desired diluted flow of analyte vapor. The gas flow rates were controlled with calibrated mass flow controllers, and analyte saturation of the flow stream that passed through the solvent bubblers was verified by measurement of the rate of mass loss of the solvent in the bubbler.[27] Solenoids and one-way valves were used to mix the gases and to direct them to a Teflon chamber that contained the conducting polymer composite detectors.[28] The performance of the entire vapor generation system was validated using a flame ionization detector to determine the concentration of a series of standard mixtures of toluene in air that were delivered to the detector chamber. These values were within 90% of the values expected based on the vapor pressure of the analyte and the dilution ratio expected from the settings of the mass flow controllers. The flame ionization detector was calibrated using standard gas calibration mixtures obtained from commercial suppliers.

The carrier gas for all experiments was oil-free air, obtained from the general compressed laboratory source which contained 1.10 ± 0.15 parts per thousand (ppth) of water vapor. The air was filtered to remove particulates but deliberately was not dehumidified or otherwise purified. When water was used as the background analyte, the carrier gas was ultrazero air. Fluctuations in laboratory temperature, 21.5 ± 1.5 °C, could cause a ~10% error in setting and controlling the vapor concentrations between nominally identical exposures over the course of the data collection analyzed in this work. No temperature control of the apparatus or of the carbon black/polymer composite detectors was performed. The flow rate of the vapor stream entering the exposure chamber (~1 L in total volume) was maintained at 15 L min^{-1}.

The 20-detector array studied in this work had 2 nominally identical copies of each of 10 different polymer composites (Table 2). To initiate an experiment, the detectors were placed into the flow chamber and a background flow of laboratory air was introduced until the resistance of the detectors stabilized. Each exposure consisted of a three-step process that began with 180 s of air flow to achieve a smooth baseline resistance. After this period, the detectors were exposed for 300 s to analyte vapor at a controlled concentration in flowing air. The analyte exposure

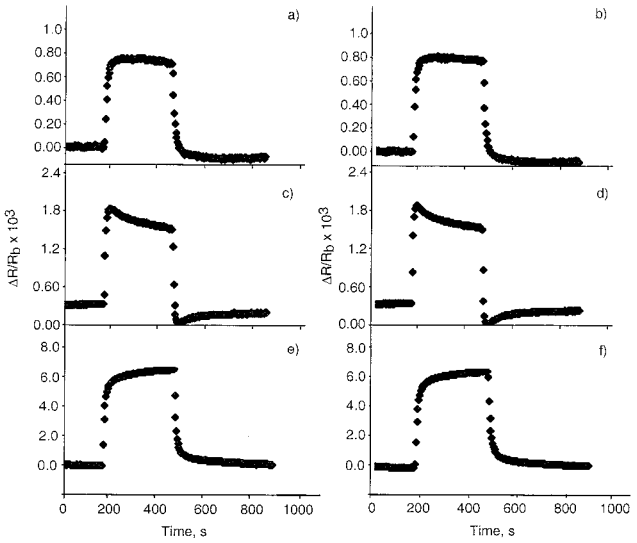


Figure 2. Relative differential resistance responses vs time for carbon black/polymer composite detectors exposed to DMMP at $P/P^o = 0.013$ in laboratory air. Replicate entries of the polymeric component of the carbon black composite indicate the responses of two different detectors of that were fabricated together and were exposed together to analyte. Each exposure consisted of 180 s of air, 300 s of analyte, and then 420 s of laboratory air. The polymer used to form the carbon black composite was (a) PEO, (b) PEO, (c) PEVA, (d) PEVA, (e) PCL, and (f) PCL.

was then followed by a 420-s flow of clean air to restore the baseline resistance values. Within each experiment, every exposure was assigned a randomly generated index number using the Microsoft Excel random number generator. The exposures were then presented to the detector array in ascending order of the assigned index values.

To obtain signatures for the different analytes of interest, each analyte was exposed to the vapor detector array in random order at a fixed value of P/P^o, where P^o is the vapor pressure of the analyte at room temperature and P is the partial pressure of the analyte. In this study, the 180-s baseline period was followed by 300 s of exposure to the test analyte at $P/P^o = 0.010$. Representative resistance versus time response signatures of selected detectors, and a summary of the output data for different carbon black/polymer detectors exposed to DMMP in background air, are presented in Figure 2 and Table 3, respectively.

A separate run was performed to probe the detector response to a test analyte in the presence of various background compounds. In this run, the exposure phase consisted of two parts. The background analyte was exposed at $P/P^o = 0.010$ for 150 s, at which time the test analyte was introduced and exposed for an additional 150 s. During the exposure of the second (test) analyte, the first (background) analyte was continually flowing. The foreground test analytes DMMP and DIMP were presented to the array at $P/P^o = 0.0017$, 0.0054, or 0.013. Each unique foreground/background combination was replicated 8 times. The total experiment thus contained 192 total exposures (2 foreground solvents each at 3 different concentrations in the presence of a selected member of the 4 background solvents, with 8 replicates of each unique exposure combination).

Although the resistance of each detector was sampled once every 3−5 s during each exposure, only the relative differential

(26) Sotzing, G. A.; Briglin, S. M.; Grubbs, R. H.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 3181−3190.
(27) Atkins, P. W. *Physical Chemistry*, 5th ed.; W. H. Freeman and Co.: New York, 1994.
(28) Severin, E. J. Ph.D. Thesis, California Institute of Technology, 1999.

Table 3. $\Delta R/R_b$ Responses of Carbon Black/Polymer Detectors to DMMP in Background Air[a]

	activity, P/P^o of DMMP		
host material	0.0017	0.0054	0.013
PEO	$(13.2 \pm 6.7) \times 10^{-5}$	$(4.35 \pm 5.6) \times 10^{-4}$	$(9.20 \pm 3.4) \times 10^{-4}$
PEO	$(12.6 \pm 3.1) \times 10^{-5}$	$(4.56 \pm 1.2) \times 10^{-4}$	$(9.03 \pm 2.3) \times 10^{-4}$
PEVA	$(28.7 \pm 9.8) \times 10^{-5}$	$(9.44 \pm 1.7) \times 10^{-4}$	$(0.188 \pm 4.9) \times 10^{-4}$
PEVA	$(24.1 \pm 7.8) \times 10^{-5}$	$(9.38 \pm 2.0) \times 10^{-4}$	$(0.190 \pm 5.1) \times 10^{-4}$
PCL	$(8.77 \pm 1.3) \times 10^{-4}$	$(0.239 \pm 4.9) \times 10^{-4}$	$(0.675 \pm 9.8) \times 10^{-4}$
PCL	$(8.81 \pm 1.0) \times 10^{-4}$	$(0.242 \pm 1.8) \times 10^{-4}$	$(0.671 \pm 4.7) \times 10^{-4}$

[a] Values are means and standard deviations for eight exposures of a carbon black/polymer composite detector to DMMP. Replicate entries of the host material denote the results for two detectors that were fabricated together and were exposed together to analyte.

resistance change, $\Delta R/R_b$, where ΔR is the steady-state resistance change of the detector during exposure to the analyte and R_b is the baseline resistance of the detector during the initial 180-s period of exposure to background air, was used in analysis of the data. The value of R_b was calculated by taking the average of the last 30 s (~10 data points) of the initial 180-s air exposure, while ΔR was calculated by averaging the last 30 s of data obtained during exposure of the detectors to the analyte of interest. In studies when a background analyte was introduced to the detectors prior to introduction of a foreground analyte, two separate $\Delta R/R_b$ values, one for the background analyte relative to clean air and the other for the foreground analyte relative to the steady-state response of the background analyte, were calculated from the data. The value of R_b for exposure to a foreground analyte was determined using the last 30 s of data recorded during exposure of the detectors to the background analyte. The $\Delta R/R_b$ values were then averaged for each set of detectors over the eight replicate exposures for each unique detector/analyte combination. The $\Delta R/R_b$ values for both members of each pair of nominally identical carbon black/polymer detectors in the array were then combined to form a single average $\Delta R/R_b$ value for that type of polymeric detector/carbon black composite in response to the particular analyte of interest. Because only four detector compositions responded well to DMMP and DIMP, for consistency, all plots and detector performance data were evaluated using only this subset of detectors. This eight-detector array consisted of two copies of carbon black/polymer composite detectors formed using poly(ethylene oxide) (PEO), poly(ethylene-*co*-vinyl acetate) (PEVA), poly(butadiene-*co*-styrene) (PBS), and poly(caprolactone) (PCL), respectively.

Principal component analysis (PCA) was performed using macros written in Excel, and the data were plotted using Delta Graph and Claris Works. Data were normalized over the eight-detector array for a given exposure rather than over a collection of exposures for a given detector. This normalization procedure corrects for different analyte concentrations that are a consequence of the differing vapor pressures of the test analytes. The data were not autoscaled prior to use in principal component analysis.

Data were analyzed for pairwise discrimination between analytes using the Fisher linear discriminant algorithm.[29,30] In this statistical approach, the resolution factor (rf) for any solvent pair is obtained along any vector, $\vec{w}$, from the vector projection onto $\vec{w}$ of the distance between the cluster centroids, $d_{\vec{w}}$, divided by the sum of the projected standard deviations, $\sigma_{a,\vec{w}}$ and $\sigma_{b,\vec{w}}$, for data arising from repeated exposures for two different analytes, *a* and *b*. The resulting numerical resolution factor along $\vec{w}$ is defined as

$$\mathrm{rf} = d_{\vec{w}}/\sqrt{\sigma^2_{a,\vec{w}} + \sigma^2_{b,\vec{w}}}$$

The Fisher linear discriminant operates by searching for the vector, $\vec{w}$, such that the rf value is maximized along this optimal discriminant vector. Assuming a Gaussian distribution relative to the mean value of the data points in a given cluster, the probabilities of correctly identifying an analyte as *a* or *b* are approximately 72, 92, and 98% from a single presentation when analytes *a* and *b* are separated with resolution factors of 1.0, 2.0, or 3.0, respectively. Data extracted from multiple exposures to an analyte estimate the statistical distributions of the clustered data, although extremely high rf factors might overestimate the actual array discrimination performance based on the analysis of a relatively small number of presentations (10−20) of the detectors to each analyte.

RESULTS

A. Differentiation of DMMP from DIMP and Other Analytes of Interest in an Air Sample. Figure 3 shows data in principal component space for DMMP, DIMP, and other analytes of interest. The vapors were presented to the detectors as single-component analytes at $P/P^o = 0.010$ in a background of laboratory air. The data indicate that, under such conditions, DMMP can be clearly resolved from the other analytes based on their distinct $\Delta R/R_b$ response patterns on the detector array. As indicated in Table 4, resolution factors for pairwise differentiation between these pure analytes produced an average resolution factor of ≈90, indicating robust resolution of DMMP from any of these other analytes under such conditions.

Table 5 shows the $\Delta R/R_b$ response for DMMP at $P/P^o = 0.0017$, 0.0054, and 0.013 to four detectors, PEO, PEVA, PBS, and PCL, that responded well to the DMMP analyte. The data displayed in this table have been normalized by the value of P/P^o and thus indicate that the pattern of response was, within experimental error, independent of the DMMP concentration. A

(29) Doleman, B. J.; Lonergan M. C.; Severin, E. J.; Vaid, T. P.; Lewis N. S. *Anal. Chem.* **1998**, *70*, 4177−4190.

(30) Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; John Wiley & Sons: New York, 1973.

Table 4. Resolution Factors for an Eight-Detector Array of Carbon Black Composites To Resolve Pairwise Each of the 10 Vapors, at Fixed Concentration, from Any Other Vapor in the Test Set[a]

	DIMP	THF	benzene	methanol	toluene	water	lighter fluid	vinegar	diesel fuel
DMMP	56	65	47	84	32	71	46	87	34
DIMP		60	33	78	56	84	29	79	40
THF			89	82	250	49	83	32	109
benzene				186	128	234	49	81	37
methanol					267	32	156	27	43
toluene						211	29	89	25
water							264	28	34
lighter fluid								87	28
vinegar									96

[a] The average and worst pairwise resolution factors are 91 and 25, respectively.

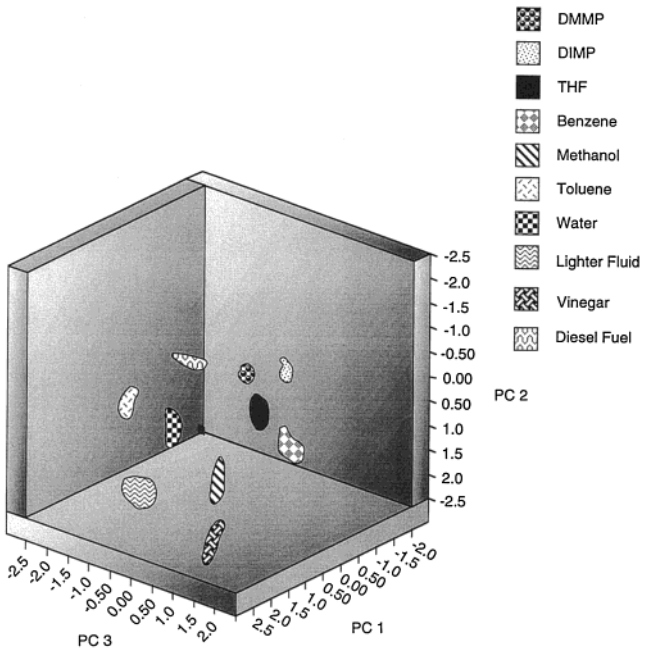


Figure 3. Data in principal component space of $\Delta R/R_b$ values produced when an eight-detector carbon black/polymer composite array was exposed to DMMP, DIMP, THF, benzene, methanol, toluene, water, lighter fluid, vinegar, or diesel fuel, each at P/P^o = 0.010, in an air background. The first three principal components contain 97% of the total variance in the data. The ellipsoids contain 95% of the data for each analyte. Each analyte was presented eight times to the array with the order of presentation randomized over all repetitions of all exposure types.

similar dependence of the $\Delta R/R_b$ detector response values versus analyte concentration has been obtained for carbon black/polymer composite chemiresistors in the presence of other vapors, for a comparable range of partial pressures of the organic vapors in a background of laboratory air.[31]

The sensitivities, $S = (\Delta R/R_b)/(P/P^o)$, of these detectors toward DMMP and DIMP, combined with the baseline noise values for the detectors, were used to obtain detection limits for DMMP and DIMP in an air sample on each detector type in our experimental configuration. Detection limits were taken as concentrations at which the signal/noise ratio (calculated from the sensitivity) was 3:1.[32] Values for these 3σ detection values for the four most responsive detectors are summarized in Tables 6 and 7.

Table 5. Normalized $\Delta R/R_b$ Responses of an Eight-Detector Array for Various Carbon Black/Polymer Composites[a]

host material	activity, P/P^o of DMMP		
	0.0017	0.0054	0.013
PEO	0.0869 ± 0.011	0.0964 ± 0.024	0.0834 ± 0.025
PEVA	0.188 ± 0.034	0.209 ± 0.023	0.170 ± 0.029
PCL	0.577 ± 0.053	0.531 ± 0.031	0.612 ± 0.035
PBS	0.146 ± 0.034	0.163 ± 0.045	0.133 ± 0.021

[a] Values represent means and standard deviations of $(\Delta R/R_b)/(P/P^o)$ for carbon black/polymer composite detectors in laboratory air. Data from two detectors of each polymer type were averaged together for this analysis. The analyte was exposed to the detectors eight times at each concentration, with the concentrations randomized in the experiment.

Table 6. Calculated Detection Limit of DMMP (in mg m^{-3}) for Various Carbon Black/Polymer Composites

host material	air	background analyte at $P/P^o = 0.010$ in air			
		THF	water	methanol	benzene
PEO	0.14	0.18	0.20	0.15	0.13
PEVA	0.050	0.055	0.068	0.053	0.047
PCL	0.059	0.051	0.048	0.062	0.057
PBS	0.19	0.22	0.16	0.24	0.18

Table 7. Calculated Detection Limit of DIMP (in mg m^{-3}) for Various Carbon Black/Polymer Composites

host material	air	background analyte at $P/P^o = 0.010$ in air			
		THF	water	methanol	benzene
PEO	0.19	0.67	0.32	0.58	0.76
PEVA	0.074	0.055	0.053	0.062	0.082
PCL	0.049	0.039	0.088	0.051	0.057

B. Differentiation of DMMP from DIMP and Other Analytes of Interest in the Presence of Varying Background Analytes. Parts a−c of Figure 4 show the $\Delta R/R_b$ response for carbon black/polymer composite detectors that contained poly(ethylene oxide), poly(ethylene-*co*-vinyl acetate), and poly(caprolactone), respectively, to DMMP at $P/P^o = 0.0017$, 0.0054, or 0.013 in the presence of various analytes that had been added at $P/P^o = 0.010$ to a laboratory air background flow. The data of Figure

(31) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 658−688.

(32) Skoog, D. A.; West, D. M. *Fundamentals of Analytical Chemistry*, 4th ed.; Saunders College Publishing: Philadelphia, 1982.

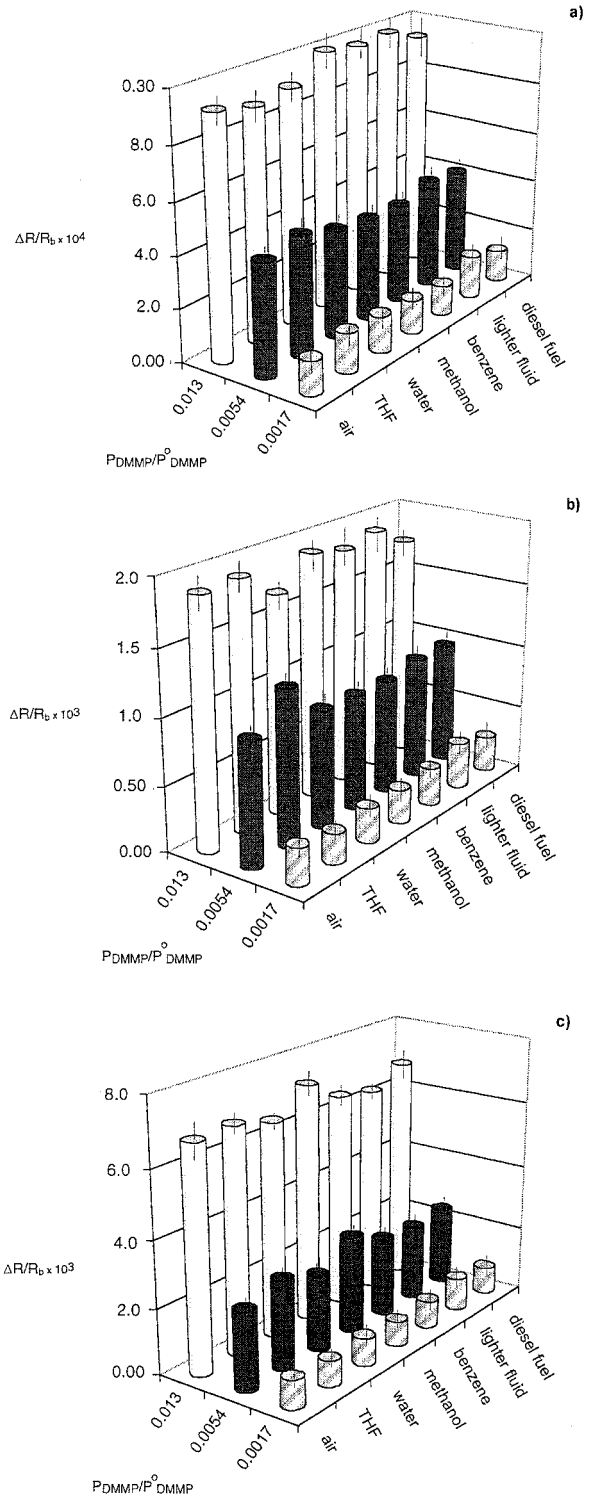


Figure 4. $\Delta R/R_b$ responses of carbon black/polymer composite detectors containing (a) PEO, (b) PEVA, or (c) PCL to DMMP at $P/P^\circ = 0.0017$, 0.0054, or 0.013 with THF, water, methanol, benzene, lighter fluid, or diesel fuel as background analytes at $P/P^\circ = 0.010$. The height of the bars represent the average values, and the error bars represents the 1σ standard deviation of the data.

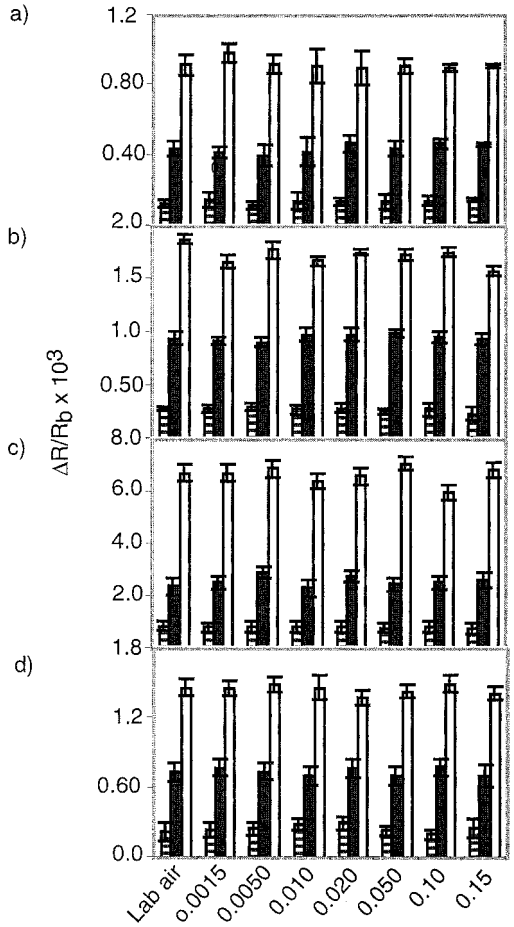


Figure 5. $\Delta R/R_b$ responses of carbon black/polymer composite detectors containing (a) PEO, (b) PEVA, (c) PCL, and (d) PBS to DMMP at $P/P^\circ = 0.0017$ (▬), 0.0054 (■), or 0.013 (□) with $P_{H_2O}/P^\circ_{H_2O} = 0.0015$, 0.0050, 0.010, 0.020, 0.050, 0.10, or 0.15 in laboratory air as the background analyte.

4 support the data of Table 5 that the $\Delta R/R_b$ response on these detectors is approximately a linear function of analyte partial pressure and furthermore indicate that the $\Delta R/R_b$ values for this range of DMMP partial pressures are, within experimental error, independent of whether the background gas was laboratory air or was laboratory air with a significant concentration of any of these other analytes. Analogous behavior was observed for DIMP. Similar behavior has also been observed previously in a general probe of the behavior of carbon black/polymer composite vapor detectors, in which $\Delta R/R_b$ response patterns to various test organic vapors were observed to be essentially independent of the composition of the background ambient for a range of analyte concentrations and analyte types under laboratory test conditions.[31]

Figure 5 shows the $\Delta R/R_b$ response of four detector types to DMMP as a function of the relative humidity of the background air, when $P_{H_2O}/P^\circ_{H_2O}$ was varied from 0 to 0.15 at room temperature. These data are of interest for detection of DMMP and DIMP in environments in which the relative humidity is not fully controlled. The data clearly show that, within experimental error, the $\Delta R/R_b$ response pattern for DMMP on carbon black composite detectors that contain either PEO, PEVA, PCL, or PBS was independent of the relative humidity over the range of values explored in this work.

Figures 6 and 7 summarize, in principal component space, the results of all experiments performed in this work involving exposures of DMMP and DIMP to an eight-detector array as a function of the composition of the background ambient gas in the flow stream. Consistent with Figures 3 and 4, the data indicate

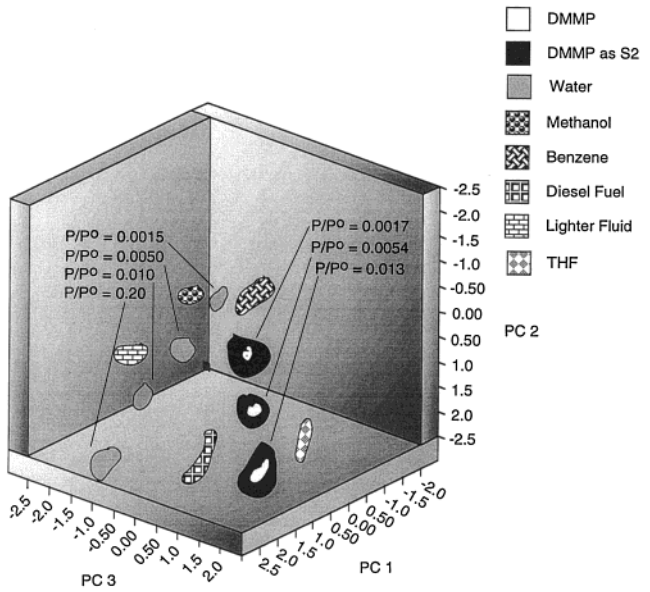


Figure 6. Data in principal component space from an eight-detector array exposed to methanol, benzene, diesel fuel, lighter fluid, or THF at $P/P^{o} = 0.010$, to water at $P/P^{o} = 0.0015$, 0.0050, 0.010, or 0.020, or to DMMP at $P/P^{o} = 0.0017$, 0.0054, or 0.013 in an background. Using the above analytes (except DMMP) as the background, the detectors were also exposed to a foreground of DMMP at $P/P^{o} =$ 0.0017, 0.0054, or 0.013. Each pure analyte or unique foreground/background combination was presented eight times to the array with the order of presentation randomized over all repetitions of all exposure types. The data obtained when DIMP was the foreground solvent in the presence of an analyte in the background is indicated by the region labeled "DMMP as S_2". The first three principal components contain 97% of the total variance in the data. The ellipsoids contain 95% of the data for each analyte.

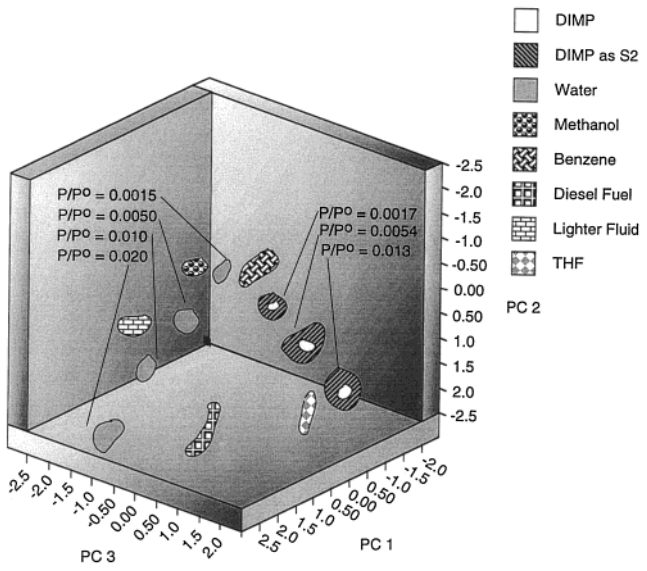


Figure 7. Data in principal component space from an eight-detector array exposed to methanol, benzene, diesel fuel, lighter fluid, or THF at $P/P^{o} = 0.010$, to water at $P/P^{o} = 0.0015$, 0.0050, 0.010, or 0.020, or to DIMP at $P/P^{o} = 0.0017$, 0.0054, or 0.013 in an background. Using the above analytes (except DIMP) as the background, the detectors were also exposed to a foreground of DIMP at $P/P^{o} =$ 0.0017, 0.0054, or 0.013. Each pure analyte or unique foreground/background combination was presented eight times to the array with the order of presentation randomized over all repetitions of all exposure types. The data obtained when DIMP was the foreground solvent in the presence of an analyte in the background is indicated by the region labeled "DMMP as S_2". The first three principal components contain 97% of the total variance in the data. The ellipsoids contain 95% of the data for each analyte.

that the pattern type for DIMP and for DMMP was preserved as the concentration of analyte increased. In addition, Figures 6 and 7 demonstrate that all of the tested concentrations of DMMP and DIMP could be clearly differentiated from the patterns produced by exposure to the other pure analytes of interest. Finally, the data of Figures 6 and 7 indicate that the DMMP and DIMP patterns were essentially unchanged regardless of the composition of the background ambient gas flow evaluated in this work. Resolution factors for the eight-detector array under various conditions of interest are presented in Tables 4 and 8.

Sensitivities and detection limits were also calculated for DMMP and DIMP in the presence of variation in the composition of the background ambient. Values for the detection limits under the various conditions studied herein are presented in Tables 6 and 7. As expected from the results above, the detection limits were relatively insensitive to the composition of the background analyte.

DISCUSSION

Under our experimental conditions, the detection limits for DMMP were lower than the EC_{50} values for sarin (Table 1). The vapor pressures of sarin and DMMP are 1.61×10^{4} and 1.45×10^{4} mg m^{-3} at 20 °C, respectively.[33] Because the equilibrium $\Delta R/R_b$ response of carbon black/polymer composite sorption detectors depends primarily on the vapor pressure of the analyte,[34] the detection limits for DMMP are expected to be very similar to those for sarin.

The data clearly show that DMMP and DIMP were robustly detected and differentiated from each other at $P/P^{o} = 0.0017$ in the presence of various other potential "interefereences" at partial pressures of $P/P^{o} = 0.010$. Because of the significant differences in vapor pressures between DMMP and DIMP and the set of interferences, the observed behavior corresponds to detection of DMMP and DIMP at the levels of 0.047−0.24 and 0.049−0.76 mg m^{-3}, respectively, in the presence of background concentrations of water, methanol, benzene, toluene, diesel fuel, lighter fluid, vinegar, and THF in the range $900-1.6 \times 10^{3}$ mg m^{-3}. The specific analyte concentrations for $P/P^{o} = 0.013$ of diesel fuel, vinegar, and lighter fluid have not been specified because these reagents are multicomponent mixtures of analytes; nevertheless, the primary component of these analytes was present at concentrations in excess of 1.7×10^{4} mg m^{-3} under our test conditions. The ability to detect rather low concentrations of DMMP in the presence of much higher concentrations of these other analytes arises primarily from the underlying thermodynamics of sorption-based detectors, as shown in detail previously,[34] which favors detection of low vapor pressure analytes based on their higher partition coefficients into polymeric detector films.

It has been shown previously that the steady-state $\Delta R/R_b$ response of carbon black/organic polymer composite detectors

(33) George, V. private communication, PM-MCD, 10205 Burbeck Rd., Fort Belvoir, VA 22060.

(34) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5442−5447.

Table 8. Resolution Factors for an Eight-Detector Array of Carbon Black Composites To Resolve Pairwise Each of the 10 Vapors at Fixed Concentration, from Any Other Vapor in the Test Set[a,b]

	$DIMP_{S2}$	THF	benzene	methanol	toluene	water	lighter fluid	vinegar	diesel fuel
$DMMP_{S2}$	39	47	68	38	28	36	33	59	47
$DIMP_{S2}$		58	36	67	52	79	36	82	42
THF			89	82	250	49	83	32	109
benzene				186	128	234	49	81	37
methanol					267	32	156	27	43
toluene						211	29	89	25
water							264	28	34
lighter fluid								87	28
vinegar									96

[a] The average and worst pairwise resolution factors are 82 and 28, respectively. [b] $DMMP_{S2}$ and $DIMP_{S2}$ are the foreground solvents with THF, benzene, methanol, toluene, water, lighter fluid, vinegar, or diesel fuel as background analytes.

to binary mixtures of organic vapors is generally a linear combination of the steady-state $\Delta R/R_b$ responses for the individual components of the mixture.[31] Thus, the presence of DMMP or DIMP could in principle be determined in a binary mixture of vapors exposed simultaneously to a carbon black/polymer composite detector array if there were some prior knowledge about the composition of the analyte. However, use of the steady-state $\Delta R/R_b$ response patterns is best-suited for situations in which the detectors are to be utilized in change detection mode, as opposed to performing an analysis of a complex analyte mixture without any prior separation steps or temporal information on changes in vapor composition with time. In principle, an array having as many as five to six components can be analyzed using the equilibrium response data of a detector array containing six to seven detectors provided that the detectors are at most partially correlated with each other.[35] Additional information can be obtained, in principle, from use of the time dependence of the detector responses; however, such data are typically not independent of the analyte background composition or linear with changes in analyte concentration. The time response of the carbon black/polymer composite detectors is also more sensitive to variations in film thickness and carbon black content than are the steady-state $\Delta R/R_b$ values used herein, so the exact performance of such a system must be evaluated for the particular task at hand and with respect to the specific implementation of the detection and sampling system used in such an application.

In this work, we have quoted limits of detection as opposed to limits of classification.[35] The former are fundamental quantities that are independent of the task of concern, while the latter are intimately linked to the nature and type of the interferences and background analytes and to how well-matched a particular algorithm is to the training and test sets of data being analyzed.[35] Under certain conditions and for certain tasks, limits of classification on polymer-coated surface acoustic wave detector arrays have been shown to be within a factor of 2−3 of the limits of detection,[35] so the values quoted herein provide a reasonable (order of magnitude) estimate of the performance that might be expected in certain tasks. We also note that the detection limits presented above are only approximate performance measures for such detectors. Signal/noise ratios, and thus limits of detection, for carbon black/polymer composite chemiresistors at equilibrium with an analyte of interest have been shown to decrease as the detector area decreases,[36] so our detection limits must be scaled accordingly if the detector area is varied. In addition, the sensitivity of sorption detectors decreases with increasing temperature,[35] so the signal/noise ratio, and thus detection limits, will decrease if the detectors are heated. In comparing the performance of these chemiresistors to prior results on polymer-coated surface acoustic wave devices for detection of DMMP and other nerve agents and nerve agent simulants,[23] it is important to note that we have not used any preconcentration of the sampled analyte and have only used steady-state $\Delta R/R_b$ amplitudes in an attempt to characterize the fundamental performance of the detectors toward the analytes of interest. Significant tradeoffs in sensitivity, time response, power, and other system properties would of course be involved in an engineering implementation of these detectors in a fielded device for nerve agent detection. Finally, we note that the detectors used in our study were not specifically designed to possess high partition coefficients toward phosphonate-containing analytes such as DMMP or DIMP. Work on polymeric coatings for surface acoustic wave devices has indicated that significant improvements in sensitivity are possible through use of detector films designed to have hydrogen-bonding acceptor groups that provide complimentarity to the hydrogen-bonding donor groups in phosphonate-containing nerve agents.[23,37] Because the sensitivity improvements in the polymer-coated surface acoustic wave devices are directly related to increases in sorption of the analytes into the polymer films,[38,39] similar improvements in detection limits toward DMMP, DIMP, sarin, and soman would be expected if such polymers were incorporated into arrays of carbon black/polymer composite vapor detectors.

In summary, generic, untailored arrays of carbon black/polymer chemiresistive vapor detectors can detect DMMP and DIMP at levels below the EC_{50} limits for the nerve agents sarin and soman. DMMP can be differentiated from DIMP and from a variety of other analytes either in laboratory air or in laboratory air that contains the presence of relatively high concentrations of various types of volatile organic vapors. Concentration-normalized response patterns for DMMP and DIMP are independent of concentration and of background analytes over the range of

(35) Zellers, E. T.; Han, M. *Anal. Chem.* **1996**, *68*, 2409−2418.
(36) Briglin, S. M.; Freund, M. S.; Tokumaru, P.; Lewis, N. S., submitted for publication.
(37) Abraham, M. H.; Whiting, G. S. *J. Chromatogr.* **1991**, *588*, 361−364.
(38) Grate, J. W.; Wise, B. M.; Abraham, M. H. *Anal. Chem.* **1999**, *71*, 4544−4553.
(39) Grate, J. W.; Abraham, M. H. *Anal. Chem.* **1996**, *68*, 913−917.

concentrations and analytes tested in this work. Further improvements in the fundamental performance parameters of these detectors toward the analytes of interest are likely to occur through the use of specially tailored polymers and better control over temperature and other system variables.

ACKNOWLEDGMENT

We acknowledge the Department of Energy (DE-FG03-98NV13367), the Army Research Office through a MURI (DAAG55-98-1-0266), and DARPA (DAAK-60-97-K9503) for support of this work. We also acknowledge helpful discussions with V. George of the Army Night Vision Laboratory regarding detection limits and background ambients of interest.

Received for review July 26, 2000. Accepted December 8, 2000.

AC0008439

A Hybrid Acoustic/Electromagnetic Technique for Locating Land Mines

Waymond R. Scott, Jr.[a], Chistoph Schroeder[a], and James S. Martin[b]
[a]School of Electrical and Computer Engineering
[b]School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332
Phone: 404-894-3048
EMail: waymond.scott@ece.gatech.edu

Abstract— A hybrid technique is presented that simultaneously uses both electromagnetic and acoustic waves in a synergistic manner to detect buried land mines. The system consists of an electromagnetic radar and an acoustic source. The acoustic source causes both the mine and the surface of the earth to be displaced. The electromagnetic radar is used to detect these displacements and, thus, the mine. An experimental and numerical model for the system has been developed.

INTRODUCTION

A technique to detect land mines that simultaneously uses both electromagnetic and acoustic waves in a synergistic manner is currently being investigated. The synergism has the potential to significantly enhance the signature of the mine with respect to the clutter and make it possible to detect a mine that would be impossible to detect by purely electromagnetic or purely acoustic means.

The configuration of the system currently being studied is shown in fig. 1. The system consists of an electromagnetic radar and an acoustic source. The acoustic source induces an acoustic (seismic) wave into the earth. The acoustic wave causes both the mine and the surface of the earth to be displaced. The displacement of the mine is different than the earth, because the acoustic properties of the mine are quite different than those of the earth. The displacement of the surface of the earth when a mine is present is different than when it is not present because of the waves scattered from the mine. The electromagnetic radar is used to detect these displacements and, thus, the mine. This idea has been discussed previously [1, 2], but it has not been seriously investigated.

In addition, the radar can be simultaneously used in a mode in which only the electromagnetic waves are used to identify (image) the mine. In this electromagnetic only mode, the radar is a conventional ground penetrating radar (GPR). Thus, both the acoustic and electromagnetic properties of the mine can be used to differentiate it from the earth. It may be possible to locate or identify mines with the GPR that are invisible to the acoustic system or visa versa. Or it may me possible to combine the output of both the GPR and the acoustic system to lower the false alarm rate of the system.

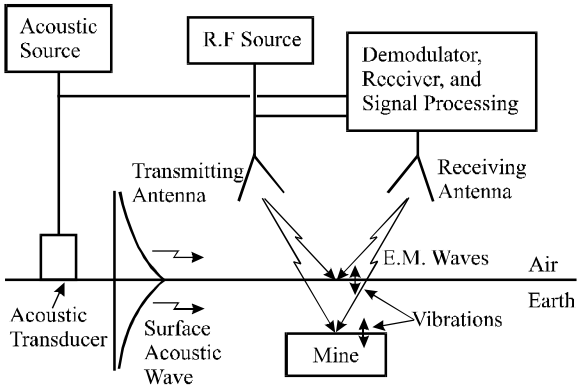


Fig. 1. Acousto-electromagnetic mine detection system with the acoustic transducer placed on the surface of the earth.

An experimental model for the system has been constructed and automated. The experimental model is being used to demonstrate the viability of the technique and to study the interactions of the acoustic and electromagnetic waves with buried mines. The technique looks promising; we have been able to detect both simulated antipersonnel mines and antitank mines buried in damp sand. However, additional investigation of the technique is needed to determine the capabilities of the technique in more varied conditions. A two-dimensional finite-difference time-domain (FDTD) model for the acoustic waves has been also developed and is being used to help understand the interactions of the acoustic waves and the mines.

EXPERIMENTAL MODEL

A radar has been designed and built to measure the acoustic vibrations of the soil and the mine. The radar radiates electromagnetic waves that are reflected off of the vibrating interface. The reflected waves are received by the radar, and a Homodyne system is used to demodulate the signals. The vibrations are determined from these demodulated signals. The two biggest challenges to make this radar perform adequately for the mine detection system

This work is supported in part by the US Army Research Office under contract DAAH04-96-1-0048.

0-7803-4403-0/98/$10.00 (c) 1998 IEEE

are 1) to make it sufficiently sensitive to be able to detect the small vibrations, and 2) to make the spot size (the area on the surface illuminated by the electromagnetic waves) sufficiently small. The radar can measure vibrations as small as 1 nm (10^{-9} m) as currently configured. To obtain this sensitivity, the radar was designed to minimize the effects of noise, such as the phase noise of the source and the electromagnetic interference from low-frequency magnetic fields. The spot size of the radar must be smaller than approximately one half of a wavelength of the acoustic waves. Currently, a small spot size is obtained by using an open-ended waveguide as the antenna for the radar. This antenna produces a sufficiently small spot size when the open end of the antenna is placed within a few centimeters of the surface. The results presented in this report are made with the radar operating at 8 GHz. The vibration of the surface of the sand has been measured with both the radar and an accelerometer; these measurements were compared and found to be in good agreement.

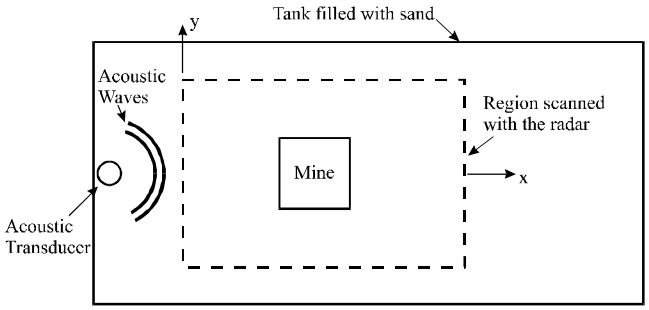
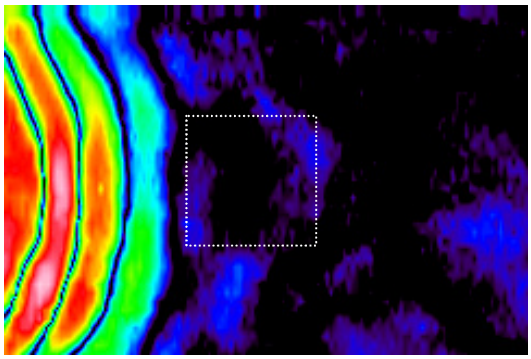


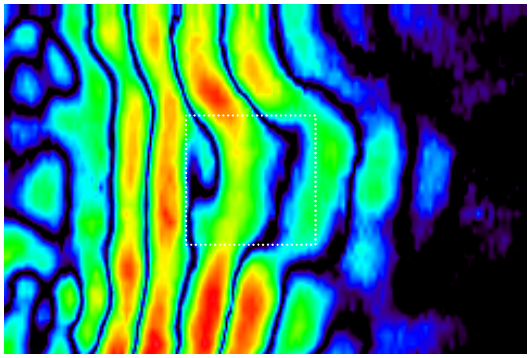
Fig. 2. Top view of experimental model.

A drawing of the top view of the experimental model is shown in fig. 2. The model consists of a tank that is filled with damp sand that has been packed to a relatively uniform density. The tank is approximately 120 cm wide, 120 cm deep, and 240 cm long. A transducer is placed on the surface of the sand and is used to launch the acoustic waves into the sand. The transducer is an electromagnetic shaker that is excited with a differentiated gaussian pulse. The pulse has a center frequency of 400 Hz. The transducer has been adjusted so that it preferentially launches acoustic surface waves. These surface waves travel across the surface of the tank and interact with a mine that is buried in the sand. The electromagnetic radar is used to measure the vibrations caused by the acoustic waves. A x-y positioner is used to scan the radar over the surface of the sand. This system is under computer control, so it can scan the radar over the surface and record the data automatically. The region scanned with the radar is indicated on fig. 2. The vibrations are measured in this region as a function of time and position. The measurements are made on a uniform rectangular grid of discrete positions in the scanned region.

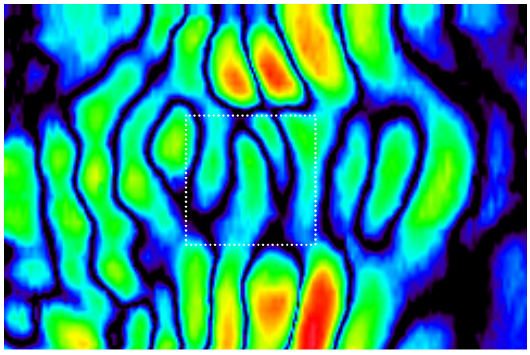
Pseudo color graphs of the amplitude of the displacement of the surface are presented in fig. 3 for three times. These results are for a simulated antitank mine buried in the sand. The simulated anti-tank mine is made out of Acrylic plastic and is 30 cm wide, 30 cm long, and 7.5 cm high. The top of the mine is 5 cm below the surface of the sand. The position of the mine is indicated by the dotted white line. In these graphs, the color scale goes from black to blue to green to yellow to red to white. The smallest displacements are in black and the largest displacements are in white. At time #1, the acoustic wave is seen traveling toward the mine. At time #2, the wave has reached the mine, and a portion of the wave



Time #1 – Acoustic pulse traveling toward the mine



Time #2 – Pulse interacting with the mine. The pulse is seen to travel faster across the mine.



Time #3 –Pulse is reflected from, transmitted through, and travels around the mine.

Fig. 3. Pseudo color graphs of the amplitude of the displacement of the surface of the sand for four times when a simulated antitank mine is buried in the sand. The top of the mine is 5 cm below the surface, and the mine is outlined by a white dotted line.

is going across the mine while the rest is going around the mine. Notice that the portion of the wave that is going across the mine is ahead of the portion that is around the mine. This is because the mine is much stiffer than is the sand; thus, the wave appears to travel faster across the mine. The mine actually moves as if it is a rigid body. The motion of the mine is essentially a rocking motion excited by the wave motion in the sand. At time #3, the wave is seen to be still going around the mine; however, a significant portion has passed across the mine. Notice that the wave that went across the mine is still ahead of the wave that went around the mine, and notice that the wave that went across the mine is smaller in amplitude than the one that went around the mine. Also notice that the displacements are smaller above the mine. The waves that are reflected from the mine can also be seen. The location and the shape of the mine are clearly evident in these graphs. We have also generated movies from the data recorded with the radar. In the movies, pseudo color graphs like those in fig. 3 are stored and played back sequentially. The interaction of the waves with the mine is clearer in the movies than it is in fig. 3; thus, is easier to see the mine.

Numerical Model

A two-dimensional FDTD model for the acoustic waves has been developed. Fig. 4 is a diagram of the FDTD model. The waves are injected with a point source, a perfectly matched layer (PML) is used to absorb the waves at the edge of the mesh, and a free surface boundary condition is used on the boundary between the soil and the air. The results from the numerical model are in fairly good agreement with those from the experimental model. The numerical model has been very useful in helping us to understand the interaction of the acoustic wave with the mines. With the numerical model, the waves can be observed below the surface; whereas, with the experimental model, the waves can only be observed on the surface.

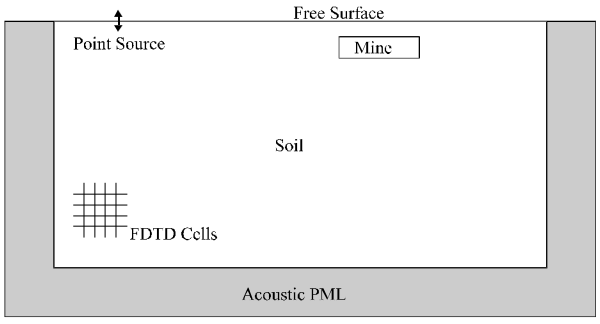
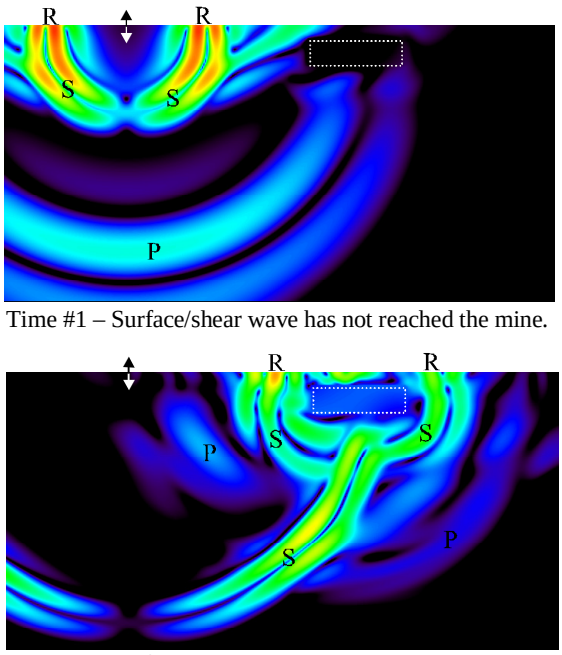


Fig. 4. Diagram of the acoustic FDTD model.

Pseudo color graphs of the amplitude of the velocity of the particles in the soil are presented in fig. 5 for the anti-tank mine buried 5 cm deep. The incident pulse is a differentiated gaussian pulse with a center frequency of 400 Hz. The soil models the sand that is used in the experimental model. At time #1, pressure (P), shear (S), and surface (R) waves are seen to be launched. The shear and surface waves overlap, because they propagate at approximately the same velocity (the surface wave propagates slightly slower than the shear wave). The surface wave is the more intense wave near the surface. The pressure wave is ahead of the surface/shear wave because it propagates faster. At time #2, the surface/shear wave is seen to have reached the mine. The portion of the wave that passed across the mine is seen to be ahead of the portion that is passing around the mine. Again this is because the wave travels faster in the mine than in the sand. Scattered pressure, shear, and surface waves are seen propagating away from the mine. Notice that a surface wave appears to be trapped above the mine, we have observed similar behavior in some of the experimental results.



Time #1 – Surface/shear wave has not reached the mine.

Time #2 – Surface/shear wave is interacting with the mine.

Fig. 5. Pseudo color graphs of the amplitude of the velocity of the particles in the soil for four times when a simulated antitank mine is buried in the sand. The top of the mine is 5 cm below the surface, and the mine is outlined by a white dotted line.

REFERENCES

[1] G.S. Smith, "Summary Report: Workshop on New Directions for Electromagnetic Detection of Non-Metallic Mines," Report for U.S. Army BRDEC and ARO, June 1992

[2] C. Stewart, *Summary of Mine Detection Research*, Vol. I, pp. 172-179, Tech. Report 1636-TR, May 1960, U.S. Army Engineering Res. and Devel. Labs, Corps. of Eng., Belvoir, VA.

Investigation of a Technique that Uses Elastic Waves to Detect Buried Land Mines

Waymond R. Scott, Jr.[a], James S. Martin[b], and Gregg D. Larson[b]
[a]School of Electrical and Computer Engineering
[b]School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332
Phone: 404-894-3048
EMail: waymond.scott@ece.gatech.edu

Abstract — A technique that uses elastic waves to detect buried mines is being investigated. In the technique, an elastic wave is launched into the earth. It propagates and interacts with a buried mine causing displacements of both the mine and the surrounding soil. Because of the interactions of the wave with the mine, the soil surface displacements near the mine are quite different from those far from it. These displacements are measured using a sensor that is suspended above the soil surface. Results using two types of sensors are presented in this paper: a specially designed radar and an array of microphones. The results from measurements with both types of sensors compare well. The effects of the mine (including resonances, reflections, and non-propagating waves) can be seen in the data. The observed resonances clearly distinguish mines from common forms of buried clutter.

INTRODUCTION

Seismic/elastic techniques show considerable promise for the reliable detection of all types of buried mines, even low-metal anti-personnel mines. The reason for this is that mines have mechanical properties that are significantly different from soils and typical forms of clutter. For example, the shear wave velocity is approximately 20 times higher in the explosive and the plastics used in typical mines than in the surrounding soil. In addition, mines are complex mechanical structures with a flexible case, a trigger assembly, air pockets, etc. This complex structure gives rise to structural resonances, non-linear interactions, and other phenomena that are atypical for both naturally occurring and man-made forms of clutter. This phenomenology can be used to distinguish a mine from clutter.

Systems have been developed at Georgia Tech that employ sensors that measure local seismic displacements without physically contacting the soil surface [1-2]. The non-contact nature of these sensors allows interrogation of the soil surface near or immediately above a mine. This substantially increases the measurable effects of the mine's presence over schemes which rely on elastic waves scattered by the mine to propagate to a remote sensor location.

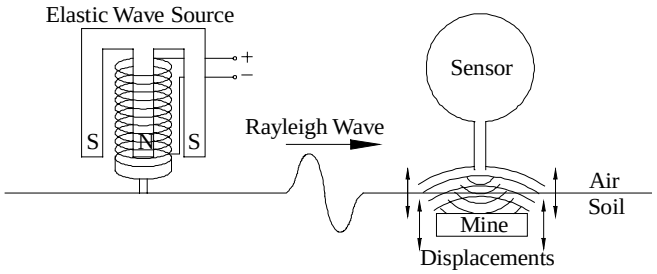


Figure 1 Diagram of the mine detection system.

Figure 1 depicts the system configuration. The system consists of a non-contact sensor and a seismic source. The source (an electrodynamic shaker coupled to the ground by a narrow foot) preferentially generates an elastic surface (Rayleigh) wave in the earth. The Rayleigh wave causes both the mine and the surface of the earth to be displaced as it propagates past the mine. Since the amplitude of Rayleigh wave displacements decreases exponentially with depth, only the soil near the surface is interrogated for the presence of mines. The depth of soil that is examined is a function of the frequency of the source. For typical mine depths and sizes, this is in the l00 to 1,000 Hz range. The motion of the mine is different from the surrounding soil, because the elastic properties of the mine are quite different than those of soil. The displacement of the surface of the earth near a mine is different than when the mine is not present because of the local and propagating waves scattered by the mine. The sensor is used to detect these displacements and, thus, the mine.

The system is currently being studied in a laboratory scale experimental model. The model, which is depicted in figure 2, consists of a wedge shaped tank filled with over 50 tons of damp compacted sand to simulate soil. The seismic source is located near the tip of the wedge and is bi-directive toward the search area and the back wall. Simulated mines, inert mines, and clutter, such as rocks and sticks, are buried within a 2 m x 2 m region in the center of the tank. The sensor can be scanned over this region with a three degree of freedom positioner fixed above the tank

This work is supported in part by the US Army Research Office under contract DAAH04-96-1-0048.

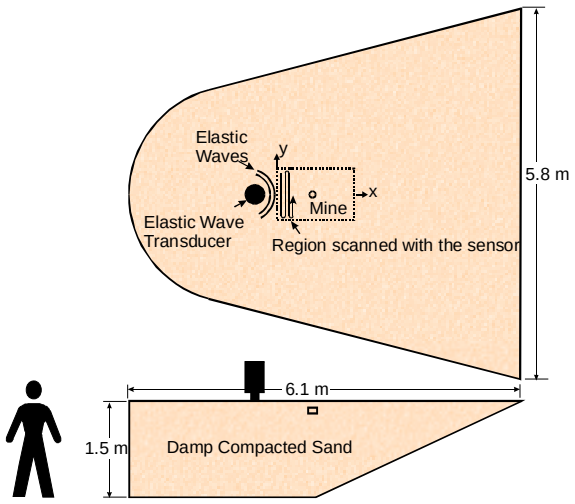


Figure 2 Diagram of the experimental model.

RADAR DISPLACEMENT SENSOR

Initial testing of the elastic wave mine detection system was conducted with a specially designed radar sensor. This idea of using elastic and electromagnetic waves concurrently has been proposed previously [3], but it has not been seriously investigated until now.

The radar sensor radiates electromagnetic waves toward the earth. These waves are reflected from the surface of the earth and the mine, where they are amplitude and phase modulated by the transient displacements of the earth and the mine. The reflected waves are received and demodulated. The resulting demodulated signals are proportional to the surface displacement. The radar operates at a frequency of 8 GHz and can measure vibrations as small as 1 nm (10^{-9} m) as currently configured. The end of a waveguide, which functions as both transmit and receive antennas, illuminates an area on the earth's surface comparable to its cross section (1cm x 2 cm) over which the measurement is averaged.

MICROPHONE ARRAY

An array of microphones is also being investigated as a possible low-cost alternative or supplement to the radar. In this scheme the acoustic pressure in the air immediately above the soil surface is measured with a microphone array. Due to the low wave speed of the Rayleigh wave, the acoustic wave that it radiates into the air is evanescent. It can, however, be measured very near the ground surface. If this measurement is made at many locations on a sufficiently large array of calibrated microphones, the pressure measurement can be inverted to determine the surface displacements. This measurement is averaged over an area proportional to the square of the height of the array. This is known as planar near-field acoustic holography and has been employed in the study of structural acoustics problems [4].

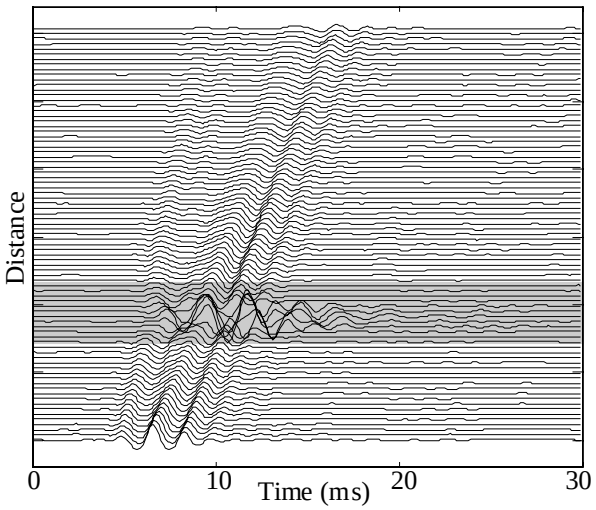


Figure 3 1-D Radar Scan Over Buried TS-50 AP Mine

EXPERIMENTAL RESULTS

The primary detection cues used for all of the inert mine types that have been studied in the experimental model have been resonances of the mine case, trigger mechanism, and overlying soil [1,2,5]. These are excited by the passage of the Rayleigh wave and characterized by large displacements that persist after the passage of the incident pulse. Although mines exhibiting resonances scatter a larger propagating wave field than similarly sized non-resonant objects, the most pronounced feature of the field scattered by these mines is its primarily local nature. This appears to result from a mostly reactive soil loading. For the mine types studied thus far, the localized resonant motion has been an excellent indicator of a mine's location and extent. Figure 3 shows a waterfall graph of a linear scan over a TS-50 antipersonnel (AP) mine with the radar sensor. Measurements were made in 1 cm increments from the source and the graphs of successive measurements are offset vertically. The presence of the mine is apparent in the figure due to the amplification of motion above the mine and ringing at this location that can be observed after the incident pulse has propagated beyond the mine.

The same scan was performed using an uncalibrated microphone 1 cm above the soil surface. The measured signals are shown in figure 4. Here the incident wave and the mine resonance are again clearly apparent. Since the 1-D scan that was performed does not contain sufficient information to invert the acoustic propagation and determine the soil surface displacements, pressure is plotted in figure 4 rather than displacement. There is, therefore, an exaggeration of the received signal contributions from the direct air acoustic path and from the faster bulk waves in the soil relative to the Rayleigh wave contributions in figure 4 versus figure 3.

An experiment was performed to address the issue of imaging a minefield containing a mine and several false targets, all covered by surface vegetation. For this study, the false targets were four mine-sized rocks and two sticks. The

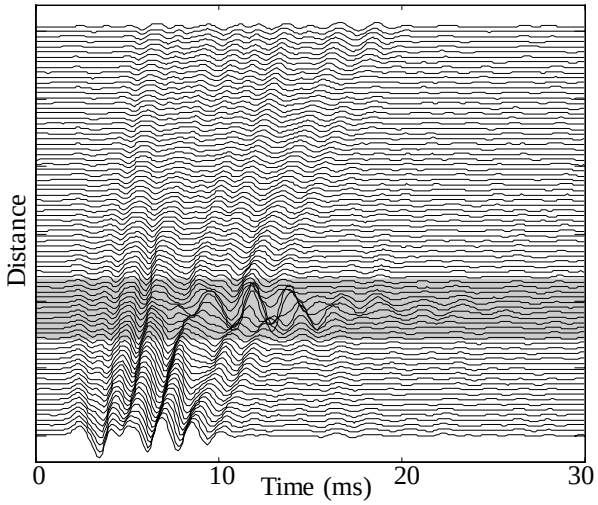


Figure 4 Microphone Scan Over Buried TS-50 AP Mine

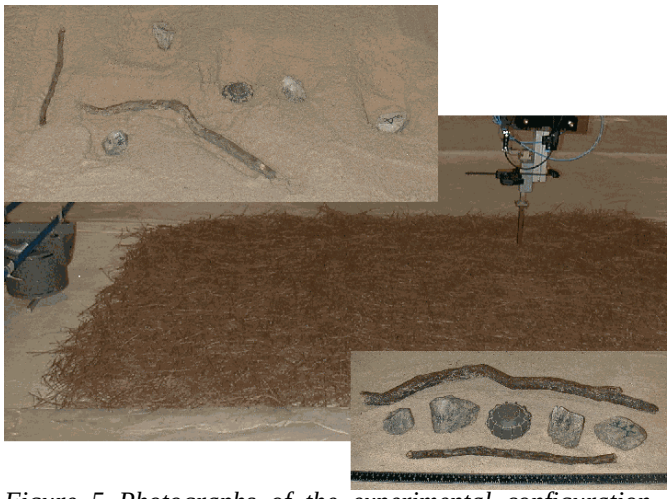


Figure 5 Photographs of the experimental configuration. Top left: Uncovered mine and clutter items. Center: the scan region covered in pine straw. Bottom Right: mine and clutter items.

mine was a TS-50 AP mine, and the ground cover was 2.3 cm of pine straw. The mine and the false targets were buried at depths of one to three centimeters within a 120 cm by 80 cm search area. The layout of this experiment and the relative scale of the buried objects are shown in figure 5. The goal was to determine whether the mine could be distinguished from the false targets and whether the radar could see through the pine straw. The image formed of this search area is shown in figure 6. The image clearly shows the location and extent of the mine and virtually no evidence of the non-resonant clutter objects. The pine straw did not prove problematic for the radar sensor. Its performance was degraded slightly by the increased surface standoff distance due to the layer of pine straw.

Imaging of mines from surface displacement measurements can be done in many ways. This image was formed by a multi-step process that filters forward travelling waves (those components directed away from the source) out of the data in the wave number domain leaving the reflected

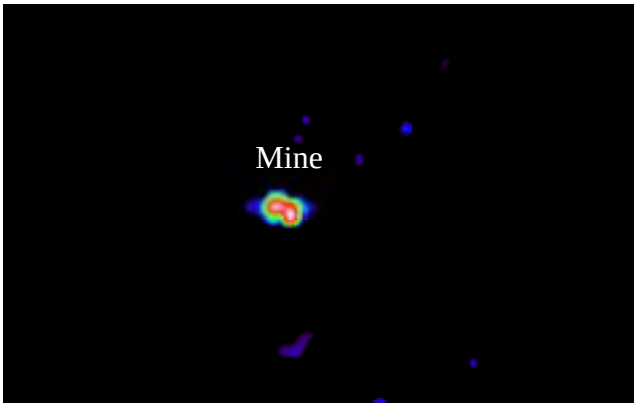


Figure 6 Mine Image Formed from 2-D Radar Scan Data

waves and a portion of the non-propagating waves. The energy in these remaining waves at times near the time of arrival of the incident Rayleigh wave is assigned to each measurement point forming an image. Unlike background subtraction, this algorithm does not rely on information that would be unavailable to a mine detection system operating in the field. Previous work has shown that this technique can be used to image several types of AP mines and distinguish them from non-resonant buried clutter [1,5].

CONCLUSIONS

Detection and imaging of inert AP mines using elastic waves and non-contact displacement sensors have been demonstrated under laboratory conditions that mimic a variety of realistic mine detection scenarios. These conditions include the presence of natural surface covering and buried clutter in close proximity to the mine. Two different sensing techniques have been demonstrated which can accomplish these detections. These should have different susceptibilities to noise and surface cover. Although not as well developed, one of these methods, the microphone array technique, offers a significant savings in the hardware requirements of the mine detection system.

REFERENCES

1. Scott W.R., Jr., Larson, G.D. and Martin J.S., "Simultaneous Use of Elastic and Electromagnetic Waves for the Detection of Buried Land Mines," *Proceedings of the SPIE: 2000 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, Vol. 4038, April 2000.
2. Scott, W.R., Jr., Schroeder, C., and Martin, J.S., "A Hybrid Acoustic/Electromagnetic Technique for Locating Land Mines," *Proceedings of the 1998 International Geoscience and Remote Sensing Symposium*, Seattle, Washington, July 1998
3. C. Stewart, Summary of Mine Detection Research, Vol. I, pp. 172-179, *Tech. Report* 1636-TR, May 1960, U.S. Army Engineering Res. and Devel. Labs, Corps. of Eng., Belvoir, VA.
4. E. Williams et. al., "Sound Reconstruction Using a Microphone Array," Journal of the Acoust. Soc. of Am., Vol. 68, No.1, pg 340-44, July 1980
5. Behboodian, A., Scott, W.R., Jr. and McClellan, J.H. "Signal Processing of Elastic Surface Waves for Localizing Buried Land Mines," *Proceedings of the 33rd Assilomar Conference on Signals, Systems, and Computers*, Assilomar, CA, October 1999

Three-Dimensional FDTD Model to Study the Elastic-Wave Interaction with Buried Land Mines

Christoph T. Schröder[(a)] and Waymond R. Scott, Jr.[(b)]
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250, USA
(a) christoph.schroeder@ee.gatech.edu, 404-894-3123
(b) waymond.scott@ee.gatech.edu, 404-894-3048

ABSTRACT

A three-dimensional finite-difference model for elastic waves in the ground has been developed and implemented. The model is used to investigate the interaction of elastic waves with buried land mines. When elastic waves interact with a buried mine, a strong resonance occurs at the mine location. The resonance can be used to enhance the mine's signature and to distinguish the mine from clutter. Results are presented for a single mine buried in the ground. The predictions of the numerical model are in fairly good agreement with experimental results.

INTRODUCTION

A new technique is being investigated at the Georgia Institute of Technology, in which buried land mines are located by using both elastic (acoustic) and electromagnetic waves in a synergistic manner [1]–[4]. Here, elastic waves interact with a buried land mine and cause the mine and the surface above the mine to vibrate. An electromagnetic radar records the vibrations and, thus, detects the mine. During the process of developing the elastic/electromagnetic sensor, a major part has been the implementation of a numerical model which simulates the interaction of the elastic waves with the buried land mines.

The numerical model is based on the finite-difference time-domain (FDTD) method. The equation of motion and the stress-strain relation, together with a constitutive relation, form a set of first-order partial differential equations that completely describes the elastic wave motion in a medium. Introducing finite differences, this set of equations can be discretized and adapted to the finite-difference time-domain modeling scheme.

The finite-difference model has been implemented in two and three dimensions. The numerical model is used to investigate the mine-wave interactions and has been very helpful for explaining and understanding the experimental results. In this paper, results obtained with the three-dimensional model are presented. The interaction of elastic waves with a buried antipersonnel mine is shown.

THREE-DIMENSIONAL NUMERICAL MODEL

In a realistic setting, one or several mines are buried in the ground, surrounded and covered by various kinds of clutter. To approximate these conditions, experiments have been performed with mines buried in a large sand box [1], [2]. In these experiments, elastic waves are launched by an electrodynamic transducer placed on the surface of the ground. The waves propagate along the surface and interact with the buried land mines. To study these mine-wave interactions, a three-dimensional finite-difference model has been developed.

Finite-Difference Model

Fig. 1 shows the three-dimensional finite-difference model. To reasonably simplify the model, the ground is assumed to be linear, isotropic and lossless. The surface of the ground is modeled as a free-surface, and a *Perfectly Matched Layer* terminates the solution space at the remaining grid edges and absorbs the outward traveling waves [5]. The solution space is discretized using a staggered finite-difference grid.

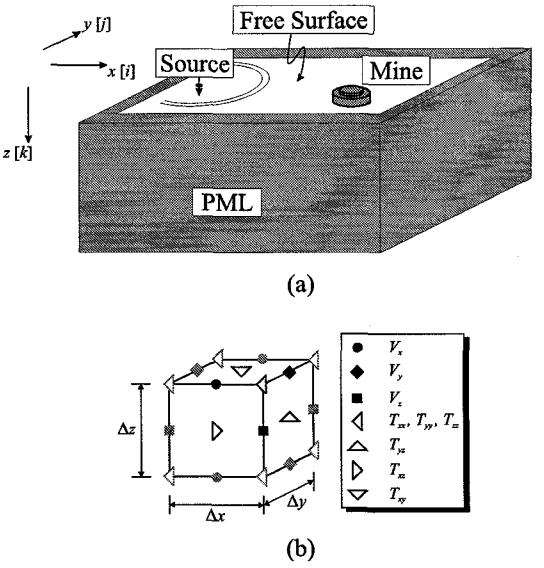


Fig. 1. Three-dimensional finite-difference model; (a) layout, (b) finite-difference cell.

The elastic wave motion in solids is described by a set of fundamental partial-differential equations: the equation of motion relating the particle velocity vector and the mechanical stress tensor, the strain-velocity relation and the elastic constitutive relation. Combining these equations, a first-order system of equations is obtained describing the elastic wave fields entirely in terms of the particle velocity and the mechanical stress. In three dimensions, nine

0-7803-6362-0/00/$10.00 (c) 2000 IEEE

equations for three unknown velocity components and six unknown stress tensor components arise.

The equations are discretized by introducing finite-differences. The discretization leads to the characteristic finite-difference grid. In this grid, the field components are staggered in space and time. The finite-difference grid can be thought of as being comprised of basis cells. The three-dimensional basis cell for the elastodynamic case resembles strongly the three-dimensional basis cell for electromagnetic finite-difference modeling, the so-called *Yee*-cell. However, due to the stress being a tensor, more field components are present in the elastodynamic case. Fig. 1 (b) shows the three-dimensional finite-difference basis cell. Note that the field components are not known at the same points in space and time. The grid is laid out such that each field component is surrounded by the field components it is dependent on.

The finite-difference model has been implemented in a fully parallel fashion. The computations for this paper's results have been performed both on a Cray T3E parallel supercomputer located at the ERDC Massively Shared Resource Center in Vicksburg, Mississippi, and on a Beowulf PC cluster located at the Georgia Institute of Technology. The Beowulf cluster has been developed and built especially for the model described in this paper.

INTERACTION OF ELASTIC WAVES WITH BURIED ANTIPERSONNEL MINES

The interaction of elastic waves with antipersonnel mines, buried in sand, is to be investigated. In the experiment, the mines are buried in a large sand-filled box. If the sand is assumed to be linear, isotropic and lossless, its properties can be described by three parameters: the material density, $\rho = 1400$ kg/m^3, the pressure wave speed, $c_p = 250$ m/s, and the shear wave speed. The shear wave speed is not constant and increases with depth.

The space step for the numerical model is chosen to be $\Delta x = \Delta y = \Delta z = 0.5$ cm; the time step is $\Delta t = 11.54$ μs and, thus, fulfills the Courant condition (the necessary condition for stability of the finite-difference algorithm). The excitation has the shape of a Gaussian pulse. To be able to compare the numerical and experimental results, the transfer function of the system is determined in a post-processing step for each point in space and convolved with the transducer motion as measured in the experiment.

To investigate the interactions with a buried land mine, a simple model for an antipersonnel mine is inserted into the numerical model. Fig. 2 shows the simple model, together with a simplified cross-sectional drawing of a real TS-50 antipersonnel mine. The simple mine model consists of a main chamber containing plastic explosives, and a smaller chamber on top of the mine's main chamber filled with air. The air-filled chamber is inserted into the model to coarsely approximate the structure of a real TS-50 mine (see Fig. 2 (a)).

Fig. 3 shows waterfall graphs of the mine-wave interaction for both experiment and numerical simulation. In these graphs, the vertical particle displacement at a number of points along a line on the surface is plotted as a function of time and offset by the distance from the source. The slope of the traveling waves in the graph indicates the wave speed. Thus, by determining the slope, the different wave types can be distinguished. The mine is buried 2 cm beneath the surface, at a distance of 70 cm from the source. A pressure wave (P) and a Rayleigh surface wave (R) arise. The waves hit the mine and are reflected (rR) and transmitted. While the interaction of the mine with the pressure wave is weak, the surface wave strongly interacts with the mine. In both experiment and numerical simulation, resonant oscillations occur at the mine location and remain even after the waves have passed the mine. For the numerical model, it can be shown that the incident waves couple into flexural waves which arise in the thin soil layer above the mine [4]. These flexural waves are confined to the thin layer and form a standing wave pattern, giving rise to the resonant oscillations. While this explains the resonance in the numerical model, it gives only one possible cause for the resonance in the experiment. A real TS-50 mine has several chambers, it has a flexible case that can support both flexural and longitudinal waves, and it contains springs that can also give rise to resonances. The authors are currently working on refining the numerical model to incorporate more details of the mine.

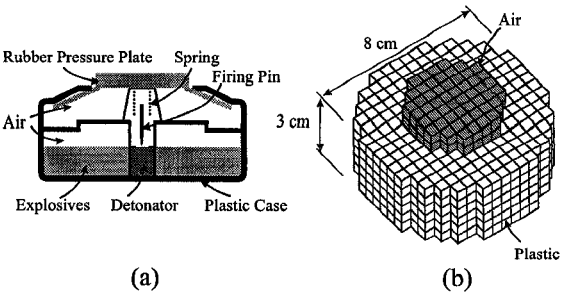
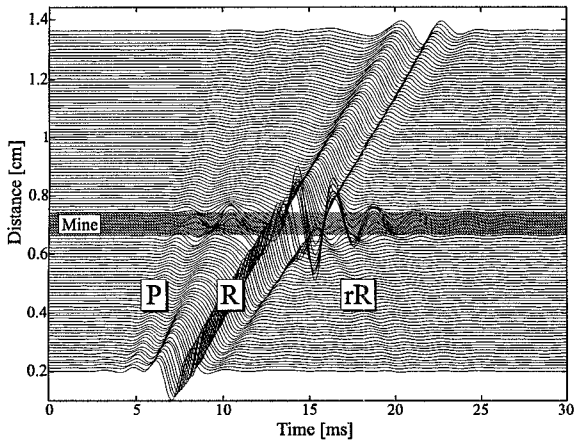
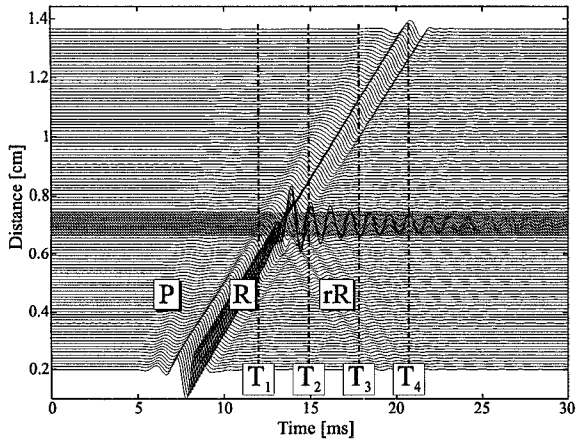


Fig. 2. (a) Cross-sectional drawing of a real TS-50 mine; (b) simple model.

In Fig. 4, the vertical particle displacement on the surface and on a cross section through the ground is shown in some pseudo-color plots as obtained with the numerical model. The upper plots show the wave fields on the surface, the lower plots correspond to a cross section through the ground. The surface plane has dimensions of 120 cm by 80 cm and the cross section has a size of 120 cm by 30 cm. The source is located on the surface, off the left edge of the plots. The dynamic range of the plots is 60 dB. The wave fields are shown at four different instances in time, corresponding to the vertical lines indicated by T_1, T_2, T_3 and T_4 in Fig. 3 (b). In the first plot, the surface wave is seen to just hit the mine. While only the surface wave (R) is visible on the surface, both the surface wave (R) and the shear wave (S) appear on the cross section. In the second plot, the surface wave has just passed the mine. The interaction of the surface wave with the mine gives rise to reflected surface waves (rR) and reflected shear waves (rS), which are clearly visible on the cross section. Pressure waves are also induced by the surface wave, but they



(a)



(b)

Fig. 3. Interaction of elastic waves with a buried antipersonnel mine; waterfall graphs of the vertical particle displacement on the surface according to (a) experiment and (b) numerical simulation.

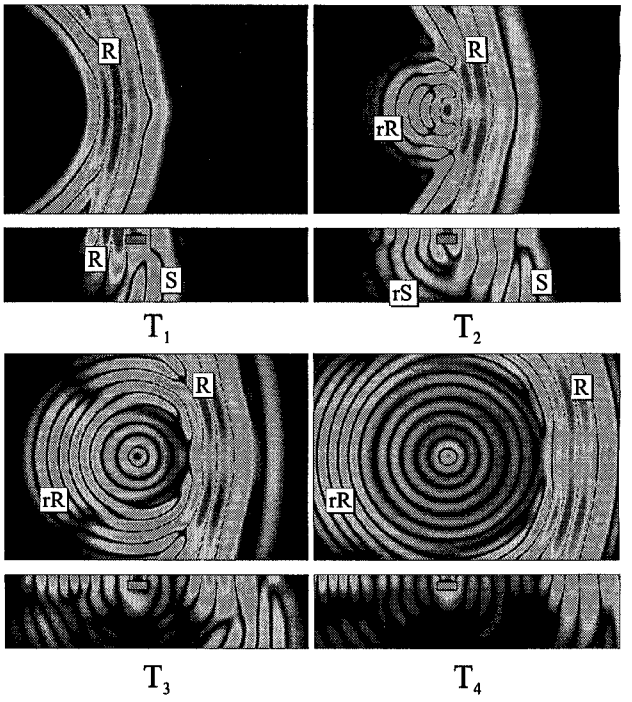


Fig. 4. Interaction of elastic waves with a buried antipersonnel mine; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instances in time, corresponding to the vertical lines in Fig. 3 (b).

are weak and not visible. In the third and fourth plot, the surface wave has passed the mine. Some energy, however, remains at the mine and causes the mine to vibrate and to radiate.

CONCLUSIONS

A three-dimensional finite-difference model for elastic waves in the ground has been developed and implemented in a fully parallel fashion. Results are obtained and compared to experimental results. When an antipersonnel mine interacts with elastic waves, a strong resonance occurs at the mine location. This resonance has been observed in both experiment and numerical simulation.

REFERENCES

[1] W. R. Scott, Jr., C. T. Schröder, and J. S. Martin, "An acousto-electromagnetic sensor for locating land mines," in Detection and Remediation Technologies for Mines and Minelike Targets III, Proc. SPIE, 1998, vol. 3392, pp. 176–186.

[2] W. R. Scott, Jr. and J. S. Martin, "Experimental investigation of the acousto-electromagnetic sensor for locating land mines," in Detection and Remediation Technologies for Mines and Minelike Targets IV, Proc. SPIE, 1999.

[3] C. T. Schröder and W. R. Scott, Jr., "Three-dimensional FDTD model to study the elastic-wave interaction with buried land mines," in Detection and Remediation Technologies for Mines and Minelike Targets V, Proc. SPIE, 2000.

[4] C. T. Schröder and W. R. Scott, Jr., "A finite-difference model to study the elastic-wave interactions with buried land mines," IEEE Trans. on Geophysics and Remote Sensing, to be published.

[5] W. C. Chew and Q. H. Liu, "Perfectly matched layer for elastodynamics; a new absorbing boundary condition," J. Comput. Acoustics, vol. 4, pp. 341–359, 1996.

0-7803-6362-0/00/$10.00 (c) 2000 IEEE

Use of Elastic Waves for the Detection of Buried Land Mines

Waymond R. Scott, Jr.[a], Christoph T. Schroeder[a], James S. Martin[b], and Gregg D. Larson[b]
[a]School of Electrical and Computer Engineering
[b]School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332
Phone: 404-894-3048
EMail: waymond.scott@ece.gatech.edu

Abstract — A prototype system that employs seismic surface (Rayleigh) waves and non-contact displacement sensors for the detection of buried landmines is being developed. The operating principle of this system is to interrogate the ground surface motion, resulting from a remote seismic source, in a region surrounding, and immediately above, a mine using a two-dimensional array of non-contact sensors. The data from this array is then processed in the space and time domains to form images. These images emphasize specific features of the ground motion, which have been linked empirically, through both numerical modeling and experimental studies, to the presence of landmines. The images formed in this way show a large contrast (~30 dB) between mines and background reverberation and a large contrast (>20 dB) between mines and mine-sized clutter objects. They also appear to be robust in simulations of realistic burial scenarios involving a variety of mine sizes and depths along with clutter objects in close proximity.

INTRODUCTION

Seismic/elastic techniques show considerable promise for the reliable detection of all types of buried mines, even low-metal anti-personnel mines. The reason for this is that mines have mechanical properties that are significantly different from soils and typical forms of clutter. For example, the shear wave velocity is approximately 20 times higher in the explosive and the plastics used in typical mines than in the surrounding soil. In addition, mines are complex mechanical structures with a flexible case, a trigger assembly, air pockets, etc. This complex structure gives rise to structural resonances, non-linear interactions, and other phenomena that are atypical for both naturally occurring and man-made forms of clutter. This phenomenology can be used to distinguish a mine from clutter.

A mine detection system has been developed at Georgia Tech that employs sensors, which measure local seismic displacements without physically contacting the soil surface [1]. The non-contact nature of these sensors allows interrogation of the soil surface near or immediately above a mine. This substantially increases the measurable effects of the mine's presence over schemes which rely on elastic waves scattered by the mine to propagate to a remote sensor location [2].

This work is supported in part by the US Army Research Office under contract DAAH04-96-1-0048 and by the US Office of Naval Research under contract N00014-99-1-0995.

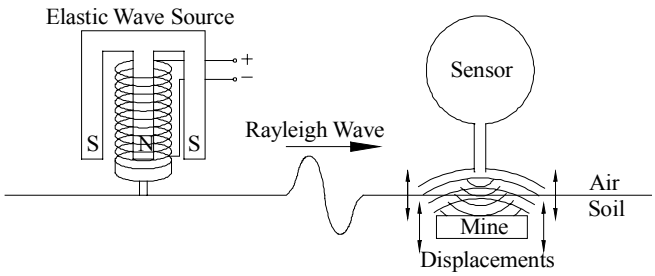


Fig. 1. Diagram of the mine detection system.

A laboratory prototype for the mine detection system is currently being studied. In the prototype, the two-dimensional array of sensors is formed synthetically by scanning a single sensor over the measurement region and exciting the ground with identical seismic signals for each measurement location. The configuration of this system is shown in Fig. 1. In the experimental model, the soil is modeled by a tank that is approximately 6 m wide, 6 m long and 1.5 m deep filled with damp compacted sand. The source depicted here is a 100 lb. electrodynamic shaker that has been coupled to the ground using an elongated foot attached to the shaker head (moving coil). The long dimension of the foot is parallel to the excited wave fronts. This was found to preferentially excite surface waves and to direct energy toward the measurement region rather than the side walls of the tank in which the tests are conducted. The shaker is free standing so that the ground is driven against the tail mass (permanent magnet).

As the Rayleigh wave propagates across the minefield, it excites motion of compliant objects such as mines and is scattered from inhomogeneties, which include both mines and clutter. The amplitude of the Rayleigh wave displacement decreases exponentially with depth and only the soil near the surface is interrogated by this signal. The depth of soil that is examined is a function of the frequency of the source. For the laboratory system this is typically in the l00 to 1,000 Hz range. Somewhat lower frequencies may be used in a field system that is not limited by reverberation in a restricted tank volume, as the laboratory system is.

The motion of the mine is different from the motion of the surrounding soil and clutter, because of the unique mechanical characteristics of a mine. Both antipersonnel (AP) and antitank (AT) mines with pressure sensing triggers exhibit structural resonances within the bandwidth of the current interrogation signal. These are excited by the passage of a Rayleigh wave. A consequence of this is that the motion above this type of mine is amplified relative to the incident motion and persists in time following the passage of the incident wave. The resonance is damped due to loss effects in the soil and energy that is radiated into the soil.

Numerical models have been developed which mimic the experimental observations of resonant mine behavior. These are three-dimensional finite-difference time-domain (3-D FDTD) models, which represent the soil as a linear, isotropic, elastic, half-space with material properties that have been determined from the experimental studies. In the numerical model, mines are represented as simplified structures with the approximate size, shape, and density of the actual mine and a closed air cavity to simulate a pressure-sensing trigger. Both the stiffness and the mass of the overlying soil and the case of the mine influence the resonant behavior. The models have also shown that the surface motion excited above a resonant mine greatly exceeds that of natural clutter objects such as rocks with a wide range of sizes and material properties.

DATA PROCESSING

The data acquired with the laboratory mine detection system is recorded as a transfer function (surface displacements relative to the input drive signal) in the frequency domain. As such, the raw data are four-dimensional (two spatial dimensions over which the synthetic array is formed and two dimensions for the amplitude of the surface displacements as a function of frequency). A meaningful image-processing algorithm must reduce this data to the two spatial dimensions and a third representing the probability of the presence of a buried land mine. Some authors have suggested that, under similar circumstances, the amplitude of measured displacement at a point is sufficient to represent this third dimension [3]. The imaging scheme that is currently employed with the laboratory system is considerably more complex than this approach as it exploits features of the observed phenomena to enhance the background contrast over that which can be achieved with this simpler scheme.

The current imaging algorithm, which has been described in a previous paper [4], first synthesizes the response of the system to a differentiated Gaussian pulse, which typically has a center frequency of 450 Hz. The data is filtered temporally and spatially to remove low and high frequency noise. Then, the forward-propagating waves are filtered out leaving only the back-scattered waves. Two data sets are created: the first data set is the energy at each spatial position in the back-scattered waves at times of arrival near the incident wave, while the second data set is the energy that propagates back toward the seismic source from each spatial position. Finally, taking the product of these two data sets forms the image.

MULTIPLE MINE BURIAL SCENARIO

Previous papers have demonstrated the feasibility of detecting both AP and AT mines using the laboratory system and of distinguishing these from buried clutter [1]. Recent investigations have focussed on imaging individual mines in close proximity to each other while distinguishing them from clutter. A common practice of mine warfare is to plant multiple AP mines in close proximity to AT mines. The AP mines thereby protect the AT mine from sappers who can more easily detect the larger object and remove it with little personal danger. This poses a unique detection problem in that it requires a system to operate with sensitivity appropriate to both mine types simultaneously. Also, the system must be capable of distinguishing individual targets and rejecting ghost images formed by multiple scattering effects.

To address this scenario experimentally, an experiment was performed in which an AT mine was surrounded by four AP mines and four AP mine-sized clutter objects. The AT mine was an inert VS-1.6 mine. It was surrounded by an improbable assortment of AP mines: M-14, TS-50, VS-50, and PFM-1 mines. The arrangement of this burial and the relative scale of the objects can be seen in Fig. 2.

The image formed from the data taken over the multiple mine burial is shown in Fig. 3. The number of mines present and their relative locations have been accurately depicted. The image of the AT mine is seen to be strongest at the back edge; this is due to the reflection at the back edge being stronger than that at the front. The effects of the rocks are much smaller than those of any of the mines. The largest rock is barely discernable with the 30-dB dynamic range used to generate the image. It can be seen in the lower left of the figure. The reason for this is that the rocks do not exhibit resonances within the frequency range of the incident signal.

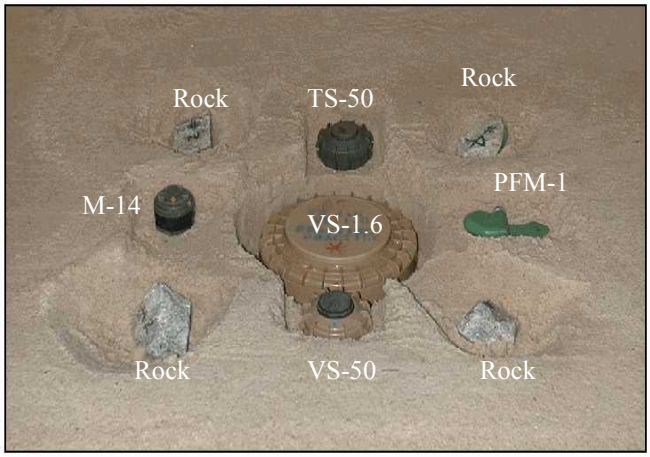


Fig. 2. VS 1.6 AT mine surrounded by TS-50, PFM-1, VS-50, and M-14 AP mines and rocks. The burial depths are 4.5 cm for the VS-1.6; 2cm for the TS-50, VS-50, and PFM-1; and 0.5cm for the M-14. The burial depths for the rocks were 3.5cm, 1.5cm, 2cm, and 1cm (clockwise, starting with the upper left rock).

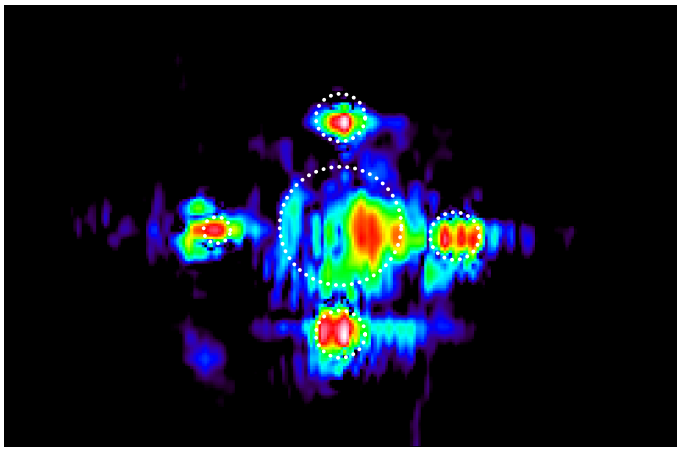


Fig. 3. Image formed of AT mine surrounded by AT mines and rocks. Dynamic range is 30 dB. Physical dimensions are 120 cm by 80 cm.

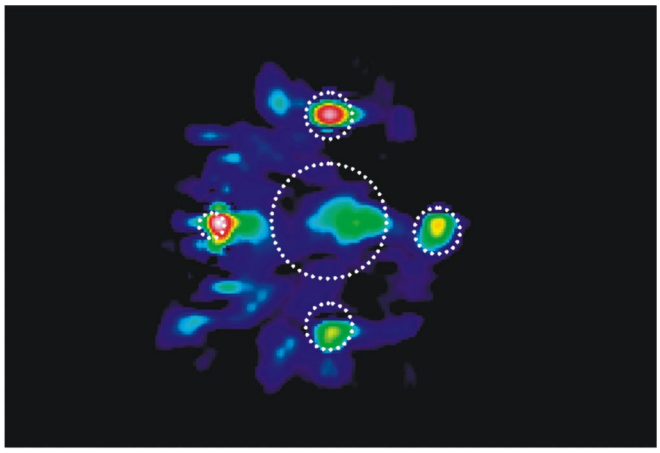


Fig. 4. Image formed from numerical simulation of data from the multiple mine burial scenario on a 30-dB scale.

Processing artifacts of the removal of forward propagating waves from the data cause the images of the individual mines to be smeared horizontally. This does not, however, mask the presence of the AP mines to the left or right of the AT mine. Nor does the seismic shadow (forward scattered signature) of the AT mine appear to degrade the image of the PFM-1 AP mine directly behind it.

This scenario was modeled numerically using the 3-D FDTD model [5]. The image in Fig. 4 was formed using these model results as input to the imaging algorithm. The similarities between this image and the experimental image are readily apparent. The model underestimates the signatures of 3 of the 5 mines. This is partially due to the lack of some important details in the model regarding the mechanical structure of these mines. Another factor may be that the depth profile of the material properties in the sand used for the model only approximately represents the actual depth profile, altering the surface motion at the mine locations. The numerical model is linear but experiments have shown that the sand is highly nonlinear and this creates increasing loss at increasing frequencies; thus, the response in the numerical model is biased toward higher frequencies enhancing the response of the smaller mines. Hence, the response of the AT mine is less pronounced in the numerical results. It can also be argued, that these numerical results show that the imaging algorithm can be significantly improved. The image from the numerical results has obvious clutter, but the numerical model has no unspecified clutter or noise. Thus, the clutter in the image must be an artifact of the algorithm that produced it.

The model does, however, predict the placement of the mine's image at the back of the true extent of the mine, just as it has occurred in the experiment. It is possible that the enhancement in the experimentally measured mine signature is due to the trenching effect of its burial. This was the focus of a previous theoretical and experimental study [6].

CONCLUSIONS

Imaging of multiple mines buried in close proximity in the presence of clutter has been demonstrated using a seismic imaging system. The results of this experiment are in good agreement with a numerical model. Both model and experiment show the system's ability to individually image small AP mines in close proximity to a much larger AT mine. Model and experiment also demonstrate that natural mine-sized clutter, in similarly close proximity, does not pose a serious problem for system operation.

REFERENCES

1. Scott, W.R., Jr., Martin, J.S, and Larson, G.D, "Experimental Model for a Seismic Landmine Detection System," *IEEE Transactions on Geoscience and Remote Sensing,* to appear July 2001.
2. "Feasibility of Acoustic Landmine Detection: Final Technical Report," *BBN Technical Report No. 7677*, May 1992.
3. Sabatier J.M. and Xiang N., "Acoustic to Seismic Coupling and Detection of Landmines," *Proceedings of the IEEE 2000 International Geoscience and Remote Sensing Symposium*, Honolulu, Hawaii, July 2000.
4. Behboodian, A., Scott, W.R., Jr. and McClellan, J.H. "Signal Processing of Elastic Surface Waves for Localizing Buried Land Mines," *Proceedings of the 33rd Assilomar Conference on Signals, Systems, and Computers*, Assilomar, CA, October 1999.
5. Schroeder, C.T. and Scott, W.R., Jr., "Three-Dimensional Finite-Difference Time-Domain Model for Interaction of Elastic Waves with Buried Land Mines," *Proceedings of the SPIE: 2000 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, Vol. 4038, April 2000.
6. Scott, W.R., Jr., Schroeder, C.T., Martin, J.S., and Larson, G.D., "Investigation of a Technique that Uses Both Elastic and Electromagnetic Waves to Detect Buried Land Mines," *Proceedings of the AP2000 – Millennium Conference an Antennas and Propagation*, Davos, Switzerland, April 2000.

A Comparison Study of the Surface Scattering Models and Numerical Model

H.T. Ewe[1], Joel T. Johnson[2] and K.S. Chen[3]
[1]Faculty of Engineering, Multimedia University,
Cyberjaya 63100, Selangor, MALAYSIA
Email: htewe@mmu.edu.my
[2]Department of Electrical Engineering and Electro-Science Laboratory,
The Ohio State University, Columbus, OH 43210 USA
[3]Center for Space and Remote Sensing Research and Institute of Space Science,
National Central University, Chung-Li, Taiwan 32054

Abstract – This paper describes a comparison study of surface scattering models and a numerical model for dielectric surfaces. Two surface scattering models namely the Integral Equation Model (IEM) and the Small Slope Approximation (SSA) model are used to calculate the backscattering coefficients of rough surfaces and the results are compared with numerical simulations based on the Moment Method (MoM). Analysis of the results obtained is also presented.

I. INTRODUCTION

In the study of the interaction of microwaves with media bounded by rough interfaces, a practical and realistic surface scattering model is required to understand the surface scattering mechanism and predict the scattering returns for a variety of surface profiles. The usability and correctness of the model are essential in microwave remote sensing and a good understanding of the validity and limitation of the models will be necessary for modeling active and passive microwave scattering mechanisms for earth terrain and also inversion models. Surface scattering models based on both the small perturbation method (SPM) and the Kirchhoff model had been widely used in the past in the theoretical modeling of microwave remote sensing [1]. Recent models such as the Integral Equation Model (IEM) [2] and the Small Slope Approximation (SSA) model [3] have shown great promise in the prediction of surface scattering returns from various research studies. In this paper, a comparison study of the IEM model and the SSA model is carried out for dielectric rough surfaces. To compare the predictions of backscattering coefficients from these two models with the actual backscattering returns from the surface, numerical simulations based on the Moment Method (MoM) are generated for the comparison. It is generally found that both the IEM and the SSA model with the inclusion of up to second order surface slope are good models in comparison with the predictions from the numerical models. The results of this study also provide useful information for the suitable applications of the two models in practical surface scattering calculations.

II. CONFIGURATION

Fig. 1 shows the configuration of the study where ε_r is the complex dielectric constant of the surface and θ_i is the incident angle. A 2-D (3-D scattering problem) isotropic Gaussian correlated surface with surface rms height σ and surface correlation length l is considered.

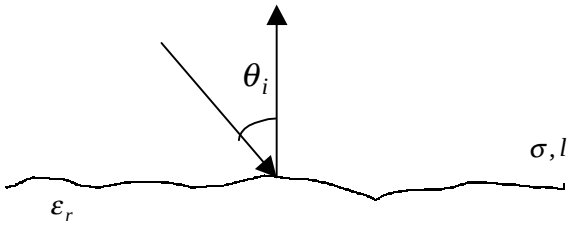


Fig 1. Surface configuration

III. SURFACE SCATTERING MODELS

a. Small Slope Approximation (SSA) Model

The Small Slope Approximation (SSA) [3] is chosen for the comparison as recent studies have shown that this is a robust model for surface scattering. The SSA is based on a series expansion in terms of surface "quasi-slope", with evaluation of average incoherent cross sections from the first term (first order in slope) requiring a calculation similar to that of the Kirchhoff approximation but retaining agreement with perturbation theory in the small height limit. Expressions for average cross sections accurate to second order in slope are also available [3] but require additional integrations. An alternate approach for evaluation of higher order corrections is described in [4] based on Monte Carlo simulation. The Monte Carlo procedure is applied in this paper, with averaged second order in slope cross sections generated from 100 realisations of 128 by 32 lambda periodic

surfaces sampled into 1024 by 256 points under plane wave incidence. Because the higher order SSA calculation can be implemented with the Fourier transform, the computation is very efficient compared to a Monte Carlo simulation with the Method of Moments.

b. Integral Equation Model (IEM)

For the Integral Equation Model (IEM) in [2], it is known that the model provides a good prediction of surface scattering coefficients for a wide range of surface profiles which include the limits of both the classical Kirchhoff model (KM) and the small perturbation model (SPM). In the expression of the IEM model, three terms of surface scattering contributions, namely the Kirchhoff term (*k*), the cross term (*kc*) and the complementary term (*c*) are included and shown in (1), respectively:

$$\sigma_{qp}^{o} = \sigma_{qp}^{k} + \sigma_{qp}^{kc} + \sigma_{qp}^{c} \quad (1)$$

where σ_{qp}^{0} is the surface backscattering coefficient and the subscripts *q* and *p* denote the scattered polarisation and the incident polarisation, respectively. In this study, an improved model of IEM is used where this model is included with the empirical reflection coefficient model [5] that provides smooth transition in the Fresnel reflection coefficient approximation across a range of angular and frequency conditions.

c. Method of Moments (MOM)

Approximate model results are compared with an "exact" numerical model based on Monte Carlo simulation with an iterative method of moments based surface scattering computation. Computational time for the numerical model is reduced through the use of the "canonical grid" algorithm for computing matrix-vector multiplies to order (N log N) where N is the number of surface sampling points. A detailed description of the algorithm is provided in [6]-[7]. To further keep computational times reasonable, numerical results use 32 realisations of 16 x 16 lambda surfaces sampled into 128 by 128 points, and the Monte Carlo simulation is performed through the use of a IBM P2SC parallel computing resources at the Maui High Performance Computing Center [8]. Due to the finite size of the surfaces modeled in the MOM, a "tapered-wave" incident field is used to avoid artificial surface edge scattering effects. The tapered wave and surface size chosen make MOM results accurate only to up to incidence angle 50 degrees, so the comparison is not continued outside this range. Numerical model results are also not included for incidence angle 0 degrees due to the dominance of the coherent scattered field at this angle for the surface statistics considered.

IV. RESULTS AND DISCUSSION

In this study, a Gaussian correlated surface with complex dielectric constant 4+j1 is used. A surface profile with $k\sigma$ = 0.5, kl = 3.0 is chosen where *k* is the wave number. Note these surface statistics have only a moderate roughness, but the relatively large slopes encountered make the SPM and Kirchhoff models inaccurate here. Backscattering coefficients are compared for the IEM and both the first and second order SSA. In the figures, only second order SSA results are shown together with the IEM and numerical calculations.

Fig. 2 show the comparison of the VV backscattering coefficients calculated from the IEM, SSA model (second order) and numerical model. It is found from the calculations that the SSA model (first order) predictions are correct only for low incident angle (< 20°). For higher angles than this, its predictions are generally lower than those of the other models. However, the second order SSA, IEM, and numerical model all show good agreement for the range of angles considered. Although the data for numerical model is only up to 50°, it is found that both the IEM and SSA model continue to show agreement up to 70°.

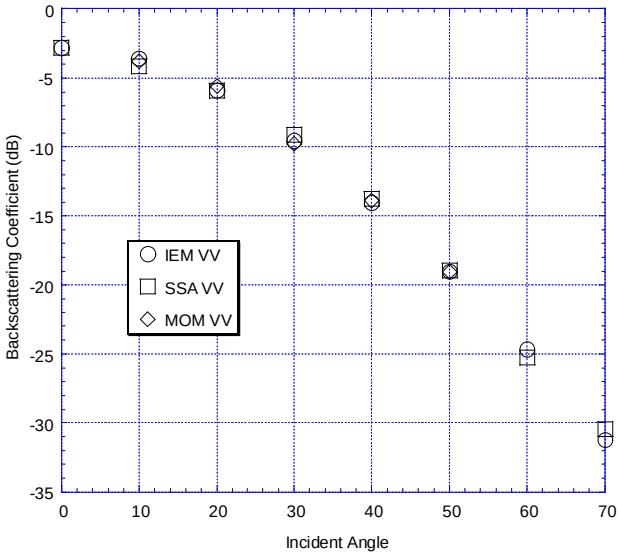


Fig 2. Comparison of the calculations of different surface models for VV polarisation ($k\sigma$ = 0.5, kl = 3.0).

In Fig. 3, the same comparison is presented for HH polarisation. It is again found that the SSA model (first order) calculations are accurate only for low incident angles and generally give higher backscattering coefficients for high incident angles as compared with the other models. The match between the SSA model (second order) calculations and the numerical model calculations is again very good and those of the IEM also compare very well with the numerical models for the particular problem considered in this example.

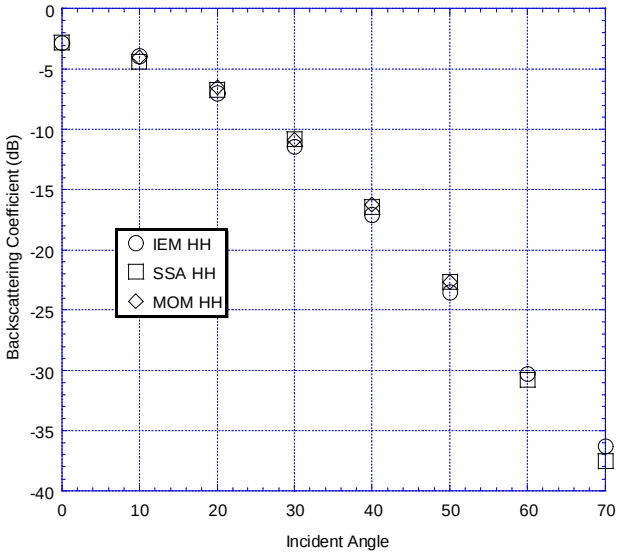


Fig 3. Comparison of the calculations of different surface models for HH polarisation ($k\sigma$ = 0.5, kl = 3.0).

V. CONCLUDING REMARKS

Generally, it is found in this example that both the SSA model with up to second order terms of surface slope and the IEM compare well with the numerical calculations for like-polarisation in the backscattering directions. Additional examples and studies of bistatic cross sections are currently in progress.

REFERENCES

[1] F. T. Ulaby, R. K. Moore and A. K. Fung, A. K., *Microwave Remote Sensing: Active and Passive (Vol. II)*. Reading, Massachusetts: Addison-Wesley, 1982

[2] A.K. Fung, *Microwave Scattering and Emission Models and Their Applications*, Artech House, Norwood, MA, 1994.

[3] A. G. Voronovich, *Wave Scattering from Rough Surfaces*. Berlin: Springer-Verlag, 1994.

[4] S. T. McDaniel, “Acoustic and radar scattering from directional seas," *Waves in Random Media*, vol. 9, no. 4, pp. 537--549, 1999.

[5] T. D. Wu, K. S. Chen, A. K. Fung and M. K. Tsay, “A Transition Model for the Reflection Coefficient in Surface Scattering”, *Proc. Of IGARSS’1998*, pp. 2375-2377.

[6] J. T. Johnson, R. T. Shin, J. A. Kong, L. Tsang, and K. Pak, ``A numerical study of ocean polarimetric thermal emission," *IEEE Trans. Geosc. Remote Sens.* vol. 37, no. 1 (part I), pp. 8--20, 1999.

[7] J. T. Johnson, R. T. Shin, J. A. Kong, L. Tsang, and K. Pak, ``A numerical study of the composite surface model for ocean scattering," *IEEE Trans. Geosc. Remote Sens.*, vol. 36, no. 1, pp. 72-83, 1998.

[8] Maui High Performance Computing Center WWW site, at http://www.mhpcc.edu

Radar Image Studies of Scattering from Random Rough Surfaces

Hyunjun Kim and Joel T. Johnson
The ElectroScience Laboratory
The Ohio State University
1320 Kinnear Rd, Columbus, OH 43212
Email : hjk@esl.eng.ohio-state.edu

Abstract—
In this paper electromagnetic scattering from deterministic rough surfaces is described through analysis of radar images. Backscatter results are considered for 1-D random rough surfaces which satisfy an impedance boundary condition (IBC). Radar images of single realizations of random surfaces are formed through back-projection tomography. Detailed investigations of the images are also provided to clarify major and secondary scattering events, angular dependencies, and polarization effects, and a ray tracing analysis is performed to predict multiple scattering images. For an ocean-like surface, effects of surface length scale components are discussed.

1. INTRODUCTION

In recent years several analytical models for rough surface scattering have shown promising results for specified ranges of surface statistics [1]. No approximate solutions, however, clearly explain all possible scattering mechanisms for a single surface realization since evaluation of analytical theories is typically based on results for cross sections averaged over surface realizations. Due to lack of information on the physical behavior of scattering from rough surfaces, a more descriptive approach is necessary to understand the existing theories. Imaging techniques offer a unique tool for analysis and understanding of rough surface scattering phenomena [2].

Recent development of efficient numerical methods for backscattering predictions enable radar image formation. Exact numerical results can provide useful information on important scattering features from single surface realizations when the existing analytical models are not valid any longer. Also, images formed from numerical scattering models can be used as a reference solution to evaluate the performance of the existing models, and to develop more accurate analytical scattering models.

In Section 2, the scattering geometry for deterministic rough surfaces is described. In Section 3, the theoretical background for an efficient numerical model and construction of 2-D inverse synthetic aperture radar (ISAR) images is included. In Section 4, various scattering features captured from the images are examined.

2. SCATTERING GEOMETRY

One dimensional realizations of Gaussian random process surfaces with Gaussian or Pierson-Moskowitz (P-M) spectra are considered in this paper. Surfaces with a Gaussian spectrum can be characterized by the rms height (σ) and correlation length (l_c). A 1.92 m surface with rms height 2.0 cm and correlation length 7.5 cm is used for the Gaussian surface case.

Unlike the Gaussian surface, the P-M ocean-like surface has a multi-scale roughness controlled by maximum and minimum scale cutoff frequencies (k_{du} and k_{dl}) in the surface spectrum. A single realization of a P-M ocean-like surface is generated for 6 m/sec wind speed with $k_{du} = 587$ rad/m and $k_{dl} = 0.001$ rad/m so that all important length scale variations at X- to Ku-bands for a 7.68 m surface length are included. Large and small scale roughness effects of the P-M surface at large incidence angles will be discussed.

The medium is assumed to be described by an impedance boundary condition (IBC) with a relative permittivity of $(39.7 + i40.2)$ to approximate sea water. The surface is truncated at $x = L/2$ and $-L/2$ and a phase-corrected tapered beam with spot size $g = L/5$ is used to confine the incident field to this area so that edge effects can be avoided.

3. BACKGROUND THEORIES

An iterative method of moments based on the forward-backward (FB) method has been successfully applied to scattering from 1-D surfaces with relatively large rms heights. Several fast techniques have been proposed to accelerate the computational efficiency for the 1-D FB method. To improve the efficiency of the 1-D FB computation, the novel spectral acceleration (NSA) algorithm has been developed [3]. A spectral domain representation of the scalar Green's function and the source current is

used to compute the coupling between points separated by wide distances. For a large number of unknowns, the resulting computation count and memory requirement reduce to $O(N)$ for fixed surface statistics.

A 2-D ISAR image of a deterministic surface is constructed from a set of frequency and angular swept complex backscatter field data. Back projection tomography using an inverse Fourier transform is employed to generate the images [4]. Backscatter data were collected over 4 GHz frequency bandwidth (10 - 14 GHz) and a 20° angular bandwidth corresponding 3.75 cm down- and 3.65 cm cross-range resolution in the image domain, respectively. Frequency and angle steps are appropriately chosen so that unambiguous ranges include the possible ranges of secondary scattering sources in cases involving multiple scattering. To reduce the side-lobe level, the Hamming window is used.

4. RESULTS AND DISCUSSIONS

First, we investigate ISAR images of the Gaussian roughness surface for aspect angles centered at normal incidence. The two images in Figure 1 show the horizontal (HH) and vertical (VV) polarization images constructed from the exact numerical results (FB/NSA). Each image is expressed within the dynamic range of 60 dB and composed of 200 × 160 pixels in a 2 m × 1.6 m range so that each pixel size is much smaller than the range resolution.

Note that the images have a maximum scattering level near the center of the surface due to the tapered wave illumination on the surface. The surface profile is also overlayed to match the scattering centers and the corresponding surface points. The scattering centers come primarily from the near specular points with maximum image levels of about -3.4 dB for both polarizations, showing very little polarization dependence as expected. The images for both polarizations show additional scattering sources below the surface, possibly from multiple scattering effects.

To study the origin of these additional scattering points, a ray tracing analysis is carried out to predict the locations of points which occur due to double reflection. In this analysis, we draw three rays which represent the incident ray on one point, the horizontally propagating ray connecting two points and the scattered ray on the other point, respectively. Once possible double reflection points are found, the corresponding time delays of each ray is calculated. The time delayed points are then placed along the down-range from the middle of two points as shown in Figure 2. Overlayed by the images of Figure 1, the predicted points match the time-delayed images from multiple scattering effects. The intensity of the secondary scattering images becomes higher as the rms height increases due to the strong near specular interactions.

As a second example, the ocean-like rough surface is investigated through radar images at incident angles centered at 70°. Due to the wide range of length scales in the surface roughness, radar images are expected to show contributions from both long scale and small scale portions. As shown in Figure 3, some polarization differences are observed unlike the Gaussian surface case at normal incidence. As indicated in the dynamic range, VV polarization is about 15 dB above HH polarization in the overall backscatter level.

Another important difference between polarizations is the scattering source distribution over the entire surface. As observed in the figure, most of the single scattering returns for HH backscatter come from portions with large local slopes and show little contribution from the backside of the surface. On the other hand, for VV returns the scattering sources are more evenly distributed over the surface showing some contributions even from the geometrically shadowed regions. These phenomena have been addressed from the statistics of ocean backscatter study by Donohue [5], in which examination of induced current distributions and local radar cross sections revealed different scattering behavior between polarizations, similar to the results observed in the radar images.

In addition to the strong single scattering returns due to the Bragg components of the surface, some multiple scattering effects are also observed below the surface for both polarizations. By filtering the spectrum contents of the surface, the origin of these multiple scattering effects can be analyzed. By changing the limit of high or low cutoff frequencies of surface spectrum, it is found that multiple scattering events and the overall backscattering levels are closely related to the Bragg scattering component of the surface in conjunction with the large scale variations.

5. REFERENCES

[1] A. G. Voronovich, *Wave Scattering from Rough Surfaces,* Springer Series on Wave Phenomena 17, 1995.

[2] H. Kim, and J. T. Johnson, "Radar images of rough surface scattering: Comparison of numerical and analytical models," *submitted to IEEE Trans. Antennas Prop.*

[3] H.-T. Chou and J. T. Johnson, "A novel acceleration algorithm for the computation of scattering from rough surfaces with the forward-backward method," *Radio Science*, vol. 33, pp. 1277–1287, 1998.

[4] D. L. Mensa, *High Resolution Radar Cross Section Imaging,* Artech House, 1991.

[5] D. J. Donahue, H.-C. Ju, and D. R. Thompson, "Polarization dependent radar backscatter calculations from cresting ocean waves," *Proc. URSI*, Boulder, CO, p. 262, 1999.

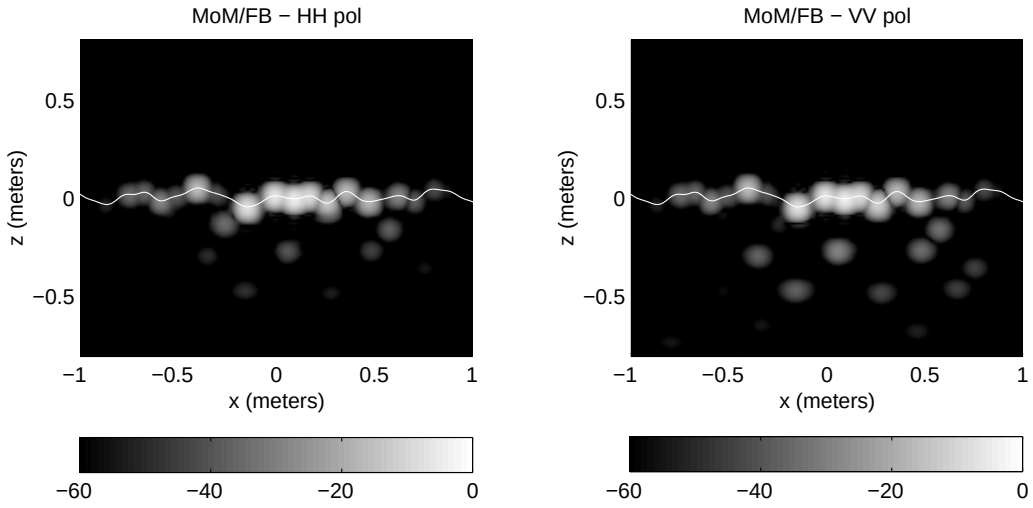


Figure 1: 2-D ISAR images for a Gaussian surface: $f = 10 \sim 14$ GHz, $\delta f = 50$ MHz, $\theta_i = -10^\circ \sim 10^\circ$ and $\delta\theta = 0.2^\circ$

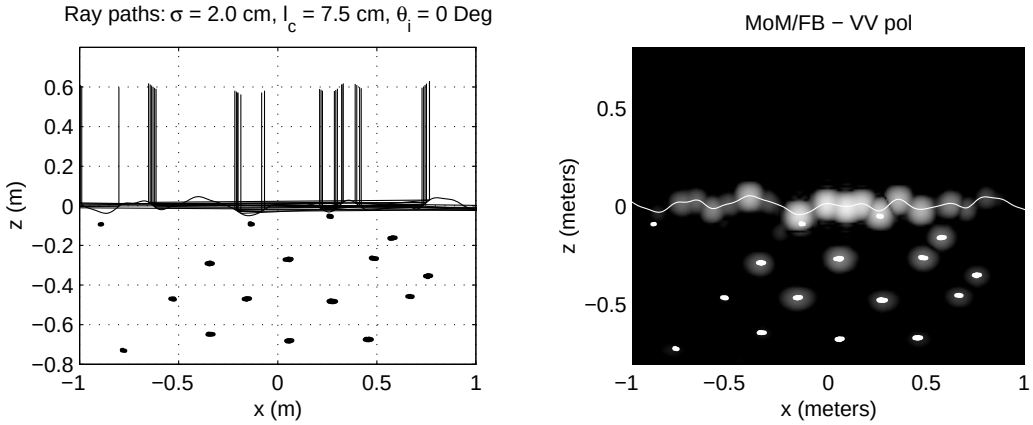


Figure 2: Ray tracing prediction for the location of image spots due to multiple scattering at normal incidence

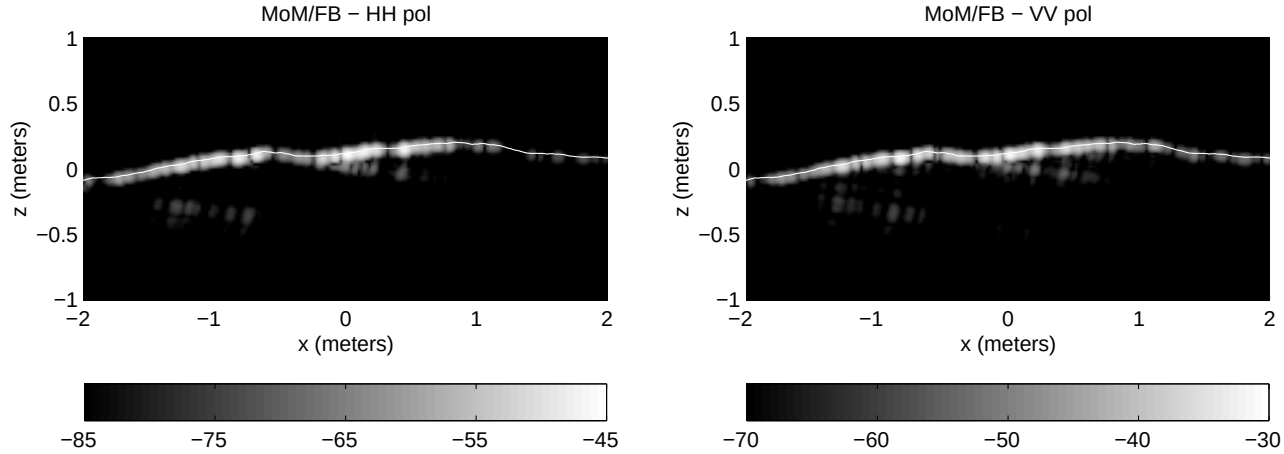


Figure 3: 2-D ISAR images for an ocean-like surface: $f = 10 \sim 14$ GHz, $\delta f = 25$ MHz, $\theta_i = 60^\circ \sim 80^\circ$ and $\delta\theta = 0.2^\circ$

On the complex conjugate roots of the Rayleigh equation: The leaky surface wave

Christoph T. Schröder[a] and Waymond R. Scott, Jr.[b]
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0250

(Received 24 May 2001; accepted for publication 11 September 2001)

In recent experiments and numerical studies, a leaky surface wave has been observed at the surface of an isotropic homogeneous elastic solid. This paper gives a detailed description of this leaky surface wave and explains its origin from the fundamental differential equations. Theoretically, the leaky surface wave arises from the complex conjugate roots of the Rayleigh equation. The complex conjugate roots give rise to a wave that propagates along the surface and is coupled to a plane shear wave in the medium. Due to the coupling, the surface wave leaks energy into the medium and is highly inhomogeneous. Its particle motion at the surface is prograde in nature, distinguishing it from the well-known Rayleigh surface wave which causes a retrograde particle motion. © *2001 Acoustical Society of America.* [DOI: 10.1121/1.1419085]

PACS numbers: 43.20.Bi, 43.35.Pt, 68.35.Ja [ANN]

I. INTRODUCTION

Elastic surface waves have been the subject of extensive research, since Lord Rayleigh first discovered the existence of elastic waves confined to the superficial region of an infinite homogeneous isotropic solid more than a century ago.[1] Elastic surface waves play an important role in various fields. In seismology, surface waves have been found to carry the bulk of the energy among the waves excited by an earthquake. In electroacoustics, surface waves are utilized to make filters and resonators. In geoscience, the propagation characteristics of surface waves are used to obtain information about the physical properties of the ground. In this paper, the theory of elastic surface waves is revisited.

A point source placed on the surface of a homogeneous isotropic medium excites five different kinds of waves: a pressure wave and a shear wave propagating in the medium, a Rayleigh surface wave that is confined to the surface of the medium, a lateral wave that is induced by the pressure wave at the surface, and a leaky surface wave that travels along the surface with a wave speed smaller than the pressure wave but larger than the shear wave. The first four of these wave types are well-known and have been treated extensively in the literature. However, the existence and theoretical foundation of the leaky surface wave has not been discussed as much.

The leaky surface wave arises from the complex conjugate roots of the Rayleigh equation. The leaky surface wave is an inhomogeneous wave that propagates along the surface with a phase velocity larger than the shear wave but smaller than the pressure wave. It couples into a plane shear wave that propagates in the medium. Due to the coupling, the surface wave loses energy and, thus, decays in its propagation direction.

The leaky surface wave has been observed by various authors in experimental and numerical studies. For example, Roth *et al.* noticed a rapidly decaying seismic surface wave in an environment with a very high Poisson ratio that had a phase velocity larger than the Rayleigh wave, but smaller than the pressure wave.[2] Smith *et al.* identified prograde and retrograde surface wave modes in a geologic study conducted on the shore of the gulf of Mexico.[3] Glass and Maradudin found a leaky surface wave to exist in the flat-surface limit of a corrugated crystal surface.[4] And Phinney provides a theoretical study of the leaky surface wave, which he calls a *Pseudo-P* mode.[5] Although not all of these authors explicitly describe a leaky surface wave, their observations are consistent with the results presented in this paper.

In this paper, the theoretical derivation of the leaky surface wave is described in some detail. In Sec. II, the governing equations are briefly outlined, leading to the Rayleigh equation. In Sec. III, the various roots of the Rayleigh equation are discussed. It will be shown that for materials with a high Poisson ratio a leaky surface wave exists, due to the complex conjugate roots of the Rayleigh equation. In Sec. IV, the waves excited by a line source on the surface are derived analytically. The method of steepest descent is applied to obtain closed-form expressions for the various waves in the far field.

II. THE RAYLEIGH EQUATION

The elastic wave fields at the surface of a semi-infinite, isotropic, lossless, homogeneous half space are to be determined. The half-space is bounded at $z=0$ by a free-surface boundary. The fields are assumed to be invariant in the y-direction and nonzero only in the x-z plane (*plane-strain case*, $u_y=\partial/\partial y=0$). Thus, the originally three-dimensional problem reduces to a two-dimensional one. The elastic wave fields in a medium may be expressed in terms of their potential functions:[6]

$$\mathbf{u}=\nabla\Phi+\nabla\times\mathbf{H}, \tag{1}$$

where $\mathbf{u}$ is the displacement vector, Φ is a scalar potential describing the longitudinal pressure wave, and $\mathbf{H}$ is a vector

[a]Electronic mail: christoph.schroeder@ece.gatech.edu
[b]Electronic mail: waymond.scott@ece.gatech.edu

potential describing the transverse shear wave. The potentials satisfy the wave equations

$$\nabla^2\Phi = \frac{1}{c_P^2}\frac{\partial^2\Phi}{\partial t^2}, \tag{2}$$

$$\nabla^2\mathbf{H} = \frac{1}{c_S^2}\frac{\partial^2\mathbf{H}}{\partial t^2}, \tag{3}$$

for the pressure and shear wave, respectively, with their corresponding wave speeds, c_P and c_S. In the *plane-strain* case, the only nonzero vector potential component is H_y and the only nonzero displacement components are u_x and u_z. The only independent stress components are τ_{xx}, τ_{zz}, and τ_{xz}.

Assuming harmonic time-dependence, the plane wave solutions for Φ and H_y satisfying Eqs. (2) and (3) are given by

$$\Phi = Ae^{j\xi x + j\alpha z - j\omega t}, \tag{4}$$

$$H_y = Be^{j\xi x + j\beta z - j\omega t} \tag{5}$$

and the wave numbers are defined by

$$\alpha^2 = \omega^2/c_P^2 - \xi^2, \tag{6}$$

$$\beta^2 = \omega^2/c_S^2 - \xi^2, \tag{7}$$

$$\xi^2 = \omega^2/c^2. \tag{8}$$

Using the well-known differential equations, the displacement and stress components are obtained:

$$u_x = (j\xi\, Ae^{j\alpha z} - j\beta\, Be^{j\beta z})e^{j\xi x}, \tag{9}$$

$$u_z = (j\alpha\, Ae^{j\alpha z} + j\xi\, Be^{j\beta z})e^{j\xi x}, \tag{10}$$

$$\tau_{xx} = \mu((2\alpha^2 - \beta^2 - \xi^2)Ae^{j\alpha z} + 2\beta\xi\, Be^{j\beta z})e^{j\xi x}, \tag{11}$$

$$\tau_{zz} = \mu((\xi^2 - \beta^2 Ae^{j\alpha z} - 2\beta\xi\, Be^{j\beta z})e^{j\xi x}, \tag{12}$$

$$\tau_{xz} = \mu(-2\alpha\xi\, Ae^{j\alpha z} + (\beta^2 - \xi^2)Be^{j\beta z})e^{j\xi x}. \tag{13}$$

At the surface, the normal stress vanishes, and thus $\tau_{zz}|_{z=0} = \tau_{xz}|_{z=0} = 0$. Using this condition, the ratio of the coefficients is determined from Eqs. (12) and (13) to be

$$\frac{A}{B} = \frac{2\xi\beta}{\xi^2 - \beta^2} = -\frac{\xi^2 - \beta^2}{2\xi\alpha}. \tag{14}$$

Re-inserting the amplitude ratio into Eqs. (12) and (13), the characteristic equation is obtained:

$$(\xi^2 - \beta^2)^2 + 4\xi^2\alpha\beta = 0. \tag{15}$$

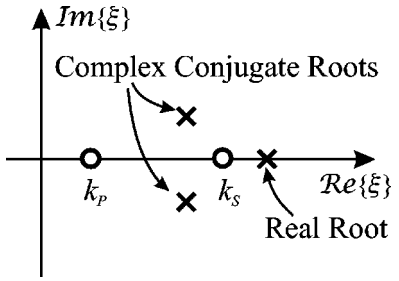


FIG. 1. Schematical arrangement of the roots in the complex ξ-plane.

Equation (15) is commonly called the *Rayleigh Equation*, because it gives rise to the well-known Rayleigh surface wave.

Using Eqs. (6)–(8), the characteristic equation can be rewritten in terms of the wave speeds:

$$\left(2 - \frac{c^2}{c_S^2}\right)^2 + 4\sqrt{\frac{c^2}{c_P^2} - 1}\sqrt{\frac{c^2}{c_S^2} - 1} = 0. \tag{16}$$

By rationalizing, this equation may be expressed as

$$\left(\frac{c}{c_S}\right)^2\left(\left(\frac{c}{c_S}\right)^6 - 8\left(\frac{c}{c_S}\right)^4 + (24 - 16(c_S/c_P)^2)\left(\frac{c}{c_S}\right)^2 - 16(1 - (c_S/c_P)^2)\right) = 0. \tag{17}$$

Equation (17) always has three solutions for c^2 (when neglecting the trivial solution). Dependent on the Poisson Ratio ν of a material, different kinds of roots arise. For $\nu < 0.263$, Eq. (17) has three real roots. For $\nu > 0.263$, Eq. (17) has one real root and two complex conjugate roots. In each case, the real root that is smallest in magnitude gives rise to the Rayleigh surface wave, which propagates along the surface and decays into the medium. The other roots have often been classified as erroneous or nonphysical roots of the Rayleigh equation.[6–8] However, it will be shown here that the complex conjugate roots of the Rayleigh equation in fact give rise to a *leaky surface wave*.

III. THE ROOTS OF THE RAYLEIGH EQUATION

Let the roots of Eq. (17) be denoted by c. In general, c will be complex:

$$c = c_r + jc_i, \tag{18}$$

TABLE I. Solutions to the Rayleigh equation.

(1)	$\xi_r > 0,\ \xi_i = 0$	$\alpha_r = 0, \alpha_i > 0$ $\Phi \sim e^{-\lvert\alpha_i\rvert z} e^{j\lvert\xi_r\rvert x}$	$\beta_r = 0, \beta_i > 0$ $H_z \sim e^{-\lvert\beta_i\rvert z} e^{j\lvert\xi_r\rvert x}$
(2)	$\xi_r > 0,\ \xi_i < 0$	$\alpha_r > 0, \alpha_i > 0$ $\Phi \sim e^{(j\lvert\alpha_r\rvert - \lvert\alpha_i\rvert)z} e^{(j\lvert\xi_r\rvert + \lvert\xi_i\rvert)x}$	$\beta_r < 0, \beta_i < 0$ $H_z \sim e^{(-j\lvert\beta_r\rvert + \lvert\beta_i\rvert)z} e^{(j\lvert\xi_r\rvert + \lvert\xi_i\rvert)x}$
(3)	$\xi_r > 0,\ \xi_i > 0$	$\alpha_r < 0, \alpha_i > 0$ $\Phi \sim e^{(-j\lvert\alpha_r\rvert - \lvert\alpha_i\rvert)z} e^{(j\lvert\xi_r\rvert - \lvert\xi_i\rvert)x}$	$\beta_r > 0, \beta_i < 0$ $H_z \sim e^{(j\lvert\beta_r\rvert + \lvert\beta_i\rvert)z} e^{(j\lvert\xi_r\rvert - \lvert\xi_i\rvert)x}$
(4)	$\xi_r > 0,\ \xi_i < 0$	$\alpha_r < 0, \alpha_i < 0$ $\Phi \sim e^{(-j\lvert\alpha_r\rvert + \lvert\alpha_i\rvert)z} e^{(j\lvert\xi_r\rvert + \lvert\xi_i\rvert)x}$	$\beta_r > 0, \beta_i > 0$ $H_z \sim e^{(j\lvert\beta_r\rvert - \lvert\beta_i\rvert)z} e^{(j\lvert\xi_r\rvert + \lvert\xi_i\rvert)x}$
(5)	$\xi_r > 0,\ \xi_i > 0$	$\alpha_r > 0, \alpha_i < 0$ $\Phi \sim e^{(j\lvert\alpha_r\rvert + \lvert\alpha_i\rvert)z} e^{(j\lvert\xi_r\rvert - \lvert\xi_i\rvert)x}$	$\beta_r < 0, \beta_i > 0$ $H_z \sim e^{(-j\lvert\beta_r\rvert - \lvert\beta_i\rvert)z} e^{(j\lvert\xi_r\rvert - \lvert\xi_i\rvert)x}$

TABLE II. Solutions to the Rayleigh equation.

	$\xi=$	$\alpha=$	$\beta=$
(1)	$2.5998k_P$	$j2.3997k_P$	$j0.8711k_P$
(2)	$(1.2184-j0.2526)k_P$	$(0.4030+j0.7636)k_P$	$(-2.1448-j0.1435)k_P$
(3)	$(1.2184+j0.2526)k_P$	$(-0.4030+j0.7636)k_P$	$(2.1448-j0.1435)k_P$
(4)	$(1.2184-j0.2526)k_P$	$(-0.4030-j0.7636)k_P$	$(2.1448+j0.1435)k_P$
(5)	$(1.2184+j0.2526)k_P$	$(0.4030-j0.7636)k_P$	$(-2.1448+j0.1435)k_P$

where c_r is the real part and c_i is the imaginary part of c. If c is purely real ($c=c_r$), ξ will be real [see Eq. (8)]. If additionally c is smaller than both the pressure wave speed and the shear wave speed, i.e., $c=c_r<c_S<c_P$, α and β will be purely imaginary. For the solution to be physical, $\mathrm{Im}\{\alpha\}=\alpha_i>0$ and $\mathrm{Im}\{\beta\}=\beta_i>0$, thus the waves described by these wave numbers *propagate* in the *x*-direction and *decay* in the *z*-direction. This solution represents the well-known Rayleigh surface wave, first explored by Lord Rayleigh more than a century ago.[1] A solution of this form always exists, independent of the Poisson ratio of a material.

For $\nu<0.263$, two more purely real roots of the Rayleigh equation exist. It can be shown that for these roots the wave speed is always larger than the pressure wave speed, c_P. These two roots describe the angles of incidence at which *complete mode conversion* occurs. In the case of complete mode conversion an incident shear wave, for example, is completely reflected as a pressure wave, without inducing a reflected shear wave component. Complete mode conversion is physically possible only for materials with a Poisson ratio smaller than 0.263 and occurs at two distinct angles of incidence, defined by the two roots of the Rayleigh equation (see also, for example, Graff[6]).

If $\nu>0.263$, one real root and two complex conjugate roots of the Rayleigh equation arise. The real root again gives rise to the Rayleigh surface wave. For the complex conjugate roots, the wave speed is complex, $c=c_r+jc_i$, and consequently also the wave numbers are complex: $\xi=\xi_r+j\xi_i$, $\alpha=\alpha_r+j\alpha_i$ and $\beta=\beta_r+j\beta_i$. It can be shown that for the complex conjugate roots the real part of the wave speed is always smaller than the pressure wave speed, but larger than the shear wave speed, $c_S<\mathrm{Re}\{c\}<c_P$.

Although c may be a solution to Eq. (17), it does not necessarily follow that also Eq. (15) is fulfilled. This is due to the manipulation of Eq. (15) to arrive at Eq. (17). In fact, the complex conjugate roots of Eq. (17) do not represent solutions to Eq. (16). It can be shown, however, that they do represent solutions to Eq. (15), if the signs of the wave numbers α and β are picked correctly. It may be recalled that according to Eqs. (6) and (7) the wave numbers α and β are functions of the square root of c^2,

$$\alpha=\pm\frac{\omega}{c}\sqrt{\frac{c^2}{c_P^2}-1}=\pm(\alpha_r+j\alpha_i), \tag{19}$$

$$\beta=\pm\frac{\omega}{c}\sqrt{\frac{c^2}{c_S^2}-1}=\pm(\beta_r+j\beta_i). \tag{20}$$

The sign in front of the square roots must be chosen according to physical and causal constraints of the underlying problem. To obtain Eq. (16), the positive sign has been assumed for both α and β. However, it turns out that Eq. (15) is only satisfied if, for the complex conjugate roots, both the real part and the imaginary part of α and β have opposite signs, i.e., $\mathrm{sign}(\alpha_r)\neq\mathrm{sign}(\beta_r)$ and $\mathrm{sign}(\alpha_i)\neq\mathrm{sign}(\beta_i)$.

Figure 1 shows schematically the arrangement of the roots in the complex ξ-plane. The possible solutions of the Rayleigh equation are summarized in Table I, giving all possible combinations of ξ, α and β. Only waves propagating in the positive *x*-direction are considered. Five possible solutions arise.

The first solution describes the Rayleigh surface wave. For the second solution, both the pressure wave potential and the shear wave potential, Φ and H_y, propagate and increase in the positive *x*-direction ($\xi_r>0$, $\xi_i<0$). However, Φ propagates and decays in the positive *z*-direction ($\alpha_r>0$, $\alpha_i>0$), whereas H_y propagates and decays in the negative *z*-direction ($\beta_r<0$, $\beta_i<0$). For the third solution, the potentials decay in the *x*-direction ($\xi_i>0$). Φ now propagates and increases in the negative *z*-direction ($\alpha_r<0$, $\alpha_i>0$), whereas H_y propagates and increases in the positive *z*-direction ($\beta_r>0$, $\beta_i<0$). For the fourth and fifth solution, the signs of α and β are reversed.

The behavior of the five possible solutions is best demonstrated by calculating the wave fields for a medium with a specific value of Poisson's ratio. Assuming Poisson's ratio to be $\nu=0.4$, the elastic wave fields are computed using Eqs. (9)–(13). The amplitude of the shear wave potential is chosen to be unity, and the amplitude of the pressure wave potential is computed with Eq. (14). For $\nu=0.4$, the pressure wave speed exceeds the shear wave speed by a factor of about 2.45: $c_P=2.4495c_S$. The roots of Eq. (17) in terms of the shear wave speed c_S are

$$c=\begin{cases}0.9422c_S\\(1.9276+j0.3996)c_S\\(1.9276-j0.3996)c_S.\end{cases} \tag{21}$$

The resulting wave numbers are shown in Table II. All wave numbers are expressed in terms of the longitudinal wave number, $k_P=\omega/c_P$.

In Fig. 2, the displacements according to the five solutions of the Rayleigh equation are plotted versus *x* and *z*. The distance on the axes is normalized to the wavelength of the Rayleigh surface wave, λ_R. The two columns show the horizontal and vertical displacements, u_x and u_z. Pseudo-color plots are used to display the wave fields, employing a logarithmic scale with a dynamic range of 50 dB. Superimposed with the horizontal displacement component is the real part

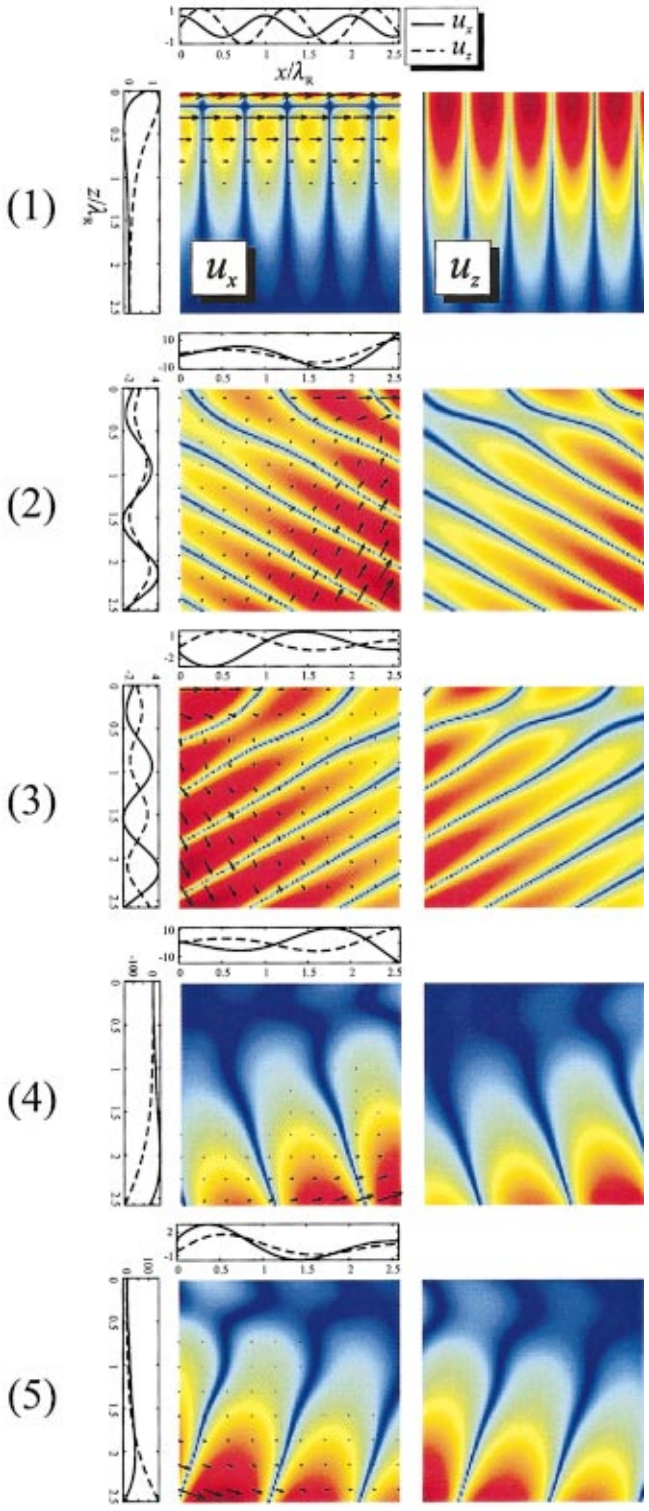


FIG. 2. Horizontal and vertical displacements according to the five solutions of the Rayleigh equation.

of the complex Poynting vector, thus indicating the direction of the energy flow. The upper edge of each plot corresponds to the free surface.

First, the Rayleigh surface wave is shown. For the Rayleigh surface wave, the energy flow is seen to be parallel to the surface. Both the horizontal and the vertical displacements decay away from the surface. The second solution describes a shear wave in which energy is carried toward the surface. Close to the surface, the energy flows parallel to the surface. The waves grow exponentially in the x- and z-direction. The third solution is the converse of the second one. This time, the energy flows from the surface into the medium, and the waves decay in the x-direction. Again, close to the surface the energy flow is parallel to the surface. The fourth and fifth solution are similar to the second and third. However, now a pressure wave propagates in the medium, and the energy flows at a different angle with respect to the surface. Also, the exponential growth is enhanced.

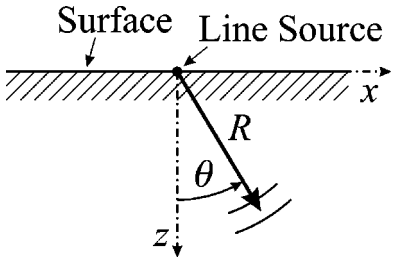


FIG. 3. Line source on the surface.

All of the five possible solutions described above can, in certain cases, represent physical solutions. For example, if a field distribution is created on the surface that matches the field distribution of Solution (5) on the surface, waves similar to the ones described by Solution (5) would be induced in the medium. If a field distribution is generated within the medium that is equal to the field distribution of Solution (2), waves propagating toward the surface would be excited that perfectly couple into a surface wave. Of course, the solutions as described here would require an infinite medium and wave fields of infinite extent that are nonzero at infinity, which violates physical as well as causal constraints. However, over a finite range all of these solutions can be excited with the appropriate field distributions.

IV. WAVES DUE TO A LINE SOURCE ON THE SURFACE

In the previous section, the solutions to the wave equation at a free-surface boundary have been described in a general form. In this section, the wave fields due to a specific excitation, a line source on the surface, are determined. The

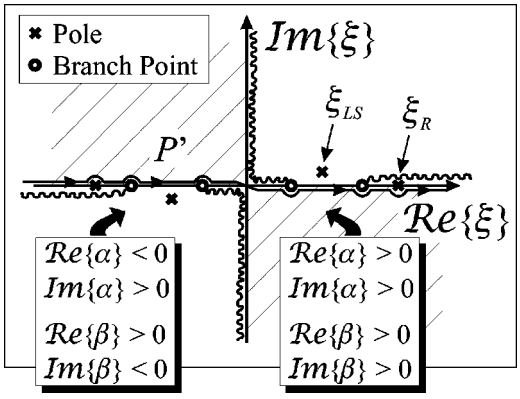


FIG. 4. Location of the poles and branch cuts in the complex ξ-plane for the line-source problem.

analysis presented here is largely based on the general mathematical description of Felsen and Marcuvitz.[9]

Figure 3 shows the underlying geometry. A line source is placed on the surface at $x=0$ and extends into the y-direction. The line source excites the normal stress component τ_{zz}. If harmonic time dependence is assumed, the displacement fields due to a line source can be written in form of an integral equation:[6]

$$u_x(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\xi}{F_0(\xi)} \times[-2\alpha\beta e^{j\beta z}+(\beta^2-\xi^2)e^{j\alpha z}]e^{j\xi x}\,d\xi, \quad (22)$$

$$u_z(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\alpha}{F_0(\xi)} \times[2\xi^2 e^{j\beta z}+(\beta^2-\xi^2)e^{j\alpha z}]e^{j\xi x}\,d\xi, \quad (23)$$

where

$$F_0(\xi) = (\xi^2-\beta^2)^2+4\xi^2\alpha\beta \quad (24)$$

is the Rayleigh equation. The amplitude of the excitation is assumed to be unity.

A. General considerations

The integrals in Eqs. (22) and (23) each represent an inverse Fourier transform from the wave number domain into the spatial domain. The integrands contain poles and branch points. The poles are due to the roots of the Rayleigh equation in the denominator. The branch points arise because of the square root dependence of α and β on ξ. They are located at the roots of α and β, at $k_P=\pm(\omega/c_P)^2$ and $k_S=\pm(\omega/c_S)^2$ [see Eqs. (6) and (7)].

To compute the integrals, contour integration in the complex ξ-plane must be applied. The integration must be performed along the real ξ-axis. Figure 4 shows the location of the poles and branch points in the complex ξ-plane. To determine the waves propagating in the positive x-direction, Eqs. (22) and (23) must be integrated along the path P'. The contour is closed at infinity. Only the poles and branch cuts for $\mathrm{Re}\{\xi\}>0$ are included in the integration contour (indicated by the indentations of P'), whereas the poles for $\mathrm{Re}\{\xi\}<0$ are excluded and, therefore, do not contribute to the integral.

Due to the branch points, the integrands are not single-valued. To make the integrands unique, a Riemann surface for the ξ-plane is necessary, with branch cuts providing the transition from one Riemann sheet to the other.[9] The location of the branch cuts in general is arbitrary, but defines the disposition of those regions in the complex ξ-plane in which for example $\mathrm{Re}\{\alpha\}>0$ or $\mathrm{Re}\{\alpha\}<0$. Figure 4 shows the top Riemann sheet for Eqs. (22) and (23). The signs of the wave numbers on the top Riemann sheet must be chosen according to physical and causal reasons. The integration along the real axis determines the shear and pressure waves excited by the line source. For the shear and the pressure waves to be causal, they must propagate away from the source and vanish at infinity. For this to be true, the wave numbers along the real axis must be chosen such that $\mathrm{Re}\{\alpha\}>0$, $\mathrm{Im}\{\alpha\}>0$, $\mathrm{Re}\{\beta\}>0$ and $\mathrm{Im}\{\beta\}>0$. It can be easily shown that in this case the wave numbers in the entire second and fourth quadrant must behave in the same way. In the first and third quadrant, the branch cuts must then be chosen such that $\mathrm{Re}\{\alpha\}<0$, $\mathrm{Im}\{\alpha\}>0$, $\mathrm{Re}\{\beta\}>0$ and $\mathrm{Im}\{\beta\}<0$. This is true because α and β must be continuous across the real axis. Thus, in the first and third quadrant, the pressure wave potential propagates and increases in the negative z-direction, whereas the shear wave potential propagates and increases in the positive z-direction. It is evident that in the first quadrant of the top Riemann sheet the wave numbers behave as described for Solution (3) of the Rayleigh equation as indicated in Table I. The poles on the top Riemann sheet correspond to physically existing waves and, therefore, the pole associated with Solution (3) of the Rayleigh equation represents a physical solution to the line-source problem. The pole in the third quadrant is the equivalent to the pole in the first quadrant, but describes a wave traveling in the negative x-direction. The two poles on the real ξ-axis are present on all sheets and, consequently, also represent physical waves. All other poles of the Rayleigh equation lie on different sheets and, thus, are nonphysical for the line-source case.

It can be seen in Fig. 4 that four poles and four branch cuts exist on the top Riemann sheet. The poles at $\pm\xi_R$ on the real ξ-axis give rise to the well-known Rayleigh surface wave. The complex poles at $\pm\xi_{\mathrm{LS}}$ in the first and third quadrant describe leaky surface waves propagating to the right and left, respectively. As described earlier, the leaky surface wave couples into a plane shear wave. Both the leaky surface wave and the shear wave that is fed from the surface wave are inhomogeneous, which is indicated by the pole being complex.

B. Steepest-descent approximation

To evaluate the integrals asymptotically in the far field, the method of steepest descent shall be applied. To simplify the procedure, the two terms of the integral are treated separately. Dividing the integrals each into a pressure wave term and a shear wave term, Eqs. (22) and (23) are rewritten as

$$u_x^S(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\xi}{F_0(\xi)}(-2\alpha\beta)e^{j\beta z}e^{j\xi x}\,d\xi, \quad (25)$$

$$u_x^P(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\xi}{F_0(\xi)}(\beta^2-\xi^2)e^{j\alpha z}e^{j\xi x}\,d\xi, \quad (26)$$

$$u_z^S(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\alpha}{F_0(\xi)}2\xi^2 e^{j\beta z}e^{j\xi x}\,d\xi, \quad (27)$$

$$u_z^P(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\alpha}{F_0(\xi)}(\beta^2-\xi^2)e^{j\alpha z}e^{j\xi x}\,d\xi. \quad (28)$$

The total displacements equal the superposition of the pressure wave component and the shear wave component:

$$u_x(x,z) = u_x^S(x,z)+u_x^P(x,z), \quad (29)$$

$$u_z(x,z) = u_z^S(x,z)+u_z^P(x,z). \quad (30)$$

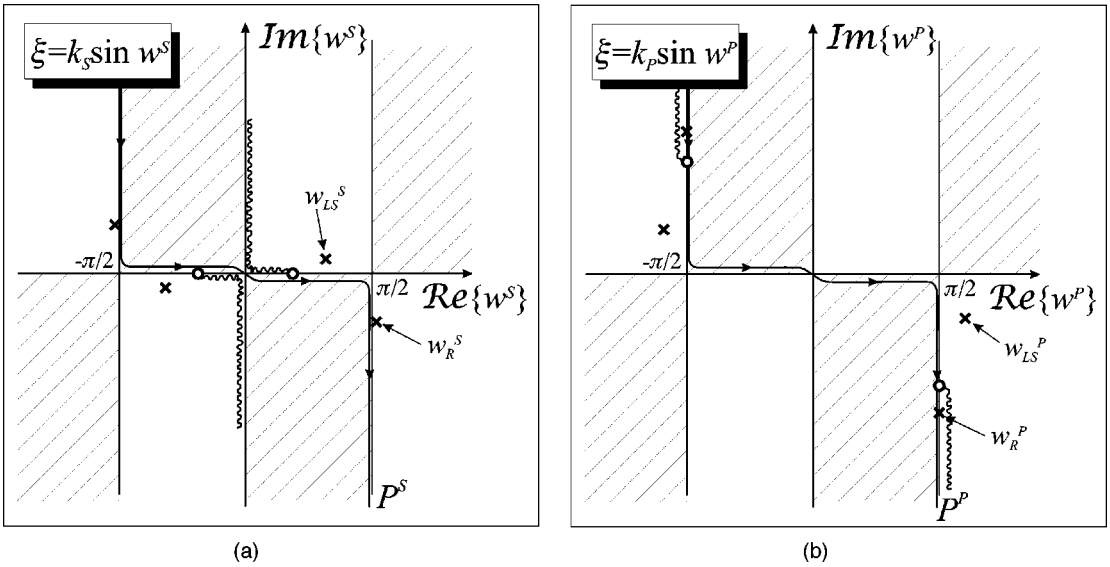


FIG. 5. Location of the poles and branch cuts (a) in the complex w^S-plane and (b) in the complex w^P-plane.

To facilitate the evaluation of the integrals, the complex ξ-plane is transformed into the complex w^S-plane for the two shear wave terms, and into the w^P-plane for the pressure wave terms:

$$\xi = k_S \sin w^S, \tag{31}$$

$$\xi = k_P \sin w^P. \tag{32}$$

These transformations are single-valued.[9] From the periodicity of $\sin w^S$ and $\sin w^P$ it is evident that multiple values for w^S and w^P correspond to a single value of ξ. Thus, the transformations can be used to map the entire ξ-plane with its multiple Riemann sheets into adjacent strips of width 2π in the w^S- or w^P-plane. The arrangement of the poles and branch cuts of the top Riemann sheet in the complex w^S-plane and w^P-plane are shown in Fig. 5. Here, the top Riemann sheet is mapped into a strip reaching from $-\pi$ to π in the complex w^S-plane for the shear wave terms, and similarly for the pressure wave terms in the complex w^P-plane. The positions of the transformed Rayleigh wave pole and the leaky surface wave pole in the complex w^S- and w^P-plane are indicated by w_R^S, w_{LS}^S, and w_R^P, w_{LS}^P, respectively. The transformed integration paths are denoted by P^S and P^P.

The separate transformations for the shear wave terms and the pressure wave terms become necessary, because, when the method of steepest descent is applied, the different terms will give rise to different steepest-descent paths. By applying the different transformations, the steepest-descent paths will have a rather simple shape for both the shear wave

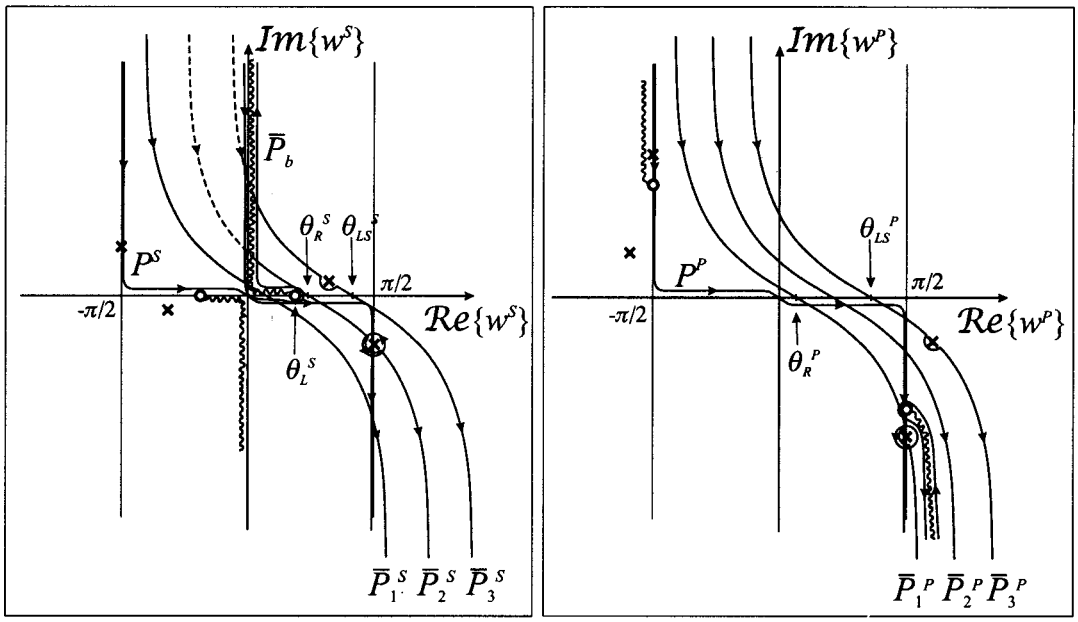


FIG. 6. Steepest-descent paths for (a) the shear wave terms in the complex w^S-plane and (b) the pressure wave terms in the complex w^P-plane.

terms and the pressure wave terms, thus, making the steepest-descent approximation considerably easier.

Applying the transformations, two of the branch cuts are eliminated in each of the integrals in Eqs. (25)–(28). For the shear wave terms the branch cuts at $\pm k_S$ vanish, whereas for the pressure wave terms the branch cuts at $\pm k_P$ are removed. For the shear wave terms, β then reduces to

$$\beta = k_S \cos w^S, \tag{33}$$

and for the pressure wave terms α becomes

$$\alpha = k_P \cos w^P. \tag{34}$$

Introducing polar coordinates,

$$x = R \sin \theta, \tag{35}$$

$$z = R \cos \theta, \tag{36}$$

Eqs. (25)–(28) are rewritten as integrals in the w^S- and w^P-plane:

$$u_x^S(x,z) = -\frac{j}{\mu\pi}\int_{P^S}\frac{\xi}{F_0(\xi)}(-2\alpha\beta) \times e^{jk_S R\cos(w^S-\theta)}\beta\, dw^S, \tag{37}$$

$$u_x^P(x,z) = -\frac{j}{\mu\pi}\int_{P^P}\frac{\xi}{F_0(\xi)}(\beta^2-\xi^2) \times e^{jk_P R\cos(w^P-\theta)}\alpha\, dw^P, \tag{38}$$

$$u_z^S(x,z) = -\frac{j}{\mu\pi}\int_{P^S}\frac{\alpha}{F_0(\xi)}2\xi^2 e^{jk_S R\cos(w^S-\theta)}\beta\, dw^S, \tag{39}$$

$$u_z^P(x,z) = -\frac{j}{\mu\pi}\int_{P^P}\frac{\alpha}{F_0(\xi)}(\beta^2-\xi^2) \times e^{jk_P R\cos(w^P-\theta)}\alpha\, dw^P, \tag{40}$$

where θ describes the polar angle measured from the surface normal toward the propagation direction (see Fig. 3). The wave numbers in terms of w^S are

$$\xi(w^S) = k_S \sin w^S, \tag{41}$$

$$\beta(w^S) = k_S \cos w^S, \tag{42}$$

$$\alpha(w^S) = \pm\sqrt{k_P^2 - \xi(w^S)^2}, \tag{43}$$

and in terms of w^P

$$\xi(w^P) = k_P \sin w^P, \tag{44}$$

$$\beta(w^P) = \pm\sqrt{k_S^2 - \xi(w^P)^2}, \tag{45}$$

$$\alpha(w^P) = k_P \cos w^P. \tag{46}$$

The signs of $\alpha(w^S)$ and $\beta(w^P)$ must be chosen as described earlier for the complex ξ-plane. Thus, in the shaded and nonshaded regions of Fig. 5, α behaves just as in the shaded and nonshaded regions of Fig. 4.

With the integrals transformed as described above, it is relatively straightforward to apply the method of steepest descent. For the method of steepest descent, the integration paths P^S and P^P are deformed into new paths, the *steepest-descent paths* $\overline{P^S}$ and $\overline{P^P}$, respectively. The new path is chosen such that the dominant contribution to the integral arises from only a small section of the path. To achieve this, the path is deformed such that it passes through the saddle point of the integrand. Away from the saddle point it follows the direction in which the integrand decays most rapidly. Along this path, the integrand will then be negligible everywhere but around the saddle point, and the integral can be approximated by the contribution from the integrand in the vicinity of the saddle point.

The path of steepest descent is a path of constant phase.[9] For integrals in the form of the ones in Eqs. (37)–(40), the steepest-descent path is given by

$$\mathrm{Re}\{w^{S,P}\} - \theta = \cos^{-1}(\mathrm{sech}(\mathrm{Im}\{w^{S,P}\})). \tag{47}$$

The procedure is the same for the shear wave terms and the pressure wave terms. In Fig. 6, three steepest-descent paths are shown each for the shear wave terms in the w^S-plane, $\overline{P_1^S}$, $\overline{P_2^S}$, $\overline{P_3^S}$, and for the pressure wave terms in the w^P-plane, $\overline{P_1^P}$, $\overline{P_2^P}$, $\overline{P_3^P}$. Each steepest descent path corresponds to a different propagation (polar) angle. The saddle point in each case is located at the intersection of the steepest-descent path with the real w^S- or w^P-axis, respectively. Physically, the contributions from the saddle points describe the pressure and the shear waves in the far field .

When the original integration path is deformed into the steepest-descent path, care has to be taken whether poles or branch cuts are crossed during the deformation. According to Cauchy's theorem, if a singularity is crossed during the deformation from one integration path into another, the contribution from the contour integral around the singularity must be included into the total integral. For example, when P^S in Fig. 6 is deformed into the steepest-descent path $\overline{P_1^S}$, no singularities are crossed during the deformation. However, for $\overline{P_2^S}$, the integrals around the branch cut, $\overline{P_b}$, and around the pole at w_R^S must be included. For $\overline{P_3^S}$, the branch cut integral as well as the integrals around the poles at w_R^S and w_{LS}^S contribute to the total integral.

The contour integrals around the singularities in both the w^S- and the w^P-plane give rise to different types of waves. The integral around the branch cut in the w^S-plane describes a lateral wave. The lateral wave is a plane shear wave induced by the pressure wave propagating along the surface. It appears only if the polar angle exceeds $\theta_L^S = \sin^{-1}(k_P/k_S)$, because, mathematically, the branch cut integral contributes to the total integral only for $\theta > \theta_L^S$. It can be shown that the integral around the branch cut in the w^P-plane is approximately zero and, thus, it does not contribute to the total integral. The integrals around the poles give rise to the Rayleigh surface wave and the leaky surface wave. They exist only for $\theta > \theta_R^S$ and $\theta > \theta_{\mathrm{LS}}^S$ in the w^S-plane, and for $\theta > \theta_R^P$ and $\theta > \theta_{\mathrm{LS}}^P$ in the w^P-plane. The total Rayleigh surface wave and the total leaky surface wave are comprised of the superposition of the contributions from the integrals around the singularities both in the w^S-plane and w^P-plane. The angles $\theta_R^{S,P}$ and $\theta_{\mathrm{LS}}^{S,P}$ are easily obtained by inserting $w_R^{S,P} = \sin^{-1}(\xi_R/k_{S,P})$ and $w_{\mathrm{LS}}^{S,P} = \sin^{-1}(\xi_{\mathrm{LS}}/k_{S,P})$ into Eq. (47).

The integrals are now approximately determined using the method of steepest-descent. Five separate wave types arise: the bulk shear wave, the bulk pressure wave, the Rayleigh surface wave, the leaky surface wave and the lateral wave. A detailed description of the steepest-descent method is given, for example, by Felsen and Marcuvitz.[9]

Using the method of steepest-descent, the shear wave in the far field comes out to be

$$u_x(R,\theta)|_{\text{Shear}} = \frac{j}{\mu\pi}\sqrt{\frac{2\pi}{k_S R}}\frac{e^{j(k_S R-\pi/4)}}{F_0(k_S \sin\theta)} \cdot 2\sqrt{k_P^2-k_S^2\sin^2\theta}\,k_S^3 \sin\theta\cos^2\theta, \quad (48)$$

$$u_z(R,\theta)|_{\text{Shear}} = \frac{-j}{\mu\pi}\sqrt{\frac{2\pi}{k_S R}}\frac{e^{j(k_S R-\pi/4)}}{F_0(k_S \sin\theta)} \cdot 2\sqrt{k_P^2-k_S^2\sin^2\theta}\,k_S^3 \sin^2\theta\cos\theta. \quad (49)$$

For the pressure wave,

$$u_x(R,\theta)|_{\text{Pressure}} = \frac{-j}{\mu\pi}\sqrt{\frac{2\pi}{k_P R}}\frac{e^{j(k_P R-\pi/4)}}{F_0(k_P \sin\theta)} \cdot [k_S^2-2k_P^2\sin^2\theta]k_P^2 \sin\theta\cos\theta, \quad (50)$$

$$u_z(R,\theta)|_{\text{Pressure}} = \frac{-j}{\mu\pi}\sqrt{\frac{2\pi}{k_P R}}\frac{e^{j(k_P R-\pi/4)}}{F_0(k_P \sin\theta)} \cdot [k_S^2-2k_P^2\sin^2\theta]k_P^2 \cos^2\theta. \quad (51)$$

The Rayleigh surface wave and the leaky surface wave are derived from the contour integral around the respective poles of the Rayleigh equation. Using contour integration, the Rayleigh wave is described by

$$u_x(R,\theta)|_{\text{Rayleigh}} = \frac{2\xi_R}{\mu F_0'|_{\xi_R}}[U(\theta-\theta_R^S) \cdot(-2\alpha_R\beta_R)e^{j\beta_R R\cos\theta} + U(\theta-\theta_R^P) \cdot(\beta_R^2-\xi_R^2)e^{j\alpha_R R\cos\theta}]\cdot e^{j\xi_R R\sin\theta}, \quad (52)$$

$$u_z(R,\theta)|_{\text{Rayleigh}} = \frac{2\alpha_R}{\mu F_0'|_{\xi_R}}[U(\theta-\theta_R^S)\cdot 2\xi_R^2 e^{j\beta_R R\cos\theta} + U(\theta-\theta_R^P)\cdot(\beta_R^2-\xi_R^2)e^{j\alpha_R R\cos\theta}] \cdot e^{j\xi_R R\sin\theta}, \quad (53)$$

where

$$F_0'|_{\xi_R} = 4\xi_R(\xi_R^2-\beta_R^2)+8\xi_R\beta_R\alpha_R-4\xi_R^3\left(\frac{\beta_R}{\alpha_R}+\frac{\alpha_R}{\beta_R}\right). \quad (54)$$

$F_0'|_{\xi_R}$ is the derivative of the Rayleigh equation with respect to ξ at $\xi=\xi_R$, and $U(\theta-\theta_R^{S,P})$ is the Heaviside unit step function; α_R, β_R, and ξ_R describe the wave numbers of the Rayleigh wave (see Table II). The result for the leaky surface wave is determined in exactly the same way and is obtained by simply replacing $\theta_R^{S,P}$, ξ_R, α_R and β_R by $\theta_{\text{LS}}^{S,P}$, ξ_{LS}, α_{LS} and β_{LS}, i.e., inserting the wave numbers for the leaky-wave pole instead of the Rayleigh wave pole.

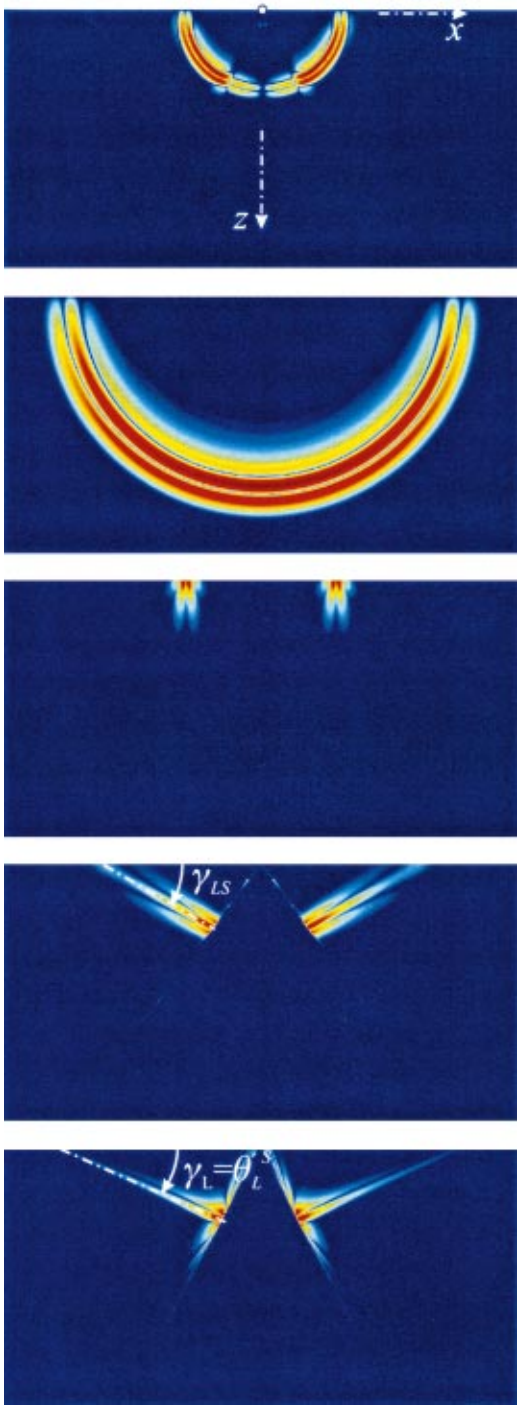


FIG. 7. Waves due to a point source on the surface. From top to bottom: shear wave, pressure wave, Rayleigh surface wave, leaky surface wave, lateral wave.

The lateral wave is defined by the integral around the branch cut, $\overline{P_b}$ in Fig. 6. Following Felsen and Marcuvitz,[9] the integral is asymptotically approximated to become

$$u_x(R,\theta)\big|_{\text{Lateral}} = U(\theta-\theta_L^S)\cdot\frac{-2^{3/2}/(\mu\sqrt{\pi})}{[k_S R|\sin(\theta-\theta_L^S)|]^{3/2}} \cdot\frac{(\sin\theta_L^S)^{3/2}(\cos\theta_L^S)^{5/2}}{(2\sin^2\theta_L^S-1)^2} \cdot e^{jk_S R\cos(\theta-\theta_L^S)+j3/4\pi}, \tag{55}$$

$$u_z(R,\theta)\big|_{\text{Lateral}} = U(\theta-\theta_L^S)\cdot\frac{2^{3/2}/(\mu\sqrt{\pi})}{[k_S R|\sin(\theta-\theta_L^S)|]^{3/2}} \cdot\frac{(\sin\theta_L^S)^{5/2}(\cos\theta_L^S)^{3/2}}{(2\sin^2\theta_L^S-1)^2} \cdot e^{jk_S R\cos(\theta-\theta_L^S)+j3/4\pi}. \tag{56}$$

C. Example

Equations (48)–(56) give the asymptotic far-field approximations for the wave fields excited by a harmonic line source on the surface. To determine the wave fields for a specific excitation in the time domain, the results must be transformed from the frequency domain into the time domain and convolved with the excitation function. The inverse-Fourier transform is given by

$$\hat{\mathbf{u}}(R,\theta,t) = \int_{\omega=-\infty}^{+\infty} G(\omega)\cdot\mathbf{u}(R,\theta,\omega)\cdot e^{-j\omega t}\,d\omega. \tag{57}$$

Here, $G(\omega)$ represents the Fourier transform of the excitation function. To obtain the particle velocity rather than the displacement, the displacement is differentiated with respect to time:

$$\hat{\mathbf{v}}(R,\theta,t) = \int_{\omega=-\infty}^{+\infty} G(\omega)\cdot\mathbf{u}(R,\theta,\omega)\cdot(-j\omega)e^{-j\omega t}\,d\omega. \tag{58}$$

In the following, a differentiated Gaussian pulse is used as excitation, with its Fourier transform

$$G(\omega) = -j\sqrt{2\pi}t_0^2\omega\cdot e^{0.5-0.5(\omega t_0)^2}, \tag{59}$$

where t_0 describes the width of the pulse. The particle velocity is determined here, because the analytical results are to be compared to numerical results, and the numerical finite-difference code that has been developed for this purpose computes the particle velocity rather than the particle displacement.

The wave fields excited by a differentiated Gaussian pulse are computed for a material with a Poisson ratio of $\nu = 0.4$. The wave fields are calculated according to Eqs. (48)–(56) and then transformed into the time domain using Eq. (58). In Fig. 7, the separate wave fields throughout the half

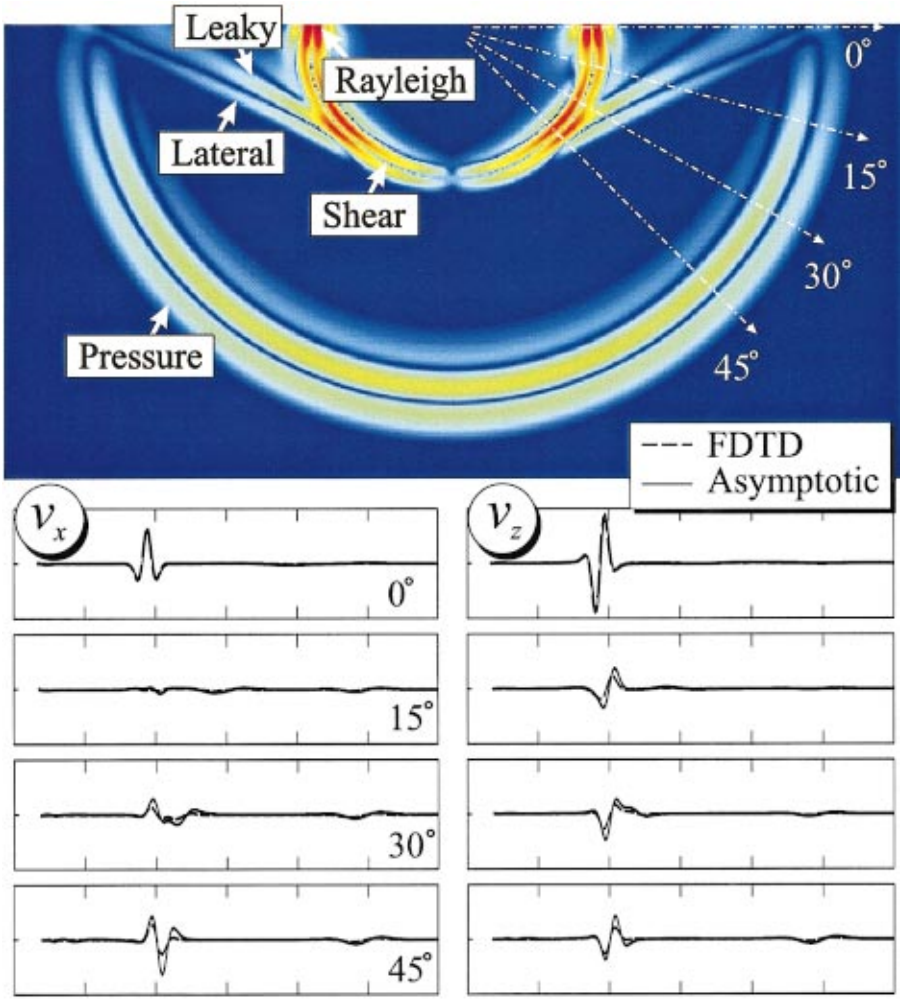


FIG. 8. Finite-difference results; comparison to asymptotic solution.

space are plotted for one instant in time, ten milliseconds after their excitation. The plots show, from top to bottom, the shear wave, the pressure wave, the Rayleigh surface wave, the leaky surface wave, and the lateral wave. A logarithmic color scale is used, ranging from dark red (0 dB) over yellow and green to blue (−40 dB). The top of each plot coincides with the surface of the medium. The source is located on the surface, at the center of each plot.

The shear wave and the pressure wave exhibit cylindrical wave fronts. They both vanish at the surface. The Rayleigh surface wave is confined to the surface and decays into the *z*-direction. The leaky surface wave propagates along the surface with a speed greater than the one of the shear wave, but smaller than the speed of the pressure wave. It feeds an inhomogeneous plane shear wave. The angle that the shear wave makes with the surface is approximately defined by

$$\gamma_{\mathrm{LS}} = \sin^{-1}(\mathrm{Re}\{\xi_{LS}\}/k_S). \tag{60}$$

Due to the coupling into the shear wave, the leaky wave ''leaks'' energy into the medium and decays in its propagation direction. The lateral wave propagates at an angle of approximately

$$\gamma_L = \theta_L^S = \sin^{-1}(k_P/k_S). \tag{61}$$

The artifacts that are especially visible for the lateral wave are due to the Fourier transform algorithm that is being used.

Figure 8 shows the wave fields due to a line source on the surface at one instant in time as determined numerically with the finite-difference time-domain (FDTD) algorithm. The wave fields are plotted at the top on a logarithmic color scale on a cross section through the half space and at the bottom on a linear scale along four radial lines, corresponding to four distinct propagation angles, as a function of the distance from the source. Again, a material with a Poisson ratio of 0.4 is assumed. With careful inspection, the five different wave types are distinguishable. The differences between the FDTD result and the asymptotic approximation are mainly due to the fact that the asymptotic approximation describes the waves in the far field, whereas the FDTD computations show the waves in the near field. Results essentially identical to the FDTD results have been obtained when integrating Eqs. (22) and (23) numerically rather than approximating the integrals asymptotically.[10]

To obtain a better picture of the behavior of the various waves at the surface, the particle motion due to the different surface waves is analyzed. For this, the wave fields of the Rayleigh surface wave, the leaky surface wave and the lateral wave are computed using Eqs. (52)–(56) for harmonic time-dependence, and hodograms of the particle motion at the surface are generated. In these hodograms, the vertical displacement along the surface is plotted versus the horizontal displacement. The hodograms are shown in Fig. 9. As it is well-known, the particle motion due to a Rayleigh surface wave is retrograde (counterclockwise) in nature [Fig. 9(a)]. This is caused by a phase shift between the horizontal and the vertical displacement component: the horizontal displacement is lacking 90 degrees in phase behind. The hodogram also indicates that the Rayleigh wave does not decay as it propagates along the surface. The particle motion

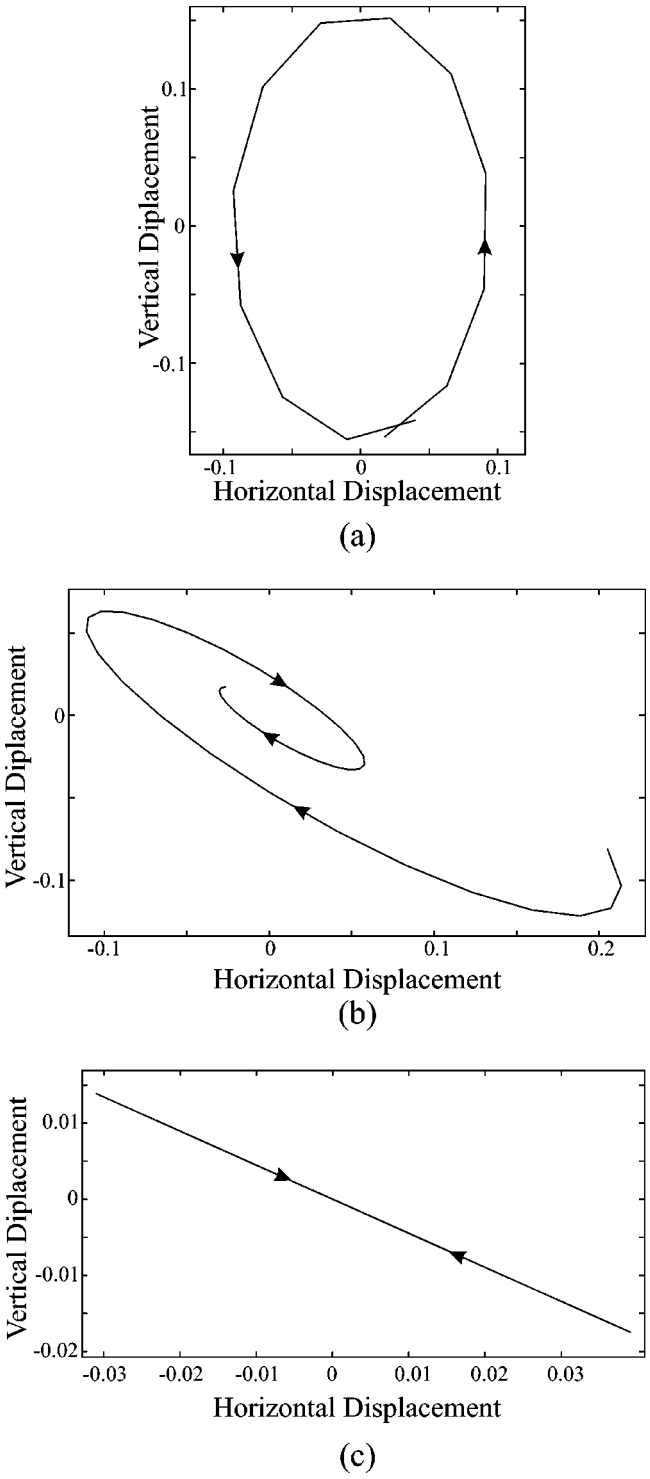


FIG. 9. Hodograms of the particle motion at the surface. Plots for the vertical displacement vs the horizontal displacement for (a) the Rayleigh surface wave, (b) the leaky surface wave, and (c) the lateral wave.

due to the leaky surface wave is prograde (clockwise), caused by the horizontal displacement being ahead in phase of the vertical displacement [Fig. 9(b)]. Clearly, the leaky surface wave decays as it travels along the surface. For the lateral wave, the displacement components are in phase, and the hodogram shows a diagonal line [Fig. 9(c)]. The lateral wave also decays as it propagates along the surface. The prograde and retrograde particle motions of surface waves

have also been observed experimentally by, for example, Smith *et al.*[3]

V. CONCLUSIONS

This paper gives a detailed theoretical description of the existence of a leaky surface wave in an isotropic homogeneous solid. The leaky surface wave is shown to arise from the complex conjugate roots of the Rayleigh equation. It exists only for materials with a Poisson ratio larger than about 0.263. The leaky surface wave propagates along the surface with a wave speed smaller than the pressure wave, but larger than the shear wave. Due to matching tangential wave vectors at the surface, it couples into a plane shear wave directed into the medium. Both the surface wave and the plane shear wave are, because of the coupling, inhomogeneous. It is demonstrated that a normal line source on the surface of an infinite half-space in fact excites the leaky surface wave. The far field expressions for the leaky surface wave are given and are compared to numerical results obtained by using the finite-difference time-domain method. The particle motion on the surface due to the leaky surface wave is prograde (clockwise), contrary to the well-known Rayleigh surface wave which induces a retrograde (counterclockwise) particle motion.

ACKNOWLEDGMENTS

The authors would like to thank Professor Glenn S. Smith, Professor Peter H. Rogers, Dr. Gregg D. Larson, and James S. Martin for their advice and consideration throughout this research and their help in obtaining support for this project. This work is supported in part under the OSD MURI program by the U.S. Army Research Office under Contract No. DAAH04-96-1-0448, and by a grant from the U.S. Office of Naval Research under Contract No. N00014-99-1-0995.

[1] J. W. S. Rayleigh, Proc. London Math. Soc. **17**, 4 (1887).
[2] M. Roth, R. Spitzer, and F. Nitsche, in *Proc. 12th Annu. Int. Meeting, Symp. Application of Geophysics to Engineering and Environmental Problems* (1999), pp. 49–55.
[3] E. Smith, P. S. Wilson, F. W. Bacon, J. F. Manning, J. A. Behrens, and T. G. Muir, J. Acoust. Soc. Am. **103**, 2333 (1998).
[4] N. E. Glass and A. A. Maradudin, J. Appl. Phys. **54**, 796 (1983).
[5] R. A. Phinney, Bull. Seismol. Soc. Am. **51**, 527 (1961).
[6] K. F. Graff, *Wave Motion in Elastic Solids* (Dover, New York, 1975).
[7] J. Miklowitz, *The Theory of Elastic Waves and Wave Guides* (North-Holland, New York, 1978).
[8] J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, New York, 1973).
[9] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves* (IEEE Press, 1994).
[10] C. T. Schröder and W. R. Scott, Jr., IEEE Trans. Geosci. Remote Sens. **38**, 1505 (2000).

A staggered-grid finite-difference method with perfectly matched layers for poroelastic wave equations

Yan Qing Zeng and Qing Huo Liu[a)]
Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708

(Received 1 November 2000; accepted for publication 3 March 2001)

A particle velocity-strain, finite-difference (FD) method with a perfectly matched layer (PML) absorbing boundary condition is developed for the simulation of elastic wave propagation in multidimensional heterogeneous poroelastic media. Instead of the widely used second-order differential equations, a first-order hyperbolic leap-frog system is obtained from Biot's equations. To achieve a high accuracy, the first-order hyperbolic system is discretized on a staggered grid both in time and space. The perfectly matched layer is used at the computational edge to absorb the outgoing waves. The performance of the PML is investigated by calculating the reflection from the boundary. The numerical method is validated by analytical solutions. This FD algorithm is used to study the interaction of elastic waves with a buried land mine. Three cases are simulated for a mine-like object buried in ''sand,'' in purely dry ''sand'' and in ''mud.'' The results show that the wave responses are significantly different in these cases. The target can be detected by using acoustic measurements after processing. © *2001 Acoustical Society of America.*
[DOI: 10.1121/1.1369783]

PACS numbers: 43.20.Bi, 43.20.Fn, 43.20.Gp [ANN]

I. INTRODUCTION

Simulation of elastic waves propagating in fluid-saturated porous media is of great importance to geophysical exploration, reservoir engineering, and military applications.

With the increasing difficulty of exploring natural resources and the growing realization that hydrocarbon reservoirs are more heterogeneous and complex than assumed in the past, it is desirable to characterize the subsurface materials as fluid-saturated porous media than perfectly elastic single phase materials.

In military applications, simulation of waves in porous media is important for underwater acoustics. More recently, acoustic waves are used for land mine detection. For plastic land mines, acoustic waves provide a much better measurement than the traditional electromagnetic induction method. Although the soil can be approximated as a single phase elastic material, it is more accurate to treat soil as two phase composite materials consisting of granular solid and pore fluid.

Simulation of wave propagation in porous, fluid-saturated media requires the analytical or numerical solution of Biot's equations.[1–3] For a heterogeneous, complex model, in general it is not possible to find analytical solutions to Biot's equations. Numerical methods have to be used to obtain these solutions. Finite-difference (FD) algorithms have been developed to simulate wave propagation in poroacoustic media,[4] and in homogeneous poroelastic media.[5] A centered-grid FD scheme has also been developed for heterogeneous poroelastic media.[6] In this article, a velocity-strain, finite-difference method is developed in a staggered grid for heterogeneous poroelastic media. In this method, Biot's equations[3] are reformulated into first-order equations to arrive at a leap-frog system in a staggered grid both in time and space domains. Numerical solutions have been validated by analytical solutions.

In order to simulate an unbounded medium, an absorbing boundary condition (ABC) must be implemented to truncate the computational domain in numerical algorithms. There are many kinds of ABCs developed for numerical simulation of wave propagation. Cerjan *et al.*[7] introduced a simple damping taper to the boundaries to attenuate the outgoing waves. Since this lossy layer is not perfectly matched to the interior region, however, it requires a substantial number of layers for the taper to be effective. Clayton and Engquist[8] (CE) use the paraxial approximation to the wave equation to make the boundary transparent to outgoing waves. The CE ABC can lead to instability when the Poisson ratio is greater than 2.[9] Since Berenger[10] proposed the highly effective perfectly matched layer (PML) as an absorbing boundary condition for electromagnetic waves, the PML has been widely used for finite-difference and finite-element methods. Chew and Liu[11,12] first proposed the PML for elastic waves in solids, and proved the zero reflections from PML to the regular elastic medium. Hastings *et al.*[13] have independently implemented the PML ABC for two-dimensional problems by using potentials. The PML has also been extended to model acoustic waves and electromagnetic waves in lossy media.[14] The PML has been applied to the second-order Biot's equation for fluid-saturated poroelastic media,[15] which requires a complicated convolution. In this article, as the PML is developed for the first-order system, incorporation of PML becomes much simpler. The effectiveness of this ABC is confirmed by examining the reflection from the boundary.

By using this numerical method, interaction of elastic waves with a buried plastic minelike object is investigated. Three cases are simulated for a plastic minelike object buried

a)Electronic mail: qhliu@ee.duke.edu

in ''sand,'' in purely dry ''sand'' and in ''mud.'' The results show that the wave responses are significantly different in these cases. After processing, the target can be detected by using surface acoustic measurements.

II. FORMULATION

A. The governing equations

The propagation of acoustic waves in porous and fluid-saturated media is different from that in single phase elastic media. In addition to the regular P waves and S waves in solid elastic media, a slow P wave which results from the relative motion between solid frame and fluid may be present in porous media. Thus the pattern of energy dissipation in porous media is different from that in solid elastic media. Based on continuum mechanics and macroscopic constitutive relationship, Biot[1–3] developed a theory of wave motion in a porous elastic solid saturated with a viscous compressible fluid. Biot's theory was confirmed by Burridge and Keller[16] based on the dynamic equations which govern the behavior of medium on a microscopic scale. Plona[17] also confirmed Biot's theory through experiments.

In an isotropic, heterogeneous porous elastic medium, the parameters describing the physical properties of the medium are as follows:

- μ shear modulus of dry porous matrix
- λ_c Lame constant of saturated matrix
- ϕ porosity
- η viscosity
- κ permeability
- ρ the overall density of the saturated medium determined by $\rho_f\phi+(1-\phi)\rho_s$
- ρ_s density of solid material
- ρ_f density of fluid
- a tortuosity of the matrix
- K_s bulk modulus of the solid
- K_f bulk modulus of the fluid
- K_b bulk modulus of the dry porous frame

The macroscopic displacements and strains are defined as

- u_i ith component of displacements of solid particle
- U_i ith component of displacements of fluid particle
- w_i ith component of relative displacement, $w_i=\phi(U_i-u_i)$
- e_{il} il component of strain tensor in porous medium, $e_{il}=[((\partial u_l/\partial x_i)+(\partial u_i/\partial x_l))/2]$
- e $e=\Sigma_{i=1,2,3}e_{ii}$
- ξ dilatation for the relative motion, $\xi=-\Sigma_{i=1,2,3}(\partial w_i/x_i)$

For a three-dimensional isotropic, heterogeneous and porous elastic medium, wave propagation is governed by Biot's equations:[3]

$$2\sum_l \frac{\partial}{\partial x_1}(\mu e_{i1})+\frac{\partial}{\partial x_i}(\lambda_c e-\alpha M\xi)=\frac{\partial^2}{\partial t^2}(\rho u_i+\rho_f w_i), \tag{1}$$

$$\frac{\partial}{\partial x_i}(\alpha Me-M\xi)=\frac{\partial^2}{\partial t^2}(\rho_f u_i+mw_i)+\frac{\eta}{\kappa}\frac{\partial w_i}{\partial t}, \tag{2}$$

where $m=a\rho_f/\phi$ and

$$M=\frac{1}{\phi/K_f+(\alpha-\phi)/K_s}, \quad \alpha=1-\frac{K_b}{K_s}.$$

Let $\mathbf{v^s}$ be the velocity of the solid particle, and $\mathbf{v^f}$ be the velocity of the pore fluid relative to the solid frame. Then the second-order equations (1) and (2) can rearranged as the first-order equations

$$(m\rho-\rho_f^2)\frac{\partial v_i^s}{\partial t}=2m\sum_l \frac{\partial}{\partial x_l}(\mu e_{il})+m\frac{\partial}{\partial x_i}(\lambda_c e-\alpha M\xi)-\rho_f\frac{\partial}{\partial x_i}(\alpha Me-M\xi)+\rho_f\frac{\eta}{\kappa}v_i^f, \tag{3}$$

$$(m\rho-\rho_f^2)\frac{\partial v_i^f}{\partial t}=\rho\frac{\partial}{\partial x_i}(\alpha Me-M\xi)-\rho\frac{\eta}{\kappa}v_i^f-2\rho_f\sum_l \frac{\partial}{\partial x_l}(\mu e_{il})-\rho_f\frac{\partial}{\partial x_i}\times(\lambda_c e-\alpha M\xi). \tag{4}$$

The time derivatives of strains e_{il} and ξ can be expressed as

$$\frac{\partial e_{il}}{\partial t}=[(\partial v_l^s/\partial x_i+\partial v_i^s/\partial x_l)/2], \tag{5}$$

$$\frac{\partial \xi}{\partial t}=-\nabla\cdot\mathbf{v^f}. \tag{6}$$

In the explicit first-order finite-difference schemes, Eqs. (3)–(6) consist of a leap-frog system for the strain field e_{il}, ξ and velocity field $\mathbf{v^s}$ and $\mathbf{v^f}$. With proper absorbing boundary conditions, these equations can be solved numerically for the wave field in an unbounded medium.

Equations (3)–(6) predict the existence of three different waves in fluid-saturated poroelastic media: a shear wave and two compressional waves with a faster and a slower propagation velocities.

B. Equations for the PML absorbing boundary condition

In this article, the perfectly matched layer (PML) will be used to truncate the unbounded medium, absorbing all outgoing waves. The artificial absorptive medium is introduced in the regular medium by modifying Eqs. (3)–(6) with complex coordinates.[12,13,15] In the frequency domain (where a time-harmonic factor $e^{-j\omega t}$ is implied and $j=\sqrt{-1}$), a complex coordinate variable is chosen as

$$\tilde{x}_i=\int_0^{x_i} e_i(x_i')dx_i', \tag{7}$$

$$e_i=a_i+j\frac{\omega_i}{\omega} \quad (i=1,2,3), \tag{8}$$

where $a_i \geqslant 1$ is a scaling factor, and $\omega_i \geqslant 0$ is an attenuation factor. The operator $\partial/\partial x_i$ can be expressed in terms of the regular coordinate,

$$\frac{\partial}{\partial \tilde{x}_i} = \frac{1}{e_i}\frac{\partial}{\partial x_i}. \tag{9}$$

The PML formulation is to replace x_i in (3)–(6) by the corresponding complex coordinate $\tilde{x}_i$. In a PML region, the real part a_i is a scaling factor, and the imaginary part ω_i represents a loss in the PML. In a regular non-PML region, $a_i = 1$ and $\omega_i = 0$.

In order to simplify the PML equations, the field variables are split as follows:

$$v_i^s = \sum_{k=1}^{3} v_i^{s(k)}, \quad v_i^f = \sum_{k=1}^{3} v_i^{f(k)},$$

where $v_i^{s(k)}$ and $v_i^{f(k)}$ represent the split field variables containing space derivative $\partial/\partial x_k$ only. For example, Eq. (3) for $i=1$ can be split into the following three equations,

$$(m\rho - \rho_f^2)\frac{\partial v_1^{s(1)}}{\partial t} = 2m\frac{\partial}{\partial x_1}(\mu e_{11}) + m\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M \xi) - \rho_f \frac{\partial}{\partial x_1}(\alpha M e - M\xi),$$

$$(m\rho - \rho_f^2)\frac{\partial v_1^{s(2)}}{\partial t} = 2m\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f \frac{\eta}{\kappa} v_1^f,$$

$$(m\rho - \rho_f^2)\frac{\partial v_1^{s(3)}}{\partial t} = 2m\frac{\partial}{\partial x_3}(\mu e_{13}).$$

The diagonal strain components e_{ii} need not be split. However, other strain components have to be split as $e_{il} = \Sigma_{k=1}^{3} e_{il}^{(k)}$ for $i \neq l$, and $\xi = \Sigma_{k=1}^{3} \xi^{(k)}$. For example, from (5) we have

$$\frac{\partial e_{ii}}{\partial t} = \frac{\partial v_i^s}{\partial x_i},$$

$$e_{12} = e_{12}^{(1)} + e_{12}^{(2)},$$

$$\frac{\partial e_{12}^{(1)}}{\partial t} = \frac{1}{2}\frac{\partial v_2^s}{\partial x_1}, \quad \frac{\partial e_{12}^{(2)}}{\partial t} = \frac{1}{2}\frac{\partial v_1^s}{\partial x_2},$$

$$\xi = \xi^{(1)} + \xi^{(2)} + \xi^{(3)},$$

$$\frac{\partial \xi^{(1)}}{\partial t} = -\frac{\partial v_1^f}{\partial x_1},$$

$$\frac{\partial \xi^{(2)}}{\partial t} = -\frac{\partial v_2^f}{\partial x_2}, \quad \frac{\partial \xi^{(3)}}{\partial t} = -\frac{\partial v_3^f}{\partial x_3}.$$

In frequency domain PML formulation, after x_i is replaced by $\tilde{x}_i$, equations for v_1^s, for example, can be rewritten as

$$(m\rho - \rho_f^2)(-j\omega)\left(1 + j\frac{\omega_{x_1}}{\omega}\right)v_1^{s(1)} = 2m\frac{\partial}{\partial x_1}(\mu e_{11}) + m\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M \xi) - \rho_f \frac{\partial}{\partial x_1}(\alpha M e - M\xi),$$

$$(m\rho - \rho_f^2)(-j\omega)\left(1 + j\frac{\omega_{x_1}}{\omega}\right)v_1^{s(2)} = 2m\frac{\partial}{\partial x_2}(\mu e_{12}) + \left(1 + j\frac{\omega_1}{\omega}\right)\rho_f \frac{\eta}{\kappa} v_1^f,$$

$$(m\rho - \rho_f^2)(-j\omega)\left(1 + j\frac{\omega_1}{\omega}\right)v_1^{s(3)} = 2m\frac{\partial}{\partial x_3}(\mu e_{13}).$$

By taking the inverse Fourier transform, the above equations yield the time-domain PML equations,

$$(m\rho - \rho_f^2)\left(\frac{\partial v_1^{s(1)}}{\partial t} + \omega_1 v_1^{s(1)}\right) = 2m\frac{\partial}{\partial x_1}(\mu e_{11}) + m\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M \xi) - \rho_f \frac{\partial}{\partial x_1}(\alpha M e - M\xi), \tag{10}$$

$$(m\rho - \rho_f^2)\left(\frac{\partial v_1^{s(2)}}{\partial t} + \omega_1 v_1^{s(2)}\right) = 2m\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f \frac{\eta}{\kappa}\left(v_1^f + \omega_2 \int_{-\infty}^{t} v_1^f d\tau\right), \tag{11}$$

$$(m\rho - \rho_f^2)\left(\frac{\partial v_1^{s(3)}}{\partial t} + \omega_1 v_1^{s(3)}\right) = 2m\frac{\partial}{\partial x_3}(\mu e_{13}), \tag{12}$$

$$\frac{\partial e_{11}}{\partial t} = \frac{\partial v_1}{\partial x_1} - \omega_1 e_{11}, \tag{13}$$

$$\frac{\partial e_{12}^{(1)}}{\partial t} = \frac{1}{2}\frac{\partial v_2^s}{\partial x_1} - \omega_1 e_{12}^{(1)}, \tag{14}$$

$$\frac{\partial e_{12}^{(2)}}{\partial t} = \frac{1}{2}\frac{\partial v_1^s}{\partial x_2} - \omega_2 e_{12}^{(2)}, \tag{15}$$

$$\frac{\partial \xi^{(1)}}{\partial t} = -\frac{\partial v_1^f}{\partial x_1} - \omega_1 \xi^{(1)}, \tag{16}$$

$$\frac{\partial \xi^{(2)}}{\partial t} = -\frac{\partial v_2^f}{\partial x_2} - \omega_2 \xi^{(2)}, \tag{17}$$

$$\frac{\partial \xi^{(3)}}{\partial t} = -\frac{\partial v_3^f}{\partial x_3} - \omega_3 \xi^{(3)}. \tag{18}$$

Similar equations can be obtained in the same way for other components v_2^s, v_3^s, v_1^f, v_2^f, v_3^f, and e_{13}, e_{23}. Within the PML region, Eqs. (3) and (4) are split into 18 equations for 3D. Equation (5) is split into nine equations and Eq. (6) is split into three equations. So the total number of the equa-

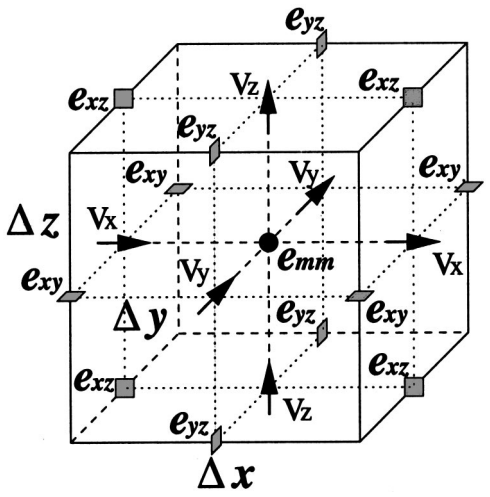


FIG. 1. The relative locations of field components in a unit cell of staggered grid.

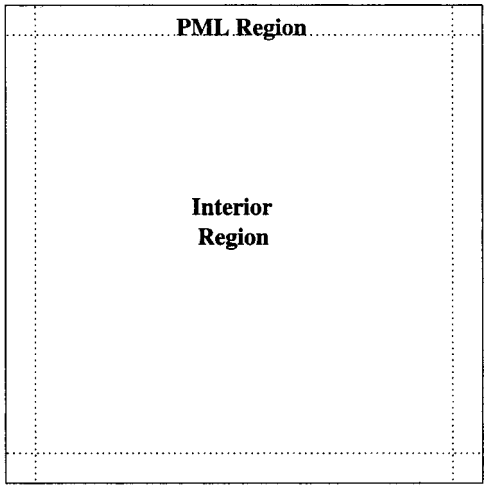


FIG. 2. Computational domain with an interior region and a PML boundary region.

tions is 30 for the PML region, compared to 13 for the regular interior region. Thus, the memory requirement within the PML region is about two and half times that required by a regular medium for three-dimensional problems. This extra memory requirement in the PML region is offset by the effectiveness of PML in absorbing the outgoing waves.

C. Finite-difference implementation

The governing equations for the PML absorbing boundary condition, such as Eqs. (10)–(18), are first-order partial differential equations for particle velocity and strain. They can be solved with different numerical methods. For the evaluation of seismic and acoustic responses of specific models, the accuracy and convenience of the numerical method are of primary concern. The finite-difference method is widely used in wave modeling because of its flexibility and accuracy. For the first-order, leap-frog system of Eqs. (3)–(6) and (10)–(18), the explicit finite-difference method is used on a staggered grid.

To implement a 3-D finite-difference solution to the equations with the PML, the material parameters and unknown field components are discretized on a regular 3-D grid at the intervals Δx_1, Δx_2 and Δx_3. The time domain is also discretized with time step Δt. There are two discretization schemes to approximate the first-order derivatives, i.e., the centered grid and the staggered grid. Because the centered-grid operator to perform first derivatives is less accurate than the staggered grid operator,[18] a staggered grid is used in this article. For the FD implementation of Biot's equations on a staggered grid in Fig. 1, the velocity field components are located at the cell's face centers, while material parameters and normal strains are located at the center of the cell, and shear strains are located the at six edge centers. Strain field is computed at $n\Delta t$ and velocity field is computed at $(n+\frac{1}{2})\Delta t$. This staggered grid is similar to that for elastic waves in a solid.[19,20]

With this discretization, the leap-frog system can be written in a time-stepping form. In order to make the layout of the formulas simple, the governing equations with PML boundary conditions can be generalized as the first-order differential equation. For examples, (11), (13) and (16) can be rewritten as

$$\frac{\partial v_1^{s(2)}}{\partial t}+c_0 v_1^{s(2)}=c_1\int_{-\infty}^{t} v_1^f d\tau + c_2\left[2m\frac{\partial}{\partial x_2}(\mu e_{12})+\rho_f\frac{\eta}{\kappa}(v_1^f)\right], \tag{19}$$

$$\frac{\partial e_{11}}{\partial t}+\omega_1 e_{11}=\frac{\partial v_2^s}{\partial x_1}, \tag{20}$$

$$\frac{\partial \xi^{(1)}}{\partial t}+\omega_1 \xi^{(1)}=-\frac{\partial v_1^f}{\partial x_1}, \tag{21}$$

where c_0, c_1 and c_2 space-dependent coefficients. The corresponding time-stepping equations can then be written as

$$v_1^{s(2)}[j_1,j_2,j_3,(n+\tfrac{1}{2})]=f_1 v_1^{s(2)}[j_1,j_2,j_3,(n-\tfrac{1}{2})] + f_2 R_1, \tag{22}$$

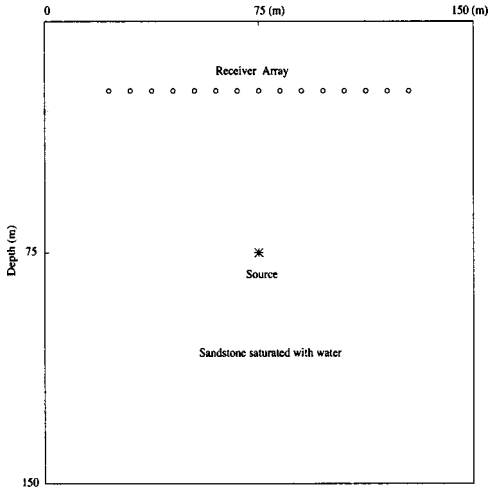


FIG. 3. A physical model showing source and receiver geometry used to generate seismograms. Receivers are 7.5 m apart.

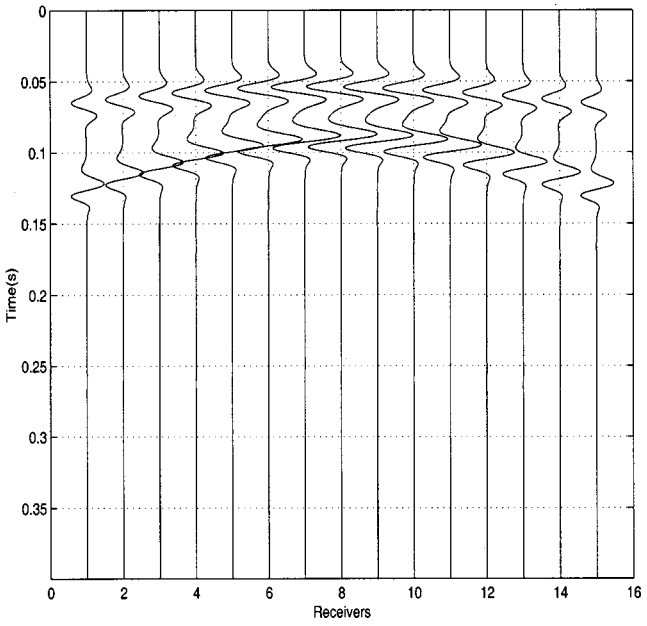


FIG. 4. Vertical velocity components waveforms of the model in Fig. 3.

$$e_{11}[j_1,j_2,j_3,(n+1)]=g_1 e_{11}[j_1,j_2,j_3,n]+g_2 R_2, \quad (23)$$

$$\xi^{(1)}[j_1,j_2,j_3,(n+1)]=-g_1\xi^{(1)}[j_1,j_2,j_3,n]-g_2 R_3, \quad (24)$$

$$f_1=\frac{c_0/2-1/\Delta t}{1/\Delta t+c_0/2}, \quad f_2=\frac{1}{1/\Delta t+c_0/2},$$

$$g_1=\frac{\omega_1/2-1/\Delta t}{1/\Delta t+\omega_1/2}, \quad g_2=\frac{1}{1/\Delta t+\omega_1/2},$$

where R_1, R_2 and R_3 are the right-hand sides of Eqs. (19), (20) and (21) respectively. It should be noted that the material parameters in the above equations must be properly averaged in order to arrive at a higher accuracy.[19] In order to incorporate the PML boundary condition, the computational domain is divided into a PML region and an interior region, as shown in Fig. 2. The absorption of outgoing waves is achieved by the PML region, which consists of several cells of mathematically defined materials with a quadratically tapered ω_i profile to increase the attenuation toward the outer boundary. In this article, ω_i of the PML region is chosen as

$$\omega_i(j_i)=\frac{(M-1/2-j_i)^2}{(M-1/2)^2}\omega_{i,\max}, \quad (25)$$

where $\omega_{i,\max}$ is the value at the center of the cell at the outermost boundary. At the outer boundary, the velocity components and shear strain are forced to be zero. For convenience, $\omega_{i,\max}$ can be expressed in terms of dominant frequency and a normalized coefficient a_0. Then Eq. (25) becomes

$$\omega_i=\begin{cases}2\pi a_0 f_0 (l_{x_i}/L_{\mathrm{PML}})^2, & \text{inside PML},\\ 0, & \text{outside PML},\end{cases} \quad (26)$$

where f_0 is the dominant frequency of the source, L_{PML} is the thickness of the PML region, and l_{x_i} is the distance from the interface between the interior region and the PML region.

III. NUMERICAL RESULTS

The finite-difference algorithm on a staggered grid has a higher accuracy than on a centered grid. Thus, the PML region can be made thinner with a staggered grid with a better absorption than with a centered grid. In this article, the length of the PML region is chosen to be 10 cells. The FD algorithm is illustrated by its two-dimensional implementation.

In the following numerical results, a pure P-wave source is used to excite the seismic wave field. The source time function is the first derivative of a Gaussian function

$$S(t)=(t-t_0)e^{-\pi^2 f_0^2 (t-t_0)^2},$$

where f_0 is the predetermined dominant frequency, and t_0 the time shift.

Unless otherwise stated, a bulk source is used in the following examples. The source energy is partitioned linearly between the solid and the fluid phases with factors

$$W_f=\phi, \quad W_s=(1-\phi), \quad W_r=\phi|W_f-W_s|,$$

where W_f is the weighting factor for the fluid motion, W_s is for the solid motion and W_r is for the relative motion between solid frame and pore fluid.

A. PML performance

The effectiveness of the absorbing boundary condition is an essential factor for the successful numerical simulation. A good absorbing boundary condition has the characteristics of effective absorption of outgoing waves without requiring a large memory. In this article, the performance of the PML is investigated on a homogeneous, fluid-saturated poroelastic

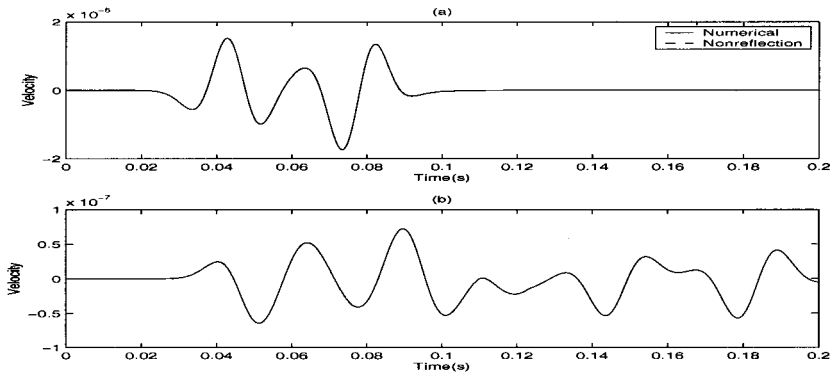


FIG. 5. (a) Comparison of numerical solution at a location 5 cells away from the PML boundary in the model in Fig. 3 with a reference solution. (b) The difference between these two solutions.

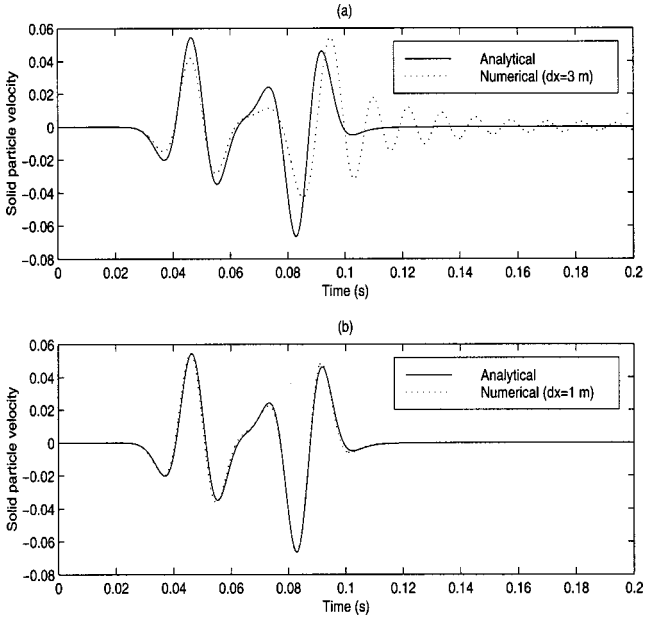


FIG. 6. Comparison of analytical and numerical solutions for the vertical velocity component in the solid for (a) $\Delta x=3$ m and (b) $\Delta x=1$ m.

medium by comparing numerical solutions of PML model with reference solutions that do not have reflections. The optimized a_0 can be obtained by examining the attenuation of the wave field in the PML region.

The homogeneous model for the PML performance test is a sandstone fully saturated with water having the properties of $\rho_s=2650$ kg/m^3, $\rho_f=1040$ kg/m^3, and porosity $\phi=0.3$. This model has the wave properties $v_{\text{fast}}=2365.7$ m/s for fast P wave, $v_{\text{slow}}=776.95$ m/s for slow P wave and $v_{\text{shear}}=960.5$ m/s for shear wave. The geometry of the physical model with receivers and source is shown in Fig. 3.

Figure 4 shows the numerical results of vertical velocity in solid from the model with the PML boundary condition. With the PML region, the reflections are eliminated from the seismograms. Another advantage of the PML absorbing boundary condition is its stability. For this particular model, $v_{\text{fast}}/v_{\text{shear}}=2.46$, which will cause an instability problem for the CE boundary condition.[10]

In order to quantify the reflections from the PML boundary, the reflection coefficient in dB is calculated for a location 5 cells away from PML interface in this model. The reflection was obtained by comparing the numerical results from the model with the PML boundary to those from a much larger model with the Dirichlet boundary in which reflections have not arrived within the time window of interest. Figure 5(a) shows these two results together, while Fig. 5(b) shows the difference. The PML result and the reference are almost indistinguishable at the signal scale. Compared to incident signal, the reflection is about 50 dB down.

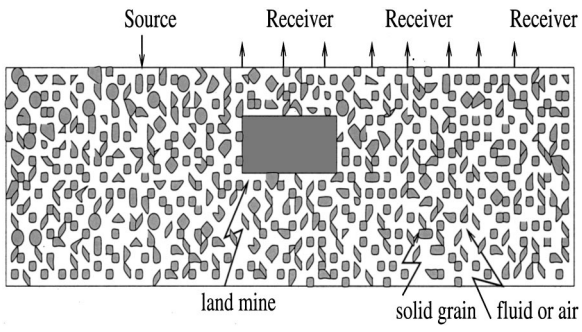


FIG. 7. Geometry of the model of a buried minelike object.

TABLE I. Properties of sand and mud (from Ref. 21).

	Sand	Mud
Porosity	0.4	0.6
Density (kg/m^3)	1990	1660
Structure factor	1.8	1.9
Permeability (m^2)	3×10^{-11}	3×10^{-13}
Bulk modulus (Pa)	4×10^7	1×107
Shear modulus (Pa)	2.2×10^7	6×10^6

B. Validation of numerical results

An analytical solution for the particle velocity field in a homogeneous, fluid-saturated poroelastic medium subject to a point source in 3-D space or a line source in 2-D space can be derived.[6] The particle velocity is obtained in a closed form via potential functions.

For Biot's equations, it is convenient to solve for the particle velocity through potential functions. The velocity of particles and body force at source can be expressed in terms of potentials as

$$\mathbf{u}=\nabla\psi_s+\nabla\times\Psi_s,$$

$$\mathbf{U}=\nabla\psi_f+\nabla\times\Psi_f,$$

$$\mathbf{f}=\nabla\Phi+\nabla\times\Psi,$$

where $\nabla\cdot\Psi_f=0$, $\nabla\cdot\Psi_s=0$ and $\nabla\cdot\Psi=0$ and describe the rotational potentials, while ψ_s, ψ_f and Φ describe the dilatational potentials. If the source is purely dilatational, then its rotational component disappears. In the time domain, for a purely P-wave point source and ideal nonviscous fluid, the potential can be expressed as

$$\psi_s(r,t)=\frac{\alpha s(t-r/V_f)+\beta s(t-r/V_s)}{4\pi r},$$

$$\psi_s(r,t)=\frac{\alpha A_f s(t-r/V_f)+\beta A_s s(t-r/V_s)}{4\pi r},$$

where r is the distance and $s(t)$ is the source time function. A_f and A_s represent the ratios between the solid and fluid motion for the fast P wave and the slow P wave. The coefficients α and β are determined by the regularity conditions. V_f and V_s are the velocities of the fast P wave and the slow P wave, respectively.

In two dimensions, for a pure P-wave line source along the y axis, the solution can be obtained by integrating the point source solution in the y direction. In the $x-z$ plane, the dilatational potentials are

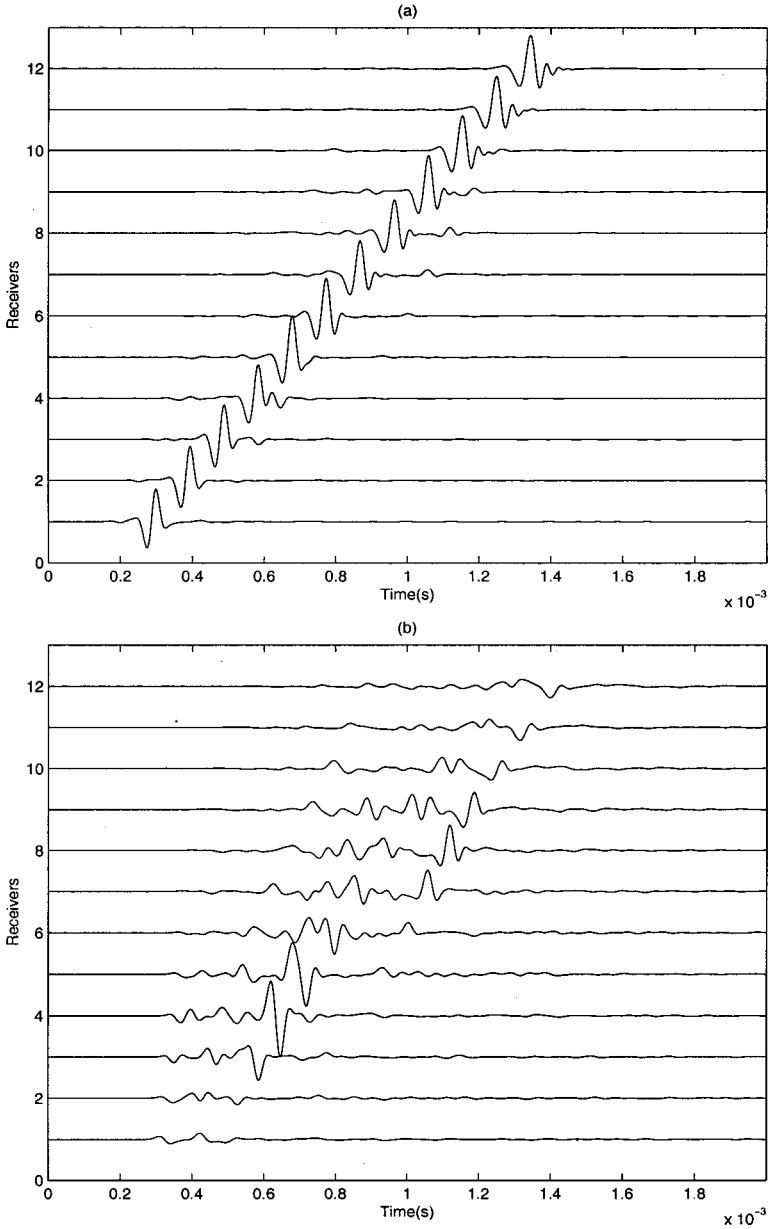


FIG. 8. Seismograms of the vertical particle velocity on the surface for a plastic minelike object in a ''dry sand'' model. (a) The total field. (b) The scattered field.

$$\psi_s(x,z,t)=\frac{\alpha \mathrm{H}(t-r/V_f)}{2\pi}\int_{r/V_f}^{t}\frac{s(t-\tau)}{\sqrt{\tau^2-r^2/V_f^2}}\,d\tau + \frac{\beta \mathrm{H}(t-r/V_s)}{2\pi}\int_{r/V_s}^{t}\frac{s(t-\tau)}{\sqrt{\tau^2-r^2/V_s^2}}\,d\tau,$$

$$\psi_f(x,z,t)=\frac{\alpha A_f \mathrm{H}(t-r/V_f)}{2\pi}\int_{r/V_f}^{t}\frac{s(t-\tau)}{\sqrt{\tau^2-r^2/V_f^2}}\,d\tau + \frac{\beta A_s \mathrm{H}(t-r/V_s)}{2\pi}\int_{r/V_s}^{t}\frac{s(t-\tau)}{\sqrt{\tau^2-r^2/V_s^2}}\,d\tau,$$

where H( ) is the Heaviside step function and $r=\sqrt{x^2+z^2}$. Once the potential functions are available, the velocity can be easily obtained by taking the gradient of potential functions.

The validation of the numerical method can be done by comparing the numerical results with the above analytical solution. A homogeneous model whose parameters are the same as the previous model is considered. A P-wave line source of the first derivative Gaussian time function with $f_0=40$ Hz is located at (0,0). Then the solution at (30 m,30 m) is calculated numerically and analytically. The numerical solutions for two different grid spacings are displayed with an analytical solution in Fig. 6. The oscillatory tails and disagreement in Fig. 6(a) for a coarser grid are caused by the dispersion of the slow P wave. When the grid spacing is decreased, the numerical solutions agree well with the analytical solution in Fig. 6(b).

C. Applications

This algorithm can be used to characterize a reservoir in a large scale as well as to investigate soil property in a small scale. In this section, this method is used to model a plastic minelike object that is difficult to detect with the more conventional electromagnetic induction sensors (i.e., metal de-

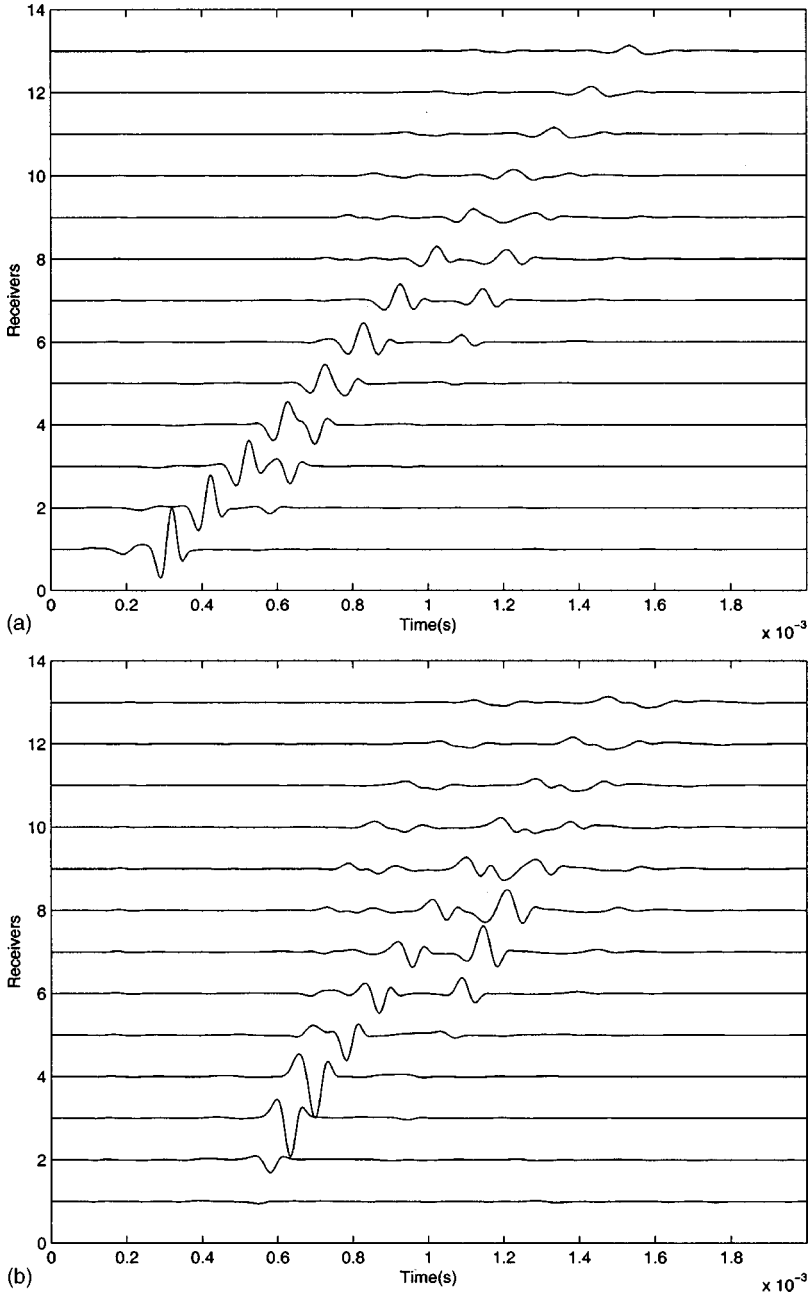


FIG. 9. Same as Fig. 8 except for the ''sand'' model.

tectors). The geometry of the model of a buried object is shown in Fig. 7.

For the plastic minelike object, the parameters are chosen as $\rho = 1330$ kg/M^3, S-wave velocity $v_{\text{shear}} = 417$ m/s and P-wave velocity $v_p = 1060$ m/s. The material in the ground may be considered as a two-phase composite material consisting of granular solid and pore fluid. The nature of this composite varies with environment, geographic location, and with depth below the surface of ground. Three different types of soil—dry ''sand,'' ''sand'' and ''mud''—are used. The first is similar to a coarse sand deposit saturated with air. The second is similar to a coarse sand deposit saturated with water. The third is similar to a fine clay mixture or mud. The elastic properties of these materials are listed next and in Table I.

Grain and fluid properties	$K_s = 3.6 \times 10^{10}$ Pa
	$K_f = 2.0 \times 10^9$ Pa
	$\rho_s = 2650$ kg/m^3
	$\rho_f = 1000$ kg/m^3
Viscosity	$\eta = 1.0 \times 10^{-3}$ Mks
Sound speed of fluid	$C_0 = 1414$ m/s

A vertical dipole line source with the first derivative of Gaussian pulse with center frequency of 15 kHz is located on the free surface. The receivers are uniformly distributed on the free surface at a distance 2 to 14.5 cm from the source. The plastic minelike object is buried in the ground with its upper edge 2 cm beneath the surface of the ground. The space step is chosen such that there are 15 grid points per minimum wavelength. The time step is chosen according to the stability condition. Because the minimum wave velocity is different, these three models have a different grid size and number of cells. The computational domain contains 600 $\times 160$ cells for the ''dry sand'' model, 400×100 cells for the

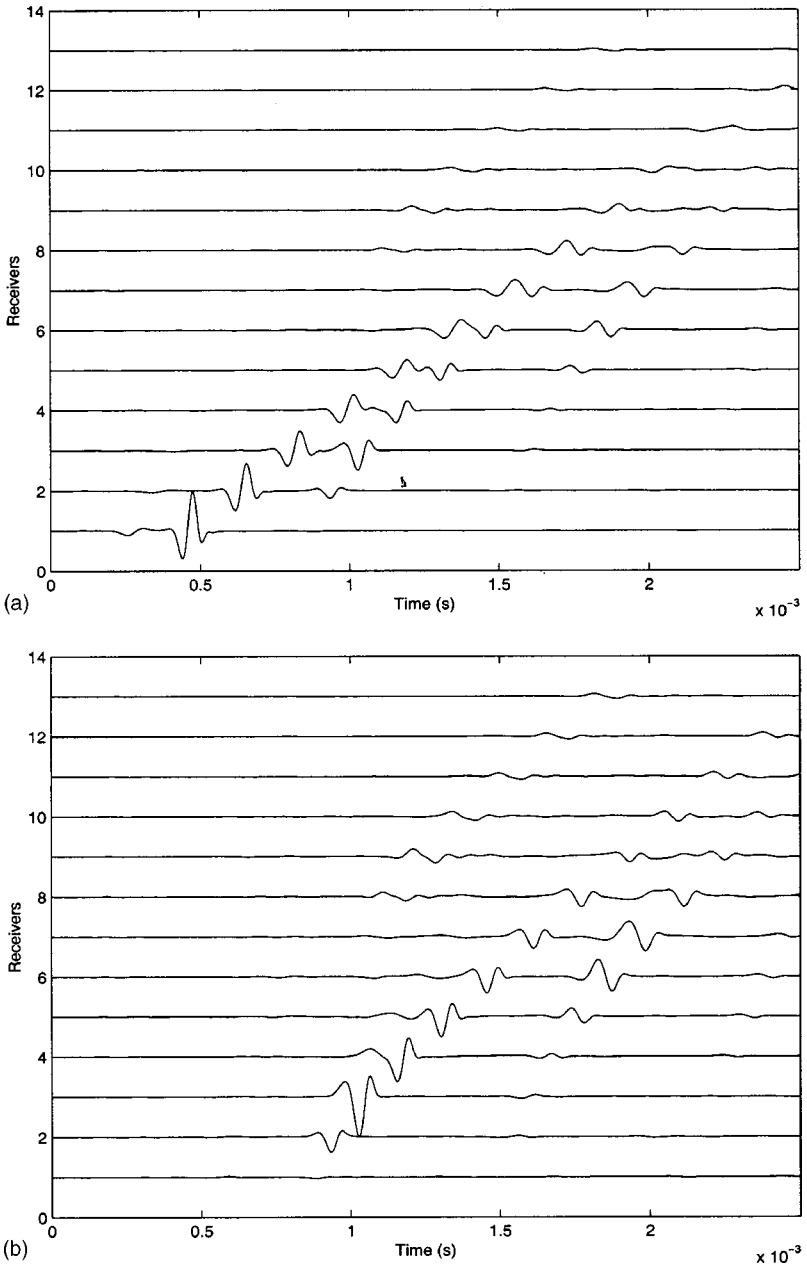


FIG. 10. Same as Fig. 8 except for the ''mud'' model.

''sand'' model, and 600×160 for the ''mud'' model, including a PML region with a thickness of 10 cells.

The numerical simulations of the soil with plastic mine-like objects were carried out to examine the effects of the plastic object on the wave field and investigate the possibility of the detection of a plastic object in the ground by using acoustic method. Figure 8(a) shows the synthetic seismograms of the vertical velocity of solid particle at the surface for the dry sand model. Here, the vertical velocity is plotted as a function of time and vertically offset by the distance from the source. The wave speed is indicated by the slope of the traveling waves in the figure. Clearly, the Rayleigh surface wave dominates the wave field. In the scattered field in Fig. 8(b) (obtained by subtracting the background field from the total field), the reflected shear wave dominates the wave field. Figures 9(a) and (b) show the total field and scattered field for the ''sand'' model. In the total field, a strong Rayleigh surface which has lower speed than the shear wave is present. The reflected shear wave can be seen clearly just behind the surface wave on the first several traces. Because it has a higher speed, it surpasses the surface wave at far offsets. In this figure, the reflected shear wave dominates other waves. Figures 10(a) and (b) show the total and scattered fields respectively for the ''mud'' model. The results are very similar to those of the ''sand'' model except that the speeds of surface waves and shear waves are much lower.

IV. CONCLUSION

A particle velocity-strain, finite-difference method combined with the perfectly matched layer (PML) has been formulated in three dimensions for the simulation of seismic waves propagating in porous media. The performance of the PML boundary in two dimensions has been studied on a

homogeneous model by calculating the reflection from the boundary. The analytical solution to Biot's equations is used to validate the numerical algorithm. The results show that the first-order, leap-frog, staggered-grid scheme has a higher accuracy than the second-order, centered-grid finite-difference scheme. Even with only a 10-cell PML region, the reflection is much less than that of second-order centered-grid scheme with 20-cell PML. The results also show the scheme is stable even if the ratio of the fast P-wave velocity to shear wave velocity is greater than 2.

This numerical method has been used to investigate the interaction of elastic waves with a buried minelike object. Three cases have been investigated: a plastic object buried in purely dry ''sand,'' in ''sand,'' and in '' mud.'' The results show that the wave responses are substantially different for different backgrounds.

[1] M. A. Biot, ''Theory of propagation of elastic waves in a fluid-saturated porous solid. 1. Low-frequency range,'' J. Acoust. Soc. Am. **28**, 168–178 (1956).

[2] M. A. Biot, ''Theory of propagation of elastic waves in a fluid-saturated porous solid. 2. Higher-frequency range,'' J. Acoust. Soc. Am. **28**, 179–191 (1956).

[3] M. A. Biot, ''Mechanics deformation and acoustic propagation in porous media,'' J. Appl. Phys. **33**, 1482–1498 (1962).

[4] S. Hassanzadeh, ''Acoustic modeling in fluid saturated porous media,'' Geophysics **56**, 424–435 (1991).

[5] X. Zhu and G. A. McMechan, ''Finite-difference modeling of the seismic response of fluid saturated, porous, elastic media using Biot theory,'' Geophysics **56**, 328–339 (1991).

[6] N. Dai, A. Vafidis, and E. R. Kanasewich, ''Wave propagation in heterogeneous, porous media: A velocity-stress, finite-difference method,'' Geophysics **60**, 327–340 (1995).

[7] C. Cerjan, D. Kosloff, R. Kosloff, and M. Reshef, ''A non-reflecting boundary condition for discrete acoustic and elastic wave equations,'' Geophysics **50**, 705–708 (1985).

[8] R. Clayton and B. Engquist, ''Absorbing boundary conditions for acoustic and elastic wave equations,'' Bull. Seismol. Soc. Am. **67**, 1529–1540 (1977).

[9] K. D. Mahrer, ''An empirical study of instability and improvement of absorbing boundary conditions for elastic wave equation,'' Geophysics **51**, 1499–1501 (1986).

[10] J. P. Berenger, ''A perfectly matched layer for the absorption of electromagnetic waves,'' J. Comput. Phys. **114**, 185–200 (1994).

[11] W. C. Chew and Q. H. Liu, ''Perfectly matched layers for elastodynamics: A new absorbing boundary condition,'' J. Comput. Acoust. **4**, 72–79 (1996).

[12] Q. H. Liu, ''Perfectly matched layers for elastic waves in cylindrical and spherical coordinates,'' J. Acoust. Soc. Am. **105**, 2075–2084 (1999).

[13] F. D. Hastings, J. B. Schneider, and S. L. Broschat, ''Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation,'' J. Acoust. Soc. Am. **100**, 3061–3069 (1996).

[14] Q. H. Liu and J. Tao, ''The perfectly matched layer for acoustic waves in absorptive media,'' J. Acoust. Soc. Am. **102**, 2072–2082 (1997).

[15] Y. Q. Zeng, J. He, and Q. H. Liu, ''The application of the perfectly matched layer in numerical modeling of wave propagation in poroelastic media,'' Research Report, Duke University (1999).

[16] R. Burridge and J. B. Keller, ''Poroelasticity equations derived from microstructure,'' J. Acoust. Soc. Am. **70**, 1140–1146 (1981).

[17] J. Plona, ''Observation of the second bulk compressional wave in a porous medium at ultrasonic frequencies,'' Appl. Phys. Lett. **36**, 259–261 (1980).

[18] G. Kneib and C. Kerner, ''Accurate and efficient seismic modeling in random media,'' Geophysics **58**, 576–588 (1993).

[19] Q. H. Liu, F. Daube, C. Randall, E. Schoen, H. Liu, and P. Lee, ''A 3D finite difference simulation of sonic logging,'' J. Acoust. Soc. Am. **100**, 72–79 (1996).

[20] Q. H. Liu, ''Some current trends in numerical methods for transient acoustic and elastic waves in multidimensional inhomogeneous media,'' Current Topics Acoustical Research **2**, 31–42 (1998).

[21] J. I. Dunlop, ''Propagation of acoustic waves in marine sediments, a review,'' Exploration Geophysics **19**, 513–535 (1988).

Combinatorial Approaches to the Synthesis of Vapor Detector Arrays for Use in an Electronic Nose

Adam J. Matzger, Carolyn E. Lawrence, Robert H. Grubbs,* and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received October 4, 1999

Introduction

The technology of "electronic noses", i.e., devices that functionally mimic the sense of olfaction, is rapidly evolving, driven by the practical demand for objective analysis of odors as well as by the intellectual challenge of mimicking the mammalian sense of olfaction.[1−3] Numerous implementations of artificial noses have emerged; however, the most prevalent devices are based on detecting a physical or chemical change in a polymer film upon exposure to a gaseous analyte.[4] The measured quantity in an individual detector can be the frequency shift of a resonating crystal in a quartz crystal microbalance (QCM)[5] or a surface acoustic wave configuration (SAW),[6,7] changes in the optical absorption or emission properties of a dye that has been impregnated into a polymer,[8−10] or changes in the electrical resistance of a conductive polymer (cp)[11] or of a carbon black/polymer composite (cb/pc) film.[12] Any individual detector does not respond highly specifically toward an individual analyte, but the pattern of responses in an array of differentially responsive detectors can be used to identify, classify, and in some cases quantify, the analyte of interest. The common threads that tie these methods together are (a) their reliance on sorption of an analyte into the polymer film to produce the physicochemical effect that leads to the signal transduction event in the detectors of the array, and (b) the use of pattern recognition to send the collection of signals to a central processing unit for analysis.

Like the artificial olfactory systems, there is mounting evidence that mammals do not employ lock-and-key type specificity to individual analytes in the broadly responsive portion of their olfactory systems.[13] Instead, humans, mice, and other mammals are believed to rely on a pattern generated from the response of many broadly tuned olfactory receptor sites, just as the individual polymer types in an electronic nose are swollen by many chemically diverse analytes. Humans are thought to have ∼1000 different olfactory genes that presumably encode for ∼1000 olfactory receptor proteins. A major goal of developing a functional mimic for the mammalian olfactory system is therefore to produce a highly diverse array of differentially responsive vapor detectors.

* Authors to whom correspondence should be addressed.

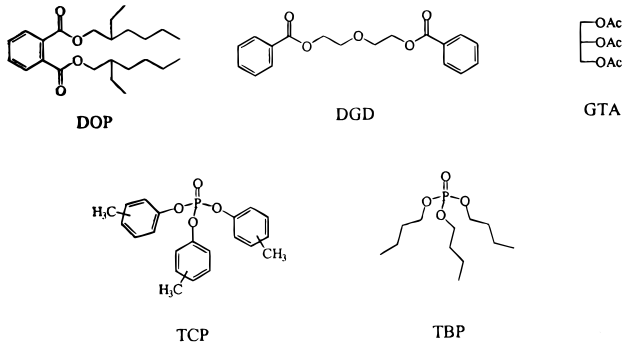


Figure 1. Plasticizers employed in this study (chlorinated paraffins omitted).

In our implementation of an artificial nose, arrays of carbon black/polymer composites comprised of 10−20 compositionally different detectors have been shown to differentiate efficiently between many organic vapors when assessed under controlled conditions in the laboratory. Swelling of the polymers leads to an increase in the dc resistance of the composite film, and the pattern of resistances is read though simple electronics and transferred to a processor for analysis. Distinctive patterns have allowed pairwise differentiation between species that differ in structure and polarity as well as between members of homologous series of, for example, alcohols or alkanes.[14] Furthermore, the response of the carbon black/polymer composite sensors is linearly related to the concentration of a given vapor.[12,14,15] However, even though hundreds of commercial polymers are potentially available for use, a major distinction between the existing conducting polymer composite electronic nose system and the mammalian olfactory system lies in sensor diversity. It would be useful if such diversity could be obtained from combining a limited number of feedstocks in a combinatorial fashion, as opposed to having to use different feed solutions for the fabrication of each detector. We report herein two implementations of a combinatorially based strategy to yield conducting polymer composite arrays having detectors with diverse vapor response properties.

Methods

Poly(methyl methacrylate), poly(vinyl chloride), poly(vinyl acetate), and polystyrene were used as received (Scientific Polymer Products, Inc.). The plasticizers di(2-ethylhexyl)phthalate (DOP), diethylene glycol dibenzoate (DGD), glycerol triacetate (GTA), tributyl phosphate (TBP), chloroparaffin (50% Cl, CP50), chloroparaffin (70% Cl, CP70), and tricresyl phosphate (TCP) (Figure 1) were used as received (Scientific Polymer Products, Inc.).

Vapor detectors were fabricated from solutions consisting of 20% carbon black (Black Pearls 2000, a furnace material produced by Cabot Co.), 20% plasticizer, and 60% polymer (by weight), or from solutions of 20% carbon black and 80% polymer, using methods described previously.[14] The detector films were spray coated onto the surface of a glass slide onto

10.1021/cc990056t CCC: $19.00 © 2000 American Chemical Society
Published on Web 04/29/2000

which Au lines had been deposited by thermal evaporation. The changes in dc electrical resistance between the Au contact lines produced by these composites upon sorption of an analyte were monitored using a Keithley model 2002 meter connected to a Keithley model 7001 multiplexer. The flow system used to produce controlled concentrations of analyte and the associated electronics used to control the analyte delivery and data acquisition functions have been described previously.[12] Resistance measurements of the composites were obtained before, during, and after exposure to controlled concentrations of hexane, toluene, chloroform, tetrahydrofuran, acetone, ethyl acetate, ethanol, or methanol. All exposures were adjusted to deliver the analyte at a concentration that was equal to 5% of the analyte's vapor pressure at room temperature. This produced analyte concentrations of (in units of parts per thousand (ppth)): methanol (6.63), ethanol (3.21), ethyl acetate (4.87), acetone (12.9), tetrahydrofuran (9.02), chloroform (10.7), and hexane (9.74).

The average steady-state dc resistance response of the composite was used to determine the differential resistance change of the detector, ΔR_{max}, in the presence of an analyte. Dividing this value by the baseline resistance, R_b, yielded the relative differential resistance, $\Delta R_{max}/R_b$, which was the key quantity used to determine the response of a given detector to an analyte. Data were collected for all detectors exposed 15 times to each analyte, with each exposure being performed for 5 min. Individual exposures to any given analyte were performed in random order during the data collection period.

The ability of each detector array to distinguish between different analytes was evaluated using the Fisher linear discriminant algorithm, which searches for the projection vector that maximizes the pairwise resolution factor for each set of analytes. The numerical metric produced by this method describes the separation of the data clusters when projected along this optimal discriminant vector.[14] Resolution factors obtained from this process represent the statistical confidence of being able to identify correctly pairwise one of the analytes in the test set of interest from another, with resolution factors of 1.0, 2.0, and 3.0 corresponding to approximately 76, 92, and 98% confidence, respectively, of correctly identifying an analyte of a given solvent pair during a single subsequent exposure to the detector array.

Results and Discussion

Two approaches to obtaining sensor diversity from a limited number of chemical feedstocks have been explored in this work. One method uses a series of block copolymers synthesized from combinations of monomer feedstocks, and the other method involves modification of the properties of a base polymer composite vapor detector through addition of a series of plasticizers.[16] In each case, a question of concern is whether detectors that contain the more compositionally-complex films have unique sorption properties for at least some members of a test set of analyte vapors.[17]

Figure 2 illustrates the effect that addition of various plasticizers had on the response of a carbon black/poly(vinyl acetate) composite vapor detector. As shown in the figure,

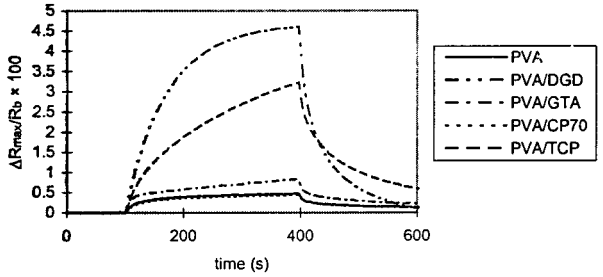


Figure 2. Response curves for a poly(vinyl acetate) (PVA)/carbon black composite detector compared to the response of plasticized poly(vinyl acetate)/carbon black detectors. In each case, the plasticizer indicated in the legend was present at 20 wt % in the solution used to form the detectors. The analyte, at 5% of its vapor pressure at room temperature (i.e., a concentration of 12.9 ppth of acetone in air), was exposed to the sensors starting on the abscissa at 100 s and ending at 400 s. The $\Delta R_{max}/R_b$ data of the ordinate, where R_b is the drift-corrected baseline resistance of the detector prior to the exposure and ΔR_{max} is the maximum differential resistance change upon exposure to the analyte relative to this baseline resistance value, clearly indicate that the plasticizer had a distinct effect on the response properties of the base polymer used in the carbon black/polymer composite detectors.

a relatively small $\Delta R_{max}/R_b$ response of the detector to 12.9 ppth of acetone was observed in the absence of an additive. In contrast, each plasticizer, despite being present in an equal mass fraction, resulted in a different response from the detector, with some of the responses being much larger than the responses of the unplasticized polymer composite film.

These differences in response permitted formation of a detector array that could distinguish between various analytes based on their differing array response patterns, even though only one base polymer was used in the formation of the detector array. The array that was evaluated consisted of different carbon black composite detectors, one formed using the unplasticized polymer and the other detectors formed from the same polymer through addition of a fixed, constant weight fraction of different plasticizers. Tables 1 and 2 present the resolution factors of such four- or five-detector arrays in differentiating pairwise between seven test analyte vapors, with the data in Table 1 for a poly(vinyl chloride)-based array and the data in Table 2 for a poly(vinyl acetate)-based array. Given that an rf of 3 corresponds to 98% confidence of correctly differentiating between a pair of analytes on a single exposure to the detector array, both arrays formed from the base polymer with a series of plasticizers were clearly able to differentiate between the test analytes with high levels of confidence. When fewer detector elements (i.e., fewer different types of plasticizers) were employed, the observed resolution factors generally decreased for the test set of analytes under consideration (see Supporting Information). In practice, columns of such base polymers would be deposited onto the pixels of a substrate having active electronics and electrical contacts to the polymer films arranged into a rectangular array of unit cells. Different plasticizers would then be deposited along the rows of this substrate. Through use of a binary mixing process, a target of 1000 compositionally different detectors could thus be achieved through use of 30 different polymers with 30 different, compatible plasticizers. Similarly, as few as 30 total polymer and plasticizer feedstocks would be needed if a

Table 1. Resolution Factors for a Four-Detector Array Consisting of Carbon Black Composites[a]

	toluene	chloroform	THF	acetone	ethyl acetate	ethanol	methanol
hexane	7.4	12.7	13.3	22.0	14.7	10.2	24.0
toluene		6.1	9.7	12.5	9.4	10.4	15.7
chloroform			4.5	8.3	6.1	9.7	19.2
THF				10.8	4.3	11.1	18.4
acetone					10.9	16.9	24.0
ethyl acetate						12.4	22.5
ethanol							20.0

[a] One detector was formed from a composite of pure poly(vinyl chloride), and three other detectors were formed from composites of poly(vinyl chloride) that had been plasticized with DOP, DGD, or TCP, respectively. Analytes were presented to the detectors at a concentration of 5% of their vapor pressure at room temperature in a background air ambient. A resolution factor of 3 represents a 98% statistical confidence for correct identification of one analyte of the chosen solvent pair relative to the other analyte in a single subsequent exposure to the detector array.

Table 2. Resolution Factors for a Five-Detector Array of Carbon Black Composites[a]

	toluene	chloroform	THF	acetone	ethyl acetate	ethanol	methanol
hexane	11.0	23.7	14.4	60.8	13.2	21.1	23.9
toluene		17.8	9.0	35.5	9.4	12.3	15.9
chloroform			9.0	8.6	7.7	9.4	10.8
THF				13.6	1.8	7.5	11.9
acetone					11.6	32.8	40.9
ethyl acetate						8.1	12.7
ethanol							6.5

[a] One detector was formed from pure poly(vinyl acetate) as the polymer; the other four detectors were formed from composites of poly(vinyl acetate) that had been plasticized with DGD, GTA, CP70, or TCP, respectively. Analytes were presented to the detectors at 5% of their vapor pressure at room temperature in a background air ambient. A resolution factor of 3 represents a 98% statistical confidence for correct identification of one analyte of the chosen solvent pair relative to the other analyte in a single subsequent exposure to the detector array.

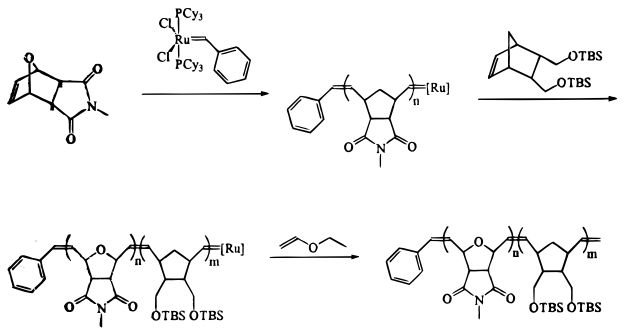
ternary mixing process (and no compositional gradients) were used to fabricate the 1000 chemically different detector films.

A second method for potentially producing diversity in conducting polymer composites involves the use of block copolymers. Unlike polymer blends, block copolymers can be made from chemically diverse monomers without obtaining macrophase separation of the components. A potential drawback is the possibility that the response of the block copolymers could be a linear combination of the two homopolymers, thus yielding no unique sorption data that could be used to identify analytes.[18]

Block copolymers were synthesized using the living polymerization process of ring-opening metathesis polymerization, as depicted in Scheme 1.[19] The monomers were chosen such that one monomer was very polar and the other monomer was relatively nonpolar, to maximize the difference in the sorption coefficient of the pure homopolymers in the presence of a range of analytes.

Figure 3 compares the response properties of carbon black composites of homopolymers formed from the two different monomers to the response properties of carbon black composite detectors formed from the block copolymerization reactions of these same two monomer units. Five different block copolymer compositions were evaluated during this experiment, and each analyte was presented to the resulting carbon black composite chemiresistors at 5% of the analyte's vapor pressure. The trend in the responses to hexane corresponds closely to what would be anticipated based on polarity of the polymers. Thus, the silyl-containing homopolymer exhibited the largest $\Delta R_{max}/R_b$ response. Also, the magnitude of the response decreased nearly linearly with increasing proportion of the polar block (Figure 3a). Similarly, methanol effectively swelled the polymers containing

Scheme 1. Synthesis of Block Copolymers for Sorption Studies[a]



[a] The first monomer formed the polar block, while the silyl-containing, second monomer formed the nonpolar block of the material.

the polar units, but the response decreased for polymers containing more than 50% of the silyl monomer (Figure 3b). Exposure to acetone, however, provided an example of an analyte in which the 50:50 block composition was the most responsive (Figure 3c). Different behavior was observed upon exposure to ethyl acetate (Figure 3d). These results demonstrate that the responses of block copolymers are not a simple function of their constitution and suggest that multiblock structures synthesized by living polymerization processes may offer an interesting approach to constructing large libraries of chemically diverse polymeric sensors.

The combination of the use of block copolymers and plasticizers suggests that a highly chemically diverse "nose-chip" detector array could be formed from a limited number of feedstock solutions, if these feedstocks were arrayed spatially across a substrate to form compositionally distinct materials at each position in the array. Such a procedure would avoid the requirement of delivering a different

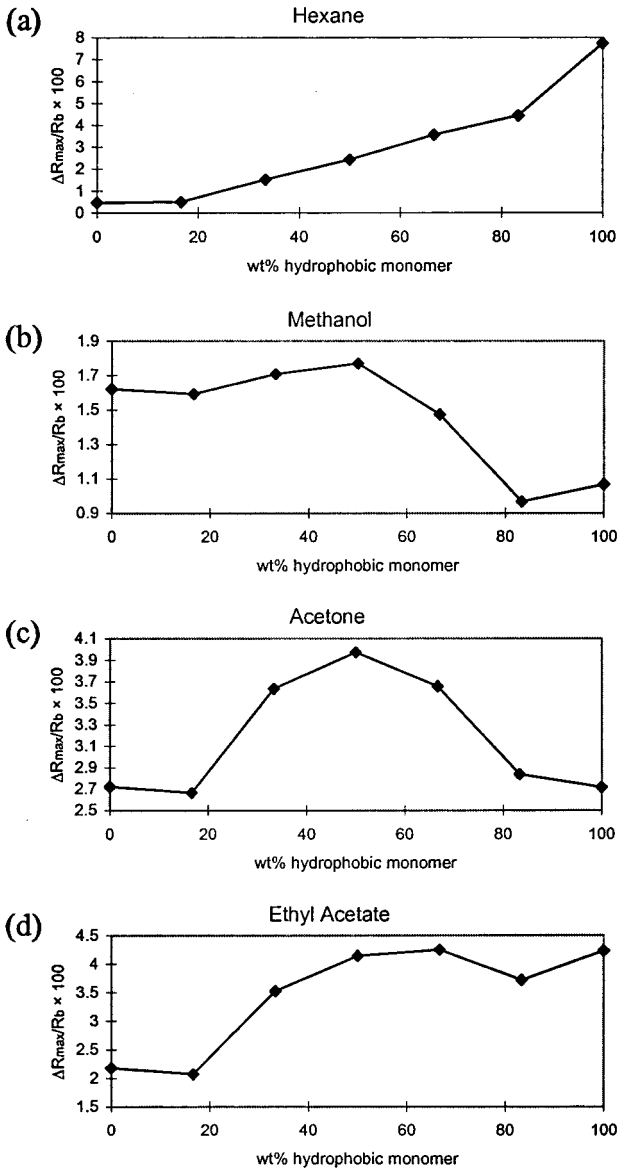


Figure 3. Response data for block copolymers as a function of polymer composition. The blocks were formed from living polymerization of a polar monomer (left) and a nonpolar (silyl-containing) monomer, as displayed in Scheme 1. Each analyte was present at 5% of its vapor pressure for a period of 300 s, with the background gas being laboratory air. The values on the ordinate are the median maximum relative differential resistance responses, $\Delta R_{max}/R_b$, of each detector during the analyte exposure period. Data were obtained from 15 exposures to each analyte, with every individual analyte exposure performed in random order.

polymer from a different solution to each pixel. The electronics that form the basis for such a "nose-chip", having thousands of individually addressable pixels, are readily fabricated using conventional CMOS technology, and the behavior of such combinatorially-formed "nose-chips" is under investigation.

Conclusions

We have demonstrated a simple technique for altering the response properties of polymer-based chemiresistor vapor detectors through use of plasticizers. The approach should be applicable to a variety of implementations of electronic noses. In addition, a series of well-defined block copolymers was found to have sorption behavior that was not a linear combination of the properties of the two separate blocks. Taken together, these methods suggest approaches for forming large numbers of compositionally different detectors from a limited number of feedstock solutions.

Acknowledgment. This work was supported by DARPA, the Army Research Office, the Department of Energy, and NASA.

Supporting Information Available. Experimental details, resolution factors utilizing arrays of different numbers of detectors, a full description of resolution factors, and a representative table of $\Delta R_{max}/R_b$ values are available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Trends Biotechnol.* **1998**, *16*, 250−258.
(2) Bartlett, P. N.; Elliott, J. M.; Gardner, J. W. *Ann. Chim.* **1997**, *87*, 33−44.
(3) Schaller, E.; Bosset, J. O.; Escher, F. *Food Sci. Technol.-Lebensm.-Wiss. Technol.* **1998**, *31*, 305−316.
(4) Harsányi, G. *Polymer films in sensor applications: technology, materials, devices and their characteristics*; Technomic Pub. Co.: Lancaster, 1995.
(5) Chang, S. M.; Iwasaki, Y.; Suzuki, M.; Tamiya, E.; Karube, I.; Muramatsu, H. *Anal. Chim. Acta* **1991**, *249*, 323−329.
(6) Grate, J. W.; Martin, S. J.; White, R. M. *Anal. Chem.* **1993**, *65*, A987−A996.
(7) Grate, J. W.; Martin, S. J.; White, R. M. *Anal. Chem.* **1993**, *65*, A940−A948.
(8) Dickinson, T. A.; White, J.; Kauer, J. S.; Walt, D. R. *Nature* **1996**, *382*, 697−700.
(9) White, J.; Kauer, J. S.; Dickinson, T. A.; Walt, D. R. *Anal. Chem.* **1996**, *68*, 2191−2202.
(10) Krech, J. H.; RosePehrsson, S. L. *Anal. Chim. Acta* **1997**, *341*, 53−62.
(11) Persaud, K. C.; Khaffaf, S. M.; Payne, J. S.; Pisanelli, A. M.; Lee, D. H.; Byun, H. G. *Sens. Actuator B−Chem.* **1996**, *36*, 267−273.
(12) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298−2312.
(13) Malnic, B.; Hirono, J.; Sato, T.; Buck, L. B. *Cell* **1999**, *96*, 713−723.
(14) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177−4190.
(15) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 658−668.
(16) Plasticizers have been employed in polymeric electrolytes, and were found to have only a secondary effect on selectivity: Buhlmann, K.; Schlatt, B.; Cammann, K.; Shulga, A. *Sens. Actuator B−Chem.* **1998**, *49*, 156−165.
(17) A combinatorial library of dye impregnated vapor detectors has been fabricated by addition polymerization: Dickinson, T. A.; Walt, D. R.; White, J.; Kauer, J. S. *Anal. Chem.* **1997**, *69*, 3413−3418.
(18) Doleman, B. J.; Sanner, R. D.; Severin, E. J.; Grubbs, R. H.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 2560−2564.
(19) Weck, M.; Schwab, P.; Grubbs, R. H. *Macromolecules* **1996**, *29*, 1789−1793.

CC990056T

Detecting Hidden Targets: A Procedure for Studying Performance in a Mine-Detection-like Task

Daniel T. Cerutti*[*], Ioan M. Chelaru, and John E. R. Staddon
DukeUniversity, Psychology: Experimental, Durham, NC, 27708

ABSTRACT

We report preliminary results from an experiment designed to study the perceptual and learning processes involved in the detection of land mines. Subjects attempted to identify the location of spatially distributed targets identified by a sweeping a cursor across a computer screen. Each point on the screen was associated with a certain tone intensity; targets were louder than "distractor" objects. We looked at the effects on target detection and false-alarm rates of the intensity difference between target and distractor signals, the number of distractors and training order. The time to detect 50% of targets (threshold detection time) was measured by a rapid adaptive technique (PEST) which generated reliable thresholds within few trials. The results are consistent with a simple model for the detection of cryptic prey by foraging predators: search was slower with more distractors, and the effect of distractors was greater when intensity ratio (IR) was lower. Although subjects got no accuracy feedback, performance improved somewhat with experience and was better in the low-IR condition when it followed the high-IR condition. The procedure seems to be a useful one for studying more complex mine-related detection tasks with a range of signal types and numbers of concurrent detection signals.

Keywords: land mine detection, simulation, threshold, adaptive staircase, training, foraging, cryptic

1. INTRODUCTION

Land mines are effective barriers to the extent that they can be concealed. Development of sensitive metal-detection equipment reduces concealment, but has led to the evolution of low-metal or no-metal mines, which pose new technical problems. Mine detection will always be subject to technical limitations. Mines must usually be detected by human operators using hand-held devices. Improving the performance of the human operator in the man-machine system is therefore just as important as perfecting the technology of detection.

Testing system performance in a naturalistic setting – real mine lanes, real mines – is expensive and difficult. The human operator soon learns the characteristics of the setup, at which point data become less useful. Changing the setup is expensive and time consuming. Two ways to deal with this problem are (a) to develop a "virtual minefield," a computer-based system that closely matches the natural one in terms of the tactile, visual and auditory information presented to the operator. Such a system would allow the effect of technical and training modifications on operator efficiency to be tested directly. And (b) to learn more about the basic psychological processes that underlie operator performance in mine-detection tasks. Option (b) is less direct and less certain to succeed than option (a), but may suggest specific technical changes and may have long-run benefits in adding to basic science. We now report preliminary results from a project aimed at understanding the psychological processes involved in mine-detection-type tasks.

An Analogy from Behavioral Ecology

Most laboratory studies of detection involve trial-by-trial procedures and allow the test subject little freedom of action. Their focus is on the sensory limitations of the subject not on the way the subject interacts with his environment. But mine detection is more like the problem faced by a predator attempting to find a concealed or cryptic (camouflaged) prey than a sensory-discrimination task. There is a substantial amount of research on the detection of cryptic prey[4,5,7,8,13]. Relevant work has looked at how rate of target (prey) detection depends on the crypticity of the target.

For example, Gendron[2] studied in a semi-natural environment how quail forage for food objects that matched (high crypticity) or did not match (low crypticity) their background. He found (a) that the animals rarely made a mistake (false alarm rate was close to zero); (b) they search faster for the less cryptic food. Based on these results, Gendron and Staddon[3] proposed a descriptive model for how detection of cryptic prey depends on crypticity and search rate. Figure 1 shows the suggested relation. Probability of detection remains high with conspicuous (low crypticity) targets at all but the highest

[*] Correspondence: Email: cerutti@psych.duke.edu; Telephone 919 660 5707; Fax 919 660 5726

search rates. With less conspicuous (cryptic) targets, probability of detection decreases rapidly as search rate increases. These relationships can be summarized in the following equation:

$$P^K_d + (S/M)^K = 1$$

or,

$$P_d = [1 - (S/M)^K]^{1/K}, \qquad K > 0. \tag{1}$$

where S is search rate, M is the maximum search rate, K is a measure of crypticity, or how well the target is masked by extraneous noise, and P_d is probability of detection. The model predicts that the probability of detecting a target, P_d is a joint function of both target crypticity and search rate, S. Moreover, the optimal search rate (at which rate of target detection is maximized) decreases as targets become more cryptic.

Figure 1 illustrates a further assumption of the model, that all curves intersect the abscissa at the maximum search rate (M). This represents an upper limit on the rate of search imposed by the observer's perceptual or physiological abilities or by the detection equipment. In more complex detection tasks, with two target types that differ in crypticity, the optimal search rate will also reflect their relative densities[3].

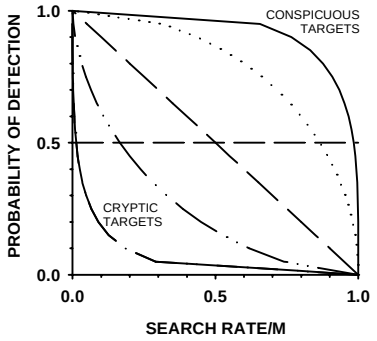


Figure 1 The probability of detecting a target as a function of search rate. Curves represent targets differing in their degree of crypticity. The dashed horizontal line at P_d = 0.5 indicates threshold detection. Search rate is plotted as a proportion of maximum rate.

Gendron and Staddon[4] did an experiment to test this model with human subjects. Subjects searched for computer-generated visual targets. The targets were characters embedded in a background of nontarget symbols ("distractors" or "clutter"). Crypticity was manipulated by varying the number of distractor characters and search rate was manipulated by varying the display duration. The main results were (a) that probability of detection, P_d, decreased as search rate increased; (b) the decrease in detection probability was more rapid for more cryptic targets,; and (c) that P_d tended to converge at one point on the abscissa, the maximum search rate[1]. All these results are in agreement with model predictions, which encouraged us to use the model as a starting point for the present study.

Our desktop-based simulation required subjects to scan a patch of simulated terrain with a computer mouse. The terrain contained both "mines" (targets) and distractor objects, with targets producing louder auditory signals. We manipulated crypticity in two ways, by changing the relative loudness of distractor signals and by increasing the density of distractors. Each subject was tested with three densities of distractors at two signal values (a total of six conditions). This permitted us to study the effect of experience in two ways, by presenting difficult detection trials before easier trials or vise versa, and by repeating the first condition after the first six conditions.

A major consideration in research and training with humans is to collect reliable data with a minimal investment of time and physical resources. The forced-choice procedures used in previous experiments on prey detection require many trials at each level of an independent variable. We therefore used an *adaptive staircase* procedure, specifically developed for rapidly determining stimulus thresholds (P_d = .5) in vision and audition, to study target detection (see review by Treutwein[12]). This procedure increases or decreases the relevant stimulus parameter trial-by-trial, depending on whether a subject made an incorrect or correct detection on the preceding trial. The increments and decrements become progressively smaller as the threshold is located, rapidly converging on a threshold. In our simulation, we used a staircase to adjust the time available to search terrain until P_d = 0.5. We could then generate psychometric functions relating the effects of signal strength, for example, to the search time a subject required to find 50% of the targets.

2. METHODS

Subjects were six undergraduate college students, five female and one male. They were paid a flat hourly wage for serving. Subjects were seated at a computer and given the following instructions to read while the experimenter read them aloud:

> *This study is designed to learn about how people find mines. We are using a computer simulation of a minefield and a metal detector to simulate real problems in mine finding.*
>
> *In a real minefield, a person searching for buried mines uses a mine detector to locate mines and mark their positions. The marks are used by another person who disarms them. A metal-detector indicates a buried mine by making a tone when it is over a mine. Unfortunately, metal detectors also indicate other objects like small metal fragments, and the operator must decide whether a tone indicates a mine or a harmless object.*
>
> *Your task involves using the computer mouse to (1) scan small patches of a simulated minefield and (2) to mark the locations of mines. The computer's speakers will make a tone when the mouse pointer is over an object. You must decide if the object is a mine and mark it by clicking the left mouse button. You must only mark mines.*

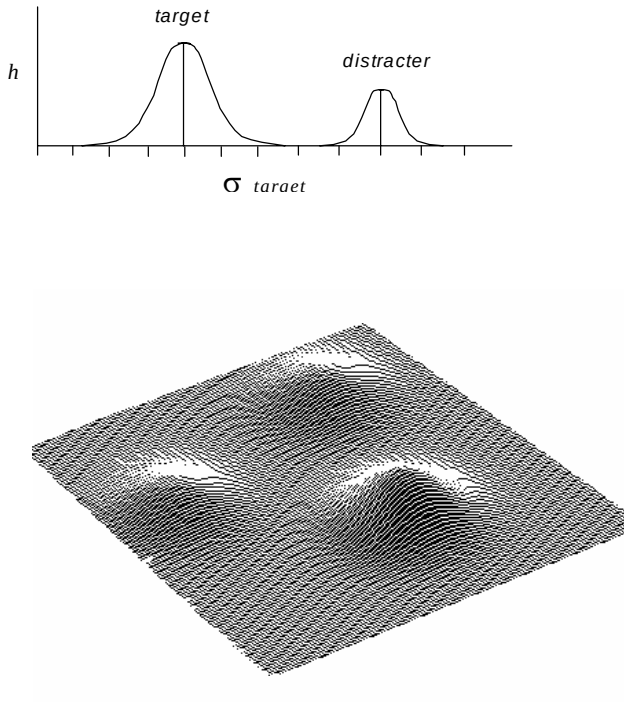


Figure 2. The upper panel shows a two-dimensional representation of target and distractor signal strength. The contour plot is an isometric representation of the spatial distribution of signal strength (z-axis) for a target (lower right) and two distractors.

Subjects were then shown a 125 mm square patch of simulated terrain (220 pixels wide by 200 high), a light brown marbleized pattern, on the computer monitor. The terrain contained two 2 mm dots, one green and one black. They were told that the black dot represented a mine (target) and the green dot represented a non-mine object (distractor). Then they were asked to sweep over the dots with the mouse pointer and listen to the tone produced by the speaker. The signal frequency was constant at 500 Hz, The volume of the signal generated by target and distractor varied according to a Gaussian spatial distribution. Figure 2 is a visual representation of a target signal (lower right) and two distractor objects, where signal volume is a function of the height of the distribution. The function used to determine signal strength for a target was the following:

$$h_m = \exp[-[(x-m_x)^2 + (y-m_y)^2]/2\sigma^2_m, \qquad (2)$$

where h_m is the magnitude of the target signal, m_x and m_y are the target's position on the terrain, x and y are the mouse-pointer position on the terrain, and σ_m was 6 pixels. The signal strength for the distractor objects, h_o, was determined in the same way except that h_o was a proportion of h_m and $\sigma_o=\sigma_m/2$. The value of h_o was $0.5h_m$ in the low target/distractor signal conditions and h_o was $0.25h_m$ in the high target/distractor signal conditions.

During this initial training, the black dot (mine) generated an 83 db tone at 100 mm from the speaker at the peak; the green dot (non-mine) generated a 65 db tone at 100 mm from the speaker at the peak. They were then presented with a patch of terrain without dots and they were told that the mine and the object were now hidden and their task was to use the mouse

to locate the target. When they located the target, they were instructed to point to it with the mouse and to click the left mouse button to mark target position. Once they put down a mark, the position of the target was revealed. Their mark appeared blue if it was within the target "halo" of approximately 10 pixels (6 mm) from the target; their mark appeared red if it was outside the halo. They were given these training-with-feedback trials until they completed ten correct detections in a row.

Subjects began experimental conditions immediately after meeting the training criterion. Trials in experimental conditions were similar to training trials except that there was no feedback about correct detections and there was a "progress bar" to the right of the terrain. The progress bar indicated the time remaining in a trial. Trial duration varied between 5 and 30 seconds and was adjusted between trials with a maximum likelihood procedure known as the "best PEST" (Parameter Estimation by Sequential Testing[6]). The PEST procedure arranged very long and very short trials at the outset of a condition (i.e., 30-s and 5-s), and then used correct detections and detection failures over successive trials to estimate the threshold display duration (correct detections and detection failures were determined as described for training). For example, assuming a subject detected the target at 30-s but not at 5-s, the following trials might be 20-s and 10-s. If the subject succeeded with first but not the second, the PEST would narrow the range of values further, and so forth. Conditions were terminated after 40 trials, considered adequate for reliable threshold estimation with the best PEST procedure[12].

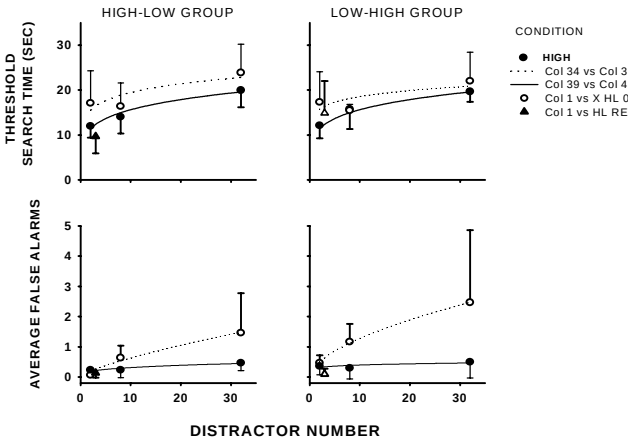


Figure 3. Upper plots show mean threshold search times for the HIGH-LOW and LOW-HIGH groups. Lower plots show average false alarms. The HIGH conditions, indicated by filled points, contained distractors producing 65 db signals; the LOW conditions, indicated by open points, contained distractors producing 81 db signals. Triangles are replications, shifted slightly on the x-axis to reduce crowding. Vertical lines are standard deviations. Points are fit with power functions.

Each subject was exposed to seven conditions, each separated by a five minute rest period or longer at the subject's request. One group of three subjects (HIGH-LOW Group) first experienced three successive conditions in which the target signal was 83 db and the distractor signal was 65 db (as in training). The first of these conditions contained 2 distractors, the second contained 8 distractors, and the third contained 32 distractors. These were followed by three identical conditions but in which the distractor signal was 81 db. The second group of three subjects (LOW-HIGH Group), completed the sequence of three conditions with the 81 db distractor signal followed by the three conditions with the 65 db distractor. In the seventh and final condition, each subject repeated his or her first condition.

3. RESULTS

Individual subject's performances are summarized in the Appendix. Group performances are displayed in Figure 3. The upper plots show threshold search times as a function of the intensity ratio and distractor number. As expected, threshold time was consistently greater for the LOW (more cryptic, lower intensity ratio) condition. Moreover, the effect of the intensity ratio was more or less constant and independent of the number of distractors: the HIGH and LOW curves are all parallel. Although the number of data points is too small to be sure, the function relating threshold search time to distractor number appears to be negatively accelerated.

The lower plots in Figure 3 show the mean false alarms per trial (FAR) as a function of the intensity ratio and distractor number. In these plots, FAR included all marks outside of a target halo. The FAR for the HIGH conditions remained low at all distractor values, while the FAR for the LOW condition increased with distractors in a nearly linear function.

The data in Figure 3 are replotted in Figure 4 to illustrate the effect of testing order, LOW-HIGH vs. HIGH-LOW, on search time. Not shown in Figure 4 are data from the last, replication condition. There was an apparent order effect: performance on the LOW condition was slightly better if it followed the HIGH condition rather than preceded it. Another order effect is seen in the right plot of Figure 4: threshold search times for the two-distractor conditions are higher than anticipated by the function for both groups of subjects. The likely explanation is that this was the first LOW condition experienced by all subjects. Finally, Figure 5 compares the threshold search time in the first condition with the final replication of that condition. Five of the six subjects showed lower thresholds in the replication, but as seen in Figure 5, the reduction was modest.

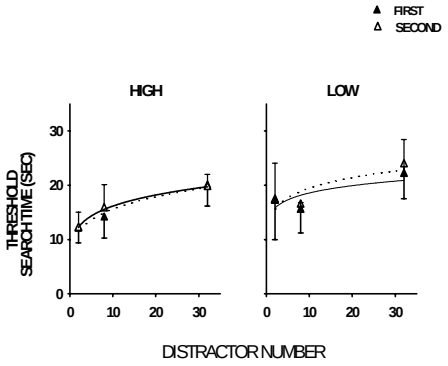


Figure 4. Performances of the HIGH (65 db distractors) and LOW (81 db distractors) conditions plotted together. Open points and dotted lines show results of the first three conditions, filled points and solid lines show results of the next three conditions. Plots are fitted

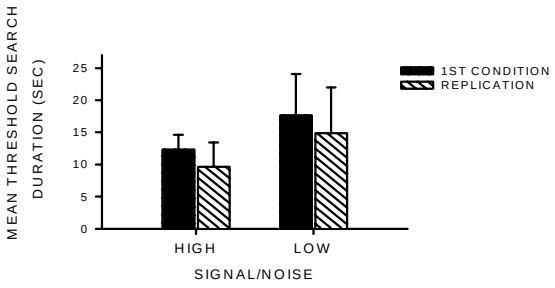


Figure 5. Comparison of results from subjects' first condition and the final replication of the condition.

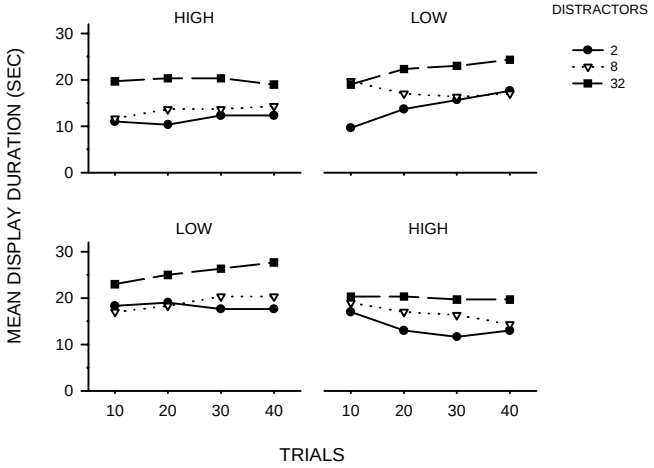


Figure 6. Display durations at ten trial intervals. The HIGH-LOW group results are shown in the upper plots and the LOW-HIGH group results are shown in the lower plots.

Figure 6 shows that the PEST procedure was almost always able to establish thresholds well within the 40 trials of a condition. On the x-axis is plotted display duration at 10-trial intervals. Successful determination of a threshold is indicated by an increasingly horizontal trend in display duration. For example, the top-row graphs show data from the High-Low Group. The solid squares show display duration with 32 distractors. In the left plot, display duration is about 20-s at trial 10 and remains close 20-s at 20, 30, and 40 trials. This indicates that the threshold was determined by the 10th trial. In the right plot, however, display duration is again at 20-s at trial 10, but climbs slightly, suggesting that the threshold may have been slightly underestimated. Visual inspection of Figure 5 suggests that the PEST produced reliable results with the possible exception of the LOW conditions, here target and distractor signals were most similar. These conditions might have benefited from additional trials.

4. DISCUSSION

The solid line in Figure 7 shows how the PEST threshold measurement is related to the search-rate measure predicted by the Gendron-Staddon[3] model. Setting P_d=0.5 and rearranging, the relation between search rate, S, and the crypticity measure K is:

$$S=100^{((LN(-EXP(LN(0.5)*K) + 1))/K)} \quad (3)$$

The search-crypticity function in Figure 7 shows that threshold search rate remains high with non-cryptic prey and then suddenly falls off as crypticity increases. The problem, of course, is that we don't know how the crypticity measure, K, is related to actual independent variables in our study, such as the number of distractors. We have provided a rough fit to our findings by plotting search rate in each condition as a proportion of maximum search rate. And, by assuming that search rate falls close to zero with a large number of distractors, we plot crypticity as a proportion of maximum distractors (distractors divided by 100). The model provides a rough fit to the data without making any provision for effects of intensity ratio. The obtained search rate appears to fall off more quickly than predicted, the number of "crypticity" values we used tried is too few to be sure.

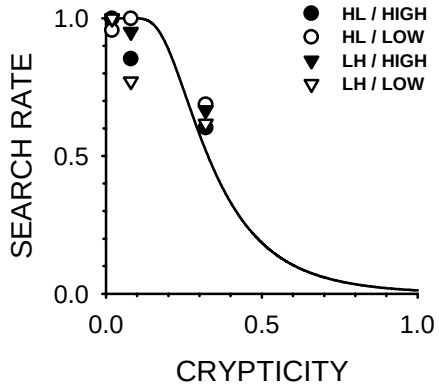


Figure 7. Search-crypticity function and mean thresholds from the present study. The solid line shows threshold search rate as a function of crypticity predicted by the Gendron-Staddon model. Data points are from conditions described in the legend. Other details in text.

The small effect of the intensity ratio on threshold search time was not anticipated. We chose the IR parameters in the experiment after we found values that produced large differences in the time it took to find mines. We did not anticipate that the main effect of decreasing the intensity ratio would be to increase FAR, although in retrospect the finding is obvious. This is a major departure from previous findings on prey detection by animals and humans, where FAR stays very low irrespective of the intensity ratio. But in fact, our simulation data closely resemble the data collected in field tests. Those data show fewer false alarms with the metal detector than with the ground-penetrating radar (GPR). The reason is that the GPR gives the same auditory signal to all buried objects, as in our low signal/noise condition.

We were also surprised by the marginal effect of training history. Many studies show that training subjects on progressively more difficult discriminations, a procedure called *fading*, results in more rapid learning and greater accuracy (e.g., Sidman & Stoddard, 1967; Terrace, 1963). Our HIGH-LOW group was exposed to a fading regimen but showed very little benefit. A likely reason is that fading requires feedback. Feedback in our study was impoverished in that our subjects had auditory feedback from the computer, but no indication of whether their marks were on target. Training is such a critical issue in performance that this question merits a more systematic investigation.

The present experiment is a first step toward the development of a valid model of the behavior involved in the detection of hidden targets. Little is known about the variables that must be included in a more complete model and how they contribute to the final performance. The ultimate value of a model will be to predict performance under novel conditions and to suggest ways to make the feedback from sensors most useful to the human operator.

5. ACKNOWLEDGEMENTS

We wish to thank Rosa Maria Cerutti for assistance in running the experiment. This work was supported by U.S.Army Research Office, Grant No. DAAH04-96-1-0448 (Multidisciplinary University Research Initiative to Demining).

6. REFERENCES

1. Blough, D. S. 1979. Effects of the number and form of stimuli on visual search in the pigeon. J. Exp. Psychol. Anim. Behav. Proc. 5:211-223.
2. Gendron, R. P. (1982). The foraging behavior of bobwhite quail searching for cryptic prey. Ph.D. diss. Duke University.
3. Gendron, R. P., and J. E. R. Staddon. 1983. Searching for cryptic prey: the effect of search rate. Am. Nat. 121:172-186.
4. Gendron, R. P., and J. E. R. Staddon. 1984. A laboratory simulation of foraging behavior: The effect of search rate on the probability of detecting prey. Am. Nat. 124: 407-415.
5. Guilford, T., & Dawkins, M. S. (1989). Search image versus search rate: Two different ways to enhance prey capture. Animal Behaviour, 37, 163-165.
6. Pentland, A. (1980). Maximum likelihood estimation: The best PEST. Perception and Psychophysics, 28, 377-379.
7. Plaisted, K. C., & Mackintosh, N. J. (1995). Visual search for cryptic stimuli in pigeons: Implications for the search image and search rate hypotheses. Animal-Behaviour, 50, 1219-1232.
8. Reid, P. J., & Shettleworth, S. J. (1992). Detection of cryptic prey: Search image or search rate? Journal-of-Experimental-Psychology:-Animal-Behavior-Processes, 18, 273-286.
9. Sidman, M., & Stoddard, L. T. (1967. The effectiveness of fading in programming a simultaneous form discrimination for retarded children. Journal of the Experimental Analysis of Behavior, 10, 3-15.
10. Staddon, J. E. R., & Gendron, R. P. (1983). Optimal detection of cryptic prey may lead to predator switching. *American Naturalist, 122*, 843-848.
11. Terrace, H. (1963). Discrimination learning with and without "errors." Journal of the Experimental Analysis of Behavior, 10, 3-15.
12. Treutwein, B. (1995) Adaptive psychophysical procedures: Minireview. *Vision Res.*, **35**, 2503–2522.
13. Tye, A. (1989). A model of search behaviour for the northern wheatear Oenanthe oenanthe (Aves, Turdidae) and other pause-travel predators. Ethology, 83, 1-18.

7. APPENDIX

Subjects' mean search threshold and false-alarm rate for the last ten trials of each condition.

		Search time(sec)		False-alarm rate	
Subj.	Distractors	High Signal	Low Signal	High Signal	Low Signal
		High-to-low group			
HL1	2	10.8	13.0	0.2	0.0
	8	10.8	19.8	0.0	0.8
	32	15.6	16.8	0.2	0.7
	2	10.4	-----	0.1	
HL2	2	10.2	25.4	0.3	0.0
	8	13.2	10.4	0.2	0.6
	32	22.2	25.8	0.4	1.0
	2	5.6	-----	0.3	
HL3	2	15.0	13.0	0.3	0.2
	8	18.2	19.0	0.6	0.9
	32	22.0	29.0	0.3	1.0
	2	13.0	-----	0.4	
		Low-to-high group			
LH1	2	25.0	9.2	0.2	0.7
	8	17.0	12.8	0.0	0.8
	32	27.4	21.0	0.3	0.5
	2	-----	22.8		0.0
LH2	2	14.6	15.0	0.7	0.5
	8	15.0	20.8	0.6	0.9
	32	23.8.	17.0	0.6	1.0
	2	-----	9.0		0.3
LH3	2	12.4	12.2	0.2	0.2
	8	29.0	13.6	0.2	0.4
	32	29.0	21.0	0.1	0.9
	2	-----	12.8		0.0

ELSEVIER

Journal of Applied Geophysics 1319 (2001) xxx

APPLIED GEOPHYSICS

www.elsevier.nl/locate/jappgeo

A tapered-permittivity rod antenna for ground penetrating radar applications

Chi-Chih Chen *, Kishore Rama Rao, Robert Lee

ElectroScience Laboratory, The Ohio State University, Columbus, OH 43212, USA

Abstract

A new dielectric rod antenna design modified from its previous version developed by Chen [IEEE Trans. Geosci. Remote Sens. (accepted for publication)] is presented. Such an antenna is useful in detecting small and shallow subsurface objects with excellent depth and spatial resolutions. These features make it useful in detecting small anti-personnel (AP) mines, pavement cracks and the surface layer. Broad bandwidth electromagnetic energy is fed into the rod from one end, guided along the rod and then radiated from the other end where the rod diameter is linearly tapered to a point. The tapered-permittivity design uses an additional permittivity taper to overcome the problem of a frequency-dependent pattern when high dielectric material is used for size reduction. This new design reduces the internal reflections at the tapered section and results in better efficiency and less antenna clutter. © 2001 Published by Elsevier Science B.V.

Keywords: Antenna; Ground penetration radar; Landmines; Modeling

1. Introduction

Ground penetrating radars (GPR's) have been used as non-intrusive instruments for subsurface investigation (Noon et al., 2000; Plumb, 1998; Sato, 1996; Redman, 1994). For deep applications, a GPR is usually operated in the MHz frequency range to achieve a greater penetration but with a poorer depth resolution. For shallow applications, such as the detection of buried landmines, the 1–6 GHz frequency range gives a reasonable trade-off between penetration and depth resolution. For all GPR systems, the antenna design plays a vital role in measurement performance. Broad bandwidth, good efficiency, low antenna clutter, and weak antenna–ground interaction are the major requirements of a good GPR antenna. Note that antenna clutter could be broadband such as feed point reflection and direct coupling, or narrowband such as ringing. Most deep-application GPR's place their antennas close to the ground surface for coupling more energy into ground, and less radiation above the ground and less surface scattering. A major drawback of this configuration is the strong antenna–ground interaction that can significantly change the characteristics of the antenna. This has been shown in the literature (Procter, 1950; Abul-Kassem, 1972; Hayes, 1983). Such an interaction produces noticeable antenna clutter and limits the antenna to be applicable to only a few ground conditions. In a well-controlled environment, the effects of the interaction can be removed via the use of a known buried calibration target; however, in reality, it is not practical to dig a hole for the placement of a calibration target since the disturbance of the soil usually changes the electrical char-

* Corresponding author. Fax: +1-614-2923403.
E-mail address: Chen.118@osu.edu (C.-C. Chen).

0926-9851/01/$ - see front matter © 2001 Published by Elsevier Science B.V.
PII: S0926-9851(01)00074-X

UNCORRECTED PROOF

acteristics of the soil as well. In addition, the performance of such a GPR system would depend heavily on the surface and soil conditions, which often vary from time to time and from place to place.

It is usually desirable to elevate the GPR antenna off the ground for shallow applications to achieve faster survey speed and less antenna–target and antenna–ground interactions. Faster speed is gained from a larger radiation beam coverage and freedom of sensor motion. However, two major problems are the ground surface scattering that reduces both radar efficiency, and sensitivity. The surface roughness creates undesirable clutter that arrives at the radar over a wide time range due to scattering from different patches of surface within the illumination spot. This clutter also masks out the desired responses from shallow targets. This situation becomes worse as the antenna height above the ground increases. One way to reduce the reflection from the ground surface is to illuminate the ground at an oblique angle when the surface is not too rough. Unfortunately, this remedy increases the illumination area and causes further spreading of the surface clutter and makes it even more difficult to detect a buried object. A focused-beam illumination can be used to achieve a greater antenna-to-ground distance while minimizing the illumination spot on the ground (Chen et al., 2000b). The minimum achievable spot size is limited to approximately one wavelength associated with the lowest operation frequency. However, this minimum spot size can only be achieved when the reflector size is several wavelengths. Although a stronger surface reflection was received from the vertical illumination, it was shown that such a reflection can be easily separated from the desired landmine responses because of the broad bandwidth and minimal ringing. Some disadvantages associated with a focused reflector system include frequency-dependent spot size, bulky reflector, high wind resistance, visual blockage and difficult calibration. These disadvantages make a focused-beam design less desirable than the new dielectric rod antenna design for field operations.

An ultra-wide bandwidth dielectric rod antenna design was introduced for improved performance in detecting anti-personnel (AP) landmines (Chen et al., accepted for publication). Chen's design is similar to previous narrowband "ployrod antennas" (Mueller and Tyrrell, 1947; Walter, 1965) but has much greater bandwidth due to the introduction of broadband excitation structure. A dielectric rod with a circular or square cross-section was used to guide broadband electromagnetic energy. The HE_{11} hybrid mode is first excited from one end of the rod using a special broadband feed structure. The electromagnetic energy then propagates along the rod with most of its energy confined within the rod when the cross-section dimension is greater then the longest operational wavelength in the material. The phase velocity is nearly constant and solely determined by the internal material. As the diameter becomes less than a wavelength, more energy begins to propagate external to the rod and causes the phase velocity to be a function of both the internal and external materials. At the other end of the rod, radiation was introduced by gradually tapering the rod dimension to a point. This approach works well for a rod made of low dielectric material. In order to further reduce the rod size, material with high permittivity must be used. This would require a long taper section such that the end reflection occurring at the material–air interface is kept small. However, a longer taper section results in undesirable spreading of the radiation center. That is, the lower frequency energy is radiated from the beginning of the taper and the higher frequency energy is radiated close to the tip. Such a spreading causes the radiation pattern to be frequency-dependent and a careful calibration is required to remove such a dependency.

Section 2 discusses modified dielectric rod antenna design with a permittivity taper in addition to the diameter taper. The field characteristics of the new rod will be presented in Section 3 using a three-dimensional (3D) numerical model based on finite difference time domain (FDTD) technique (Yee, 1966). Final conclusions will be given in Section 4.

2. Tapered-permittivity dielectric rod antenna

The ultra-wide bandwidth dielectric rod antenna developed by Chen et al. (accepted for publication) utilizes a circular dielectric cylinder made of homogeneous material to guide the electromagnetic energy

UNCORRECTED PROOF

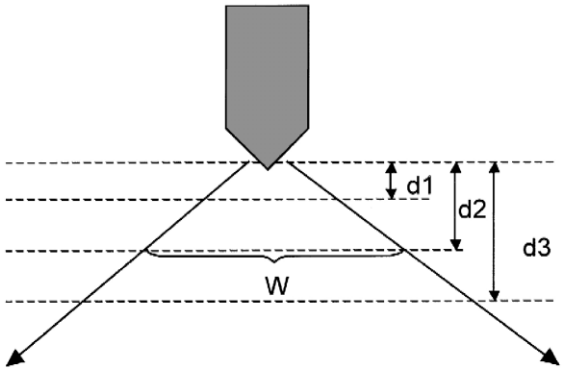


Fig. 1. Illumination spot size as a function of tip height above the ground.

using the fundamental hybrid mode. The electromagnetic radiation was then generated by removing the guiding structure in the end taper section where the diameter was gradually reduced to zero to minimize internal reflections occurring at the material–air interface. The design of this tapered section is very important. It serves as both a smooth terminator as well as an efficient broadband radiator. If the tapering is done too rapidly, large internal reflections would occur due to impedance mismatch. This increases the antenna clutter as well as reduces the antenna efficiency. The radiated fields propagate mainly in the forward direction with spherical phase fronts. The illumination spot size on the ground can

PROOF

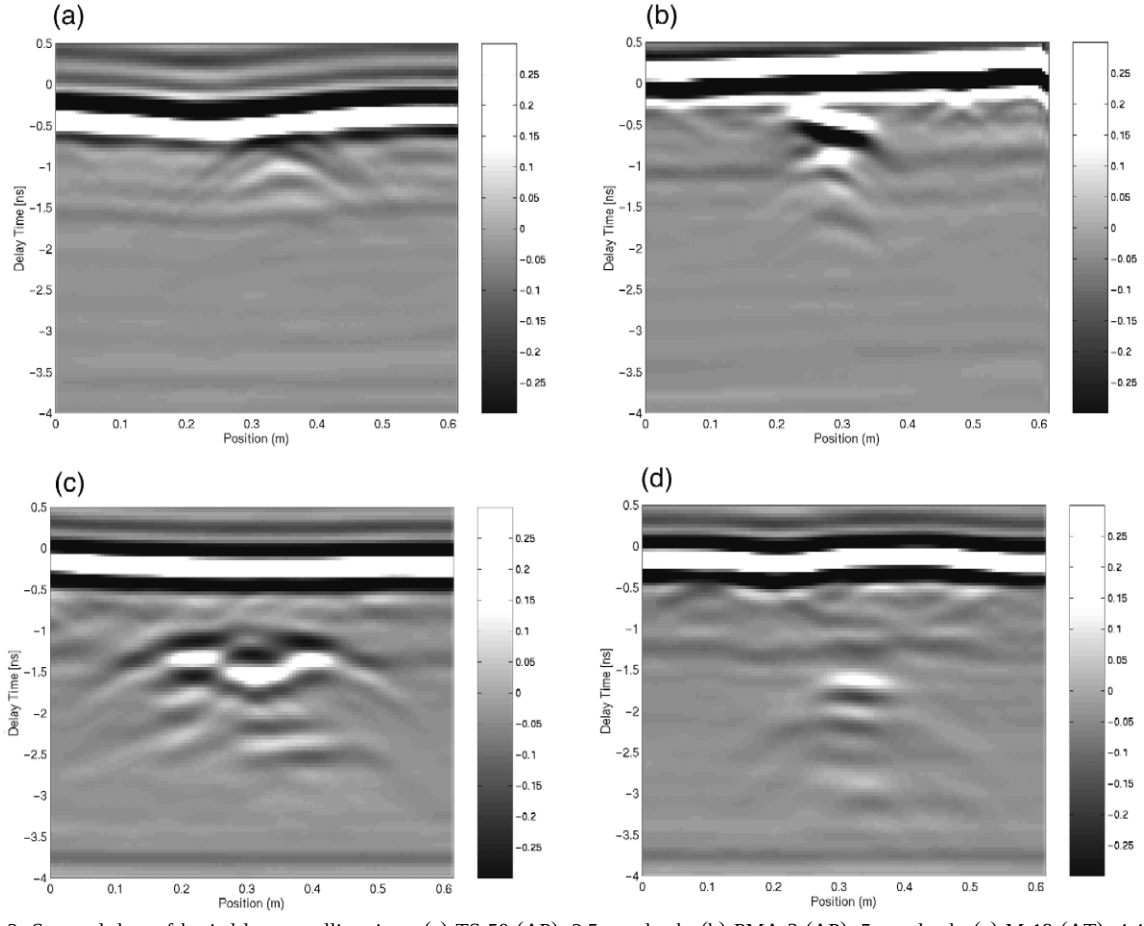


Fig. 2. Scanned data of buried low-metallic mines: (a) TS-50 (AP), 2.5-cm depth; (b) PMA-3 (AP), 5-cm depth; (c) M-19 (AT), 4.4-cm depth; (d) VS-2.2 (AT), 8.3-cm depth.

be controlled by the antenna height as illustrated in Fig. 1. By placing the antenna tip close to the ground surface, a small illumination spot can be obtained to reduce the undesired clutter due to rough surface scattering, as well as to improve the spatial resolution that is important in detecting a small target.

The rod antenna design has been applied to the detection of buried, small, non-metallic anti-personnel landmines with good success. Some measurement data obtained from a government landmine test site, Fort A.P. Hill, are shown in Fig. 2 for various mine types and depths. The soil in the site contains a mixture of sand, clay and small gravels. The relative permittivity and conductivity at the site was measured to be approximately 8 and 0.01 S/m, respectively. Each figure plots the time-domain radar responses converted from the original swept-frequency data (2–6 GHz) as the rod antenna moved across the ground. The delay time of the received signal and the antenna position are indicated by the vertical and horizontal scales, respectively. The measured field amplitude is indicated by grayscale where the positive and negative amplitudes are represented by light and dark colors, respectively. These are unprocessed data. The strong surface reflections are clearly seen at earlier time (top of the figure). This surface reflection can be significantly reduced, if desired, using processing techniques (Salvati et al., 1998). It is also noticed that different mine types show different spatial–temporal features that can be used for further target identification. These results demonstrate the fine spatial and time resolution achieved by the low antenna clutter and broadband characteristics of the new dielectric rod antenna.

In the previous design (Chen et al., accepted for publication), the rod was made of a homogeneous material that has a constant permittivity. A rod made of higher permittivity is highly desirable since it can be small, light and have finer spatial resolution. However, a high-permittivity rod requires a longer taper section in order to keep the internal reflection to a minimum. A longer taper also results in undesired dispersion and frequency-dependent radiation patterns similar to that found in narrowband "polyrod antennas" in the early days (Mueller and Tyrrell, 1947; Watson and Horton, 1948). This is because the lower frequency components radiate earlier in the taper region and the higher frequency components radiate near the end. Such a frequency-dependent behavior requires additional complicated procedures to calibrate the measured data.

The new tapered-permittivity rod antenna design introduces an additional permittivity taper to the tapered section. Notice that the previous design attempted to vary wave impedance by gradually reducing the rod dimension to match that of free space. This is similar to the problem of matching two transmission lines with different characteristic impedance investigated by Collin (1956) and many other people. The new design attempts to vary the wave impedance by varying both the diameter and the permittivity simultaneously.

3. Numerical model of the tapered-permittivity rod antenna

A three-dimensional numerical model was setup using the FDTD technique to study the field characteristics of the new design. Fig. 3 shows the sideview of the tip portion of the 3D model. The model was enclosed by a perfectly matched layer (PML) (Zhao and Cangellaris, 1996) to avoid artificial reflections caused by space truncation. The diameter and relative permittivity of the rod are 7.62 cm and 3, respectively. Different permittivity layers are shown by different grayscales in the tip section. The actual profile of the relative permittivity is also plotted in Fig. 4. The taper section was made of seven layers with relative permittivity varying from 3 to 1.15 at the tip. Each layer has the same thickness that was chosen to be approximately a quarter of the waveguide wavelength at the center frequency. The diameter of the rod in the taper section decreases linearly to nearly a point at the tip.

Two snap shots of the calculated electromagnetic fields propagating down the rod are plotted in Fig. 5. The amplitude of the electric fields is indicated by the grayscale. This simulation uses a transient current pulse in the form of a differentiated Gaussian given by Eq. (1):

$$I_0(t) = \sqrt{2}\, e^{0.5} \alpha e^{-\alpha^2} \qquad (1)$$

UNCORRECTED PROOF

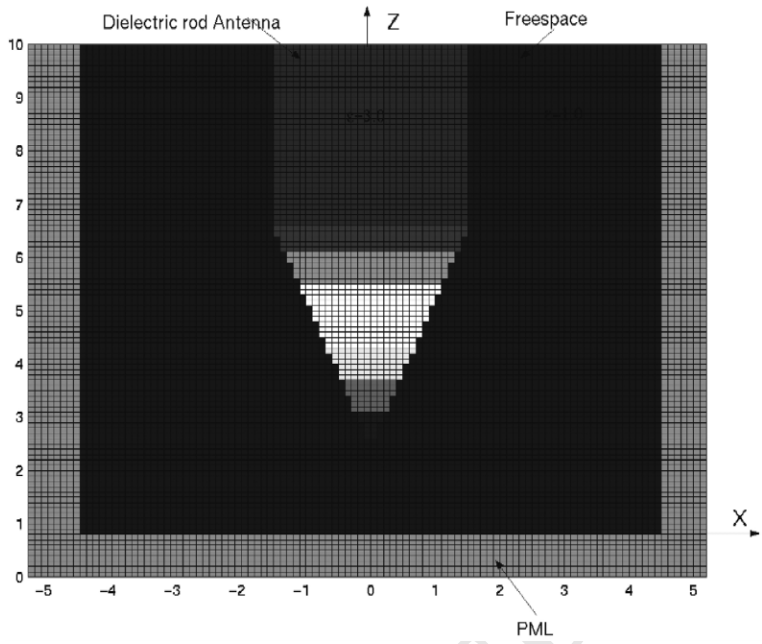


Fig. 3. Partial view of the three-dimensional FDTD model of the new dielectric rod antenna with tapered permittivity.

where $\alpha = (t - 3.2T_b)/T_b$ and $T_b = 7.2 \times 10^{-11}$ s. This pulse was chosen to contain significant frequency content over the desired 2–6 GHz frequency range. It is noticed that the electromagnetic energy is effectively guided within the rod before radiating out. One can also see the electromagnetic fields radiating away from the taper section with spherical phase fronts. These results are similar to those obtained from a rod made of constant permittivity (Chen et al., accepted for publication). A closer investigation reveals that the new design generates much less internal reflections in the taper section. This is shown in Fig. 6 by comparing the reflected responses from constant-permittivity and tapered-permittivity cases. It is observed that instead of having a large reflection near the tip (see dashed line) for the constant-permittivity case, many small reflections occur at the interfaces between the layers for the tapered cases. The total reflected energy from the tapered-permittivity case was found to be 16 dB less than that from the constant-permittivity case with the same cone geometry.

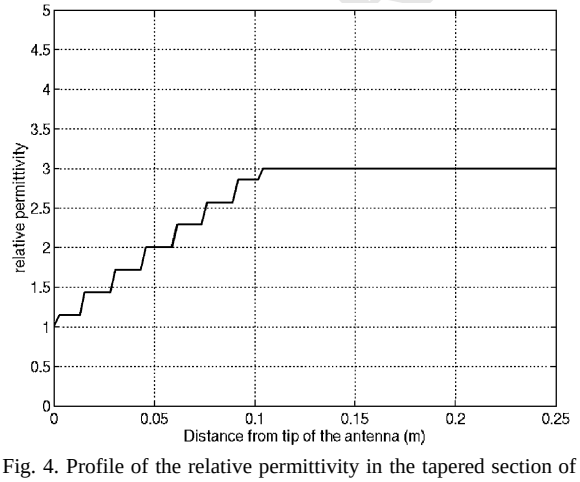


Fig. 4. Profile of the relative permittivity in the tapered section of the new dielectric rod.

It should be noted that no optimization has been applied yet to the above design. One can certainly utilize many existing optimization algorithms to obtain an optimal taper design in both dimension and permittivity to meet a tolerable reflection level with a minimal taper length. This goal is important in minimizing the movement of the radiation center over the operational frequency range such that the radiation pattern is relatively insensitive to frequency variation.

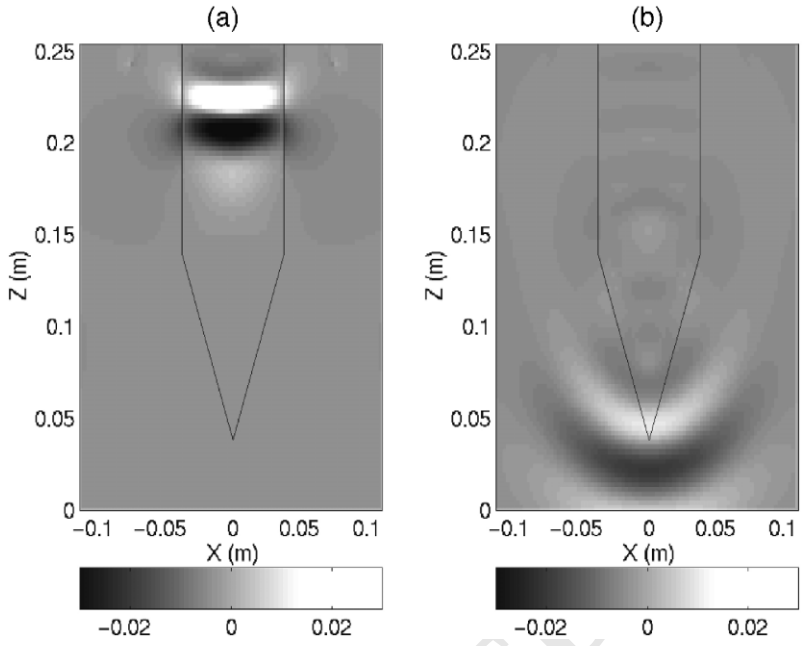


Fig. 5. Snap shots of the FDTD simulated field distributions: (a) guided fields; (b) field radiating out from the tapered section.

Fig. 7 plots the magnitude of the radiated field as a function of both frequency and position along a straight line 2.5 cm away from the rod tip and parallel to the electric field. It is observed that the radiation beamwidth varies only a little over a broad frequency range (above 2.5 GHz). It is noted that for frequencies below approximately 2.4 GHz, the diameter of the rod, i.e. 7.62 cm, becomes less than one wavelength (in material) and the electromagnetic energy becomes loosely guided in this low frequency region. It was also found that the above pattern from the tapered-permittivity design is more stable than

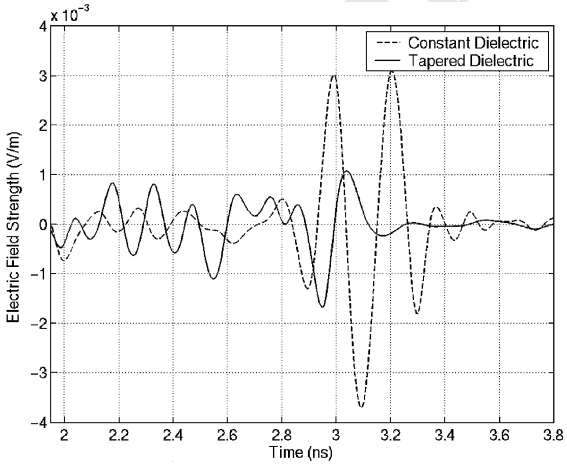


Fig. 6. Comparison of the responses reflected from the tapered section between the constant-permittivity and tapered-permittivity design.

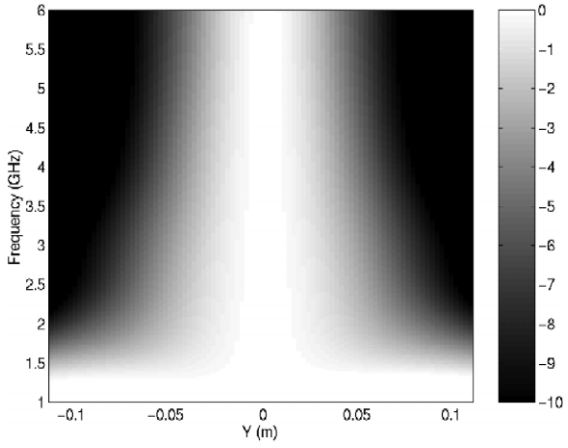


Fig. 7. Radiated field distribution plotted as a function of frequency and position on the ground surface 2.5 cm away from the rod tip.

that from the constant-permittivity design. Fig. 8 plots the amplitude contour of the radiated field at 4 GHz on the ground surface 2.5 cm away from the rod tip. Fig. 8(a) and (b) corresponds to the parallel and perpendicular field components with respect to the polarization of the excitation. These patterns also show improved rotational symmetry compared to that obtained from a constant-permittivity rod.

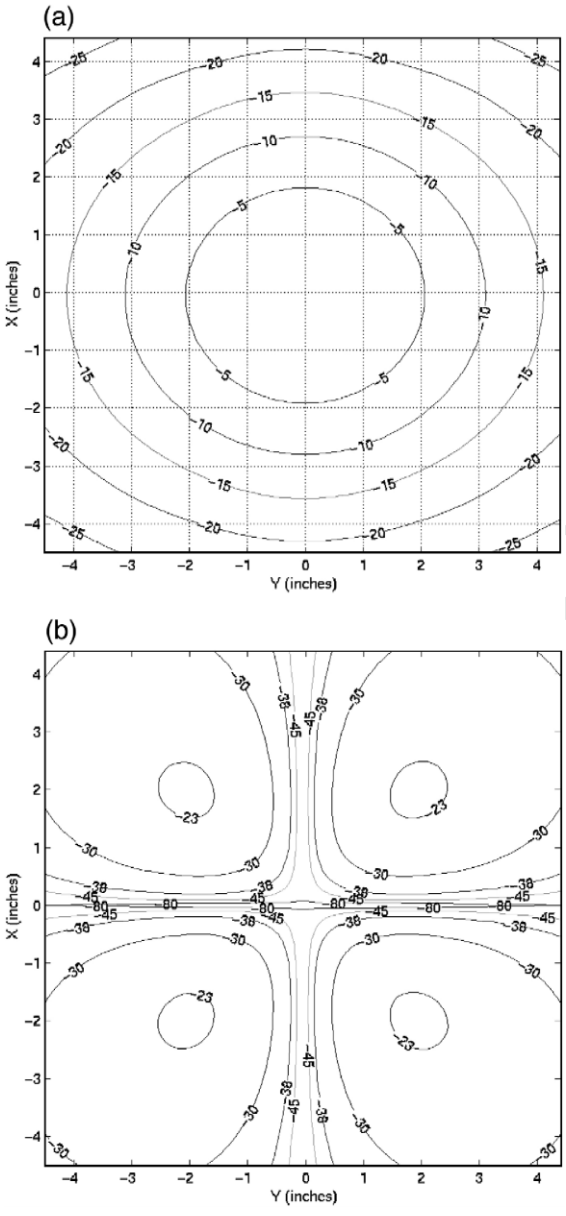


Fig. 8. Radiated field distributions on the ground surface 2.5 cm away from the rod tip: (a) co-polarization component; (b) cross-polarization component.

4. Conclusions

A new broadband dielectric rod antenna with a tapered-permittivity radiation head was presented. The original constant-permittivity taper design works well for low-permittivity rod. However, it faces a potential dispersion problem for high-permittivity rod due to the requirement of a long taper length. The new tapered-permittivity rod antenna design significantly reduced the internal reflections from the radiation head (tapered section) for the same taper length. This means a shorter taper length can be achieved for a similar reflection level obtained from the constant-permittivity rod. It was also found that the radiated pattern on the ground surface from the new design is relatively frequency-independent and rotationally symmetric. The 10-dB spot size on the ground 2.5 cm away was found to be approximately one wavelength at the lowest frequency, i.e. 2 GHz. Such a small spot size is important in detecting small objects. The antenna size of the new rod antenna is significantly smaller than a focusing reflector in achieving a similar spot size. The new tapered-permittivity design allows one to design a rod antenna with high permittivity material such that one can achieve a smaller size, lighter weight and finer spatial resolution.

In addition to the landmine application, this new antenna design will also be useful for other non-destructive inspections that require high spatial and temporal resolutions. Fully polarimetric version of the broadband dielectric rod antenna is also being developed.

References

Abul-Kassemm, A.S., 1972. Experimental study of the characteristics of a horizontal antenna above a dissipative homogeneous earth. Master Thesis, The Colorado University, Colorado, USA.

Chen, C.-C., Nag, S., Rama Rao, K., Lee, R., 2000a. A new ultra-wide bandwidth dielectric rod antenna for ground pene-

UNCORRECTED PROOF

tration radar applications. IEEE Trans. Geosci. Remote Sens., Accepted for publication.

Chen, C.-C., Nag, S., Burnside, W.D., Halman, J., Shubert, K., Peters Jr., L., 2000b. A stand-off focused-beam land mine radar. IEEE Trans. Geosci. Remote Sens. 38 (1), 507–514.

Collin, R.E., 1956. The optimum tapered transmission line matching section. Proc. IRE, 539–548 April.

Hayes, P., 1983. Electromagnetic behavior of transmission lines and resonant wires very near a lossy dielectric interface. PhD Dissertation, the Ohio State University, Ohio, USA.

Mueller, G.E., Tyrrell, W.A., 1947. Polyrod antennas. The Bell System Technical Journal, XXVI, American Telephone and Telegraph, New York.

Noon, D.A., Stickley, G.F., Longstaff, D., 2000. Proceedings of the 8th International Conference on Ground-Penetrating Radar. University of Queensland, Gold Coast, Australia.

Plumb, R., 1998. Proceedings of the 7th International Conference on Ground-Penetrating Radar. University of Kansas, Lawrence, USA.

Procter, R.F., 1950. Input impedance of a horizontal dipole aerial at low height above ground. Proc. IEE (London), Part III, 188–190.

Redman, D., 1994. Proceedings of the 5th International Conference on Ground Penetrating Radar. University of Waterloo, Kitchener, Ontario, Canada.

Salvati, J.L., Chen, C.-C., Johnson, J.T., 1998. Theoretical study of a surface clutter reduction algorithm. Proceedings of International Geoscience and Remote Sensing Symposium. IEEE, Piscatway, NJ, USA, pp. 1460–1462.

Sato, M., 1996. Proceedings of the 6th International Conference on Ground Penetrating Radar. Tohoku University, Sendai, Japan.

Walter, C.H., 1965. Traveling Wave Antennas. McGraw-Hill, New York, p. 214.

Watson, R.B., Horton, C.W., 1948. The radiation pattern of dielectric rods—experiment and theory. J. Appl. Phys. 19, 661–670.

Yee, K.S., 1966. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. IEEE Trans. Antennas Propag. 14, 302–307.

Zhao, L., Cangellaris, 1996. GT-PML: generalized theory of perfectly matched layers and its application to the reflectionless truncation of finite-difference time-domain grids. IEEE Trans. Microwave Theory Tech. 44 (12), 2555–2563.

UNCORRECTED PROOF

Optimizing fusion architectures for limited training data sets

B. A. Baertlein and A. H. Gunatilaka

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

A method is described to improve the performance of sensor fusion algorithms. Data sets available for training fusion algorithms are often smaller than desired, since the sensor suite used for data acquisition is always limited by the slowest, least reliable sensor. In addition, the fusion process expands the dimension of the data, which increases the requirement for training data. By using structural risk minimization, a technique of statistical learning theory, a classifier of optimal complexity can be obtained, leading to improved performance. A technique for jointly optimizing the local decision thresholds is also described for hard-decision fusion. The procedure is demonstrated for EMI, GPR and MWIR data acquired at the US Army mine lanes at Fort A.P. Hill, VA, Site 71A. It is shown that fusion of features, soft decisions, and hard decisions each yield improved performance with respect to the individual sensors. Fusion decreases the overall error rate (false alarms and missed detections) from roughly 20% for the best single sensor to roughly 10% for the best fused result.

Keywords: Sensor fusion, statistical learning theory, support vector machines, pattern classification

1. INTRODUCTION

Sensor fusion has been proposed as a means of meeting stressing requirements in detection of land mines. Fusion offers the potential for increased probability of detection, decreased false alarm rates, and operation in a broader range of environments. A very large number of sensor technologies have been proposed for mine detection,[1] and that number continues to grow.

Mine detection, whether done with a single sensor of a fused suite, can be viewed as a classification problem in which sensor data must be classified as representing mines or clutter. Suppose that we are given sensor data $\mathbf{x}$ with which to determine the truth of hypotheses H_k that describe the presence of absence of a mine. We do so by choosing H_k to maximize the Bayes' risk or, when only the probability of error is important, by choosing the hypothesis that maximizes the a posteriori probability $\Pr(H_k|\mathbf{x})$. A fundamental tenet of sensor fusion (or of any classification process) is that more information cannot degrade performance. It is easy to prove that additional independent sensor data $\mathbf{y}$ that is positively correlated with H_k (i.e., for which $\Pr(\mathbf{y}|H_k) \geq \Pr(\mathbf{y})$) will produce $\Pr(H_k|\mathbf{x},\mathbf{y}) \geq \Pr(H_k|\mathbf{x})$.

The probabilities appearing in the above expressions are unknown and must be estimated from data. All classification problems require training data from which the desired classification process is learned.* The quantity and quality of these training data have a strong influence on the performance of the classifier. Although many classifiers are known to produce optimal performance asymptotically (i.e., when the amount of training data is infinite), performance based on a finite training set may fail to meet expectations. The problem of limited training data is acute for mine detection in general, and for sensor fusion in particular. When multi-sensor data collections are performed, the data set useful for fusion is limited by the performance of the slowest or least reliable sensor.

The usefulness of a sensor-fused system (or any classifier) depends on its ability to generalize, i.e., to detect mines in data not seen previously. It is well known that problems with limited training data manifest themselves during generalization.[2] Classifiers that perform well on their training data but generalize poorly typically employ an architecture or an approximation to a decision surface that is too complex for the training data. Problems of this type arise in many classifiers, including k-nearest neighbor and neural networks.

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

*A sufficiently accurate physical model can supply the equivalent of training data, but for many mine detecting sensors the random environment has a strong influence on the sensor data, leading to an impractical number of model variables.

In this work we describe methods of optimizing the design of a classifier for a given training data set. For feature-level fusion we employ recent developments in statistical learning theory (SLT),[3] which permits us to bound the performance of classifiers designed with limited training data. Some essential aspects of SLT as it relates to mine detection are described in Section 2. The architecture of the fusion algorithm (classifier) determines its complexity, which can be bounded by SLT. The Vapnik-Chervonenkis (VC) dimension of the classifier provides a means of assessing the classifier complexity. For any classification problem, there is an optimum complexity. Unfortunately, it is largely impossible to estimate the VC dimension of most classifiers. One can implement SLT using a new form of classifier known as the support vector machine (SVM), for which the complexity is readily controlled. It has been observed that SVMs exhibit performance that meets or exceeds that of other classifiers. SVMs are described in Section 3. We describe in Section 4 a method of jointly optimizing the individual decision thresholds for decision-level fusion of hard decisions. In Section 5 we apply these methods to multi-sensor demining data collected at The US Army mine lanes at Fort A.P. Hill, VA (Site 71A). The data comprise samples of surrogates and buried mines with known positions. The sensor suite used included an EMI sensor, a GPR, and an MWIR camera. Concluding remarks appear in Section 6.

2. STATISTICAL LEARNING THEORY

Statistical learning theory, also referred to as Vapnik-Chervonenkis (VC) theory, has been under development since the 1970s. In this section we summarize the relevant parts of that theory. Descriptions of SLT have been given by Vapnik at an overview level[3] and at a deeper level.[4] Review works by Schölkopf et al.[5] and by Burges[6] may also be consulted for details not presented here.

Consider the following problem: Given N i.i.d. samples of training (sensor) data $\mathbf{x}_i$ with true classification y_i (e.g., $y_i = 1$ if sample $\mathbf{x}_i$ corrresponds to a mine and $y_i = 0$ otherwise), we wish to discover the classifier (function) $y = F(\mathbf{x})$ that will return the true identity y when presented with an input sample $\mathbf{x}$. We approximate this function by using training data to estimate parameters $\boldsymbol{\alpha}$ for a family of functions $F(\mathbf{x}; \boldsymbol{\alpha})$. The risk or expected loss for the classifier is given by

$$\mathcal{R}(\boldsymbol{\alpha}) = \int d\mathbf{x} \int dy L(y, F(\mathbf{x}; \boldsymbol{\alpha})) \Pr(\mathbf{x}, y) \tag{1}$$

where L is a loss function. The function $L(y, y')$ provides a measure of the "distance" between the true output y and the estimate $y' = F(\mathbf{x}, \boldsymbol{\alpha})$. For mine detection we are primarily interested in a loss function of the form

$$L(y, F(\mathbf{x}, \boldsymbol{\alpha}) = \begin{cases} C_{01} & y = 0,\ F(\mathbf{x}, \boldsymbol{\alpha}) = 1 \qquad \text{(false alarm)} \\ C_{10} & y = 1,\ F(\mathbf{x}, \boldsymbol{\alpha}) = 0 \qquad \text{(missed detection)} \\ 0 & y = F(\mathbf{x}, \boldsymbol{\alpha}), \qquad \text{(correct decision)} \end{cases} \tag{2}$$

where C_{01} and C_{10} are the costs of a false alarm and missed detection respectively. Clearly $C_{01} \ll C_{10}$, but it is not obvious how these costs should be assigned. In this work, we assume $C_{01} = C_{10} = 1$, in which case $\mathcal{R}$ is the probability of error. The extension to other cases is straightforward.

Since $\Pr(\mathbf{x}, y)$ in Eq. (1) is unknown, we are forced to estimate the true risk $\mathcal{R}$ from the available training data. We define the empirical risk

$$\mathcal{R}_{emp}(\boldsymbol{\alpha}) = \frac{1}{N} \sum_{n=1}^{N} L(y_i, F(\mathbf{x}; \boldsymbol{\alpha})) \tag{3}$$

Note that $\mathcal{R}_{emp}$ does not involve the probability density $\Pr(\mathbf{x}, y)$. Taking $C_{01} = C_{10} = 1$ makes $\mathcal{R}_{emp}$ an estimate of the classifier error rate.

Statistical learning theory addresses the relation between the true risk $\mathcal{R}$ and the empirical risk $\mathcal{R}_{emp}$. For a fixed classifier (fixed $\boldsymbol{\alpha}$) the empirical risk will always be less than the true risk, since one can develop a classifier that fits a finite set of training data arbitrarily well. The following bound can be derived

$$\mathcal{R}(\boldsymbol{\alpha}) \leq \mathcal{R}_{emp}(\boldsymbol{\alpha}) + \Phi(h/N) \tag{4}$$

where Φ is a confidence interval and h is the so-called "VC dimension." The confidence interval has the remarkable property that it is independent of the unknown probability distribution $\Pr(\mathbf{x}, y)$. The precise definition of h is somewhat technical, but it can be regarded as the number of training samples that can be correctly classified by

$F(\mathbf{x}, \boldsymbol{\alpha})$, i.e., the learning capacity of the classifier. A plot of these quantities appears in Figure 1. As the complexity of the classifier increases, the empirical risk (as measured by classifier performance on the training data) decreases. Simultaneously, there is less confidence (and a greater error rate) for more complex classifiers. In practice, we can use the resubstitution performance of the classifier to estimate $\mathcal{R}_{emp}$ and a validation set to estimate $\mathcal{R}$. In general, there exists an optimal value of h/N. Classifiers that have too little capacity are unable to learn the training data. Classifiers that are too complex will learn the training data well, but will have poor generalization capability. The objective of this work is to determine the level of classifier complexity that will minimize the true risk.

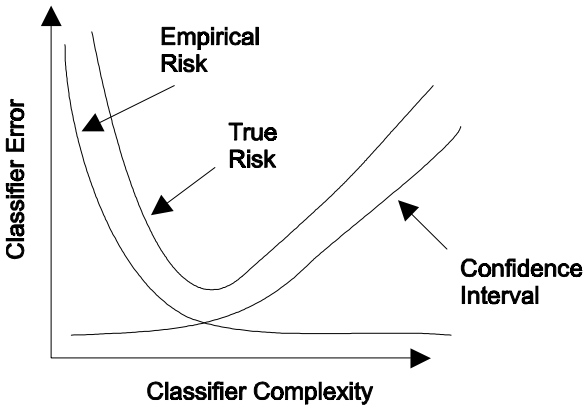


Figure 1. Components of classifier error.

The confidence interval $\Phi(h/N)$ is determined from the convergence rate of $\mathcal{R}_{emp} \rightarrow \mathcal{R}$ as $N \rightarrow \infty$. Previous work has produced a variety of expressions for this bound. For a loss function L bounded above by unity, Vapnik[4] has shown that with confidence η the value of Φ is given by

$$\Phi = \frac{\epsilon^2}{2}\left(1 + \sqrt{1 + \frac{\mathcal{R}_{emp}}{\epsilon^2}}\right) \tag{5}$$

where ϵ is given by

$$\epsilon = 2\sqrt{\frac{h}{N}\left[\ln\left(\frac{2N}{h}\right) + 1\right] - \frac{1}{N}\ln\eta} \tag{6}$$

For small empirical risk $\mathcal{R}_{emp}$ we have $\Phi \approx \epsilon^2$ while for large empirical risk we have $\Phi \sim \epsilon/2$. This last approximation is often cited in the literature, because of its simple form. It predicts that the confidence interval Φ decreases with increasing training data N and with decreasing classifier capacity h. Note also that Φ is an upper (worst case) bound. It may well be possible to design a classifier whose performance beats this bound.

Finding the optimum classifier consists in balancing the empirical risk (which decreases with increasing classifier capacity) and the confidence interval (which increases with increasing classifier capacity). There exist a variety of methods for designing classifiers, but in general optimizing classifier performance is a trial and error process. Although one is guaranteed that an optimum classifier exists (since the risk is bounded below), there is little guidance on how it can be constructed.

The concept illustrated in Figure 1 suggests a principled method for designing a classifier with optimal performance on a given training data set. If one can develop a family of classifiers $F_1, F_2, \ldots$ with increasing VC dimension $h_1 < h_2 < \ldots$, then by testing these classifiers we can determine the value of h that minimizes the true risk, i.e., we can identify the minimum in the true risk curve of Figure 1. This procedure is known as structural risk minimization (SRM),[7] and it is used most effectively in concert with support vector machines, described below.

3. SUPPORT VECTOR MACHINES

The existence of an optimal value of VC dimension h is of considerable theoretical interest, but it is of limited practical value since h is essentially impossible to estimate for any but the simplest classifiers. One classifier for which h can be determined is a linear hyperplane classifier

$$F(\mathbf{x}, \mathbf{w}, b) = \text{sgn}\,(\mathbf{x} \cdot \mathbf{w} + b) \tag{7}$$

where $\mathbf{w}$ is a vector that defines the normal to the hyperplane and b is a bias term.

Support vector machines (SVMs) are based on the concept of the optimal separating hyperplane, in which the value of $\mathbf{w}$ is selected to maximize class separation (also known as "margin"). This is achieved by minimizing $||\mathbf{w}||$ while simultaneously requiring that correct decisions are produced for the training data. We require

$$\text{minimize} \quad ||\mathbf{w}||^2 \qquad \text{subject to} \quad y_i(\mathbf{x}_i \cdot \mathbf{w} + b) \geq 1, \qquad i = 1, 2, \ldots, N \tag{8}$$

A Lagrangian formulation for the problem leads to

$$L(\mathbf{w}, b) = \frac{1}{2}||\mathbf{w}||^2 - \sum_{i=1}^{N} \alpha_i [y_i(\mathbf{x}_i \cdot \mathbf{w} + b) - 1] \tag{9}$$

where the α_i are Lagrange multipliers. This quantity must be minimized with respect to $\mathbf{w}$ and b, and maximized with respect to the α_i. The minimization requirements impose linear constraints as follows:

$$\frac{\partial}{\partial b} L = 0 \qquad \Rightarrow \sum_{i=1}^{N} \alpha_i y_i = 0 \tag{10}$$

$$\frac{\partial}{\partial w_j} L = 0 \qquad \Rightarrow \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i = \mathbf{w} \tag{11}$$

It can be shown that for a linearly separable problem, all but a few of the α_i will be identically zero. The training data $\mathbf{x}_i$ corresponding to these nonzero weights are known as the support vectors. They define the hyperplane, and the remaining training data are superfluous to the classifier. Substituting equations (10) and (11) into the Lagrangian allows us to eliminate $\mathbf{w}$ in favor of the α_i, leading to the dual form of Eq. (8). Using a vector notation for the α_i we obtain the following quadratic programming problem:

$$\text{maximize} \quad W(\boldsymbol{\alpha}) = \mathbf{e}^T\boldsymbol{\alpha} - \frac{1}{2}\boldsymbol{\alpha}^T\mathbf{Q}\boldsymbol{\alpha} \qquad \text{subject to} \quad \mathbf{y}^T \cdot \boldsymbol{\alpha} = 0; \qquad \alpha_i \geq 0 \tag{12}$$

where $\mathbf{e} = [1\ 1\ \cdots\ 1]^T$ and

$$[\mathbf{Q}]_{ij} = y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j \tag{13}$$

Numerical optimzation is used to determine $\boldsymbol{\alpha}$, and the decision function is

$$F(\mathbf{x}; \boldsymbol{\alpha}, b) = \text{sgn}\left(\sum_{i=1}^{N} [y_i \alpha_i \mathbf{x} \cdot \mathbf{x}_i + b]\right) \tag{14}$$

When the classes overlap and the data are not separable by a hyperplane, one can show that the appropriate formulation is identical to that presented above if we introduce "slack" variables $\zeta_i \geq 0$ such that the problem becomes

$$\text{minimize} \quad ||\mathbf{w}||^2 + C\sum_{i=1}^{N} \zeta_i \qquad \text{subject to} \quad y_i(\mathbf{x}_i \cdot \mathbf{w} + b) \geq 1 - \zeta_i, \qquad i = 1, 2, \ldots, N \tag{15}$$

where $C > 0$ is a user-defined constant. Taking $C \to \infty$ produces the separable case, while $C \to 0$ reduces the penalty for class overlap. The minimization proceeds as described above for the separable case, except that the α_i are now constrained as

$$0 \leq \alpha_i < C \tag{16}$$

It is instructive to compare SVMs with Fisher's linear discriminant,[8] another common hyperplane classifier. The normal to the Fisher hyperplane $\mathbf{w}_f$ for a two-class problem having means ($\mathbf{m}_0$,$\mathbf{m}_1$) and covariance matrices ($\mathbf{C}_0$,$\mathbf{C}_1$) satisfies the generalized eigenvector problem[9]

$$[\mathbf{C}_B - \lambda(\mathbf{C}_1 + \mathbf{C}_1)]\mathbf{w}_f = 0 \tag{17}$$

where $\mathbf{C}_B$ is the between-class scatter matrix

$$\mathbf{C}_B = \frac{N_0 N_1}{N}[\mathbf{m}_0 - \mathbf{m}_1][\mathbf{m}_0 - \mathbf{m}_1]^T \tag{18}$$

Since $\mathbf{C}_B$ has rank one, this equation has one non-zero eigenvector, and we find

$$\mathbf{w}_f = (\mathbf{C}_1 + \mathbf{C}_1)^{-1}[\mathbf{m}_0 - \mathbf{m}_1] \tag{19}$$

which is parallel to the path between the mean vectors. Although the Fisher discriminant is known to be optimal for Gaussian distributions with equal covariances, it does not necessarily produce an optimal separation for a finite data set. In general, the SVM hyperplane is not parallel to the Fisher hyperplane.

Linear classifiers are seldom optimal in practice, and higher-order approximations to the decision surface are commonly employed. The hyperplanes used in SVMs can also be extended to non-planar surfaces. To do so, the input data $\mathbf{x}_i$ are projected into higher dimensions by using a nonlinear tranformation $\Psi(\mathbf{x})$. In these high dimensional spaces, the data are more likely to be separable by hyperplanes. The problem formulation parallels that given above with the substitution $\mathbf{x} \rightarrow \Psi(\mathbf{x})$.

A classical problem with higher-order classifiers is dimensionality. Consider a simple polynomial transformation of the form $\Psi(\mathbf{x}) = \{1, x_1, x_2, ...x_m, x_1^2, x_1x_2, x_1x_3, ...\}$. The product terms are each inputs to the classifier. It is evident that the resulting input set has high dimensions, for which the memory, computation, and training data requirements are large. An ingenious method is used to avoid this for SVMs. Note that the SVM requires only inner products of the projected data of the form $\Psi(\mathbf{x}) \cdot \Psi(\mathbf{y})$. Mercer's theorem[10] implies that (with certain mild restrictions) such inner products can be written in terms of a symmetric kernel function $k(\mathbf{x}, \mathbf{y})$. Hence, given a transformation Ψ, we can replace the inner product $\Psi(\mathbf{x}_i) \cdot \Psi(\mathbf{x}_j) \rightarrow k(\mathbf{x}_i, \mathbf{x}_j)$. Since appropriate transformations Ψ are seldom evident, it is attractive to use Mercer's theorem in "reverse", i.e., to assume a convenient form for $k(\mathbf{x}_i, \mathbf{x}_j)$ without regard for the implied function Ψ. Thus, the above-described formulation still applies if the matrix $\mathbf{Q}$ becomes

$$[\mathbf{Q}]_{ij} = y_i y_j k(\mathbf{x}_i, \mathbf{x}_j) \tag{20}$$

Functions used for $k(\cdot, \cdot)$ include polynomials, radial basis functions, splines, and others.

4. OPTIMIZING FUSION OF HARD DECISIONS

In general, optimal fusion of hard decisions is challenging. Each sensor produces a binary decision u_i on the basis of a sensor-specific threshold t_i of a test statistic (e.g., the log-likelihood ratio). For an optimal fusion strategy one must simultaneously define a relation among the decision thresholds (t_1, t_2, t_3) (i.e., the local decision strategies) and a fusion rule $\Pr(H_k|u_1, u_2, u_3)$ (the global decision strategy). In general, numerical optimization is required,[11] although some simple expressions are available if one assumes that the sensors are independent. In a previous fusion study, we found the performance of the latter approach to be poor,[12] and in this work we explored alternative methods.

A straightforward (but not necessarily optimal) local decision strategy is to remap all of the individual sensor test statistics to a common range and then use a common threshod $t \equiv t_1 \equiv t_2 \equiv t_3$. Although this technique can work well in some situations, it is adversely affected by outliers in the data or drastically different performance from sensor to sensor.

A somewhat better approach is to define a monotonic function $G(t; \boldsymbol{\beta}_m)$ with parameters $\boldsymbol{\beta}_m$ to remap the test statistics for each sensor m. A common threshold τ is used for the remapped thresholds $G(t; \boldsymbol{\beta}_m)$ for all sensors. Since (by assumption) the data sets are small, numerical optimization of the $\boldsymbol{\beta}_m$ is practical. In addition, since the number of sensors involved is also small, one can simultaneously conduct an exhaustive search over all possible fusion rules $\Pr(H_k|u_1, u_2, u_3)$. The SRM technique is not used explicitly in this design of the hard decision fusion algorithm, but the classifiers for the individual sensors may be optimized using SRM.

5. RESULTS

We used structural risk minimization to optimize the performance of mine detectors based on both single sensors and a fused sensor suite. The threshold-remapping methods were also used to optimize hard decision fusion. In this section we describe the data set, review some sensor-specific processing, and present the results.

During July 1999 we acquired multisensor data at Fort A.P. Hill, VA. Data were collected at the calibration lanes of Site 71A, which comprises a 25 m by 5 m area with mines or clutter objects buried at the center of grid cells having dimensions 1 m by 1 m. Some important features of this test site are (1) extraneous metal has been largely removed from the area, which reduces EMI clutter; (2) the location of the mine (or clutter) item can be accurately located, which obviates problems with sensor positioning; and (3) although real mines have been emplaced, in most cases the explosive was removed from the mines, which affects its thermal signature.

We acquired data over a portion of the site that comprised 27 deactivated mines and 32 clutter objects. The sensors used included a Schiebel AN 19/2, an OSU-developed ground penetrating radar,[13] a MWIR sensor, and a LWIR sensor. In the results that follow, we have replaced the Schiebel data with data from the GEM-3 sensor,[14] collected by Duke University.[15] During most of our data collection, conditions were not conducive to IR data collection. The weather was overcast and rain fell occasionally. Nonetheless, we did manage to acquire some useful MWIR data during one night-time collection. In the discussion that follows, we do not consider the LWIR data.

5.1. Supporting Processing

The data collected by each sensor have a different format and require different processing to suppress clutter and to extract features. In this section we briefly describe that processing.

5.1.1. GEM-3 Sensor

The GEM-3 data comprise samples taken at ten points spaced 2 inches apart in a "+" pattern over each putative target, with 5 samples taken in a left-right path and 5 samples taken in a fore-aft path. At each point, in-phase and quadrature magnetic field measurements were acquired at 20 frequencies logarithmically spaced from 270 Hz to 23.79 kHz. For the 7/7/98 data set used here, data were acquired at 44 locations in the Site 71A calibration area, where the identify of the targets are known. Background data, required to correct for sensor baseline drift, were acquired between measurements. An extensive discussion of the GEM-3 data has been presented by Gao et al.,[16] which can be consulted for details not discussed here.

Processing of the GEM-3 data included correcting for the sensor background and converting the sensor output to a quantity proportional to the complex field amplitude. The resulting data set (200 complex values at each target) was reduced by computing the energy at each spatial position, leading to ten values for each target. An example EMI data vector is shown in Figure 2. For perfectly centered, symmetric targets these signatures should display symmetry about samples 3 and 8 (i.e., about the center of the left-right and fore-aft scans) and samples 1-5 (the left-right scan) and 6-10 (the fore-aft scan) should be identical.

5.1.2. GPR Sensor

The OSU GPR is a down-looking sensor using a novel dielectric rod antenna.[13] The antenna was mounted to a linear positioner, which was scanned over the target location. Data were acquired at 1-2 cm sample intervals over each target cell. Linear scans were acquired over 59 mines and clutter sites. The system uses a wide bandwidth (1-6 GHz), and after pulse compression data similar to Figure 3(a) are observed. The strong near-horizontal band in this image is the ground reflection. Using a recently developed technique,[17] that reflection can be significantly reduced, leading to the data shown in Figure 3(b). In the latter result, the characteristic hyperbolic arcs generated by the target are evident. The data used in processing comprised a subsampled version of the time-domain output acquired directly through the center of each signature.

5.1.3. MWIR Sensor

The MWIR sensor used in this collection was a COTS camera (Cincinnatti Electronics, IRRIS 160ST) operating in the spectral band 2.2-4.6 μm. The instrument's focal plane comprises 160×120 pixels, with an NEΔT of 0.025K. To avoid clutter produced by reflected sunlight, the sensor was operated at night. During data acquisitioin the sensor was positioned at a fixed height and distance from each putative target site. The camera was aimed at a ground location a known distance from the camera's ground-projected center. Using the (known) field of view, one can determine

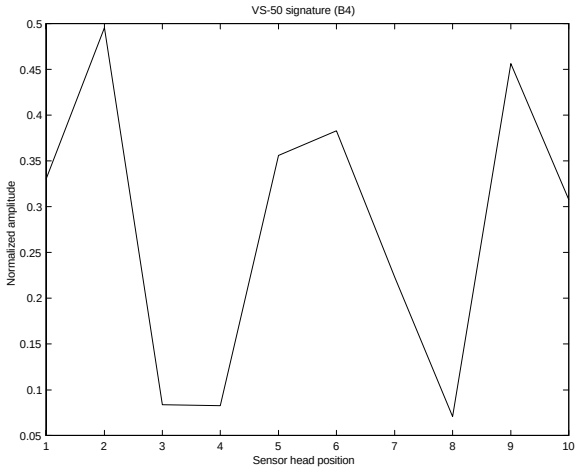
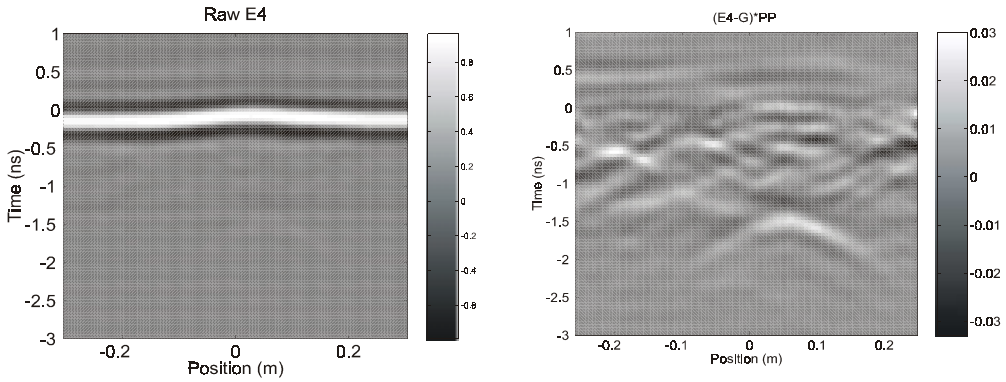


Figure 2. Example GEM-3 data.



(a) Raw GPR image (b) Clutter-reduced image

Figure 3. Comparison of raw and clutter-reduced data generated from measurements over a three-inch deep VS-50 mine.

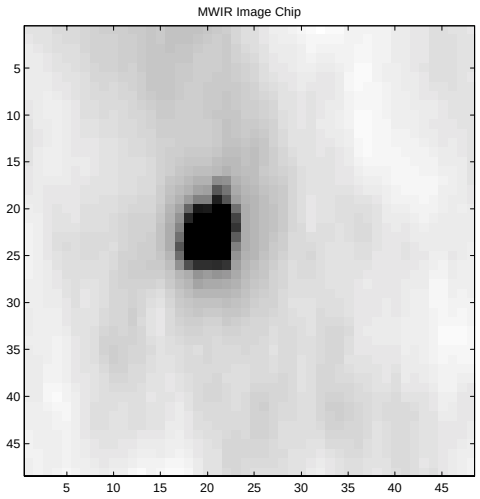


Figure 4. Example MWIR data. A M14 mine flush buried is shown.

the position of each pixel on the ground. After remapping the image to eliminate the effects of perspective, data similar to Figure 4 are obtained. Features were extracted from these data using a model-based technique described previously.[18] IR clutter is normally distributed, and a maximum likelihood (nonlinear least squares) technique was used to estimate model parameters used as features.

5.2. Individual Sensor Performance

A review of the data acquired by the GEM-3, GPR, and MWIR camera revealed that data from all three sensors were available at 42 grid cells, which comprised equal number of target and clutter samples. Features were extracted from the data acquired by each sensors as described above.

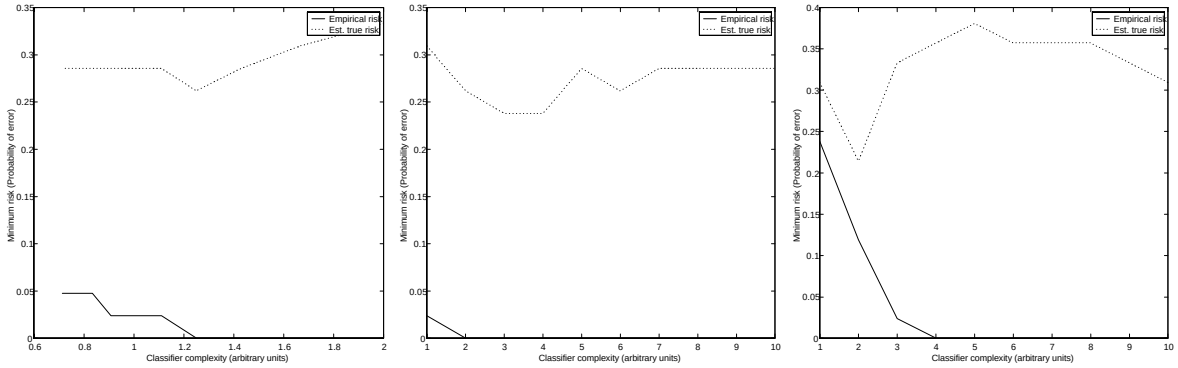
The SRM method was used to identify an optimum classifier complexity for each sensor. For this small data set we used the "leave-one-out" method (a form of cross-validation, or resampling) to estimate the true risk for a range of classifier complexity. The resubsitution method was used to determine the classifier empirical risk. Both polynomial classifiers and radial basis function classifiers[†] were examined, and the design producing the best performance was used. In general, the radial basis function classifiers had a small advantage, but the performance of other types of nonlinearities produced comparable results. This finding is similar to that reported in the literature for other SVM applications. In our tests we found that the slack variable weight C had a minimal effect on classifier performance. We used $C = 10$ for all tests. The risk calculation for each sensor is shown in Figure 5. In each case the empirical risk decreases with increasing complexity, and a minimum occurs in the estimated true risk (cf. Figure 1). The minimum risk (equal to the error rate) for the best sensor (MWIR) is slightly larger than 20%.

Using the classifier design that produced the minimum true risk, we computed the ROC curves as shown in Figure 6. We observe that no sensor has an overwhelming advantage in detection.

5.3. Fusion Performance

We explored fusion of sensor features, soft decisions (classifier outputs produced for individual sensors), and hard decisions (the result of thresholding individual sensor classifier outputs). Feature-level fusion is a straightforward process. We concatenated the feature vectors for all sensors and trained a SVM classifier. Fusion of soft decisions can be performed using a number of techniques.[12] We opted to form a hierarchical classifier in which the outputs of the single sensor classifiers are supplied to another SVM.

[†]Although the VC dimension of a radial basis function classifier is infinite, the width of the basis functions provides a degree of control over its complexity.[6]



(a) GEM-3 risk (b) GPR risk (c) MWIR risk

Figure 5. Risk calculations for the individual sensors.

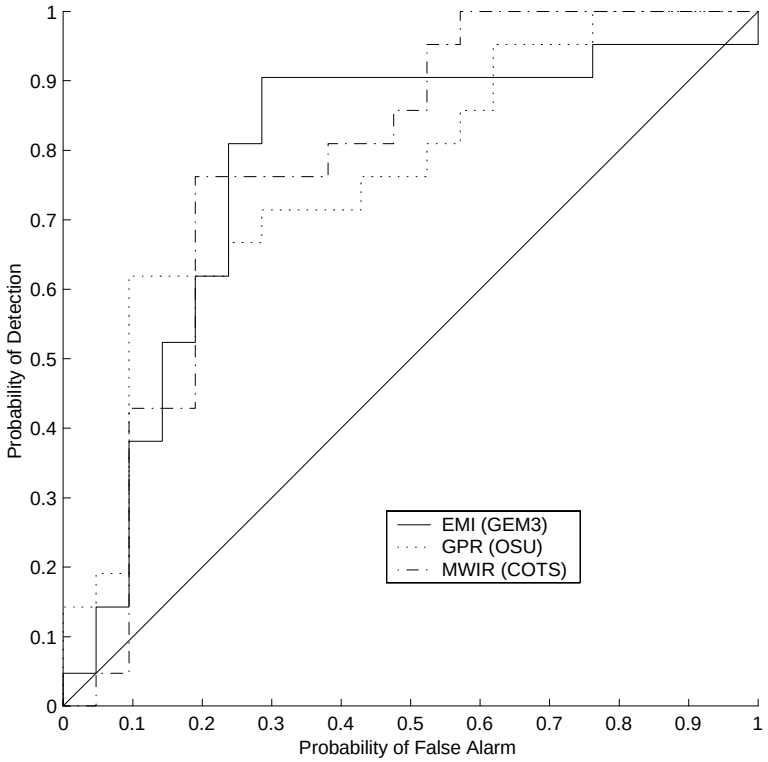
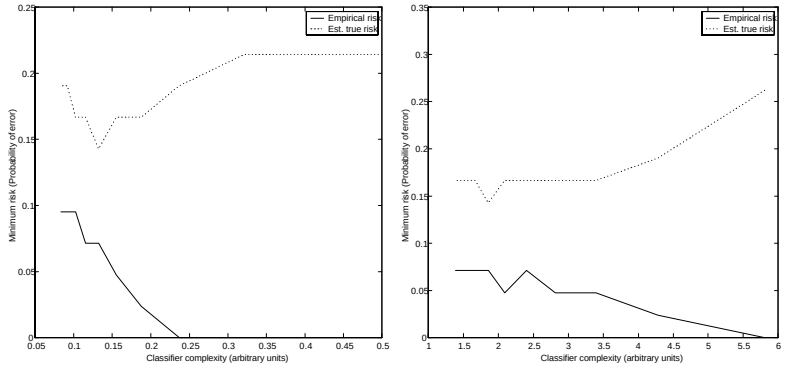
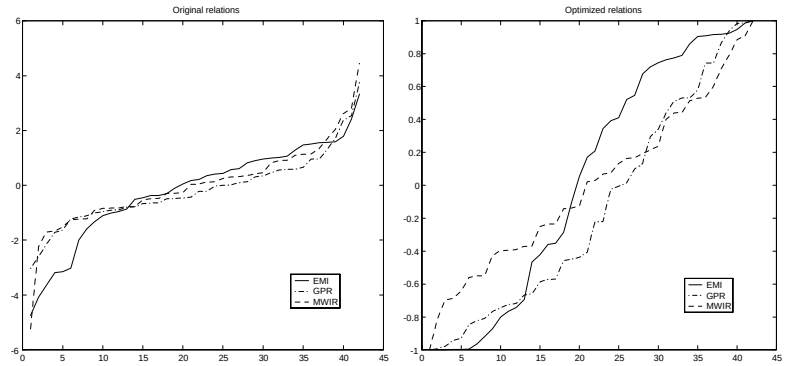


Figure 6. ROC curves for the individual sensors.



(a) Feature-level fusion risk (b) Soft-decision fusion risk

Figure 7. Risk calculations for the fused sensor suite.



(a) Original test statistics (b) Test statistics after remapping

Figure 8. Test statistics (a) before and (b) after the optimization process.

The SRM approach was used to optimize both the feature-level and soft decision-level fusion classifiers. The resulting risk data are shown in Figure 7. Again, we note the presence of minima, which denote the optimum classifier complexity for this data set. The minimum risk is approximately 15% for these classifiers.

The SVM-optimized results shown in Figure 6 were the basis for optimizing hard decision fusion. After sorting the classifier outputs into ascending order, the test statistics for these sensors trace out the curves given in Figure 8. That figure also includes the remapped test statistics after the optimization process described in Section 4. The remapping functions G were the product of a logistic function (followed by scaling to a common range) and Chebyshev polynomials with numerically optimized weights $\boldsymbol{\beta}_m$. The optimization process tends to increase the slope of these curves near the decision boundary (the zero crossing), which tends to increase the decision margin. Because the performance of these sensors are comparable, a majority voting technique was used for $\Pr(H_k|u_1, u_2, u_3)$. That rule is near optimal for more than two sensors of comparable performance. The true risk estimated for this classifier is 9.5%. It is interesting to note that for this data set, a simple linear remapping of the raw test statistics does nearly as well (12% error).

ROC curves are shown in Figure 9 for feature-level, soft-decision level, and hard-decision level fusion. We observe that feature-level and soft decision fusion produce comparable performance. In contrast to the findings of our previous

work,[12] hard-decision fusion performs better than either feature-level fusion or hard-decision fusion. We attribute this improvement to (1) our decision to reject the independent sensor hypothesis and (2) the fact that no one sensor has an overwhelming performance advantage (which tends to reduce the benefit of fusion under the independent sensor hypothesis).[11] These findings bear further investigation. The benefit of fusion can be quantified by comparing the risk for the best fused suite (hard decision fusion at roughly 10%) and the best individual sensor (MWIR at roughly 20%).

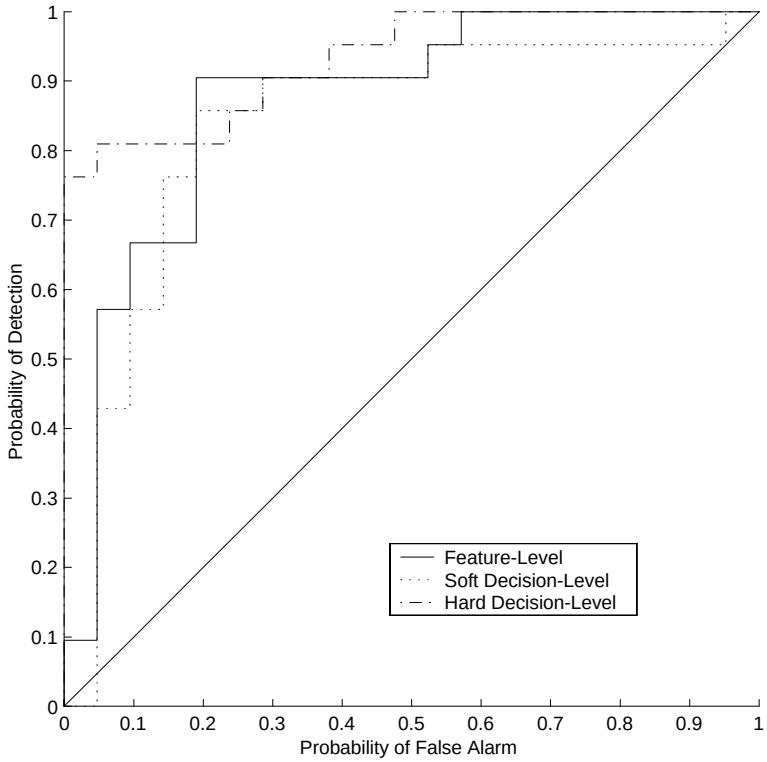


Figure 9. ROC curves for three forms of fusion.

6. CONCLUDING REMARKS

Data sets available for the design of sensor-fused mine detection algorithms are typically small, which has an adverse effect on fusion algorithm design and detection performance. We have described an approach based on structural risk minimization that permits us to extract the best possible fusion performance from these small data sets.

Tests of the algorithm described here for 42 samples of EMI, GPR and MWIR data suggest that fusion produces a measureable benefit in performance. We found that the net error rate (missed detections and false alarms) decreased by roughly a factor of two when the best individual sensor is compared to the best fused result.

ACKNOWLEDGMENTS

The authors would like to thank the staff of E-OIR Measurements, Inc. and the Joint UXO Coordinating Office for their support during data collections at Fort A. P. Hill, Site 71A. We would also like to thank L. Collins at Duke University for her assistance in obtaining and interpreting the GEM-3 data.

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the authors and are not necessarily those of Duke University or the ARO.

REFERENCES

1. C. Stewart, "Summary of mine detection research (U)," technical report 1636-TR, U.S. Army Engineer Research and Development Laboratories, Corps of Engineers, Fort Belvoir, VA, May 1960. Volume 1, DTIC AD320124.
2. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley, New York, NY, 1973.
3. V. Vapnik, *The Nature of the Statistical Learning Theory*, Springer Verlag, New York, NY, second ed., 2000.
4. V. Vapnik, *Statistical Learning Theory*, Wiley, New York, NY, 1998.
5. B. Schölkopf, C. J. C. Burges, and A. J. Smola, *Advances in Kernel Methods: Support Vector Learning*, MIT Press, Cambridge, MA, 1999.
6. C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery* **2**(2), pp. 1–49, 1998.
7. V. Vapnik, *Estimation of Dependences Based on Empirical Data (in Russian)*, Nauka, Moscow, 1979. English translation: Springer, New York, 1982. Also see Ch. 4, Vapnik, 2000.
8. R. A. Fisher, "The use of multiple measures in taxonomic problems," *Ann. Eugenic* **7**, pp. 179–188, 1936.
9. J. J. W. Sammon, "An optimal discrimination plane," *IEEE Trans. Computers* , pp. 826–829, Sept. 1970.
10. J. Mercer, "Functions of positive and negative type and their connection with the theory of integral equations," *Philos. Transactions of the Royal Society of London* **A209**, pp. 415–446, 1909.
11. P. K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag, New York, NY, 1997.
12. A. H. Gunatilaka and B. A. Baertlein, "Comparison of predection and postdetection fusion for mine detection," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **3710**, pp. 1212–1223, 1999.
13. S. Nag, L. Peters, I. J. Gupta, and C.-C. Chen, "Ramp response for the detection of anti-personnel mines," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **3710**, pp. 1313–1322, 1999.
14. I. J. Won, D. A. Keiswetter, and D. R. Hansen, "A monostatic broadband electromagnetic induction sensor," *Journal of Environmental and Engineering Geophysics* **2**, pp. 53–64, Aug. 1997.
15. L. M. Collins, "Statistical signal processing for demining: Experimental validation," tech. rep., Duke University, 1998.
16. P. Gao, L. Collins, J. Moulton, L. Makowsky, R. Weaver, D. Keiswetter, and I. J. Won, "Enhanced detection of landmines using broadband EMI data," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **3710**, pp. 2–13, April 5-9 1999.
17. A. H. Gunatilaka and B. A. Baertlein, "A subspace decomposition technique to improve GPR imaging of anti-personnel mines," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **4038-111**, April 24-28 2000.
18. I. K. Sendur and B. A. Baertlein, "Techniques for improving buried mine detection in thermal IR imagery," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **3710**, pp. 1272–1283, 1999.

A subspace decomposition technique to improve GPR imaging of anti-personnel mines

A. Gunatilaka and B. A. Baertlein

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

Ground-reflected clutter is often a performance-limiting factor in ground-penetrating radar (GPR) detection of near-surface targets including anti-personnel mines. When a down-looking antenna is scanned across the surface this reflection produces a strong band in the image, which obscures shallow targets. Imperfections in the system impulse response (e.g., antenna ringing and cable reflections) can produce similar bands. Radar images of buried targets can be degraded by these forms of clutter.

A multi-step process is described to improve imaging of near-surface objects. After range compression of the transmitted pulse, an estimate of the ground-reflected signal is obtained. Slow variations with scan position in the (complex) amplitude and time delay of this reflection are estimated in overlapping piecewise segments using a polynomial model. The resulting reflection estimate is subtracted from the data. Residual clutter is further reduced by using a subspace decomposition technique based on the generalized singular value decomposition (GSVD). In effect, the data are projected onto a space that is orthogonal to the reflection. Both surface reflections and ringing in the system impulse response are removed in this manner. The resulting cleaned data are then used with a conventional imaging algorithm (migration). The procedure is illustrated with data acquired by an OSU-developed GPR at the US Army Fort A.P. Hill, Site 71A.

Keywords: ground penetrating radar, land mines, generalized singular value decomposition, subsurface imaging

1. INTRODUCTION

Radar imaging is often an attractive technique for processing ground penetrating radar (GPR) data. When a downlooking antenna is scanned over a small target, a characteristic hyperbolic arc appears in the time-domain data. Using various algorithms one can refocus the arc energy back to the target location, thereby improving the signal-to-noise ratio and better localizing the target.

Imaging with GPR data is frequently hampered by clutter. This is particularly true for shallow targets. Attempts to detect near-suface buried targets (e.g., anti-personnel mines) using GPR are often limited by the ground reflected wave. When a downward-looking antenna is scanned across the surface this reflection produces a strong band in the time-domain data, which obscures shallow targets. Imperfections in the system impulse response (e.g., antenna ringing and cable reflections) can produce similar bands. Removing these clutter contributions is an essential preprocessing step in obtaining useful data.

In this work we describe a new technique for eliminating ground reflected clutter and system ringing from GPR data. In Section 2 we show that the reflection from uneven ground can be suppressed by subtracting an estimate of the spatially varying return. Once this dominant contribution is removed, residual system clutter can be suppressed using a subspace decomposition technique described in Section 3. In Section 4 we show that images obtained from the clutter-suppressed data are significantly improved. Finally, examples of the clutter reduction and imaging processes are presented in Section 5. The data presented in this work were acquired at the US Army Ft. A.P. Hill (Site 71A) using a GPR developed by the Ohio State University.[1]

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

2. SUPPRESSING THE GROUND REFLECTION

Several clutter reduction techniques have been discussed in the literature, but none of these methods perform satisfactorily when applied to GPR data collected over shallowly buried mines in nonuniform ground. Perhaps the simplest technique is time gating, a concept that is well known in radar and needs no explanation here. Time gating has been used to remove the ground reflection from GPR measurements of deep targets such as buried pipes, unexploded ordnance and tunnels, but it is inappropriate when the targets are very shallow.

For ensemble average subtraction[2] one averages the time-domain signatures acquired as the antenna is scanned over (mostly target-free) ground, and subtracts that average from the data. This approach is effective if the ground is flat and uniform over the antenna path - a situation rarely encountered in the field. It also tends to remove parts of the target if the scan path includes the target response.

In the early-time subtraction technique[3] a band-limited specular reflection is synthesized in the time-domain and subtracted from the data. This technique was reformulated in the frequency domain[4] to improve its performance. In that modification the frequency-domain GPR measurements are represented by a parametric model consisting of a sum of damped exponential terms. The unknown parameters are estimated using a total-least-squares (TLS) Prony technique.

Van der Merwe[5] improved the TLS-Prony method[4] by using an iterative technique. That method involves frequency-domain data collected at each antenna position. It removes spectral components different from a reference (target) signature. Clutter components distinct from that signature are successfully removed both in early time and late time. The need for a priori knowledge of the target signature is not unrealistic, since there usually exists some information regarding the mine population in a given region. For practice, however, a number of reference signatures must be examined to ensure that a detection is not missed. Further, Van der Merwe has shown that even when the reference signature corresponding to the measured mine is utilized, the algorithm can inadvertently remove it if the measured signature is sufficiently distorted due to tilting of the mine, interactions with the ground interface, or differences in the ground properties.

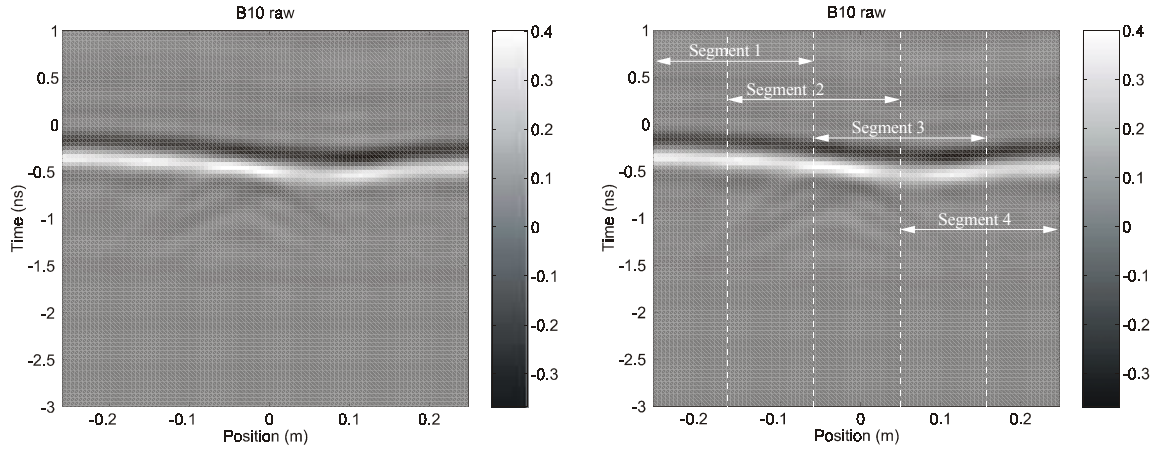
Brunzell[6] proposed a method in which the time delay of each GPR scan was estimated by using nonlinear optimization. By stacking a sequence of system impulse responses with these estimated delays a ground reflection was synthesized and subsequently subtracted from the measured data. This method works well, even with GPR data taken over soil having an uneven profile, provided the magnitude of the ground reflection is uniform across the GPR map. Because of factors such as variations in soil moisture content, variations in the magnitude of the ground reflection are often encountered in data collected over real ground and, hence, Brunzell's method is often unsatisfactory in practice.

Limitations in the existing techniques led us to develop the method proposed here. Our work is an extension of the method of Brunzell, in which spatial variations in both the ground reflection magnitude and delay are simultaneously estimated using nonlinear optimization. Because these variations are relatively smooth functions of position, we model them using low-order polynomials. These polynomials cannot track the rapid spatial variations that occur when a buried mine is encountered and, as a result, the ground reflection estimates are only weakly affected by the presence of targets.

To illustrate the process, consider the response of a one-inch deep TS-50 mine shown in Figure 1(a). This figure shows a plot of signal amplitude as functions of time/depth (vertically) and antenna position (horizontally). The reflection from the air-ground interface, which occurs around -0.5 ns, dominates the data, while the hyperbolic arc produced by the mine is barely visible. The smooth spatial variations in the ground reflection magnitude and the time delay observed in Figure 1(a) are typical of GPR data collected in real ground.

To proceed, split the measured data into N_{seg} spatial segments as indicated in Figure 1(b). Over each segment we approximate the spatial variation in both the amplitude and time delay of the ground reflection by weighted sums of low-order polynomials. The quality of the model fit improves as the segment length is reduced, but when the segment length is comparable to the dimension of a mine, then the presence of a mine improperly biases the model. To avoid this effect the segment length should be restricted to dimensions larger than a mine. To minimize redundancy in the estimate it is preferable to use orthogonal polynomials. In this work the domain of the segment was normalized to the interval [-1,1], and Chebyshev polynomials were used.

We formulate the approximate reflection as follows: Let $A_i(x)$ denote the spatially varying ground reflection amplitude and $B_i(x)$ the time delay of the ground reflection peak over segment i. We approximate $A_i(x)$ and $B_i(x)$



(a) Raw GPR data

(b) Spatial segments

Figure 1. Raw GPR map generated from data collected over a one inch deep TS-50 mine.

as the sums

$$A_i(x) = \sum_{n=0}^{4} a_{in} T_n(x) \tag{1}$$

$$B_i(x) = \sum_{n=0}^{4} b_{in} T_n(x) \tag{2}$$

where $T_n(x)$ are the Chebyshev polynomials defined by the recursive relation

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \qquad n \geq 1 \tag{3}$$

with $T_0(x) = 1$ and $T_1(x) = x$. Because the reflection coefficient at the air-ground interface is typically complex, the coefficients a_{in} in (1) are represented by real and imaginary parts. The coefficients

$$\mathbf{a}_i = \left[a_{i0}^{real}, a_{i0}^{imag}, a_{i1}^{real}, a_{i1}^{imag}, \ldots\right] \tag{4}$$

$$\mathbf{b}_i = [b_{i0}, b_{i1}, \ldots] \tag{5}$$

are unknown parameters to be estimated.

We estimate the frequency-domain ground reflection in spatial segment i as

$$G_i^{est}(x, f) = A_i(x) H(f) \exp^{-j2\pi f B_i(x)} \tag{6}$$

where $H(f)$ is the windowed frequency response of the radar, i.e.,

$$H(f) = \bar{H}(f) W(f) \tag{7}$$

in which $\bar{H}(f)$ is the average frequency spectrum of the raw GPR data within the segment and $W(f)$ is a window applied when creating the time-domain data. In our work a Hanning window was used. The time-domain reflection estimate g_i^{est} becomes

$$g_i^{est}(x, t, \mathbf{a}_i, \mathbf{b}_i) = A_i(x) h(t - B_i(x)) \tag{8}$$

where $h(t)$ is the inverse Fourier transform of $H(f)$. The measured ground reflection over segment i can then be expressed as

$$g_i^{meas}(x, t) = g_i^{est}(x, t, \mathbf{a}_i, \mathbf{b}_i) + n_i(x, t) \tag{9}$$

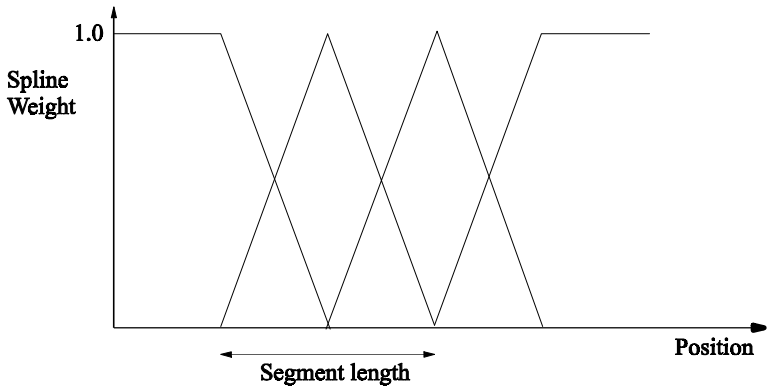


Figure 2. Linear weighting functions used for blending $A_i(x)$ and $B_i(x)$ functions.

where $n_i(x,t)$ represents the modeling error. The coefficients $\mathbf{a}_i$ and $\mathbf{b}_i$ are found by a least-square error process, which requires a nonlinear optimization of

$$[\hat{\mathbf{a}}_i,\ \hat{\mathbf{b}}_i] \ = \ \arg\min_{\mathbf{a}_i,\mathbf{b}_i} \quad \|g_i^{meas}(x,t) - g_i^{est}(x,t,\mathbf{a}_i,\mathbf{b}_i)\|^2 \tag{10}$$

Once the weighting coefficients $\mathbf{a}_i$ and $\mathbf{b}_i$ are estimated, the amplitude term $A_i(x)$ and the delay term $B_i(x)$ of segment i are evaluated using (1) .

To synthesize the ground reflection over the entire antenna scan, it is necessary to merge estimates obtained within each segment. Discontinuities at the edges of these overlapping segments are avoided by blending the $A_i(x)$ and $B_i(x)$ functions using the set of linear weighting functions shown in Figure 2.

Parameter estimation was performed using a non-linear least square error minimization function, *fminu*, in the Matlab Optimization Toolbox. We observed that the optimization routine would not converge for several data sets. Performance was improved by using a recursive approach. The parameters a_{i0}^{real}, a_{i0}^{imag} and b_{i0} were estimated for a zero-order approximation. These estimates were then used as the initial guesses for a_{i0}^{real}, a_{i0}^{imag} and b_{i0} in a first-order approximation, with zero initial values for the remaining parameters, a_{i1}^{real}, a_{i1}^{imag} and b_{i1}. This procedure was repeated until the parameters at the desired order were estimated.

Because our goal is accurate estimation of the ground bounce term, the above procedure was applied only to the depth interval that contained the ground bounce. To identify that interval, the ensemble sum of energy in all range profiles was computed by summing the signature energy in the along-scan direction. Data lying within time indices that encompassed 99% of the total energy were used for the estimation process.

3. REMOVING STATIONARY CLUTTER VIA SUBSPACE DECOMPOSITION

The above-described polynomial model eliminates most of the ground reflected clutter, which dominates GPR data acquired by down-looking antennas. This algorithm, however, does not eliminate other clutter due to antenna ringing, cable mismatch, etc., which typically appear at later times in the form of horizontal bands. Although such clutter can be eliminated, in principle, by proper system design and calibration, residual amounts are invariably present in any real system. These clutter contributions are small compared to the ground reflection, but they are undesirable because they may mask weak returns from deeper mines.

To remove these clutter bands, which are nearly spatially independent, we adopt a subspace decomposition technique proposed by Marinovich[7] for detecting non-stationary signals in stationary noise. For completeness, we summarize that formulation below.

Let W_N $(N \times M)$, $M \leq N$ be a stationary noise (clutter) matrix generated by replicating M copies of a column vector comprising a GPR impulse response waveform. A signal matrix W_S $(N \times M)$ contains the GPR response acquired when an antenna scans over a buried point scatterer. W_S has the familiar hyperbolic arc-shaped response.

Because the noise is assumed spatially stationary (i.e., it has equal columns), W_N has unit rank. Conversely, W_S is non-stationary and has rank greater than unity.

Now consider the generalized singular value decomposition (GSVD) of the signal W_S with respect to the noise W_N. In general, the GSVD yields a matrix representation of the form

$$W_S = U_S D_S X^{-1} \tag{11}$$
$$W_N = U_N D_N X^{-1} \tag{12}$$

Matrices D_S and D_N are diagonal with the elements

$$D_S = \mathrm{diag}(\alpha_i), \qquad 0 \le \alpha_1 \le \alpha_2 \le \ldots \le \alpha_M \tag{13}$$
$$D_N = \mathrm{diag}(\beta_i), \qquad \beta_1 \ge \ldots \ge \beta_q \ge \beta_{q+1} = \ldots = \beta_M = 0 \tag{14}$$

where $q = \mathrm{rank}(W_N)$. Columns of X, denoted by $x^{(i)}$, are the generalized singular vectors of W_N and W_S satisfying

$$\beta_i^2 W_S^T W_S x^{(i)} = \alpha_i^2 W_N^T W_N x^{(i)} \tag{15}$$

Because $\mathrm{rank}(W_N) = q = 1$ for stationary noise, only $\beta_1 \neq 0$. Now isolate the noise subspace in X, U_S, U_N and D_S as follows:

$$X = \begin{bmatrix} x^{(1)} & X_1 \end{bmatrix}, \quad U_S = \begin{bmatrix} u_S^{(1)} & U_{S_1} \end{bmatrix}, \quad U_N = \begin{bmatrix} u_N^{(1)} & U_{N_1} \end{bmatrix} \tag{16}$$

$$D_S = \begin{bmatrix} \alpha_1 & 0 \\ 0 & D_1 \end{bmatrix}, \quad D_N = \begin{bmatrix} \beta_1 & 0 \\ 0 & 0 \end{bmatrix} \tag{17}$$

Post-multiplying Eqns. (11) and (12) by X and using the relations in (16) and (17) yields

$$W_S X = U_S D_S = \begin{bmatrix} \alpha_1 u_S^{(1)} & U_{S_1} D_1 \end{bmatrix} \tag{18}$$
$$W_N X = U_N D_N = \begin{bmatrix} \beta_1 u_N^{(1)} & O \end{bmatrix} \tag{19}$$

which can be expanded to produce

$$W_S x^{(1)} = \alpha_1 u_S^{(1)} \tag{20}$$
$$W_S X_1 = U_{S_1} D_1 \tag{21}$$
$$W_N x^{(1)} = \beta_1 u_N^{(1)} \tag{22}$$
$$W_N X_1 = O \tag{23}$$

Eqns. (20)-(23) imply that all stationary noise energy and a small fraction of signal energy lie in a space defined by the first generalized singular vector $x^{(1)}$. Because α_1 is small, W_S is mapped in (21) with minimum loss of signal power. There is no noise energy in the subspace defined by the remaining generalized singular vectors (X_1). Because X^{-1} can be ill-conditioned, Marinovich suggests using a QR-decomposition

$$(X^{-1})^T = QR \tag{24}$$

with which (11) and (12) can be put in the form

$$W_S Q = U_S D_S R^T \quad \text{and} \tag{25}$$
$$W_N Q = U_N D_N R^T \tag{26}$$

If Q is expressed as $Q = \begin{bmatrix} q^{(1)} & Q_1 \end{bmatrix}$, then a projection matrix $P = Q_1 Q_1^T$ can be defined such that

$$W_S P \approx W_S \quad \text{and} \tag{27}$$
$$W_N P = 0 \tag{28}$$

For each clutter-reduced data set we evaluate the above projection matrix P and apply it to that data. This projection is a useful method for enhancing the detection of small anti-personnel mines, which usually have weak, localized signatures that are easily masked by clutter. The projection reduced stationary clutter, but (as shown below) did not significantly affect the signatures of small mines. Mine signatures with relatively slow spatial variation (for example, signatures of large anti-tank mines such as the TM-46) are adversely affected by this clutter removal step. Because anti-tank mines typically have strong returns, they can often be detected easily and do not require the use of this processing.

4. SUBSURFACE IMAGING

Imaging techniques can be used to focus the energy present in a point target's hyperbolic arc back to a single point. Synthetic aperture radar is based on a closely related concept. The seismic community has develped a technique known as migration for performing this processing. Migration is well suited to GPR processing and is explored here.

Several techniques have been developed for migration. We will employ an approach based on the scalar wave equation. In this work we employ the "exploding reflector" model.[8] Suppose that a monostatic radar collects data $d(x,t)$, where x is the scan dimension. Since the total signal propagation path is twice the one-way path, the signal reflected back to a monostatic radiator is the same as that received by a passive sensor when (1) the reflector becomes a radiator and (2) the velocity of propagation is reduced by a factor of two. The concept simplifies the analysis that follows.

We will employ frequency-wavenumber (ω-k) migration, which is an efficient spectral technique. The derivation of this standard technique appears in several works[9–11] and only the final result appears below. The time-dependence $e^{+i\omega t}$ is implied and suppressed in what follows.

Consider a scalar 2-D field $f(x,z,t)$ produced by some source distribution in a homogeneous medium with propagation velocity v. The spectral representation of our measurements $f(x,z=0,t)$ can be computed as

$$F(k_x, z=0, \omega) = \int dx \int dt f(x, y=0, z=0, t) e^{-i\omega t + ik_x x} \tag{29}$$

Under the exploding reflector hypothesis, the field distribution of interest is $f(x,z,t=0)$. It can be shown that this distribution is given by

$$f(x,y=0,z,t) \quad = \quad \frac{1}{(2\pi)^2} \int d\omega \int dk_x e^{+i\omega t} e^{-ik_x x} F(k_x, y=0, z=0, \omega) e^{-iK_z z} \tag{30}$$

where

$$K_z = \sqrt{k^2 - k_x^2} \tag{31}$$

where $k = 2\omega/v$ is the wavenumber in the medium (under the exploding reflector hypothesis). We convert the integral over ω to an integral over K_z by deforming the contour of integration. We find

$$f(x,y=0,z,t) = \frac{1}{4\pi^2} \int dK_z \int dk_x e^{+i\omega(K_z)t} e^{-ik_x(x-x')} F(k_x, y=0, z=0, \omega(K_z)) e^{-iK_z z} \frac{d\omega(K_z)}{dK_z} \tag{32}$$

where

$$\begin{aligned} \omega(K_z) &= \frac{v}{2}\sqrt{k_x^2 + K_z^2} \\ &= (K_z v/2)\sqrt{1 + (k_x/K_z)^2} \\ &= (K_z v/2) + (K_z v/2)(\sqrt{1 + (k_x/K_z)^2} - 1) \end{aligned} \tag{33}$$

Assuming there are no singularities in F we can deform the path of integration in K_z to the real line, which allows equation (32) to be written as a 2-D Fourier transform

$$f(x,y=0,z,t=0) = \int dK_z \int dk_x e^{-ik_x(x-x')} G(k_x, K_z) e^{-iK_z z} \tag{34}$$

where

$$G(k_x, K_z) = \frac{1}{4\pi^2} \frac{d\omega(K_z)}{dK_z} F(k_x, 0, 0, \omega(K_z)) \tag{35}$$

Thus, for the case of homogeneous ground, we can find the reflectivity function as a 2-D Fourier transform of the measured data, suitably interpolated.

5. EXAMPLE RESULTS

The effectiveness of our new technique is best judged from examples, which we present here.

When the polynomial-based model for surface reflection is applied to a one-inch deep TS-50 mine (cf. Figure 1(a)), it produces the estimated ground reflection shown in Figure 3. Subtracting this estimate from the raw data gives the map of Figure 4. The ground reflection term, which appeared near -0.5 ns in the raw data, is no longer visible in the clutter removed data. The most important observation to make from Figure 4 is that the signature of this very shallow mine is preserved and enhanced, something that is difficult to achieve for such shallow targets with previously described clutter removal techniques. Figure 4 contains a weak band of residual clutter between -1.5 ns and -1.0 ns. Applying the subspace-decomposition technique described above produces the result shown in Figure 5. The clutter band noted above is diminished but not totally eliminated, because this clutter is not entirely spatially invariant as assumed in the formulation.

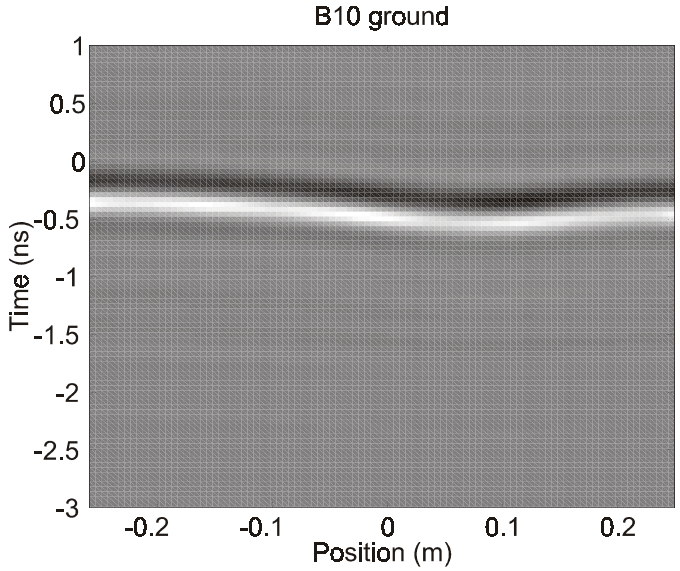


Figure 3. The ground reflection clutter profile estimated by applying the polynomial model-based approach to measured GPR data.

Next, consider a VS-50 mine buried at a three-inch depth. We will compare the proposed approach and ensemble average subtraction. Figure 6(a) is the raw data, in which the mine signature is barely visible, and Figure 6(b) was obtained after ensemble average subtraction. Because the ground reflection profile is nonuniform, average subtraction is unsatisfactory. When the raw GPR data of Figure 6(a) are processed by applying the new clutter reduction technique, the data in Figure 6(c) are obtained. The improvement in signal-to-clutter ratio with the proposed method is clear.

Some examples of migration imaging are shown in Figures 7 and 8 for data acquired over a flush-buried VS-50 mine. The benefits of the clutter reduction process are evident upon comparing Figures 7 with the raw data and Figures 8 with the clutter-reduced data. Although migration will improve the detection of shallow targets, even when the ground clutter is present, the presence of the mine is much more evident in the clutter-reduced data.

6. CONCLUDING REMARKS

We have described a technique for suppressing ground-reflection clutter that appears in down-looking GPR systems. This technique greatly increases the signal-to-clutter ratio of near-surface targets, and it improves imaging of such mines using migration. We presented examples of both clutter reduction and imaging before and after clutter reduction.

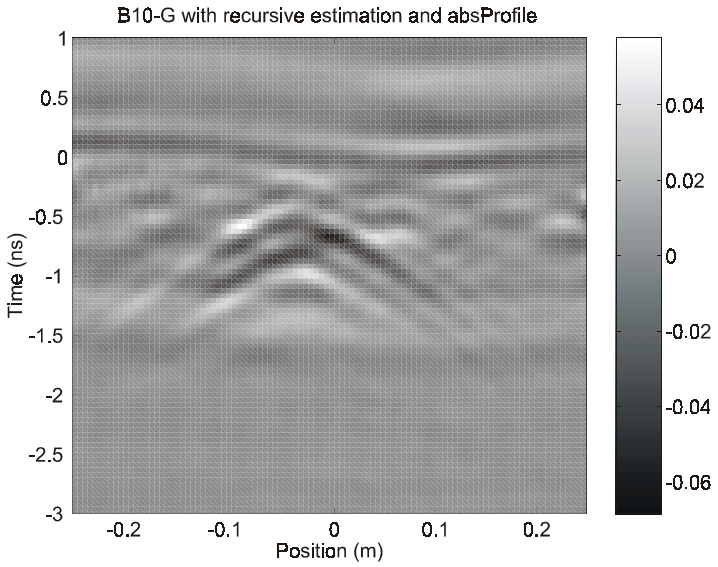


Figure 4. Data obtained by removing the estimated ground-reflection from the measured data. This data corresponds to a one inch deep TS-50 mine.

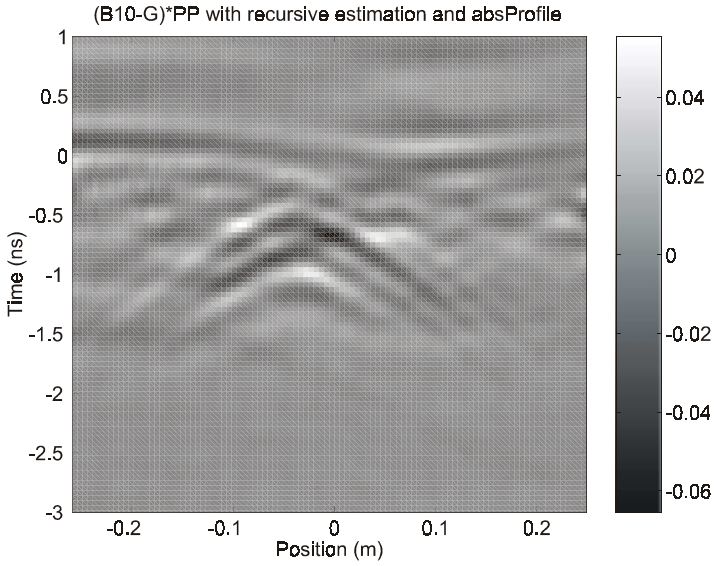
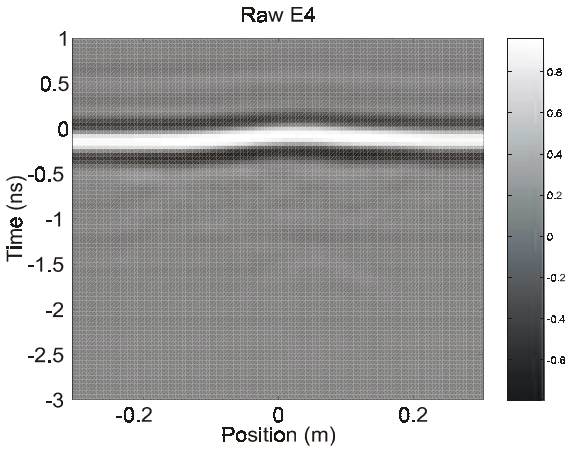
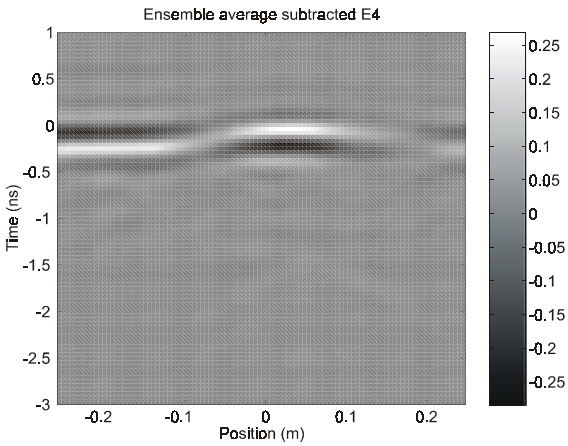


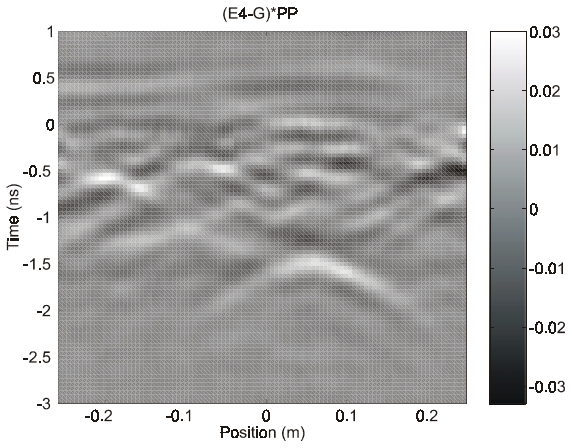
Figure 5. Data obtained by projecting the ground-reflection removed data to reduce the spatially stationary clutter.



(a) Raw GPR data

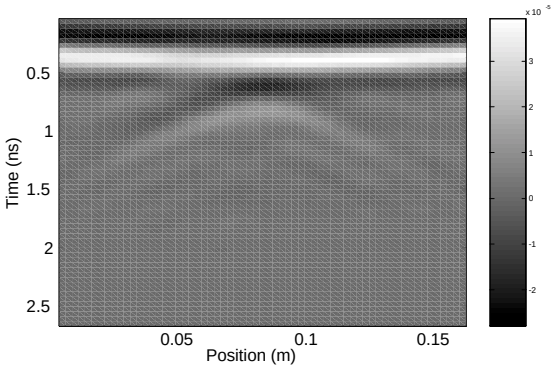


(b) Data obtained using ensemble average subtraction

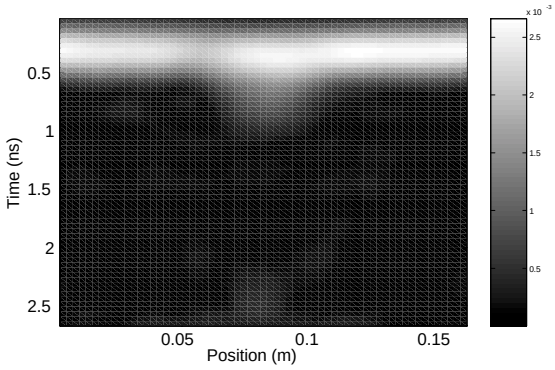


(c) Data obtained using the proposed approach

Figure 6. Comparison of raw and clutter-reduced data generated from measurements over a three inch deep VS-50 mine.

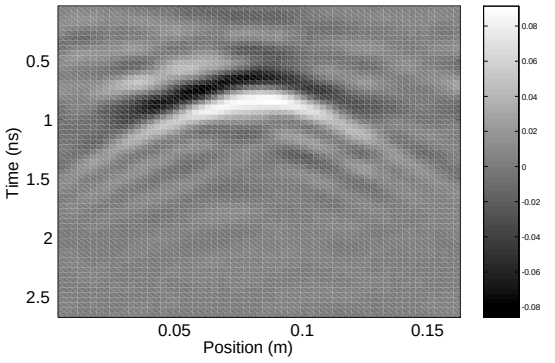


(a) Raw GPR signature (no clutter redution) of flush VS-50 mine (B2).

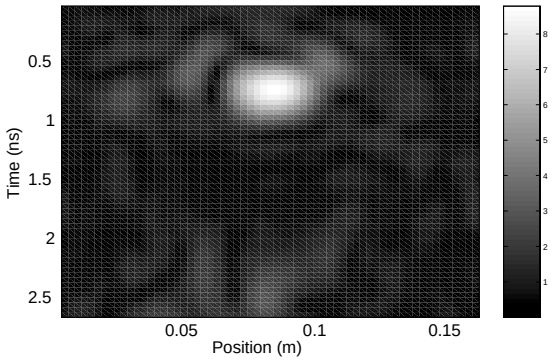


(b) Migration image of flush VS-50 without clutter reduction.

Figure 7. Migration of GPR data without clutter reduction.



(a) Clutter-reduced GPR signature of flush VS-50 mine (B2).



(b) Migration image of flush VS-50 with clutter reduction.

Figure 8. Migration of GPR data without clutter reduction.

The proposed technique is a two-phase method. In the first phase the surface reflection is estimated using low-order polynomals for the reflection complex amplitude and time delay. After subtracting the reflection estimate, the residual clutter, some of which arises from defects in the radar impulse response, is suppressed using a subspace decomposition technique. The second phase should permit the GPR to detect objects at greater depth.

The proposed method has a few limitations. Although the clutter reduction technique can remove the ground reflection without affecting the signatures of small, shallow or surface targets, it will remove slowly varying signatures due to large, shallow mines. This limitation is not overly restrictive, since large anti-tank mines are usually deeply buried. In addition, a very rough surface will produce a rapidly varying surface reflection, which may be proorly modeled by the low-order polynomials employed here. Finally, we noted previously that large flat mines can have a large spatially stationary component, which can be improperly removed by the subspace decomposition technique.

7. ACKNOWLEDGMENTS

The authors would like to thank the staff of E-OIR Measurements, Inc. and the Joint UXO Coordinating Office for their support during data collections at Fort A. P. Hill, Site 71A.

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

REFERENCES

1. S. Nag, L. Peters Jr., I. J. Gupta, and C. C. Chen, "Ramp response for the detection of anti-personel mines," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, R. E. Dugan, Editors, Proceedings of SPIE, vol. 3710, pp. 1313-1322, 1999.
2. R. Caldecott, J. D. Young, J. P. Hall, and A. J. Terzuoli, "An underground obstacle detection and mapping system," Tech. Rep. EL-3984, The Ohio State University, ElectroScience Laboratory, Columbus, Ohio, March 1985.
3. C. C. Chen, J. D. Young, L. Peters Jr.,, and F. Paynter, "Radar signature measurements of buried, unexploded ordnance targets," Tech. Rep. 727388-2, The Ohio State University, ElectroScience Laboratory, Columbus, Ohio, Feb. 1994.
4. I. J. Gupta, A. van der Merwe, and C. C. Chen, "Extraction of complex resonances associated with buried targets," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, J. T. Broach, Editors, Proceedings of SPIE, vol. 3392, pp. 1022-1032, 1998.
5. A. van der Merwe, *A Novel Signal Processing Technique for Clutter Reduction in GPR Measurements of Small, Shallow Land Mines*. PhD thesis, The Ohio State University, Columbus, OH, 1999.
6. H. Brunzell, *Signal Processing Techniques for Detection of Buried Landmines Using Ground Penetrating Radar*. PhD thesis, Chalmers University of Technology, Sweeden, 1998.
7. N. M. Marinovich, "Detection of non-stationary signals in colored noise via time-frequency subspaces," in *Time-Frequency Signal Analysis*, B. Boashash, ed., ch. 13, Longman Cheshire, Melbourne, Australia, 1992.
8. J. F. Claerbout, *Imaging the Earth Interior*, Oxford: Blackwell Scientific Publications, 1985.
9. R. Stolt, "Migration by Fourier transform," *Geophysics* **43**, pp. 23–48, 1978.
10. J. Gazdag and P. Sguazzero, "Migration of seismic data," *Proceedings of the IEEE* **72**, pp. 1302–1325, October 1984.
11. C. Cafforio, C. Prati, and E. Rocca, "SAR data focusing using seismic migration techniques," *IEEE Trans. Aerospace and Electronic Systems* **27**, pp. 194–206, March 1991.

Numerical Simulation of Thermal Signatures of Buried Mines over a Diurnal Cycle

İbrahim Kürşat Şendur and Brian A. Baertlein

The Ohio State University, ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

Three-dimensional thermal and radiometric models have been developed to study the passive IR signature of a land mine buried under a rough soil surface. A finite element model is used to describe the thermal phenomena, including temporal variations, the spatial structure of the signature, and enviromental effects. The Crank-Nicholson algorithm is used for time-stepping the simulation. The mine and the surroundings are approximated by pentahedral elements having linear interpolation functions. The FEM grid for the soil includes a random rough surface having a normal probability density and specified covariance function. The mine is modeled as a homogeneous body of deterministic shape having the thermal properties of TNT. Natural solar insolation (both direct and atmospheric-scattered components) and the effects of convective heat transfer are represented by linearized boundary conditions. The behavior over a periodic diurnal cycle is studied by running the simulation to steady state. Finite element solutions for the thermal emissions are combined with reflected radiometric components to predict the signatures seen by an IR camera. Numerical simulations are presented for a representative target, a 25 cm anti-tank mine simulant developed by the US Army. The temporal evolution of the temperature distribution and IR signature are presented for both smooth and rough surfaces.

Keywords: Thermal infrared imagery, land mines, finite element method, heat transfer, radiometry, surface roughness, thermal model, numerical simulation

1. INTRODUCTION

In this paper three-dimensional thermal and radiometric models are developed for soil containing a buried land mine. The objective of this work is to improve our understanding of the physical processes that produce the thermal mine signature. It is our hope that this model will also improve comprehension and interpretation of thermal IR imagery, suggest more efficient usage of IR sensors, and lead to the development of more robust signal processing techniques.

IR detection of land mines has a long history, and recently there has been extensive work in this area. Most existing studies, however, are concerned with signal processing algorithms or the performance achieved by a particular sensor and not with the underlying physical bases of the signatures. Although many experimental and empirical studies of these signatures have been performed, the underlying phenomena are still poorly understod. Prior studies and anecdotal evidence have implicated several physical phenomena including thermal effects related to the mine, soil disturbances, and changes in soil moisture. The presence of the mine also indirectly influences the vitality of overlying vegetation, which may have a role in the signature. Among the physics-related papers on thermal imaging we note the work of LeSchack and Del Grande[1] who describe the effects of emissivity on target signatures in the context of a two-color IR system. Janssen et al.[2] reported a study of basic phenomenology in which several sensors were used to examine the time evolution of signatures for both buried and surface mines. A related study is that of Maksymomko et al.[3] who measured the temperatures of both live and surrogate mines through multiple diurnal cycles. Simard[4] found a linear relation between the apparent temperature of buried mines and the soil temperature gradient on gravel roads. Russel et al.[5] showed that this gradient could be predicted from remote soil temperature measurements, making it possible to assess the potential effectiveness of an IR sensor prior to use. The possibility of using polarization information was examined briefly by Barbour et al.[6] and Larive et al.[7] DiMarzio et al.[8] present spectroscopic data collected by a non-imaging sensor. Pregowski and Swiderski[9] present a qualitative discussion of the effects of soil cover and water content. Another work, obliquely related to the effort proposed here, is that of Del Grande et al.,[10] who discuss the use of computed tomography and a finite difference thermal analysis code

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

(TOPAZ3D) for imaging structural flaws in materials. Discussions of the physics of non-thermal hyperspectral sensors are rare. Among these works we note a paper by DePersia et al.[11] in which the spectral signatures of various types of soil disturbances are discussed. A simulation tool for surface mines in the UV to near IR is described by Flynn et al.[12]

While models for a mine's thermal signature are not well developed, there has been progress in modeling the temperature of homogeneous soil over a diurnal cycle.[13–18] In a previous unpublished work[19] the authors extended those efforts to develop one-dimensional models for natural solar heating of soil and mines. Those models address both homogeneous earth and layered earth containing strata with mine-like properties. We developed analytical models to predict the temperature distribution on the surface and at depth. In spite of the crude approximations inherent in a 1-D model, the results predict many phenomena observed in thermal mine imagery, including the presence of cross-over times, the time lag between solar flux and mine signatures, signature differences between thermally conducting and insulating mines, and the effects of depth on IR detection.

This work begins with a discussion of the underlying physics in Section 2. The mathematical formulation of the thermal model is given in Section 3, which also presents the FEM formulation and assumptions regarding the boundary conditions (whereby the source is introduced). The radiometric models are discussed in Section 4. Models for the sunlight and skylight, the rough surface approximation for the soil surface, and the radiometric quantities are discussed. Simulation results for a representative simulant anti-tank mine are presented in Section 5. Concluding remarks appear in Section 6.

2. PHYSICAL DESCRIPTION OF PROCESSES

Thermal emission from the soil over a buried mine depends on the surface temperature and emissivity. The surface temperature is a result of (1) the solar-driven, diurnal flow of thermal energy into and out of the soil, (2) the effect of the buried mine on that flow, and (3) IR radiation from the surface. Three heat transfer mechanisms are involved in these processes. Conduction occurs in the soil and at the soil-mine interface. At the soil-air interface convective and radiative mechanisms are present. Convection involves the transfer of heat between the soil and air and is greatly affected by wind. Radiative heat transfer involves direct solar radiation, atmosphere-scattered solar radiation, and thermal radiation between the earth and atmosphere. All of these processes are time varying (and herein, assumed periodic) over the diurnal cycle. In this work we measure time t with respect to midnight local time. Mathematical descriptions of the phenomena are as follows:

The temperature distribution in the soil and mine is described by the three-dimensional heat flow equation

$$\rho(\mathbf{r})C_p(\mathbf{r})\frac{\partial T(\mathbf{r},t)}{\partial t} = \nabla\cdot(\mathcal{K}(\mathbf{r})\nabla T(\mathbf{r},t)), \tag{1}$$

where $T(\mathbf{r},t)$ [K] is the temperature distribution, $\mathcal{K}$ [W m^{-1} K^{-1}] is the thermal conductivity of the material, C_p [J kg^{-1} K^{-1}] is the specific heat at constant pressure and ρ [kg/m^3] is the density of the material. Convective and radiative heat transfer drive the thermal flow and are introduced via a boundary condition at the soil-air interface. Prior one-dimensional analyses of the problem have addressed this topic and are relevant here.[13–15,17,18,20] We write

$$\mathcal{F}_{net}(t)\big|_{\text{ interface}} = \mathcal{F}_{sun}(t) + \mathcal{F}_{sky}(t) - \mathcal{F}_{sh}(t) - \mathcal{F}_{gr}(t), \tag{2}$$

where $\mathcal{F}_{net}$ is the net heat flux into the ground, $\mathcal{F}_{sun}$ is the incident solar radiation reduced by cloud extinction, atmospheric absorption, soil albedo and the cosine of the zenith angle, $\mathcal{F}_{sky}$ is the sky brightness with a correction for cloud cover, $\mathcal{F}_{sh}$ is the sensible heat transfer from land to atmosphere due to convection, and $\mathcal{F}_{gr}$ is the gray body emission from soil's surface.* Watson[13] used

$$\mathcal{F}_{sun}(t) = S_0(1-A)(1-C)\mathcal{H}(t) \tag{3}$$

to approximate the short-wavelength solar flux. In Eq. (3) $S_0 = 1353$ [W/m^2] is the solar constant, A is the ground albedo, C is a factor that accounts for the reduction in solar flux due to cloud cover, and $\mathcal{H}(t)$ is the local insolation function. Explicit expressions for the local insolation $\mathcal{H}(t)$ can be found in the literature.[13,22] The quantity $\mathcal{F}_{sky}$ is the long-wavelength radiation from the atmosphere given by Stefan's law, $\mathcal{F}_{sky}(t) = \sigma T_{sky}^4(t)$

*In formulating this expression we have ignored changes in the state in the medium, which would appear as a latent heat term. Such terms have been considered previously by Kahle[15] and England.[21]

where $\sigma = 5.67 \times 10^{-8}$ [W m^{-2} K^{-4}] is the Stephen-Boltzman constant and T_{sky} [K] is an effective sky radiance temperature. The heat loss due to ground radiation is given by a similar expression, $\mathcal{F}_{gr}(t) = \mathcal{E}\sigma T_{gr}^4(t)$ where $\mathcal{E}$ is the mean emissivity of the surface and $T_{gr}(t)$ is the soil temperature at the soil-air interface. Kahle[15] modeled sensible heat transfer between the surface and atmosphere as

$$\mathcal{F}_{sh}(t) = \rho_a c_{pa} C_d (W(t) + 2)(T_{air}(t) - T_{gr}(t)). \tag{4}$$

where $\rho_a = 1.16$ [kg/m^3] is the density of air, $c_{pa} = 1007$ [J kg^{-1} K^{-1}] is the specific heat of air, C_d is the wind drag coefficient (chosen to be 0.002), and $W(t)$ is the wind speed. Kahle,[15] drawing on measurements by Kondratyev,[23] formulated an empirical model for the air temperature, which was later modified by England[17] to yield the expression

$$T_{air}(t) = T_{0,air} - T_{del} \cos(2\pi(t-2)/24), \tag{5}$$

where t is the local time (in hours) measured from midnight, and $T_{0,air}$ and T_{del} are estimated from local meteorological data available from the National Weather Service. Kahle also proposed a model for the sky temperature, which was modified by England[17] using Brundt's formula[15]

$$T_{sky}(t) = T_{air}(t)(0.61 + 0.05\sqrt{w})^{0.25}, \tag{6}$$

where w is the water vapor pressure in mmHg.

The radiative components seen by the IR camera include (1) thermal radiation from the soil surface, (2) soil-reflected sunlight, (3) soil-reflected skylight (thermal emission from the atmosphere and sunlight scattered by particles and air molecules), and (4) a negligible amount of thermal emission from the atmosphere directly to the camera. Thermal emission from the soil surface depends on the surface temperature and emissivity. The surface temperature distribution can be obtained from the solution of Eq. (1) with the boundary condition given by Eq. (2). The power emitted by the soil surface is determined from the soil's spectral radiance function. The radiance function for that graybody emitter is given by the product of the surface emissivity $\mathcal{E}$ and the blackbody spectral radiance $L_\lambda(T)$ described by Planck's radiation law

$$L_\lambda(T) = \frac{2c^2h}{\lambda^5} \frac{1}{e^{h\nu/kT} - 1} \qquad [\text{W m}^{-2}\text{sr}^{-1}\mu\text{m}^{-1}], \tag{7}$$

where $h = 6.63 \times 10^{-34}$ [J s] is Planck's constant, ν [Hz] is the frequency of the optical radiation, λ [m] is the wavelength of that radiation, $k = 1.38 \times 10^{-23}$ [J K^{-1}] is Boltzmann's constant, and c is the speed of light [m s^{-1}]. We write the radiation emitted by the surface as the product

$$L_{surf}(\lambda, \theta, \phi, \mathbf{r}) = \mathcal{E}(\lambda, \theta, \phi, \mathbf{r}) L_\lambda(T(\mathbf{r})), \tag{8}$$

where $\mathcal{E}(\lambda, \theta, \phi, \mathbf{r})$ is the spectral directional emissivity. In this work we make the approximation that soil is a diffuse emitter, i.e., the emissivity is independent of direction.

From the above expressions it is clear that emissivity has a direct effect on the thermal signature. In addition, grass and other forms of ground cover have a strong effect on the heat flow process and on the reflection and radiation of thermal energy. A study of the emissivity of natural materials has been made by Salisbury and D'Aria.[24,25] The reflectivity of these materials can be obtained from the emissivity values by using the Kirchoff's law for an opaque body, $\mathcal{R} = 1 - \mathcal{E}$. Using the diffuse emitter approximation, the received radiance can be written as

$$L_{rec}(\lambda, \mathbf{r}) = \mathcal{E}(\lambda, \mathbf{r}) L_{surf}(\lambda, \mathbf{r}) + \mathcal{R}(\lambda, \mathbf{r}) \left[L_{sun}(\lambda) + L_{sky}(\lambda) \right]. \tag{9}$$

This radiance expression permits us to calculate the power incident on an IR detector surface D from the radiating soil surface S. The power incident on the detector can be found by

$$\Phi = \int_{\lambda_1}^{\lambda_2} d\lambda \int_D dD \int_S dS L_{rec}(\lambda, \mathbf{r}) \frac{\cos\theta_1 \cos\theta_2}{R^2}, \tag{10}$$

where R is the distance between a point on surface S and a point on the detector D, $\cos\theta_1$ and $\cos\theta_2$ are the projection of normal vectors for the surfaces D and S in the direction of the radiation, and the spectral band of interest is λ_1-λ_2. The IR image is formed by Eq. (10).

3. THERMAL MODELING USING THE FINITE ELEMENT METHOD

The finite element method (FEM) is an efficient computational technique, widely used for the solution of heat transfer problems. The essential feature of this technique is the spatial discretization of the computational domain. In what follows we denote the computational volume by Ω and its boundary by Γ. We take this volume to have a parallelepiped shape (with the exception of the rough surface at the soil-air interface). The boundary planes in x and y of the computational volume are denoted Γ_x and Γ_y respectively. The lower boundary plane is denoted Γ_z. An individual element of the subdivided volume and its boundary are represented by Ω_e and Γ_e, respectively. The element boundaries on the soil-air interface are represented with $\Gamma_{e,air}$.

We require the steady state solution of Eq. (1) with suitable boundary conditions when driven by a time-varying source. The temperature ranges of interest are relatively small (at most, a few tens of Kelvins), and over this limited range the thermal properties of soil and mine are assumed to be independent of temperature and of time. As a result, phenomena such as the drying of soil and the movement of soil water are neglected in the current model. We also assume that the temporal and spatial dependencies of $T(\mathbf{r},t)$ are separable within an element.

Over each element Ω^e we approximate the temperature distribution as

$$T^e(\mathbf{r},t) = \sum_{i=1}^{N_e} T_i^e(t)\phi_i^e(\mathbf{r}), \tag{11}$$

where $\phi_i^e(\mathbf{r})$ are user-prescribed interpolation functions, N_e denotes the number of nodes over the subdivided finite element, and $T_i^e(t)$ are the nodal temperatures which comprise the unknown coefficients in our representation. Using this expression in Eq. (1) and enforcing weak equality with the weighting function $w(\mathbf{r})$ we obtain

$$\int_{\Omega_e} w(\mathbf{r})\left[\rho(\mathbf{r})C_p(\mathbf{r})\frac{\partial T^e(\mathbf{r},t)}{\partial t} - \nabla\cdot(\mathcal{K}(\mathbf{r})\nabla T^e(\mathbf{r},t))\right]d\mathbf{r} = 0 \qquad \forall\Omega^e \in \Omega. \tag{12}$$

Continuity in temperature and flux are enforced on the surface Γ^e shared by adjacent elements. On Γ external boundary conditions (described below) are imposed. We use the Galerkin formulation, in which the weighting functions are chosen to be identical to the interpolation functions. This formulation minimizes the mean error over the volume and is the best approximation in the variational sense.

The boundary conditions imposed on the problem are determined from physical considerations. For the transverse coordinates x and y, we assume the presence of an infinite rectilinear array of mines (e.g., a mine field). As a result of symmetry arguments, we can set the normal derivatives of T equal to zero on Γ_x and Γ_y. Although an infinite array of mines is being analyzed, for a sufficiently large computational volume, there is negligible interaction between mines, and the results are a good approximation to the response of an isolated mine. For the boundary condition on Γ_z (at the bottom face of the computational volume) we observe that at sufficiently large depths (a few tens of centimeters) the temperature is independent of time over a diurnal cycle. This typical "diurnal depth" can be predicted from one-dimensional models of soil[19] using thermal characteristics found in the literature.[26–28] On the basis of these findings, we can set T equal to a constant on Γ_z.

It remains to specify the boundary condition at the soil-air interface. As noted in Eq. (2) this condition involves a nonlinear function of T. Over the limited temperature range of interest, however, the result can be linearized using a technique described by Watson.[13] We write

$$\frac{\partial T(\mathbf{r},t)}{\partial z} \approx T(\mathbf{r},t)(\rho_a c_{pa} C_d(W(t)+2) + 4\epsilon\sigma T_{sky}^3(t)) - \mathcal{F}_{sun}(t) - \rho_a c_{pa} C_d(W(t)+2)T_{air}(t) - 4\epsilon\sigma T_{sky}^4(t). \tag{13}$$

This change renders the entire problem linear.

A boundary condition at $t = 0$ is also required. Although the solution procedure described below will converge to a steady state value using $T(\mathbf{r}, t = 0) = 0$ as the initial condition, the computation time can be reduced if the average depth-dependent temperature of homogeneous soil is used as a starting value. A suitable estimate is given by

$$T(\mathbf{r}, t=0) = \frac{\frac{1}{L_p}\int_{L_p}\mathcal{F}_{sun}(t)dt + 4\epsilon\sigma\overline{T}_{sky}^4 + \rho_a c_{pa} C_d(\overline{W}+2)\overline{T}_{air}}{4\epsilon\sigma\overline{T}_{sky}^3 + \rho_a c_{pa} C_d(\overline{W}+2)} \tag{14}$$

In Eq. (14) L_p denotes the 24 hour period and $\overline{T}_{sky}$, $\overline{T}_{air}$, and $\overline{W}$ correspond to the average values of sky temperature, air temperature and wind speed, respectively, over this period. A similar expression neglecting the convection term was derived by Watson.[13]

After substituting Eq. (11) in Eq. (12), incorporating the boundary conditions and performing some mathematical manipulations, the unknown nodal coefficients $T_i(t)$ can be expressed as a matrix equation

$$\overline{\mathbf{M}}\,\dot{\mathbf{T}} + \overline{\mathbf{K}}\,\mathbf{T} = \mathbf{F}, \tag{15}$$

in which boldface letters represent vectors, overlined boldface letters represent matrices and the superposed dot on $\mathbf{T}$ denotes the derivative with respect to time. The elements of these matrices and vectors are given by

$$M_{ij} = \int_{\Omega_e} \rho(\mathbf{r}) C_p(\mathbf{r}) \phi_i^e(\mathbf{r}) \phi_j^e(\mathbf{r}) d\mathbf{r}, \tag{16}$$

$$\begin{aligned} K_{ij} &= \int_{\Omega_e} \mathcal{K}(\mathbf{r}) \left[\frac{\partial \phi_i^e(\mathbf{r})}{\partial x} \frac{\partial \phi_j^e(\mathbf{r})}{\partial x} + \frac{\partial \phi_i^e(\mathbf{r})}{\partial y} \frac{\partial \phi_j^e(\mathbf{r})}{\partial y} + \frac{\partial \phi_i^e(\mathbf{r})}{\partial z} \frac{\partial \phi_j^e(\mathbf{r})}{\partial z} \right] d\mathbf{r} \\ &+ \int_{\Gamma_{e,air}} \left[4\epsilon\sigma T_{sky}^3(t) + \rho_a c_{pa} C_d (W(t)+2) \right] \phi_i^e(\mathbf{r}) \phi_j^e(\mathbf{r}) dS, \end{aligned} \tag{17}$$

and

$$F_i = \int_{\Gamma_{e,air}} \left[4\epsilon\sigma T_{sky}^4(t) + \rho_a c_{pa} C_d (W(t)+2) \sigma T_{air}(t) + \mathcal{F}_{sun}(t) \right] \phi_i^e(\mathbf{r}) dS. \tag{18}$$

The spatial discretization was accomplished using pentahedral elements. A sample spatial discretization is shown in Fig. 1. This choice of element allows us to represent an arbitrary rough surface and it is amenable to the calculations required in our radiometric model. Linear interpolation functions $\phi_i^e(\mathbf{r})$ were chosen.

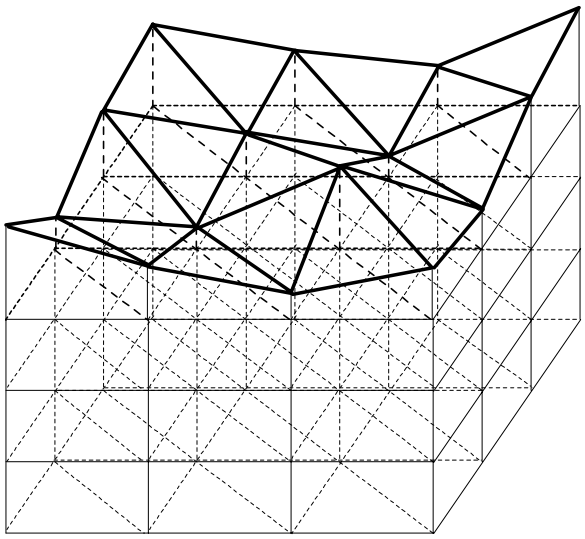


Figure 1. Sample spatial discretization of the computational domain. The volume is subdivided into pentahedral elements resulting in a triangular model for the soil surface.

Temporal discretization of Eq. (15) leads to a finite difference formulation, which is solved using the Crank-Nicholson scheme. We have

$$(2\,\overline{\mathbf{M}} + \Delta t\,\overline{\mathbf{K}}_{t+\Delta t})\mathbf{T}_{t+\Delta t} = \Delta t\,(\mathbf{F} + \mathbf{F}_{t+\Delta t}) + (2\,\overline{\mathbf{M}} + \Delta t\,\overline{\mathbf{K}}_t)\mathbf{T}_t. \tag{19}$$

Eq. (19) is solved at each time step to evaluate the nodal temperatures $\mathbf{T}_{t+\Delta t}$. Employing an LU decomposition technique for this solution is inefficient because the matrices $\overline{\mathbf{M}}$ and $\overline{\mathbf{K}}$ are sparse. Banded matrix storage provides a partial remedy for this problem, but limitations in node numbering usually results in ineffecient memory use. An efficient sparse solver is employed in our work.

4. RADIOMETRIC MODELING

To predict the mine signatures seen by an IR camera, a radiometric model must be combined with the thermal FEM model. The radiometric model describes the spatial and spectral characteristics of the fluence transmitted from the soil surface to the IR detector. The model requires the spectral behaviour of the source (sunlight and skylight), a rough surface model to describe the soil-air interface, and the spectral behaviour of the soil. A flat spectral model is assumed for soil in this paper. A more realistic model is being investigated.

4.1. Sunlight and Skylight Models

The radiance due to sunlight and skylight comprise the source function for our themal model, and reflections from the soil are a radiometric component of interest. The total radiation over the entire spectrum due to sunlight and skylight are required for the thermal model and are given by $\mathcal{F}_{sun}$ and $\mathcal{F}_{sky}$, respectively. More detailed spectral radiance models are required to describe the reflected components seen by the IR camera.

The radiance L_{sun} can be modeled with reasonable accuracy by a blackbody radiator. Neckel and Labs[29] show that a blackbody source at ≈5800 K represents a good numerical fit to exoatmospheric measured radiance data in the visible and near IR bands. Thekaekara and Drummond[30] summarized data regarding the components of the total radiation and their spectral distribution, and they proposed standard values for them. Measured solar radiation at different sites and models for meteorological and climatic factors are presented in a summary by Dogniaux[31] prepared for the European Community Programme on Solar Energy. The measurement techniques and instruments are described in Coulson.[32] Our radiometric model of L_{sun} comprises the blackbody source approximation mentioned above. Atmospheric transmittance is included in our model using the data given by Allen[33] and approximated in a functional form by Watson.[13]

The quantity L_{sky} depends on the composition of the local atmosphere and is more difficult to model accurately. It comprises wide-angle Rayleigh scattering by molecular constituents (very weak at IR wavelengths), small-angle Mie scattering by aerosols, and thermal radiation from the warm atmosphere. The thermal contribution is most important at long wavelengths, and can be approximated by a blackbody radiator at the local air temperature. Although there has been extensive work in developing models for sunlight and skylight, local changes in atmospheric particulates and water vapor content can greatly affect model accuracy. In addition, a number of molecular species (H_2O, CO_2, CO, N_2O, O_3, and CH_4) have absorption bands in the infrared, which affect the incident solar radiation.[26] Our model for L_{sky} comprises a blackbody approximation at the local air temperature.

4.2. Rough Surface Approximation

Thermal emission and reflection of ambient light from a rough surface are of considerable importance in predicting the signature seen by an IR sensor. Variations in the surface normal direction affect the natural solar insolation falling on the surface, which influences both thermal heating and surface reflections. IR imagery of rough surfaces may also contain self-shadowing of the surface as a result of the topology. To address these effects we incorporated surface roughness into our thermal and radiometric models.

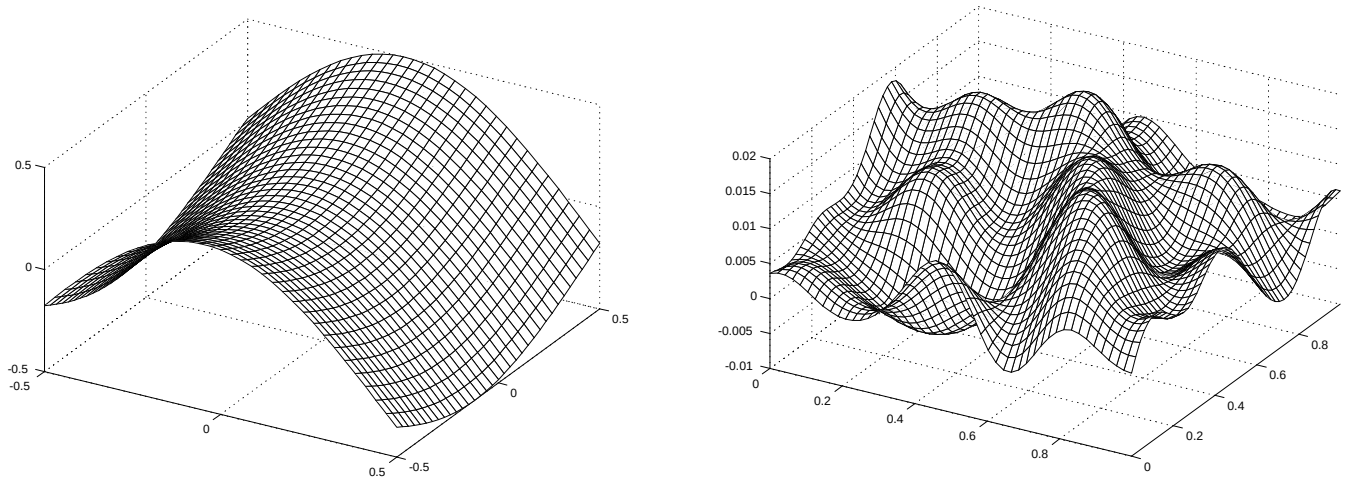
Rough surfaces were constructed by specifying random elevations for points in a rectilinear grid defined by the top surface of our FEM computational volume. The surface is then represented by fitting triangles to the grid points, as shown in Fig. 1. The same representation is used by both the thermal and radiometric models, which permits the modeling work to be decoupled to a degree. Predicting IR imagery over a rough soil surface is a computationally expensive process, because of the visible surface determination problem. Techniques devised for computer graphics[34] offer significant reductions in the computational complexity.

Consider now the problem of generating a rough surface with a random height profile $z = f(x, y)$ having zero mean. For a Gaussian distributed surface this problem can be addressed by expressing the surface in the spatial spectral domain (k_x, k_y). The spectrum $W(k_x, k_y)$ (or equivalently, the covariance or structure function) of the surface is assumed known. We also assumed a Gaussian form for the spectrum, namely:

$$W(k_x, k_y) = \frac{L_x L_y h^2}{4\pi} \exp\left(- \frac{L_x^2 k_x^2}{4} - \frac{L_y^2 k_y^2}{4} \right), \tag{20}$$

where L_x and L_y are correlation lengths in the x and y directions respectively, and h is the surface rms height. In what follows we take $L_x = L_y = L$. The image is formed via an inverse Fourier transform of a matrix of complex

random numbers with uniformly distributed phase and amplitude $\sqrt{W(k_x, k_y)}$. Some examples of random surfaces generated in this manner are shown in Fig. 2.



(a) L=100 cm. (b) L=20 cm

Figure 2. Examples of rough surfaces realized using for a Gaussian spectrum.

5. RESULTS

The models described above were used to simulate the temporal and spatial signatures of a mine buried under smooth and rough surfaces. In all simulations a steady wind speed of $W(t) = \overline{W} = 2$ [m s^{-1}] and an average air temperature of $\overline{T}_{air} = 289$ K were used. The thermal diffusivity (κ) and conductivity ($\mathcal{K}$) of soil were taken to be 5.0×10^{-7} [m^2 s^{-1}] and 2.6 [W m^{-1} K^{-1}], respectively. For TNT we used $\kappa = 9.25 \times 10^{-8}$ [m^2 s^{-1}] and $\mathcal{K}$ =0.234 [W m^{-1} K^{-1}]. It is noteworthy that TNT is a better thermal insulator than soil. Mine dimensions and construction details vary significantly. As a representative target, we selected a simulant anti-tank mine[35] developed by the US Army. The simulant mine has a diameter of 25 cm and a height of 8.33 cm. The mine is assumed to be composed of TNT as shown in Fig. 3.[†]

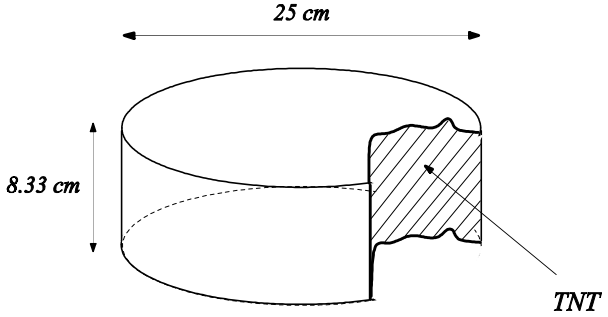


Figure 3. Model of the simulant anti-tank mine used in the simulations.

The computational volume has the dimensions $0.9984 \times 0.9984 \times 0.3154$ m^3 and was discretized with pentahedral elements of size $2.08 \times 2.08 \times 1.66$ cm^3 in the x, y, and z directions, respectively. This discretization results in 87552 finite elements and 48020 nodes, which also determines the size of the resulting matrix equation. The memory requirement of the FEM code for this discretization is about 250 MB. The Crank-Nicholson scheme is unconditionally stable, which permits us substantial freedom in choosing the time increment. The time increment was selected to

[†]The anti-tank simulants contain small metal inserts located at the center of the bottom face of the mine. This insert was not modeled in our work. We do not expect to incur a significant error as a result, because of the insert's size and location.

be 360 seconds, which gives good resolution over a 24 hour period. The FEM solution was observed to reach steady after only two simulated diurnal cycles. The simulation takes about 2 hours on a 300 MHz Pentium II machine.

The radiometric model requires the triangular rough surface representation and the output of the thermal model. Using these data and the geometry given in Fig. 4, we construct a virtual IR camera. The virtual sensor is assumed to be a LWIR camera operating near 10 μm. A pixel array of 160 (horizontally) by 120 (vertically) is assumed with an instantaneous field of view of 1 mrad.

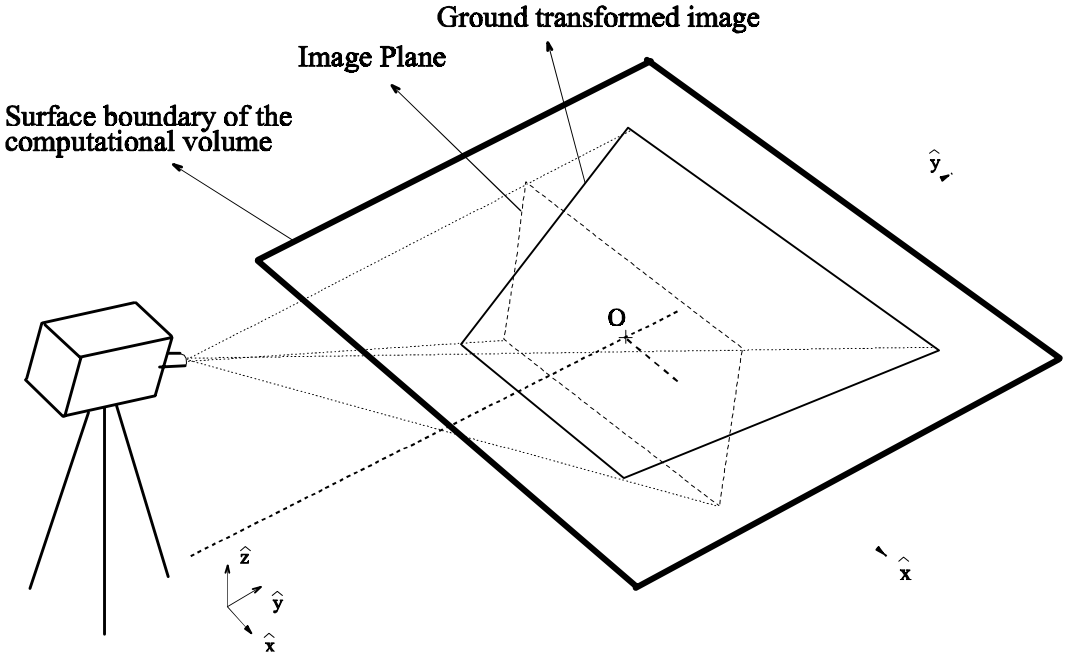


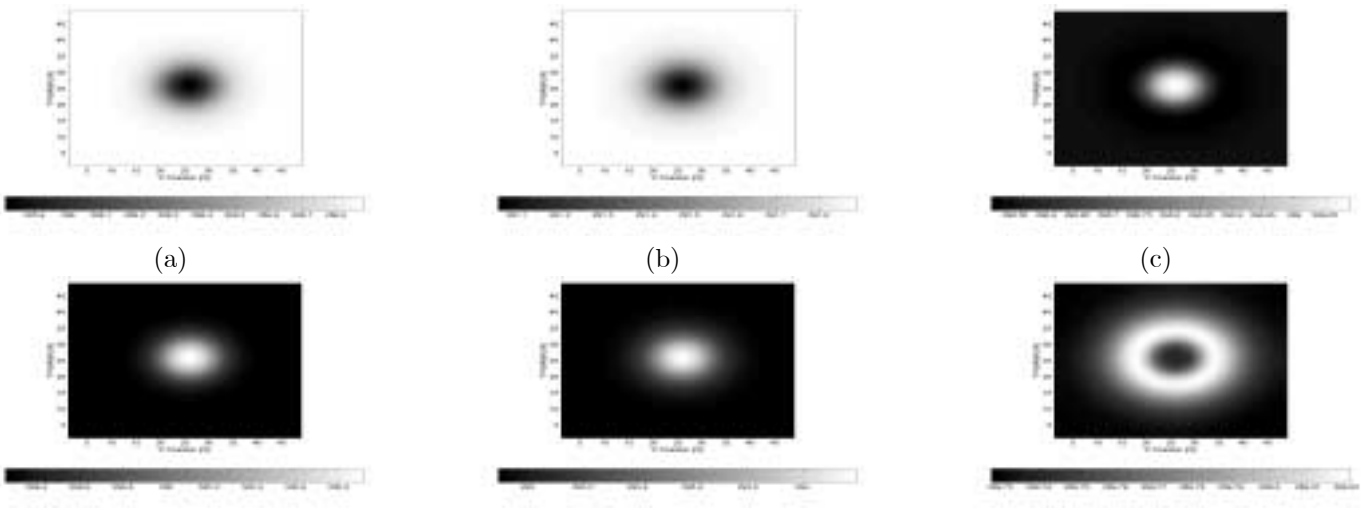
Figure 4. Sketch of the region of interest including the IR camera. Point O is both the center of the field of view and the origin of coordinates. The image plane, the ground transformed image, the surface boundaries of the computational volume, and the global axis are illustrated.

The IR camera is aimed at point O as shown in Fig. 4. Point O is also the global origin for both the thermal and radiometric models. To cover the surface above the mine with sufficient resolution, an appropriate camera locations and height must be identified. For this work we selected a camera vertical height of 2 meters, and a horizontal standoff (camera center to point O) of 3 meters. The resulting ground-projected field of view encompasses most of the computational volume.

Figure 5 shows the surface temperature distribution over the mine. The results are presented as a sequence of images evaluated at three hour time increments starting from sunrise. It is experimentally observed that contrast changes occur twice daily at the thermal cross-over times, and these events appear in our simulations. The results suggest that the physical temperature differences during the day peak at roughly 1.73 K occuring about 2 hours before sunset. The temperature difference has a maximum at the center of the mine and diminishes as we move away from the center. The temperature distribution as a function of depth is presented in Fig. 6. The location of the mine is emphasized with a rectangle. The temperature distribution is again given every three hours starting at sunrise. The figures present the evolution of heat flow into the ground and the influence of the mine on this heat flow. The surface temperature over the mine is cooler at dawn, and it warms as time proceeds. The mine tend to block the flow of heat into the soil, causing the overlying soil to become hotter during the day. During the night, the mine blocks the upward flow of energy in the soil, permitting the layer of soil above the mine to cool more rapidly. These observations and our previous findings using one-dimensional models[19] are consistent with the behavior of insulating mines.

Simulations of the 25 cm anti-tank mine simulant appear in Figs. 7 and 8 for the case of a smooth surface and a surface having a 3 cm peak-to-peak variation. The rough surface decorrelation length was 20 cm, which leads to reasonably large surface slopes. While the smooth surface case is substantially the same as a perspective-transformed

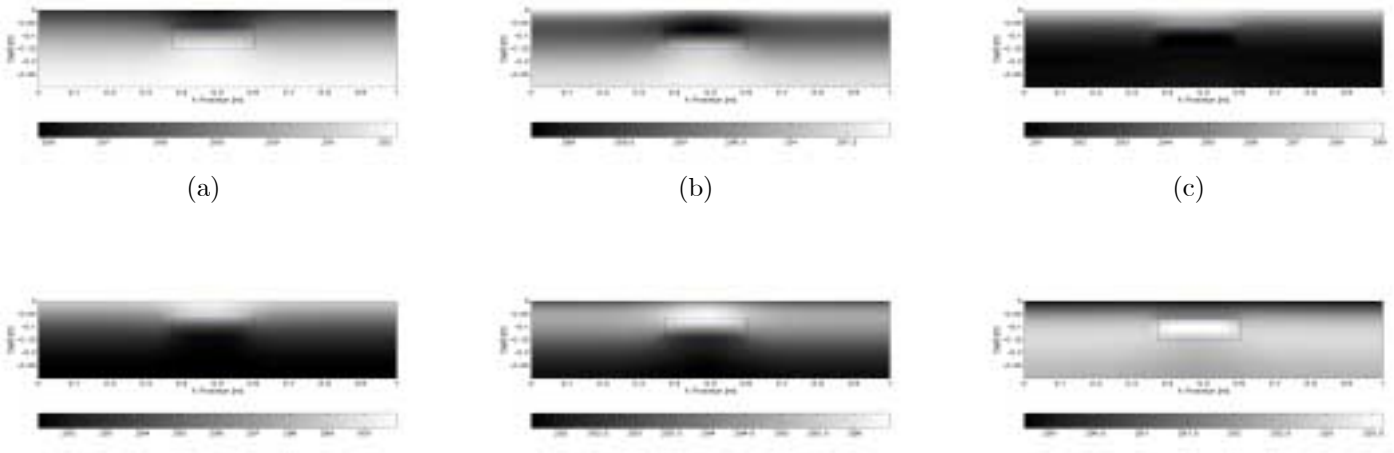
copy of the thermal surface images shown in Fig. 5, the rough surface case shows significant clutter-like variations, which could easily be mis-detected as false alarms.



(a) (b) (c)

(d) (e) (f)

Figure 5. Simulation of the soil surface temperature difference over the anti-tank mine simulant buried 6.66 cm under a smooth soil surface. The temperature distribution is evaluated at different times of the day (a) At dawn. (b) 3 hours after sunrise. (c) At noon. (d) 3 hours after noon. (e) At sunset. (f) 3 hours after sunset.



(a) (b) (c)

(d) (e) (f)

Figure 6. Simulation of the temperature distribution at depth taken through the middle of the anti-tank mine simulant buried 6.66 cm under a smooth soil surface. The temperature distribution is evaluated at different times of the day (a) At dawn. (b) 3 hours after sunrise. (c) At noon. (d) 3 hours after noon. (e) At sunset. (f) 3 hours after sunset.

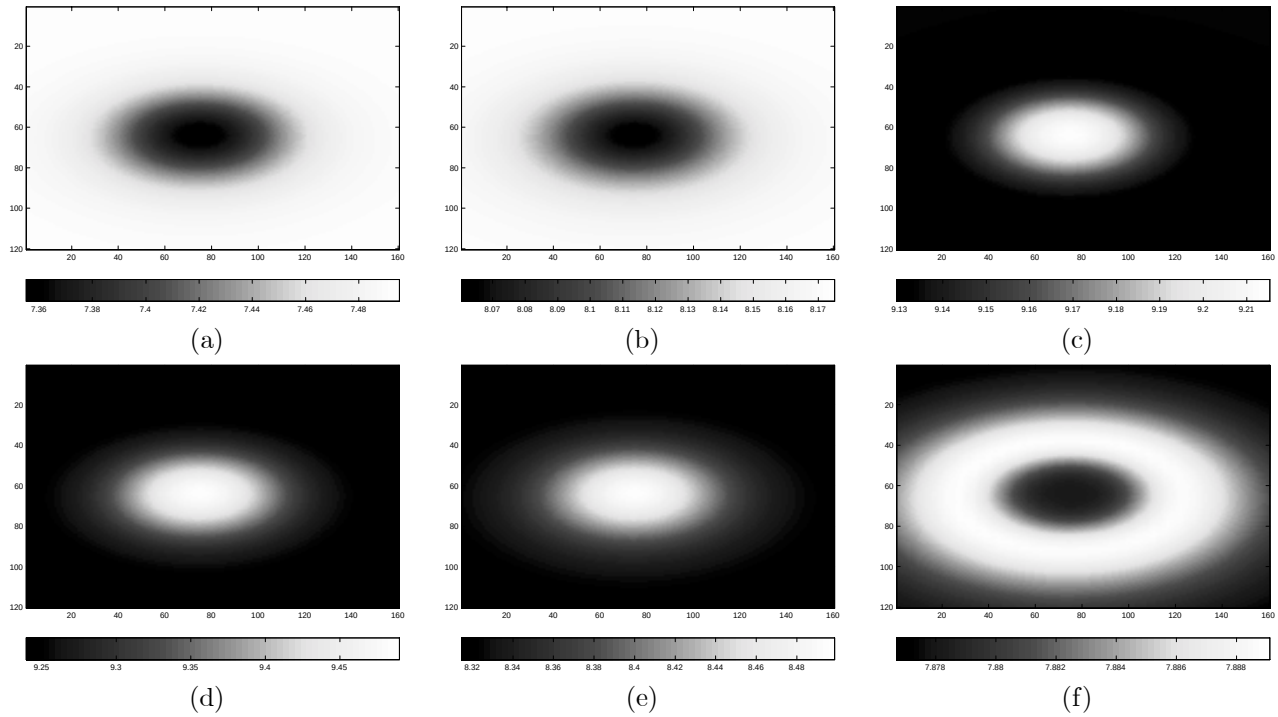


Figure 7. Simulation of the IR camera response for the SIM-25 mine buried 6.66 cm under a smooth soil surface. The radiometric model is used to predict the IR camera response at different times of the day (a) At dawn. (b) 3 hours after sunrise. (c) At noon. (d) 3 hours after noon. (e) At sunset. (f) 3 hours after sunset.

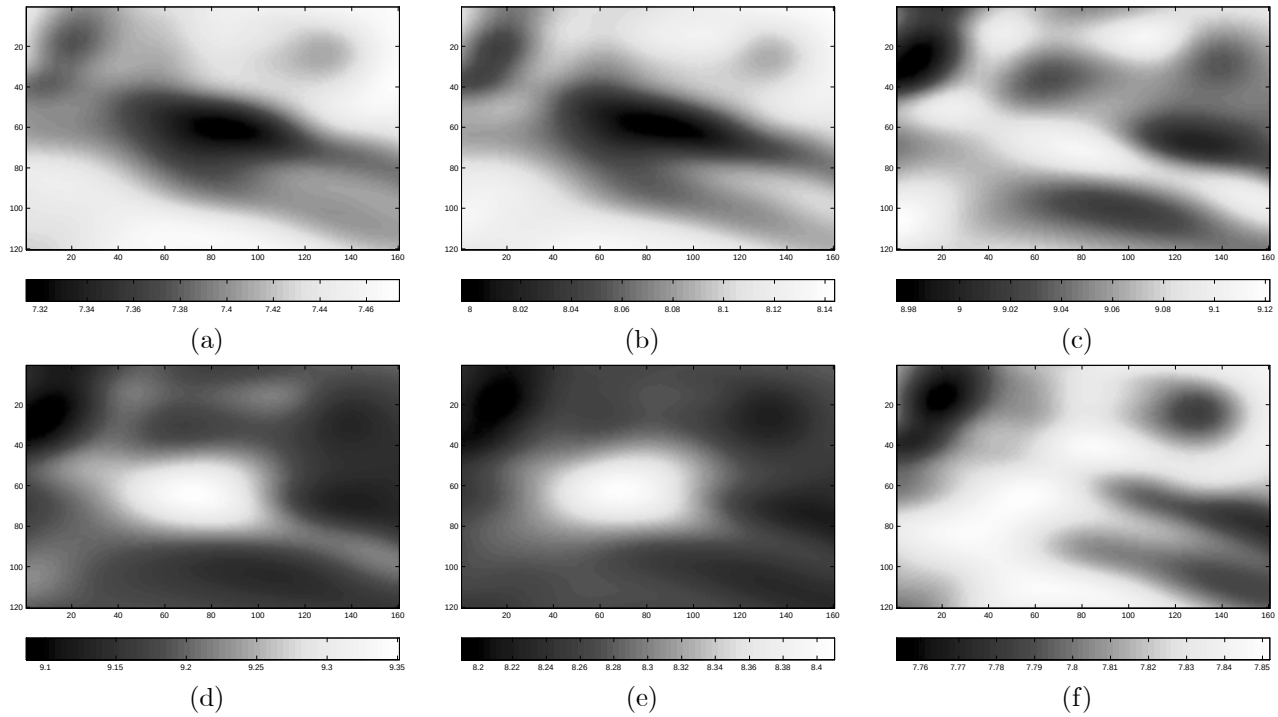


Figure 8. Simulation of the IR camera response for the SIM-25 mine buried 6.66 cm under a rough soil surface with peak-to-peak height 3 cm. The radiometric model is used to predict the IR camera response at different times of the day (a) At dawn. (b) 3 hours after sunrise. (c) At noon. (d) 3 hours after noon. (e) At sunset. (f) 3 hours after sunset.

6. SUMMARY AND CONCLUDING REMARKS

A model for thermal and radiometric effects has recently been developed for predicting the IR signature of buried land mines. The model employs a 3-D numerical solution of the thermal problem based on the finite element method (FEM). This approach permits inhomogeneous soil and the internal structure of the mine to be considered in modeling, although we have not exercised those capabilities in this initial study. In addition, the model incorporates a rough surface for the soil-air interface, which has implications for both thermal heating and for reflected radiometric terms. The temperature distribution computed using the thermal model is combined with surface-reflected radiometric components to produce the image seen at a virtual IR sensor.

Several extensions of this work are called for. Example calculations were presented, which show trends observed experimentally, but careful experimental validation of the code is necessary. The spectral content of natural sources and the spectral dependence of nature materials were largely ignored in this work, but will be addressed. Soil inhomogeneity has not been considered, but it presents no significant challenges to our formulation. A Gaussian surface model was assumed. We plan to acquire more realistic surface models using a laser surface profiler, and we intend to incorporate such data in a future work. Surface mines present additional challenges, which must also be addressed.

7. ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

REFERENCES

1. L. A. LeSchack and N. K. Del Grande, "A dual-wavelength thermal infrared scanner as a potential airborne geophysical exploration tool," *Geophys.* **41**(6), pp. p. 1318–1336, 1976.
2. Y. H. L. Janssen, A. N. de Jong, H. Winkel, and F. J. M. van Putten, "Detection of surface laid and buried mines with IR and CCD cameras, an evaluation based on measurements," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, R. L. Barnard, C. J. Lowe, and J. E. McFee, eds., *Proceedings of SPIE* **2765**, pp. 448–459, 1996.
3. G. Maksymomko, B. Ware, and D. Poole, "A characterization of diurnal and environmental effects on mines and the factors influencing the performance of mine detecting ATR algorithms," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, eds., *Proceedings of SPIE* **2496**, pp. 140–151, 1995.
4. J. R. Simard, "Improved lanmine detection capability(ILDC): Systematic approach to the detection of buried mines using passive IR imaging," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, R. L. Barnard, C. J. Lowe, and J. E. McFee, eds., *Proceedings of SPIE* **2765**, pp. 489–500, 1996.
5. K. Russel, J. McFee, and W. Sirovyak, "Remote performance prediction for infrared imaging of buried mines," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, A. C. Dubey and R. L. Barnard, eds., *Proceedings of SPIE* **3079**, pp. 762–769, 1997.
6. B. A. Barbour, M. W. Jones, H. B. Barnes, and C. P. Lewis, "Passive IR polarization sensors: A new technology for mine detection," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *Proceedings of SPIE* **3392**, pp. 96–103, 1996.
7. M. Larive, L. Collot, S. Breugnot, H. Botma, and P. Roos, "Laid and flush-buried mines (sic) detection using 8-12 μm polarimetric imager," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *Proceedings of SPIE* **3392**, pp. 115–120, 1996.
8. C. DiMarzio, T. Vo-Dinh, and H. E. Scott, "Some approaches to infrared spectroscopy for detection of buried objects," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *Proceedings of SPIE* **3392**, pp. 158–166, 1996.
9. P. Pregowski and W. Swiderski, "Using of (sic) comparison method in IR thermal detection of buried mines," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *Proceedings of SPIE* **3392**, pp. 1241–1248, 1996.

10. N. K. Del Grande, K. W. Dolan, P. F. Durbin, M. R. Gorvad, B. T. Kornblum, D. E. Perkins, D. J. Schneberk, and A. B. Shapiro, "Three-dimensional thermal imaging of structural flaws by dual-band infrared computed tomography," in *Underground and Obscured Object Imaging and Detection*, N. K. Del Grande, I. Cindrich, and P. B. Johnson, eds., *Proceedings of SPIE* **1942**, pp. 207–215, 1993.
11. A. T. DePersia, A. Bowman, P. Lucey, and E. M. Winter, "Phenomenology considerations for hyperspectral mine detection," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, eds., *Proceedings of SPIE* **2496**, pp. 159–167, 1995.
12. D. S. Flynn, D. A. Vechinsk, B. T. Blume, A. C. Dubey, and N. H. Witherspoon, "Minefield image synthesis tool," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, eds., *Proceedings of SPIE* **2496**, pp. 208–221, 1995.
13. K. Watson, "Geologic application of thermal infrared images," *Proc. IEEE* **63**(1), pp. 128–137, Jan. 1975.
14. A. W. England, J. F. Galantowicz, and M. S. Schretter, "The radiobrightness thermal inertia measure of soil moisture," *IEEE Trans. Geosci. Remote Sensing* **30**(1), pp. 132–139, Jan. 1992.
15. A. B. Kahle , "A simple thermal model of the earth's surface for geologic mapping by remote sensing," *Journal of Geophysical Research* **82**, pp. 1673–1680, 1977.
16. K. Watson, "Periodic heating of a layer over a semi-infinite solid," *Journal of Geophysical Research* **78**(26), pp. 5904–5910, Sept. 1973.
17. A. W. England, "Radiobrightness of diurnally heated, freezing soil," *IEEE Trans. Geosci. Remote Sensing* **28**(4), pp. 464–476, July 1990.
18. Y.-A. Liou and A. W. England, "A land-surface process/radiobrightness model with coupled heat and moisture transport for freezing soils," *IEEE Trans. Geosci. Remote Sensing* **36**(2), pp. 669–677, March 1998.
19. I. K. Sendur and B. A. Baertlein, "Thermal analysis of IR signatures for buried land mines," Progress Report to US Army Research Office on Contract 97-SC-ARO-1015, The Ohio State University ElectroScience Laboratory, Columbus, OH, March 1999.
20. Y.-A. Liou and A. W. England, "A land-surface process/radiobrightness model with coupled heat and moisture transport in soil," *IEEE Trans. Geosci. Remote Sensing* **36**(1), pp. 273–286, Jan. 1998.
21. Y.-A. Liou and A. W. England, "Annual temperature and radiobrightness signatures for bare soils," *IEEE Trans. Geosci. Remote Sensing* **34**(4), pp. 981–990, July 1996.
22. B. A. Baertlein and I. K. Sendur, "Thermal infrared detection of buried land mines," Progress Report to US Army Research Office on Contract 97-SC-ARO-1015, The Ohio State University ElectroScience Laboratory, Columbus, OH, December 1997.
23. K. Ya. Kondratyev, *Radiation in the Atmosphere*, Academic Press, New York, NY, 1969.
24. J. W. Salisbury and D. M. D'Aria, "Emissivity of terrestrial materials in the 8-14 μm atmospheric window," *Remote Sensing Environment* **42**, pp. 83–106, 1992.
25. J. W. Salisbury and D. M. D'Aria, "Emissivity of terrestrial materials in the 3-5 μm atmospheric window," *Remote Sensing Environment* **47**, pp. 345–361, 1994.
26. L. W. Wolfe and G. J. Zissis, *The Infrared Handbook*, revised edition, Environmental Research Institute of Michigan (ERIM), Ann Arbor, MI, 1989.
27. W. M. Rohsensow, J. P. Hartnett, and E. N. G. (eds.), *Handbook of Heat Transfer Fundementals*, 2nd ed., McGraw Hill, New York, NY, 1985.
28. R. G. Reeves, A. Anson, and D. L. (eds.), *Manual of Remote Sensing*, 1st ed., American Society of Photogrammetry, Falls Church, VA, 1973.
29. H. Neckel and D. Labs, "The solar radiation between 3300 and 12500 å," *Solar Phys.* **90**, pp. 205–258, 1984.
30. M. P. Thekaekara and A. J. Drummond, "Standard values for the solar constant and its spectral components," *Nature Physical Science* **229**(4), pp. 6–9, 1971.
31. R. Dogniaux (ed.), *Prediction of Solar Radiation in Areas with a Specific Microclimate*, First edition, Kluwer Academic Publishers, Inc., The Netherlands, 1994.
32. K. L. Coulson, *Solar and Terrestrial Radiation*, First edition, Academic Press, Inc., New York, NY, 1975.
33. C. W. Allen, *Astrophysical Quantities*, First edition, The Athlone Press., London, England, 1963.
34. J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics*, Addison-Wesley Publishing Company, Inc., 1990.
35. Anon., "Simulant mines (sims)," tech. report, US Army Countermine and Demolition (Countermine Division), Fort Belvoir, VA, October 21 1998.

Wavelet-based higher-order neural networks for mine detection in thermal IR imagery

Brian A. Baertlein and Wen-Jiao Liao

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

An image processing technique is described for the detection of mines in IR imagery. The proposed technique is based on a third-order neural network, which processes the output of a wavelet packet transform. The technique is inherently invariant to changes in signature position, rotation and scaling. The well-known memory limitations that arise with higher-order neural networks are addressed by (1) the data compression capabilities of wavelet packets, (2) projections of the image data into a space of similar triangles, and (3) quantization of that "triangle space." Using these techniques, image chips of size 28×28, which would require $O(10^9)$ neural net weights, are processed by a network having $O(10^2)$ weights. ROC curves are presented for mine detection in real and simulated imagery.

Keywords: land mines, IR imagery, wavelets, neural networks

1. INTRODUCTION

It has long been recognized that surface-laid and buried mines can be detected using infrared (IR) sensors.[1] Although significant strides have been made in both the sensors and image processing, detection of mines in natural clutter remains a challenging problem. The challenges arise, in part, because of a large number of unknown random signature variables including amplitude, contrast polarity, rotation, and size.

There exist several techniques for dealing with unknown signature variables. When a sufficiently accurate model for the signature exists and the distribution of the unknown variables is known, one can perform an integration over the variables, leading to an optimal Bayesian detector.[2] For many IR targets (especially buried mines), such models are impractical since the signature is strongly affected by the environment (surface topology, solar illumination history, etc.), and the number of unknown variables becomes excessive.

As an alternative, one can use a maximum likelihood technique to estimate the unknown variables, leading to a generalized likelihood ratio test (GLRT). Detection is then based on the error between the model (using the estimated parameters) and the data. In a recent work[3] a nonlinear optimization technique was used to estimate parameters for an empirical signature model. These parameters and the resulting model error were both used in a classifier to detect the presence of the mine. That technique, which we refer to as an "estimator-classifier" (by analogy with the "estimator-correlator" technique used for random Gaussian-distributed signals), is capable of better performance than the GLRT, since it makes direct use of the model parameters. A significant limitation of all such approaches is the extensive processing required for nonlinear optimization. Because imaging systems tend to produce massive data sets, computational efficiency is a concern.

In this work we emphasize detection of surface mines and mine fields. An important aspect of these targets is that they tend to have a small number of simple shapes (round, square, etc.). Thus, for surface mines and some buried mines, the signature of the mine is largely known with the exception of a translation (due to its position in the image), a scale factor (due to the image resolution, the size of the mine, or the burial depth), and a rotation (for non-circular mines). In IR detection of mine fields the (near) regular spatial pattern of the mines is an important discriminant. The pattern, however, will have an unknown rotation and translation, and it may vary in scale as a result of changes in image resolution or mine-laying practice.

Detection of objects with unknown position, scale and rotation (PSR) is an important general topic, and the subject has been explored by many researchers. It is well known that the magnitude of the two-dimensional Fourier transform $\mathcal{F}$ is invariant to translational shifts. Mehanian and Rak[4] showed that a log-polar mapping of the Fourier

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

transform magnitude could produce a representation that was position, scale and rotation invariant (PRSI). A limitation of this approach is that the polar coordinates must be defined with respect to a "stable" point in the image. The centroid of the signature has been used in that role, but the centroid is an unreliable reference for targets in a highly cluttered environment.

The magnitude of the Mellin transform $\mathcal{M}$ is invariant to changes in scale, and this fact has been used to develop scale-invariant detection systems. During the 1970s it was observed by several groups[5–8] that the combination of $|\mathcal{F}|$ followed by $|\mathcal{M}|$, the so-called Fourier-Mellin (FM) transform, could be used for this purpose. The use of wavelets for scale invariance has been investigated by several groups. Baertlein and Casey[9] developed an analog of the Fourier-Mellin transform noted above, in which the continuous wavelet transform of the Fourier transform modulus was computed. Szu et al.[10] used a similar technique in which the continuous wavelet transform was computed optically. Yoon et al.[11] used the CWT of the target boundary to obtain a scale-invariant representation for image recognition. Although all of these approaches provide translation and scale invariance, they are not rotationally invariant.

In this paper we propose a technique that employs higher-order neural networks to perform PSRI detection. The memory required by a higher-order network can be significant, but the image decomposition and data compression capabilities of the wavelet transform are used to mitigate this problem.

The paper is organized in three major sections. Higher-order neural networks are described in Section 2. A wavelet packet transform is required in this work, and that topic is described in Section 3. Classification results for real and simulated imagery are provided in Section 4. Concluding remarks appear in Section 5.

2. NEURAL NETWORKS AND PSR INVARIANCE

Several approaches have been developed for incorporating PSR invariance into neural networks.[12] Feed-forward multilayer perceptrons (FFMLPs) can be trained to recognize signatures that have been shifted, rotated, and scaled, but the amount of training data required is large, and accuracy is limited.[13,14] The neocognitron[15,16] is a novel neural network architecture that attempts to achieve PSRI performance by mimicking the structure of the human visual system. Modest deformations and small rotations (at most a few degrees) are tolerated in the input, but rotational invariance over a wider range of angles has not been demonstrated.

Higher-order neural networks (HONNs) offer perhaps the best approach to the PSRI problem. The theoretical nature of invariant pattern classification has been discussed by Minsky and Papert.[17] Lee et al.[18] describe the concept of correlations among higher-order products in neural networks. Giles and Maxwell[19] showed that by a suitable sharing of weights one could structurally induce translation invariance in both associative memories and FFMLPs while simultaneously reducing the number of weights to be learned. The key benefit of such networks is that the translation invariance property need not be learned from data. Examples acquired at a fixed position are sufficient to train the network to recognize shifted data. A combination of HONNs and log-polar mappings has also been investigated.[20]

The computation performed by a higher-order neuron output is straightforward. Consider an image of size $N \times N$ in which the pixels x_{ij} have been indexed as $x_{i_n j_n}$, $n = 1, 2, \ldots, N^2$. The higher-order neuron output has the form

$$y = f\left(\sum_{n=1}^{N^2} x_{i_n j_n} w_n^{(1)} + \sum_{n=1}^{N^2}\sum_{m=1}^{N^2} x_{i_n j_n} x_{i_m j_m} w_{nm}^{(2)} + \sum_{n=1}^{N^2}\sum_{m=1}^{N^2}\sum_{k=1}^{N^2} x_{i_n j_n} x_{i_m j_m} x_{i_k j_k} w_{nmk}^{(3)} + \cdots\right) \tag{1}$$

where $f(\cdot)$ is the activation function (nominally a sigmoid-like function) and the weights are given by $w_n^{(1)}$, $w_{nm}^{(2)}$ and $w_{nmk}^{(3)}$ for first, second, and third orders, respectively. As an example, the architecture of a third-order neuron is shown in Figure 1.

The linear sum is analogous to the output of a conventional neuron. The sum of second-order products involves correlations, and the third-order sum involves bicorrelations. The higher-order products explicitly provide the non-linear relations on which the network operates. Note that this is a single-layer network, i.e., the output neuron is connected directly to the input data. As a result, training may be performed with the simple Widrow-Hoff algorithm. For example, for the third-order weights, the update equation for a network that produced output s for a desired output t is

$$\Delta w_{nmk}^{(3)} = (t - s) x_{i_n j_n} x_{i_m j_m} x_{i_k j_k} \tag{2}$$

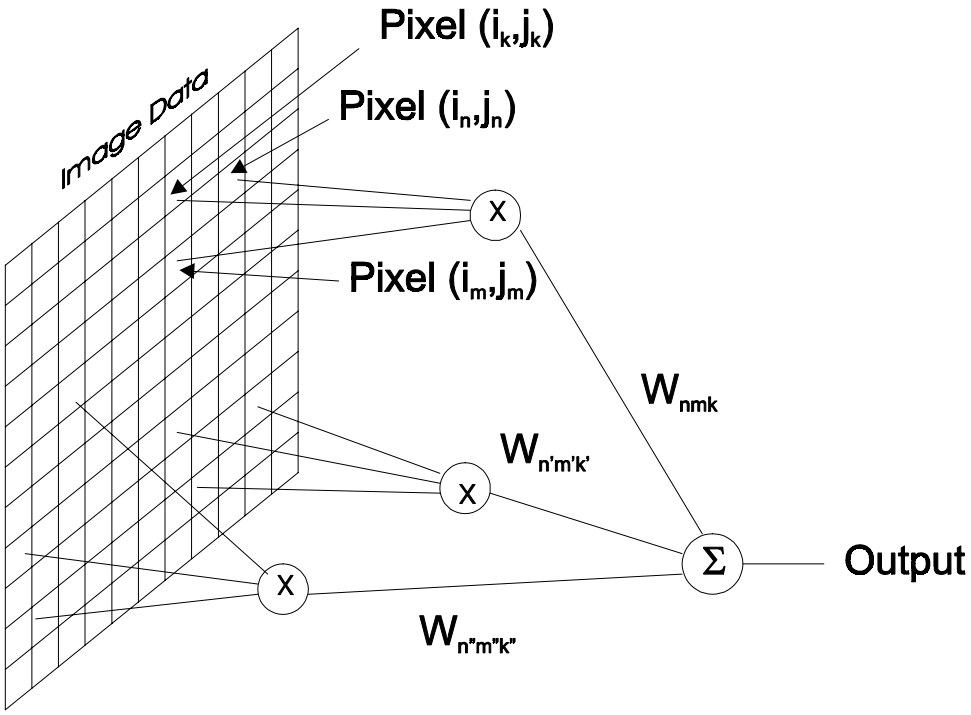


Figure 1. The architecture of a third-order neural network. Only three weight connections are shown.

Spirkovska and Reid[21] showed that a weight-sharing technique could be used to produce second-order FFMLPs with translation and scale invariance. Such a HONN is useful for detection of rotationlly symmetric signatures, which comprise most surface mines and essentially all buried mines at depths greater than a few inches. To achieve this capability we equate weights corresponding to line pairs whose lengths have a similar ratio. Explicitly, the equivalence we require is

$$w_{nm}^{(2)} = w_{k\ell}^{(2)} \text{ if } |i_n - i_m|/|j_n - j_m| = |i_k - i_\ell|/|j_k - j_\ell| \tag{3}$$

This equivalence maps greatly reduces the number of weights and the training data requirements for the network.

Recognition of objects that have been simultaneously scaled, rotated and translated requires a third-order network,[21] the memory and computational requirements of which are substantial. For an image of size $N \times N$, a network or order M requires N^{2M} words to store the network weight values. For image chips of size $N = 32$ and a third-order HONN ($M = 3$), this leads to a requirement for 1.1×10^9 words of storage. For a focal plane with 128×128 pixels, the requirement grows to 4.4×10^{12} words.

The weight sharing required for PSRI performance in a third-order HONN mitigates this need to a degree and proceeds as follows:[22,21] Triads of points share a weight if they can be related by a translation, rotation and scaling. Thus, pixels triads that form geometrically similar triangles, such as those shown in Figure 2, are assigned the same weight. The input image is mapped to this "triangle space" by combining the bicorrelation products corresponding to a given triangle equivalence class.

In principle, a third-order network having been trained on a single target example is capable of recognizing all translated, scaled, and rotated examples of that signature. In practice, however, finite pixel sizes make it difficult to recognize rotated and scaled signatures. To overcome this problem, several techniques are used. First, Spirkovska and Reid[21] showed that one can improve performance and simultaneously reduce memory requirements by quantizing the internal angles of the triangles formed. We find it more convenient to quantize the lengths of the triangle sides, having first normalized the longest side to unit length. Perantonis and Lisboa[22] showed that it was most effective to quantize the largest side of the triangle more agressively. It is also usually beneficial to use a small number of scaled and rotated signatures in training.

Although a single-layer network is sufficient to implement a HONN, Perantonis and Lisboa found that by using an additional (hidden) layer, performance could be improved somewhat.[22] The resulting three-layer network is trained

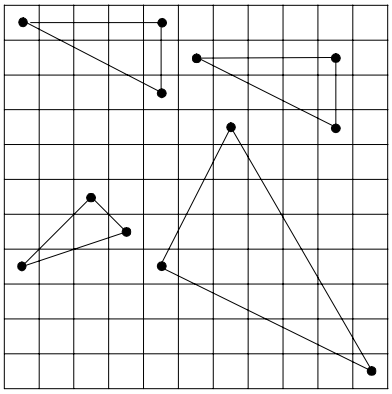


Figure 2. Examples of some pixel triples that define (approximately) similar triangles. Such pixel triples share the same weight in a PSRI network.

using the backpropagation algorithm.

An interesting generalization of the HONN concept is the "functional-link" network of Pao.[23] In a functional-link network the pixel products are replaced by more general operations (including nonlinear functions) of the form $g_\ell(x_{i_1 j_1}, x_{i_2 j_2}, \ldots)$, viz:

$$y = f\left(\sum_{\ell=1}^{N} w_\ell g_\ell(\mathbf{x})\right) \tag{4}$$

where $\mathbf{x}$ is a vector of inputs to the neuron. It is clear that if we enumerate by ℓ the geometrically similar triangle shapes that appear in an image, and if we define the function

$$g_\ell(\mathbf{x}) = \sum_{\substack{(n,m,k) \text{ form} \\ \text{triangles of shape } \ell}} x_{i_n j_n} x_{i_m j_m} x_{i_k j_k} \tag{5}$$

then the PSRI HONN is equivalent to a functional-link network.

3. WAVELET PACKET TRANSFORM

The processing proposed in this paper begin with a wavelet decomposition of the imagery. This step has two benefits. First, it permits us to reject clutter at scales not of interest. In effect, the wavelet decomposition performs the function of a spatial bandpass filter. Nonlinear editing of wavelet coefficients for noise suppression or image compression has been extensively used in the literature.[24] (See Dohono et al.[25] for a more recent review.) The use of wavelets for estimating trends is a closely related function and has also received attention including use in a recent paper on biomedical signal processing.[26] Recently, both techniques were used in a demining context.[27]

A second benefit of the wavelet transform is that it can perform image compression, which further reduces memory requirements. We use 2-D wavelet packets[28] for this purpose. In a wavelet packet analysis the detail coefficients are fully decomposed, producing output packets with a common size. The approximation coefficients output by a conventional two-dimensional wavelet transform are identical to one of the packets. The compression offered by the wavelet transform provides a dramatic reduction in HONN computational requirements. To see this, note that for a conventional dyadic wavelet transform based on an octave (factor of two) change in scale at each level, the number of pixels output in each packet by a 2-D level-K transform is reduced by the factor 2^{2K}. Thus, the number of operations required to form the third-order HONN bicorrelation products is reduced by the factor 2^{6K} if a single wavelet packet is processed and 2^{4K} if all wavelet packets are processed. In practice, edge effects in the wavelet transform cause the true improvement to be slightly less than these factors.

As an example of the wavelet packet transform, consider the surface-laid M19 mine shown in Figure 3. The transform of this image at level two is shown in Figure 4. The packet in the upper left corner represents the compressed chip, which is useful in subsequent HONN processing. Other chips contain additional information. The third chip (from the left) on the top row is the result of processing equivalent to a high-pass filter in the vertical

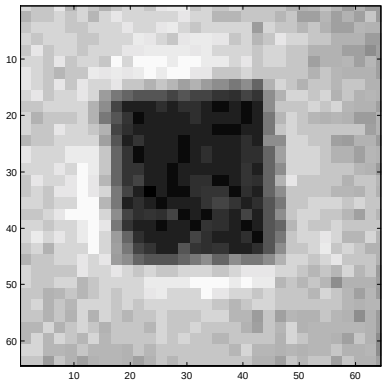


Figure 3. An LWIR image of a surface-laid M19 mine.

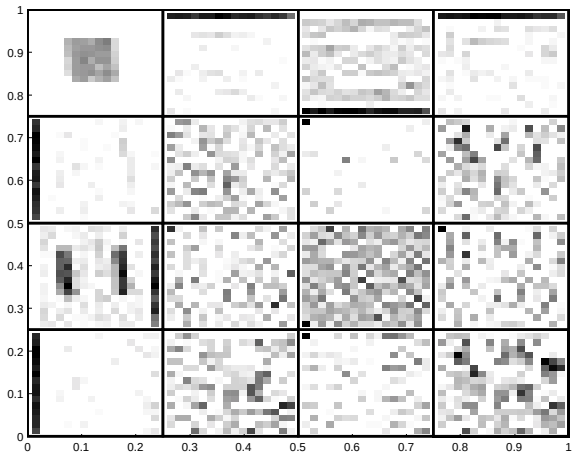


Figure 4. An example of a wavelet packet transform of order two. The pixel count of each output packet is reduced by roughly the factor $2^4 = 16$.

direction. Similarly, the third chip (from the top) in the left-most column is the product of high-pass filtering in the horizontal direction. Other image chips contain primarily small-scale noise-like components that can be neglected.

4. EXAMPLE RESULTS

In this section we review three example applications of the wavelet-HONN. In all three cases we apply the algorithm to the detection of surface mines. We describe two examples using simulated images and one using real LWIR data.

The processing used in these examples has both off-line and on-line components. For the off-line processing, one first selects the size of the image chip to be processed by the HONN. In this work we have used $N = 16$, for which $N^6 = 1.6 \times 10^7$. For an image of this size and a fixed distance quantization, pixel index triples for all similar triangles are identified, and these indices are saved. Next, training data are processed using the on-line operations described below, and the neural network is trained.

On-line processing begins by dividing the image into chips of size $N' \times N'$, and computing the wavelet packet transform of those chips. The chip size $N' > N$ is selected such that a wavelet packet transform at some level K will produce packets of size N. In our current implementation only the compressed wavelet packet is used in subsequent processing. The packet is then projected into triangle space by summing the bicorrelation products $x_{i_k j_k} x_{i_m j_m} x_{i_n j_n}$ for indices (determined off-line) comprising similar triangles (cf. Eq. (5)). The resulting sums are the HONN inputs, and the HONN output is used in detection. In our work we used a FFMLP with a single hidden layer and three hidden-layer neurons. The most computationally burdensome part of this process is forming and summing the bicorrelation products. Since this portion of the algorithm is independent of the image and extremely simple, it could be done with specialized hardware. In our implementation all processing is currently done in software.

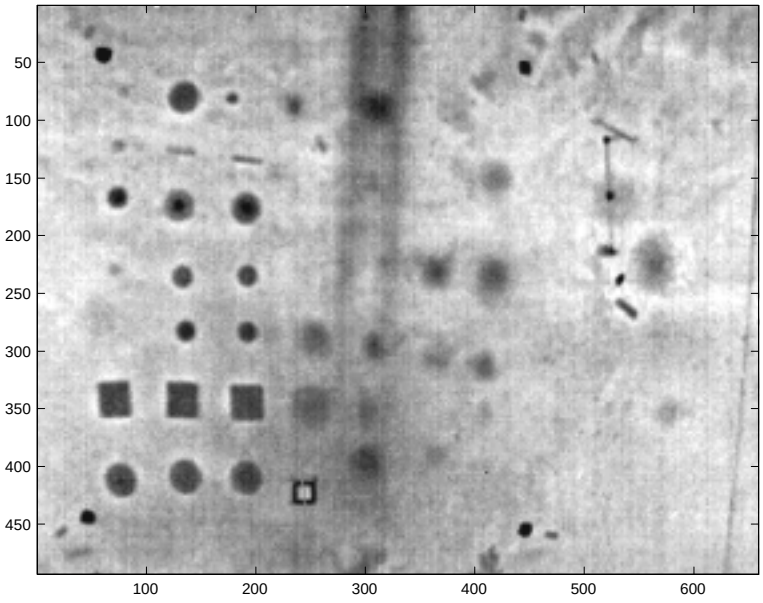


Figure 5. LWIR data used in testing the algorithm.

4.1. Mine discrimination for well-resolved targets

A preliminary test of the wavelet-HONN was made using experimental data. IR imagery acquired by E-OIR Measurements, Inc. in July 1998 over a set of mines emplaced at Fort Huachucha, AZ was used for this purpose. An example image is shown in Figure 5. The data were acquired with an Agema 570 LWIR camera looking directly down on the scene. The left half of the scene shows a number of surface-laid anti-tank mines including M19s, M15s and VS1.6s. The center and right half of the image contain a variety of buried mines and clutter artifacts. The prominant smear running through the image top to bottom is a vehicle track. The concentric light and dark squares near the bottom center of the image is a blackbody used for calibration.

We extracted from this scene a number of image chips. The data are shown in Figure 6, where we present 30 examples of surface mines and clutter. In this test we took both square (M19) and circular (M15 and VS1.6) mines to be the targets of interest, which are indicated by the surrounding boxes in Figure 6. The remaining chips show a variety of clutter objects.

Image chips of 28×28 pixels were extracted from the data and processed. The naive application of these chips to a 3D-HONN would require $28^6 \approx 5 \times 10^8$ input nodes. Using the proposed approach this memory load is reduced greatly. A level-one wavelet packet decomposition produces 16×16 compressed chips. The projection of these smaller chips onto "triangle space" followed by 5% distance quantization produces a neural network of 122 inputs.

For this small data set of 30 chips we used the leave-one-out method to train the network and assess its performance. The ROC curve produced through this process is shown in Figure 7. The performance of the network on these data is reasonably good, but the data set is quite small and it contains no rotated targets. The test does, however, offer evidence of scale invariance.

4.2. Shape discrimination for simulated well-resolved targets

One potential discriminant for surface anti-tank (AT) mines involves the shape of these targets. In this example we demonstrate the ability of the algorithm to discriminate circular and rectangular shapes. The prototype shapes are shown in Figure 8.

We tested this network with the scaled, rotated, and translated data shown in Figure 9. These data were generated by randomly scaling one of the prototypes by a factor between 3/5 and 5/3. Rotations of -45° to 45° were used, as were random translations. To these images we added spatially white, Gaussian noise to produce a signal to noise ratio of roughly 10:1. In principle, the HONN can be designed to recognize all the exemplars in Figure 9 when

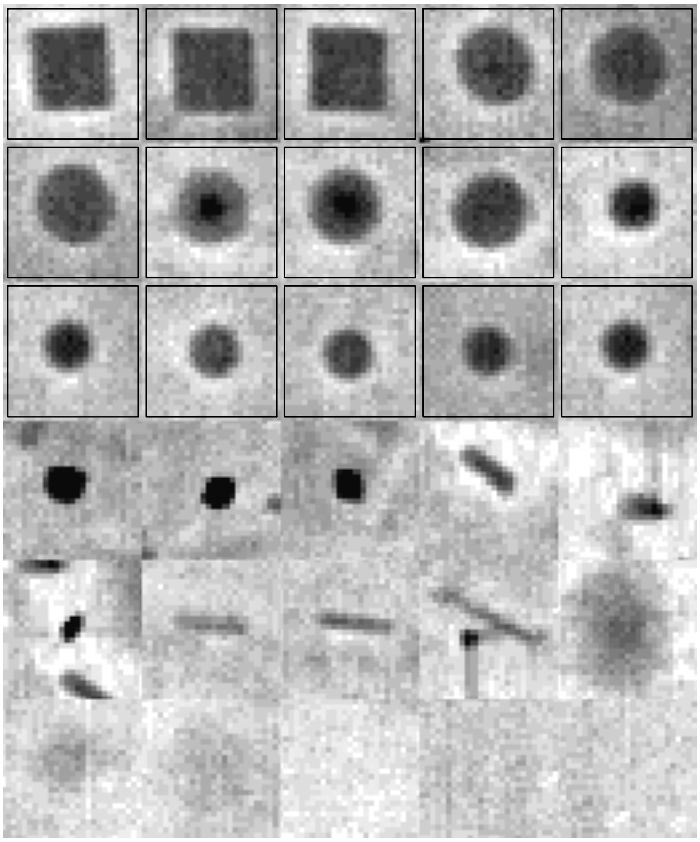


Figure 6. Example target and clutter chips extracted from the LWIR image. The boxed chips are designated targets.

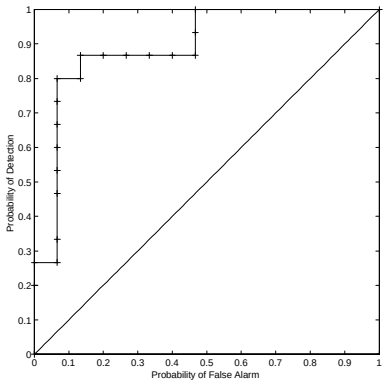


Figure 7. ROC curve generated by the LWIR surface mine data.

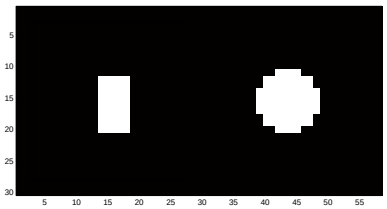


Figure 8. Prototype signatures for rectangular and circular targets.

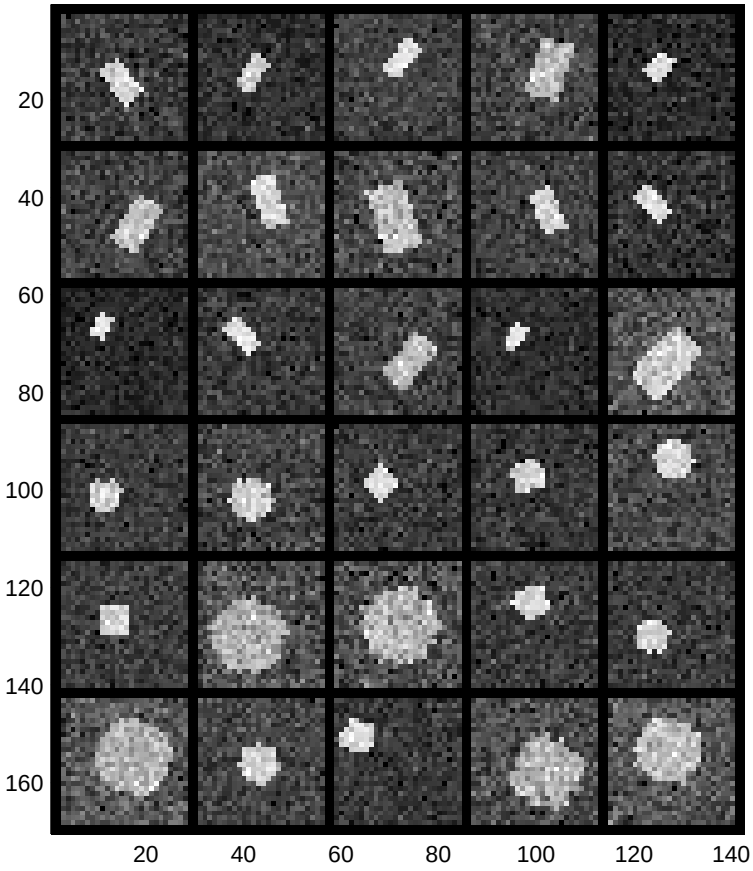


Figure 9. Examples of scaled, translated, and rotated targets used to test the network. The top half of the image contains rectangular targets, and the bottom half contains circular targets.

trained on the prototypes in Figure 8 only, but finite pixel size effects (noted above) limit this capability. We found performance was improved if we used a small number of scaled and rotated examples to train the network. The network trains quickly (within 10 epochs).

The ROC curve produced for this data set is shown in Figure 10. We find that the performance is good, even though for very small scaling some of the circular targets have a near-square shape.

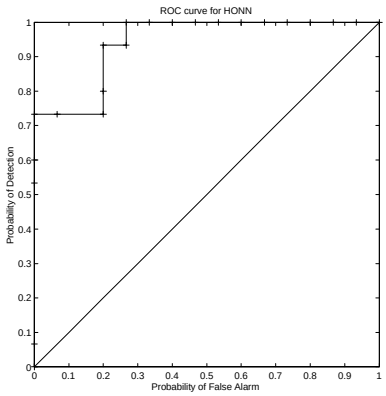


Figure 10. ROC curve generated by classifying scaled, rotated, and translated prototypes.

4.3. Pattern detection in simulated imagery

For the final example we consider detection of point targets arranged in a uniform linear array to simulate mine fields. Example data are shown in Figure 11, which contains 30 rotated, translated and scaled scenes. The simulated minefields contain five points in a row with mine-to-mine position errors. The simulated clutter scenes contain four to six points positioned randomly over the scene. A signal-to-noise ratio of roughly 20:1 was employed in these data.

The ROC curve produced by the detection of these targets is shown in Figure 12. Again, relatively good performance is noted, although the data set is small.

5. CONCLUDING REMARKS

We have described a new technique for the detection of mines and minefields. The method uses a wavelet packet decomposition as the front-end to a third-order neural network. Use of the wavelet transform and a subsequent projection into "triangle space" greatly reduces the well-known memory problems associated with higher-order neural networks. Tests of the proposed algorithm using simulated imagery and a modest amount of real data have produced encouraging results. More extensive testing using real imagery is planned.

ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the authors and are not necessarily those of Duke University or the ARO.

REFERENCES

1. C. Stewart, "Summary of mine detection research (U)," technical report 1636-TR, U.S. Army Engineer Research and Development Laboratories, Corps of Engineers, Fort Belvoir, VA, May 1960. Volume 1, DTIC AD320124.
2. H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. 1, Wiley, New York, NY., 1968.
3. I. K. Sendur and B. A. Baertlein, "Techniques for improving buried mine detection in thermal IR imagery," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **3710**, pp. 1272–1283, 1999.
4. C. Mehanian and S. J. Rak, "Bi-directional log-polar mapping for invariant object recognition," in *Automatic Object Recognition*, *SPIE* **1471**, pp. 200–209, 1991.

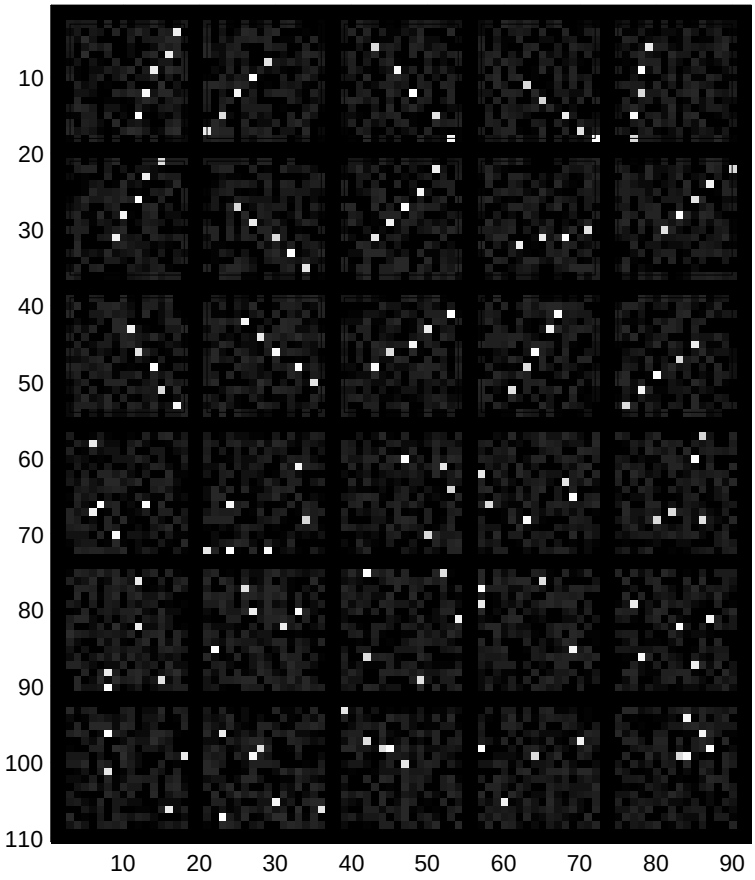


Figure 11. Simulated mine fields and clutter scenes.

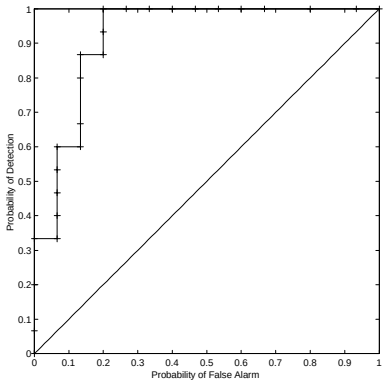


Figure 12. ROC curve generated by the simulated minefields.

5. D. Casasent and D. Psaltis, "New optical transforms for pattern recognition," *Proc. IEEE* **65**(1), pp. 77–84, 1977.
6. R. A. Altes, "The Fourier-Mellin transform and mammalian hearing," *J. Acoust. Soc. Am.* **63**(1), pp. 174–183, 1978.
7. J. Altmann and H. J. P. Reitbock, "A fast correlation method for scale- and translation-invariant pattern recognition," *IEEE Trans. Pattern Analysis and Mach. Intell.* **PAMI-6**(1), pp. 46–57, 1984.
8. P. E. Zwicke and I. Kiss, Jr., "A new implementation of the Mellin transform and its application to radar classification of ships," *IEEE Trans. Pattern Analysis and Mach. Intell.* **PAMI-5**(2), pp. 191–199, 1983.
9. B. A. Baertlein and K. F. Casey, "Magnetometric detection of buried objects," in *Proceedings of EUROEM*, (Bordeaux, France), June 1994.
10. H. H. Szu, X.-Y. Yang, B. A. Telfer, and Y. Sheng, "Neural network and wavelet transform for scale-invariant data classification," *Physical Review E* **48**(2), pp. 1497–1501, 1993.
11. S. H. Yoon, W. E. Alexander, and J. H. Kim, "Scale-invariant shape representation base on the wavelet transform," in *Proceedings*, *SPIE* **2232**, pp. 142–153, 1994.
12. C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, New York, NY, 1995.
13. S. E. Troxel, S. K. Rogers, and M. Kabrisky, "The use of neural networks in PSRI recognition," in *Proc. Joint Int. Conf. Neural Networks*, pp. 593–600, (San Diego, CA), July 1988.
14. B. Widrow and R. Winter, "Neural nets for adaptive filtering and adaptive pattern recognition," *Computer* , pp. 25–39, March 1988.
15. K. Fukushima, "Analysis of the process of visual pattern recognition by the neocognitron," *Neural Networks* **2**, pp. 413–420, 1989.
16. K. Fukushima, "A neural network for visual pattern recognition," *Computer* , pp. 65–75, March 1988.
17. M. Minsky and S. Papert, *Perceptrons: An Introduction to Computational Geometry*, MIT Press, Cambridge, MA, 1972.
18. Y. C. Lee, G. Doolen, H. H. Chen, G. Z. Sun, T. Maxwell, and H. Y. Lee, "Machine learning using a higher order correlation network," *Physica* **22D**, pp. 276–306, 1986.
19. C. L. Giles and T. Maxwell, "Learning, invariance, and generalization in higher-order neural networks," *Applied Optics* **26**(23), pp. 4972–4978, 1987.
20. Q. Xu, R. M. Inigo, and E. S. McVey, "A combined approach for large-scale pattern recognition with translational, rotational and scaling invariance," in *Automatic Object Recognition*, *SPIE* **1471**, pp. 378–393, 1991.
21. L. Spirkovska and M. B. Reid, "Higher order neural networks in position, scale, and rotation invariant object recognition," in *Fast Learning and Object Recognition*, B. Soucek, ed., pp. 153–184, John Wiley, 1992.
22. S. J. Perantonis and P. J. G. Lisboa, "Translation, rotation, and scale invariant pattern recognition by high-order neural networks and moment classifiers," *IEEE Trans. Neural Networks* **3**, pp. 241–251, March 1992.
23. Y.-H. Pao, *Adaptive Pattern Recognition and Neural Networks*, Addison Wesley, Reading, MA, 1989.
24. R. R. Coifman and M. V. Wickerhauser, "Wavelets and adapted waveform analysis," in *Wavelets: Mathematics and Applications*, J. Benedetto and M. Frazier, eds., CRC Press, Boca Raton, FL, 1993.
25. D. Dohono, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: asymptotia," *J. Royal Stat. Soc., Series B* **57**(2), pp. 301–337, 1995.
26. U. Wiklund, M. Akay, and U. Niklasson, "Short-term analysis of heart-rate variability by adapted wavelet transforms," *IEEE Engineering in Medicine and Biology Mag.* **16**, pp. 113–118, Sept./Oct. 1997.
27. T.-H. Chao, B. Lau, A. Yacoubian, and G. Henderson, "Mine detection using wavelet processing of electro-optic active sensor data," in *Detection Technologies for Mines and Minelike Targets*, A. Dubey, I. Cindrich, J. Ralston, and K. Rigano, eds., *SPIE Proc.* **2496**, pp. 433–441, 1995.
28. M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, A. K. Peters, Wellesley, MA, 1994.

The Role of Environmental Factors and Mine Geometry in Thermal IR Mine Signatures

Brian A. Baertlein and İbrahim Kürşat Şendur

The Ohio State University, ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

Thermal IR signatures of buried land mines are affected by various environmental conditions as well as the mine's composition, size and burial geometry. In this work we present quantitative relations for the effect of those factors on the signature's peak contrast and apparent diameter. We begin with a review of the relevant phenomena and the underlying physics. A three-dimensional simulation tool developed by the authors is used to simulate signatures for the case of a static water distribution. We discuss efforts to validate the model using experimental data collected at Fort A.P. Hill, VA. Using this simulation tool a variety of factors are considered, including soil water content, soil sand content, wind speed, mine diameter and mine burial depth.

Keywords: IR, thermal, mines, heat flow, modeling, simulation

1. INTRODUCTION

Anecdotal evidence and a growing body of experimental measurements suggest that buried mines are more detectable under certain environmental conditions. An analysis of the physics that describe these signatures suggests that a number of environmental factors as well as the local meteorological history play important roles. In addition, mines vary in size and burial depth, and it is well known that these mine-related factors affect the surface temperature distribution.

In this work we employ numerical modeling to quantify the aforementioned relations. A limitation of the simulations shown here is that the moisture distribution is assumed to be static, i.e., we ignore the effect of water and vapor movement on heat transfer and we ignore surface cooling through evaporation. In addition, we have linearized the analysis to simplify the simulations. In spite of these limitations, we are able to consider a number of important variables and to quantify their effect. Insight drawn from this work was used to improve signal processing techniques for buried mine detection, as described in a companion paper.[1]

The work is organized in three major parts. In Section 2 we review the physical processes that produce thermal mine signatures, including the role of moisture transport. These processes are controlled by a number of fundamental parameters including the soil thermal conductivity and diffusivity, the surface convection coefficient, the air temperature, and the incident solar radiance. We review models for these fundamental parameters that involve readily measured quantities. In Section 3 we briefly review our simulation tool and the simplification implied by a static moisture distribution. Work toward validating that tool is also described. In Section 4 we present a parameter study involving key environmental variables. The effects of those variables on surface temperature distribution are examined through simulation. We report both the shape of the surface temperature distribution and the time-dependence of the peak signature. The influence of the mine's diameter and depth on the signature are also studied. Concluding remarks appear in Section 5.

2. COUPLED THERMAL-MOISTURE MODELING

Heat transport is the primary phenomenon governing thermal IR signatures, but soil moisture content is also thought to have a significant role. Because of their importance in agriculture and remote sensing, the topics of heat and moisture transport in soils have an extensive literature. Seminal works in this area are those of Phillip and de Vries[2] and de Vries[3](hereafter, referred to as PdV) who jointly presented a set of coupled partial differential equations that describe the transport phenomena. Numerous reviews, applications and extensions of that work have appeared.

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

Because of the difficulties and computational expense involved in dealing with the resulting coupled nonlinear partial differential equations, all of the extant works are restricted to one-dimensional geometries.

Recent work with the PdV model includes that of Liou and England,[4] who simulated surface temperatures during a 60-day dry-down period for use in modeling surface radio-brightness. Camillo et al.[5] used a modified PdV formulation to predict measured soil temperatures and compared them to measurements. Scanlon and Milly[6] used a new formulation of the PdV model[7] to simulate water and heat transport in desert soils. Thomas and Sansom,[8] drawing on prior work by Thomas and King,[9] described a 1-D finite element model for the coupled transport problem. The only work in this area directly related to mine detection is apparently that of Pregowski et al.[10] who presented a combination of numerical modeling and experimental measurement to assess the influence of soil moisture on buried mine signatures.

2.1. Fundamental Equations

Our description of the physics begins with a discussion of the partial differential equations that describe both heat and moisture transport in soil.

2.1.1. Moisture Transport

Consider first the case of isothermal conditions (i.e., no temperature gradient exists), in which all of the moisture is in the liquid phase (i.e., no vapor is present). Let $\theta_\ell(\mathbf{r},t)$ be the volume fraction of liquid [m^3/m^3], and let ρ_ℓ [kg/m^3] be the density of liquid, a quantity assumed invariant with space and time. We write

$$X_m(\mathbf{r},t) \quad = \quad \rho_\ell \theta_\ell \tag{1}$$

where X_m is the total water mass in a unit volume of soil [kg/m^3]. Darcy's equation allows us to write

$$\mathbf{q}_m = -\rho_\ell K \nabla(\psi + z) \tag{2}$$

where $\mathbf{q}_m$ is the moisture flux vector [kg m^{-2} s^{-1}], z [m] is elevation, ψ is the matric potential [m] and K [m/s] is the hydraulic conductivity. We have $\psi < 0$ for unsaturated soils and $\psi = 0$ for saturated conditions. The law of continuity

$$\frac{\partial}{\partial t} X_m = -\nabla \cdot \mathbf{q}_m \tag{3}$$

leads to a form of the Richards equation

$$\frac{\partial \theta}{\partial t} = \nabla \cdot [K \nabla(\psi + z)] \tag{4}$$

The presence of water vapor and temperature gradients introduces additional terms into this equation. The effects of these terms become more important near the soil surface, where the temperature gradients are higher. Since thermal mine signatures are based on soil surface temperature, these effects are highly relevant. Phillip and de Vries[2] and de Vries[3] generalized the Richards equation to address thermal and vapor transport. Let θ_v be the volume fraction of air [m^3/m^3] (assumed to be water vapor) in the soil, and let ρ_v [kg/m^3] be the density of vapor. We write

$$X_m(\mathbf{r},t) \quad = \quad \rho_\ell \theta_\ell + \rho_v \theta_v \tag{5}$$

Let S [m^3/m^3] be the porosity of the soil (i.e., the fraction of a volume that can be occupied by liquid or air). We have

$$S = \theta_\ell + \theta_v \tag{6}$$

We can define a total water content θ as follows

$$\theta(\mathbf{r},t) = \frac{X_m(\mathbf{r},t)}{\rho_\ell} = \theta_\ell + \frac{\rho_v}{\rho_\ell}(S - \theta_\ell) \tag{7}$$

Phillip and de Vries modified equation (2) by adding a vapor flux term

$$\mathbf{q}_m/\rho_\ell = -K \nabla(\psi + z) - D_v \nabla \rho_v(\theta, T)/\rho_\ell \tag{8}$$

where D_v [m^2/s] is the isothermal water vapor diffusion coefficient and ρ_v is the absolute humidity of air in the soil matrix. In the original PdV formulation the matric potential ψ and ρ_v are both taken to be functions of θ and T. Applying the chain rule of differential calculus to the ∇ operator leads to an equation for $\mathbf{q}_m/\rho_\ell$ in terms of θ and T.

Noting deficiencies in the PdV formulation, most notably that it fails to address the bulk flow of the air phase, Milly[7] developed an alternative formulation using ψ instead of θ.* Starting with $\rho_v = \rho_v(\psi, T)$ and applying the chain rule for the spatial derivative of ρ_v one obtains

$$\begin{aligned}
\mathbf{q}_m/\rho_\ell &= -(K + D_{\psi v})\nabla\psi(\mathbf{r},t) - D'_{Tv}\nabla T(\mathbf{r},t) - K\mathbf{k} \\
&= -K_{\psi m}\nabla\psi - K_{Tm}\nabla T - J_m\mathbf{k} && (9)\\
K_{\psi m} &= K + D_{\psi v} && (10)\\
K_{Tm} &= D'_{Tv} && (11)\\
J_m &= K && (12)
\end{aligned}$$

in which we have defined the following diffusion coefficients

$$\begin{aligned}
D_{\psi v} &= D_v \frac{\partial \rho_v}{\partial \psi}\bigg|_T && (13)\\
D'_{Tv} &= D_v f \frac{\partial \rho_v}{\partial T}\bigg|_\psi && (14)
\end{aligned}$$

In these expressions the factor f has been introduced, following PdV, to account for larger temperature gradients in local liquid "islands" and in the air phase. Applying the conservation law leads to an analog of the Richards equation

$$\frac{1}{\rho_\ell}\frac{\partial X_m}{\partial t} = \frac{1}{\rho_\ell}\frac{\partial}{\partial t}\left(\rho_\ell\theta_\ell + \rho_v\theta_v\right) = \frac{\partial\psi}{\partial t}M_{\psi m} + \frac{\partial T}{\partial t}M_{Tm} \qquad (15)$$

where

$$\begin{aligned}
M_{\psi m} &= \left(1 - \left(\frac{\rho_v}{\rho_\ell}\right)\right)\frac{\partial\theta_\ell}{\partial\psi}\Big|_T + \frac{(S-\theta_\ell)}{\rho_\ell}\frac{\partial\rho_v}{\partial\psi}\Big|_T && (16)\\
M_{Tm} &= \left(1 - \left(\frac{\rho_v}{\rho_\ell}\right)\right)\frac{\partial\theta_\ell}{\partial T}\Big|_\psi + \frac{(S-\theta_\ell)}{\rho_\ell}\frac{\partial\rho_v}{\partial T}\Big|_\psi && (17)
\end{aligned}$$

Some of the partial derivatives appearing in Eq. (16) and (17) can be evaluated in closed form. When a local equilibrium exists between the soil liquid and vapor, the vapor density is given by

$$\rho_v(\psi, T) = \rho_0(T)h_r(\psi, T) \qquad (18)$$

where ρ_0 is the saturation vapor density and h_r is the relative humidity, given by

$$h_r(\psi, T) = \exp(\psi g/RT) \qquad (19)$$

in which g=9.81 [m s^{-2}] is the acceleration due to gravity and R = 461.5 [J kg^{-1} K^{-1}] is the ideal gas constant for water vapor. Milly[7] presents arguments that such an equilibrium will exist in soil under all conditions with the possible exception of rapid liquid infiltration through coarse soil (e.g., heavy rain on gravel).

Collecting these results, the following equation is obtained for the moisture transfer equation

$$\frac{\partial\psi}{\partial t}M_{\psi m} + \frac{\partial T}{\partial t}M_{Tm} = \nabla\cdot\left[K_{\psi m}\nabla\psi + K_{Tm}\nabla T + J_m\mathbf{k}\right] \qquad (20)$$

*The PdV formulation also fails to address soil moisture hysteresis. The formulation of Milly attempted to address this problem, but it was later found to be flawed.[6]

2.1.2. Thermal Transport

The heat content X_h [J/m^3] of a unit volume of soil can be expressed in several forms, most of which can be related to that derived by de Vries[3]

$$X_h(\mathbf{r},t) = (T-T_0)C + L_{v_0}\rho_v\theta_a - \rho_l\int_0^{\theta_\ell} W d\theta \tag{21}$$

$$C = C_d + c_\ell\rho_\ell\theta_\ell + c_v\rho_v\theta_v \tag{22}$$

where $T(\mathbf{r},t)$ [K] is the local temperature, T_0 [K] is a reference temperature, C [J m^{-3} K^{-1}] is the volumetric heat capacity of moist porous soil, C_d [J m^{-3} K^{-1}] is the volumetric heat capacity of dry porous soil, c_ℓ is the specific heat of liquid water at constant pressure, c_v is the specific heat of water vapor at constant pressure, L_v is the latent heat of vaporization at T_0, and W is the differential heat of wetting. Following Edlefsen and Anderson[11] we have

$$W = -g(\psi - Td\psi/dT) \tag{23}$$

Enforcing the conservation of energy leads to

$$\frac{\partial X_h}{\partial t} = C\frac{\partial T}{\partial t} + L_0\rho_\ell\frac{\partial\theta_v}{\partial t} + c_p\rho_\ell(T-T_0)\frac{\partial\theta_\ell}{\partial t} - \rho_\ell W\frac{\partial\theta_\ell}{\partial t} = -\nabla\cdot\mathbf{q}_h \tag{24}$$

where $\mathbf{q}_h$ is the thermal flux [J m^{-2} s^{-1}].

The thermal flux has components due to conduction, latent heat in the vapor phase, and sensible heat in the liquid phase. Milly[7] writes the formulation of de Vries[3] in the form

$$\mathbf{q}_h = -\lambda\nabla T - \rho_\ell L_{v_0} D_{\psi v}\nabla\psi + c_\ell(T-T_0)\mathbf{q}_m \tag{25}$$

where λ [W m^{-1} K^{-1}] is the thermal conductivity of the porous, moist soil. He then adds to this flux the advection of the heat of wetting due to pressure head gradients, as suggested by Groevenelt and Kay[12] to find

$$\mathbf{q}_h = -K_{Th}\nabla T - K_{\psi h}\nabla\psi - J_h\mathbf{k} \tag{26}$$

where

$$K_{Th} = \lambda + c_\ell(T-T_0)\rho_\ell D'_{Tv} \tag{27}$$

$$K_{\psi h} = -\rho_\ell(L_{v_0}D_{\psi v} + gTD_{Ta}) + c_\ell(T-T_0)\rho_\ell(K + D_{\psi v}) \tag{28}$$

$$J_h = c_\ell(T-T_0)\rho_\ell K \tag{29}$$

In this result D_{Ta} is the diffusion coefficient for adsorbed water under the influence of a temperature gradient. Using $\theta_v = \theta_v(\psi,T)$ and $\rho_v = \rho_v(\psi,T)$ can derive expressions for the LHS of Eq. (24). Omitting the details, and combining the result with Eq. (26) we obtain for the heat transfer equation

$$\frac{\partial}{\partial t}X_h(\mathbf{r},t) = M_{Th}\frac{\partial T}{\partial t} + M_{\psi h}\frac{\partial\psi}{\partial t} = -\nabla\cdot\mathbf{q}_h = \nabla\cdot[K_{Th}\nabla T + K_{\psi h}\nabla\psi + J_h\mathbf{k}] \tag{30}$$

in which

$$M_{Th} = \left[C + H_1\frac{\partial\rho_v}{\partial T} + H_2\frac{\partial\theta_\ell}{\partial T}\right] \tag{31}$$

$$M_{\psi h} = \left[H_1\frac{\partial\rho_v}{\partial\psi} + H_2\frac{\partial\theta_\ell}{\partial\psi}\right] \tag{32}$$

$$H_1 = \theta_v\left[L_0 + c_p(T-T_0)\right] \tag{33}$$

$$H_2 = (c_\ell\rho_\ell - c_p\rho_v)(T-T_0) - (\rho_\ell W + \rho_v L_0) \tag{34}$$

2.2. Surface Boundary Condition

The boundary condition for moisture transport at the air-soil interface must address both evaporation and precipitation. We have[13]

$$\left.\frac{\mathbf{q}_m \cdot \mathbf{k}}{\rho_\ell}\right|_{z=0} = E - P \tag{35}$$

where E is the evaporation rate and P is the precipitation. The evaporation rate is determined from the aerodynamic diffusion relation

$$E = \frac{k^2 U_a \phi}{\rho_\ell [\ln(z_a/z_0)]^2} (\rho_v\big|_{z=0} - \rho_{va}) \tag{36}$$

where k is von Karman's constant (=0.4), z_0 is the surface roughness length, U_a is the wind speed measured at a reference height z_a, and ρ_{va} is the absolute vapor density measured at the same height.

The boundary condition for thermal transport comprises an energy balance. Approaches to this problem have been described by Camillo et al.,[5] Liou and England,[14] and Milly,[13] all of which are similar to

$$\mathbf{q}_h \cdot \mathbf{k}\Big|_{z=0} = -(1-A) I_{sun} - \epsilon I_{atm} + \epsilon\sigma[T(z=0,t)]^4 + \rho_\ell \left[L + c_\ell (T(z=0,t) - T_0)\right] E - \rho_\ell c_\ell (T_a - T_0) P + H \tag{37}$$

where A is the surface albedo at solar wavelengths, I_s is the incoming solar radiation [W/m^2], ϵ is the soil emissivity at thermal wavelengths, I_{atm} [W/m^2] is the longwave emission from the atmosphere to the earth, σ is the Stefan-Boltzman constant, and H is the convective exchange of sensible heat with the atmosphere. The heat absorbed by the soil is given by

$$\mathbf{q}_h \cdot \mathbf{k}\Big|_{z=0} = -\lambda\Big|_{z=0} \frac{\partial T(z=0,t)}{\partial z} \tag{38}$$

The convective losses H are given by

$$H = h_c [T(0,t) - T_a(t)] \tag{39}$$

where

$$h_c = \frac{C_a k^2 U_a \phi}{[\ln(z_a/z_0)]^2} \tag{40}$$

is the heat transfer coefficient for convection. In this expression C_a is the volumetric heat capacity of the air at constant pressure, $T_a(t)$ is the temperature of the atmosphere at the surface, and ϕ is a stability factor related to the Richardson number.[13] It is assumed that meteorological measurements of U_a, I_{atm}, and T_a are available. One can derive a similar form for the radiation terms. We have

$$-\epsilon I_{atm} + 2\epsilon\sigma[T(0,t)^4] = \epsilon\sigma\left[T_a(t)^4 - T(0,t)^4\right] = h_R(T)[T_a(t) - T(0,t)] \tag{41}$$

where

$$h_R(T) = \epsilon\sigma[T_a(t)^2 + T(0,t)^2][T_a(t) + T(0,t)] \tag{42}$$

is a temperature-dependent heat transfer coefficient for radiation losses.

2.3. Parameter Models

The foregoing relations involve numerous parameters, many of which depend on the variables (ψ, T). This dependence leads to nonlinearities in the model. Unlike the transport of heat and moisture, which are based on well-known physical laws, parameter models are often based on measurements or empirical data.

Although the moisture model development is somewhat more straightforward that the thermal transport equations, its parameters are much more complex. Since our primary interest in this work lies with thermal transport, we briefly summarize some useful models for those parameters. In papers published during the 1970s, Mualem[15,16] has provided what is perhaps the most widely used model for the volume fraction of water $\theta(\psi, T)$ and for the hydraulic conductivity K. Some useful approximate forms are also described by Milly.[13] Expressions for the diffusion coefficients are also required and are described in the original PdV formulation. Drawing on prior work by de Vries,[3] Milly[13] presents expressions for $D_{\psi v}$ and for D'_{Tv}. The correction factor f is provided by deVries.[3] A model for the saturated water vapor density has been presented by Kimball et al.[17]

Many of the thermal parameters are derived from mixture models. Following de Vries,[18] we write for the volumetric heat capacity C

$$C = C_\ell \theta_\ell + C_s(1-S) + C_v \theta_v \tag{43}$$

where C_u is the heat capacity of component u. Some typical values for these constants have been tabulated.[18]

The volumetric thermal conductivity λ is estimated from a similar mixture model[18]

$$\lambda = \frac{\lambda_\ell \theta_\ell + \lambda_s(1-S)k_s + (\lambda_a + \lambda_{vap})\theta_v k_a}{\theta_\ell + (1-S)k_s + \theta_v k_a} \tag{44}$$

where k_u is the ratio of the space average of ∇T for particles of type u to ∇T for water. Values of k_u for spherical and cylindrical particles have been given.[18] For spheres we have

$$k_u = \frac{2}{3}\left[1 + \left(\frac{\lambda_u}{\lambda_w} - 1\right) g_i\right]^{-1} + \frac{1}{3}\left[1 + \left(\frac{\lambda_u}{\lambda_w} - 1\right)(1 - 2g_u)\right]^{-1} \tag{45}$$

in which the g_u are shape factors for the constituents of soil. The value of g_u has been tabulated for some solid soil constituents, and g_ℓ is not needed since it multiplies zero in k_ℓ. The value of g_a and values for λ_u are given by Kimball et al.[17]

$$\lambda_v = 0.4186(0.0566 + 0.000153(T - 273)) \tag{46}$$

$$\lambda_\ell = 0.4186(1.32 + 5.59 \times 10^{-3}((T - 273) - 2.63 \times 10^{-5}(T - 273)^2) \tag{47}$$

Those authors also present an appraisal of methods for calculating k_a, a parameter with no physical meaning. It was proposed by De Vries that the thermal conductivity of air in the pore space be calculated as a mixture of air and water vapor vapor

$$\lambda_{av} = \lambda_a + \lambda_v \tag{48}$$

The vapor conductivity is given by[3]

$$\lambda_v = h_r L D_{atm} \nu \frac{d\rho_0}{dT} \tag{49}$$

where L [J/kg] is the latent heat of vaporization for water,[17]

$$L = 4186(595.9 - 0.548(T - 273)) \tag{50}$$

D_{atm} is the molecular diffusivity of water vapor in the air[17]

$$D_{atm} = 0.229 \times 10^{-4}(T/273)^{1.75} \tag{51}$$

and ν is a mass flow factor that accounts for the mass movement of the soil air due to the unequal diffusion rate of air and water vapor molecules

$$\nu = \frac{P}{P - p} \tag{52}$$

In this expression P is the barometric pressure and p is the partial pressure of water vapor, given by

$$p = h\rho_0 RT/M_w \tag{53}$$

in which R is the gas constant (8314.3 [J kg-mole^{-1} K^{-1}]) and M_w is the molecular weight of water (18.016 kg/kg-mole). A model for the saturated water vapor density ρ_0 has been presented by[17]

$$\rho_0(T) = 10^{-3} \exp(19.819 - 4975.9/T) \tag{54}$$

The boundary conditions required additional parameters. Idso et al.[19] note that soil albedo is a function of water content of the form

$$A = \begin{cases} A_d + (A_w - A_d)(2\theta_\ell/S) & 2\theta_\ell/S < 1 \\ A_w & 2\theta_\ell/S < 1 \end{cases} \tag{55}$$

and they present values for the constants. Values of the emissivity at IR wavelengths have been presented by Salisbury and D'Aria.[20,21] The solar radiance I_s and the attenuation of that radiance by the atmosphere can be computed using MODTRAN4,[22] but a convenient closed form approximation is also available.[23]

3. THERMAL MODELING FOR STATIC MOISTURE DISTRIBUTIONS

3.1. Prior Work

The authors have previously described a combined thermal-radiometric model for the IR signature of a buried land mine.[24] Three-dimensional thermal and radiometric models were developed to study the passive IR signature of a mine buried under a rough soil surface. That work is based on a linearized version of the model equations described in Section 2, and is further restricted to the case of a static, uniform water distribution. Evaporation and pressure gradients are ignored, but we retain the dependence of the parameters on the environmental factors, including θ_ℓ. In this case, the classical heat transfer equation is obtained

$$\frac{\partial T(\mathbf{r},t)}{\partial T} = \kappa(\mathbf{r})\nabla^2 T(\mathbf{r},t) \tag{56}$$

in which $\kappa = \lambda/C$ is the diffusivity, a coefficient that can be determined from environmental conditions. The solution of Eq. (56) also involves boundary conditions at the soil surface, which are given by the following linearized version of Eq. (37)

$$-\lambda\Big|_{z=0}\frac{\partial T(z=0,t)}{\partial z} \approx -(1-A)I_{sun} - h_R[T_a(t) - T(z=0,t)] + h_c[T(z=0,t) - T_a(t)] \tag{57}$$

We see that the boundary condition involves additional parameters including λ, the solar insolation I_{sun}, the convection heat transfer coefficient h_c, the radiative heat transfer coefficient h_R, and the local air temperature $T_a(t)$.

The numerical solution of Eq. (56) is based on the finite element method. The mine and soil are approximated by pentahedral elements, and T is described by linear interpolation functions over those elements. The Crank-Nicholson algorithm is used for time-stepping the simulation. The behavior over a periodic diurnal cycle is studied by running the simulation to steady state. Finally, the finite element solution for the surface temperature is combined with reflected radiometric components to predict the signatures seen by an IR camera.

3.2. Model Validation

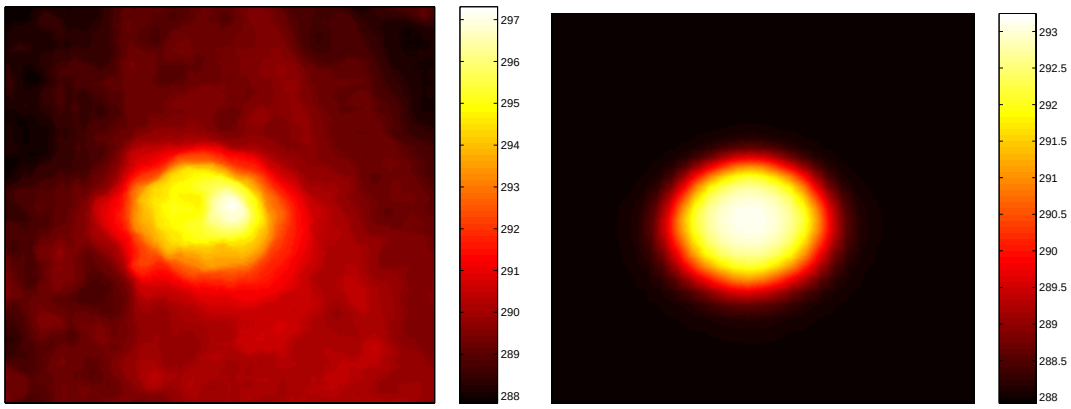
Recently, two efforts were undertaken to validate the model. In the first, we confirmed the validity of several approximations in the FEM formulation by developing an alternative solution, based on an integral equation formulation. Of particular interest were effects related to spatial discretization and boundary conditions, which are obviated in an integral equation formulation. The results of that study are documented separately,[25] where it is shown that the FEM and integral equation code are in good agreement. Since the FEM formulation and the integral formulation have only the heat transfer equation in common, this agreement increases our confidence that both codes are correct.

We have also undertaken a comparison of the model with experimental data. Work to date has used 8-12 μm calibrated images provided by TRW. Those data comprise images of AT mines at times near 10:00A and 12:00 noon. The collections were done at the Fort. A.P. Hill, Site 71A mine lanes #13 (dirt) and #14 (gravel). Collateral information provided with the imagery included soil temperature at multiple depths (0.5", 2", 4" and 8"), wind speed and direction, soil moisture content (one depth), air temperature, barometric pressure, and downwelling and upwelling radiance in the 0.3-3 μm and 3-50 μm bands.

To predict the measured data, it was necessary to make certain assumptions about unknown environmental parameters. Specifically, we assumed values for soil thermal conductivity=1.8 [W/mK] and diffusivity=10^{-6} [m^2/s]. For the mine under consideration here, an EM12 mine surrogate, we assumed a 30 cm diameter and 5.2 cm height. The mine's internal contents are not known, but some sources suggest that a combination of Styrofoam pellets and carnauba wax was used, which we modeled as a good thermal insulator. The resulting FEM mesh was comprised of pentahedral elements with cell dimension of 1.27 cm height and 2.56 cm base. A surface emissivity of 0.98 was assumed. MODTRAN4 was used to estimate the incident radiance as a function of time, and it was found to replicate measured values to an accuracy of about 10%. Other environmental parameters were derived from the available data and models.

The measured data appear in Figure 1(a) after low-pass filtering to reduce small-scale image clutter related to surface roughness and emissivity variations. The results of our modeling effort appear in Figure 1(b), the shape of which is seen to compare well with Figure 1(a). A one-dimensional cut through the data is shown in Figure 2, which helps to quantify the agreement. We see that the signature shape and temperature contrast are fairly well

predicted, but the model prediction is approximately 1 K below the measured data and it shows approximately 1 K less contrast. There are several possible sources for this error including our use of periodic boundary conditions in time. Warmer conditions the previous night can produce an offset that is not easily accounted for without additional data.



(a) Measured data after low-pass filtering to remove surface clutter.

(b) Model prediction.

Figure 1. Measured and modeled results used in validation. The measurements were acquired at 12 noon on a dirt road over an EM12 mine surrogate.

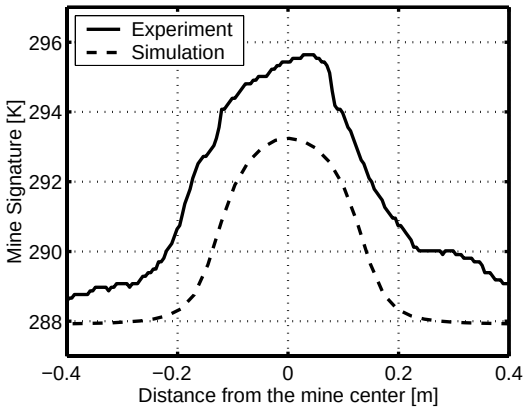


Figure 2. Comparison of model and experiment.

4. PARAMETER STUDIES

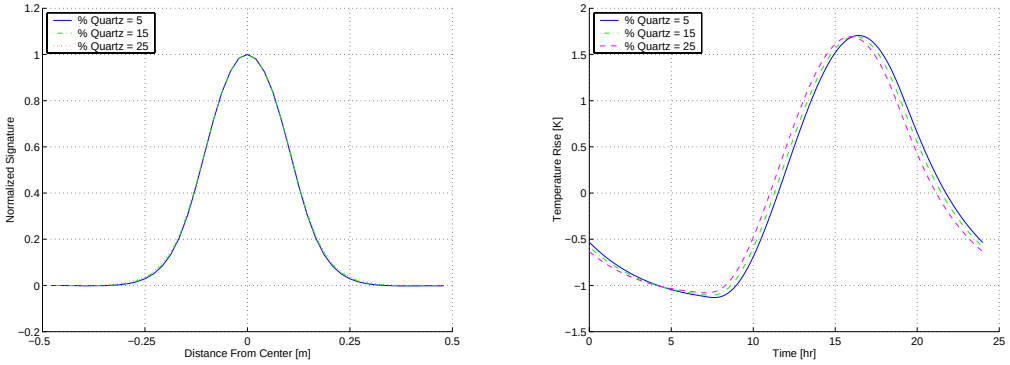
To quantify the effects of various environmental and mine-related parameters we have undertaken a numerical study of the model's sensitivity to these parameters. We used a baseline soil conductivity of 2.6 W m^{-1} K^{-1} and a diffusivity of 5×10^{-7} m^2/s. For the mine parameters we used a conductivity of 0.23 W m^{-1} K^{-1} and a diffusivity of 9.3×10^{-8} m^2/s. The dependence of the mine signature on some fundamental soil parameters is reviewed in Figures 3 and 4. Some additional parameters are studied in a companion paper.[1] In these results we have attempted to quantify the mine's surface signature using two parameters: the signature's surface temperature distribution and the maximum thermal contrast as a function of time. Additional studies reported in the companion paper suggest that, with the exception of very weak signatures that appear near the thermal cross-over time, the shape of the signature does not

change substantially with time. For this reason, all plots of the spatial temperature distribution are normalized to unit magnitude as determined at a time of 1600, when the signature attained its peak.

Thermal conductivity and diffusivity (or heat capacity) are the fundamental parameters which control thermal transport in soil. These parameters depend, in turn, on a number of environmental variables. We have selected soil water and quartz (sand) content as the parameters to be examined here. As noted above, water content has an important role in the physics. Quartz has a large thermal conductivity, which makes sand content a useful surrogate for soil thermal conductivity.

Figure 3 shows the shape and thermal contrast as a result of variations in soil sand content. The effect of varying water content is shown in Figure 4. It is evident that the spatial signature is essentially invariant to these parameters, while the contrast shows a weak effect. Variations in sand and water content have the effect of delaying the mines signature (or shifting its phase). This effect arises primarily as a result of changing the diffusivity of the soil.

Wind speed is another environmental variable that can affect the thermal signature via its effect on the convection coefficient. This parameter is considered in Figure 5. We see that the wind speed has little effect on the shape of the signature, but it has a more pronounced effect on the contrast. Unlike changes in soil parameters, however, it has no effect on the phase of the signature.



(a) Signature spatial dependence

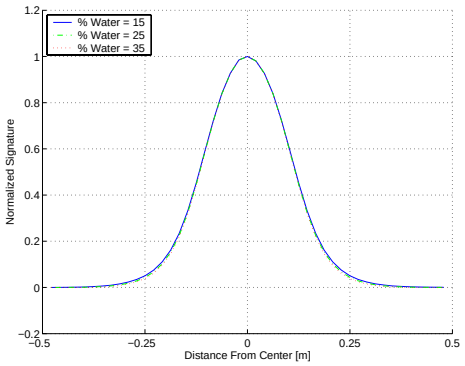
(b) Temporal variation in contrast over center of mine

Figure 3. Apparent size and temperature contrast produced by variations in soil sand content.

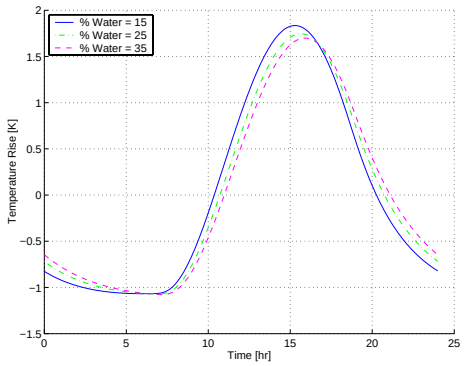
The effects of mine-related variables are reviewed in Figure 6 and 7. Variations in the mine diameter for a fixed depth are shown in Figure 6. It is found that the apparent signature diameter changes strongly with the mine diameter. A three-fold increase in the mine diameter from 10 cm diameter (AP mines) to 30 cm diameter (AT mines) produces less than a two-fold increase in the diameter of the signature at the surface. The temperature rise over the center of the mine, however, shows relatively little change over this range. Variations in burial depth (cf. Figure 7 appear to have little effect on the shape or size of the signature at the surface. The depth, however, has a significant effect on the temperature.

5. SUMMARY AND CONCLUDING REMARKS

We have reviewed the physics that relate environmental effects to thermal IR mine signatures. A model developed for the case of a static homogeneous moisture distribution was exercised to develop a quantitative relation among the signature shape, peak contrast, and various environmental and mine-related variables. We found that soil parameters have little effect on the shape of the signature, but they can make a more significant change in the amplitude and timing of the contrast peak. The wind speed can also affect the signature's peak contrast, but it has no effect on the signature shape or its temporal behavior. The depth of the mine has a significant effect on the amplitude of the contrast but, surprisingly, it has little influence on its shape. As expected, the mine diameter is tightly coupled to the shape of the surface temperature distribution.

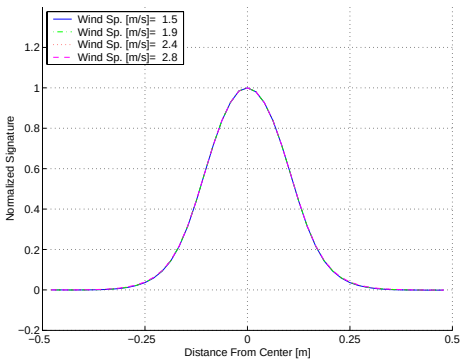


(a) Signature spatial dependence

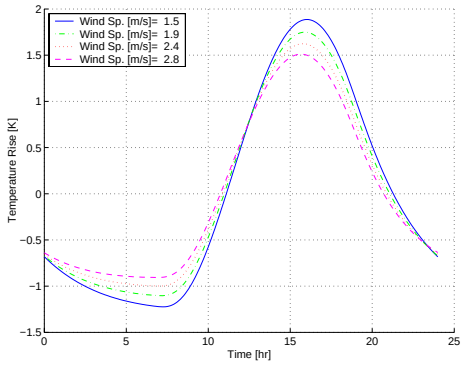


(b) Temporal variation in contrast over center of mine

Figure 4. Apparent size and temperature contrast produced by variations in soil water content.



(a) Signature spatial dependence



(b) Temporal variation of contrast over center of mine

Figure 5. Apparent size and temperature contrast produced by variations in wind speed.

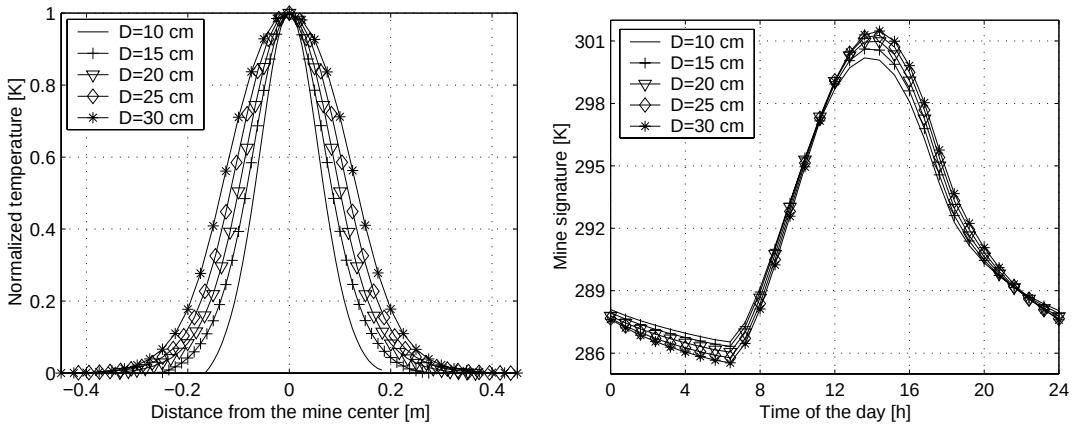
6. ACKNOWLEDGMENTS

The authors would like to thank TRW for providing the experimental results, with special thanks to Ms. Sofia Monawer for her help in interpreting the data.

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

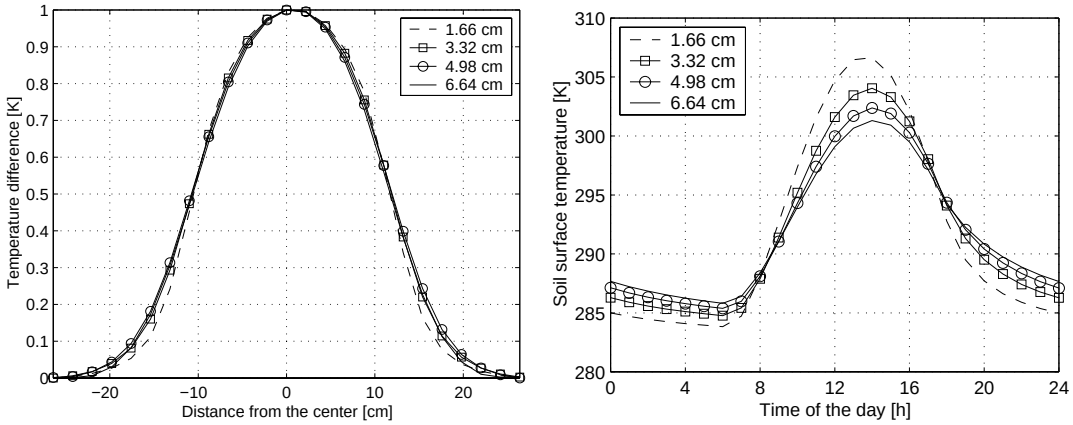
REFERENCES

1. D.-H. Chen, I. K. Sendur, W.-J. Liao, and B. A. Baertlein, "Using physical models to improve thermal ir detection of buried mines," in *Detection and Remediation Technologies for Mines and Minelike Targets VI*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **4394-26**, 2001.
2. J. R. Phillip and D. A. de Vries, "Moisture movement in porous materials under temperature gradients," *Trans. American Geophysical Union (Eos)* **38**(2), pp. 222–232, 1957.



(a) Signature spatial dependence (b) Temporal variation over center of mine

Figure 6. The effect of mine diameter on apparent signature size and thermal time history.



(a) Signature spatial dependence (b) Temporal variation over center of mine

Figure 7. The effect of burial depth on apparent signature size and thermal time history.

3. D. A. de Vries, "Simultaneous transfer of heat and moisture in porous media," *Trans. American Geophysical Union (Eos)* **39**(5), pp. 909–916, 1958.
4. Y.-A. Liou and A. W. England, "A land surface process/radiobrightness model with coupled heat and moisture transport in soil," *IEEE Transactions on Geoscience and Remote Sensing* **36**(1), pp. 273–286, 1998.
5. P. J. Camillo, R. J. Gurney, and T. J. Smugge, "A soil and atmospheric boundary layer model for evapotranspiration and soil moisture studies," *Water Resources Research* **19**(2), pp. 371–380, 1983.
6. B. R. Scanlon and P. C. D. Milly, "Water and heat fluxes in desert soils 2. numerical simulations," *Water Resources Research* **30**(3), pp. 721–733, 1994.
7. P. C. D. Milly, "Moisture and heat transport in hysteretic, inhomogeneous porous media: a matric head-based formulation and a numerical model," *Water Resources Research* **18**(3), pp. 489–498, 1982.
8. H. R. Thomas and M. R. Sansom, "Fully coupled analysis of heat, moisture, and air transfer in unsaturated soils," *J. Eng. Mech.* **121**, pp. 392–405, 1995.

9. H. R. Thomas and S. D. King, "Coupled temperature/capilary potential variations in unsaturated soil," *Journal of Engineering Mechanics* **117**(11), pp. 2475–2491, 1991.
10. P. Pregowski, W. Swiderski, R. T. Walczak, and K. Lamorski, "Buried mine and soil temperature prediction by numerical model," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, T. J. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **4038**, pp. 1392–1403, 2000.
11. N. E. Edlefsen and A. B. C. Anderson, "The thermodynamics of soil moisture," *Hilgardia* **16**, pp. 31–299, 1943.
12. P. H. Groevenelt and B. D. Kay, "On the interaction of water and heat in frozen and unfrozen soils, 2. the liquid phase," *Soil Sci. Soc. Am. Proc.* **38**(3), pp. 400–404, 1974.
13. P. C. D. Milly, "A simulation analysis of thermal effects on evaporation from soil," *Water Resources Research* **20**(8), pp. 1087–1098, 1984.
14. Y.-A. Liou and A. W. England, "Annual temperature and radiobrightness signatures for bare soils," *IEEE Transactions on Geoscience and Remote Sensing* **34**(4), pp. 981–998, 1996.
15. Y. Mualem, "Extension of the similarity hypothesis used for modeling the soil water characterstics," *Water Resources Research* **13**(4), pp. 773–780, 1977.
16. Y. Mualem, "A conceptual model of hysteresis," *Water Resources Research* **10**(3), pp. 514–520, 1974.
17. B. A. Kimball, R. D. Jackson, R. J. Reginato, F. S. Nakayama, and S. B. Idso, "Comparison of field-measured and calculated soil-heat fluxes," *Soil Science Society of America Journal* **40**, pp. 18–25, 1976.
18. D. A. de Vries, "Heat transfer in soils," in *Heat and mass transfer in the biosphere: Part 1. Transfer processes in the in the plant environment*, D. A. deVries and N. H. Afgan, eds., ch. 1, pp. 5–28, Hemisphere Press, 1975.
19. S. B. Idso, R. D. Jackson, R. J. Reginato, B. A. Kimball, and F. S. Nakayama, "The dependence of bare soil albedo on soil water content," *J. Applied Meteor.* **14**(1), pp. 109–113, 1975.
20. J. W. Salisbury and D. M. D'Aria, "Emissivity of terrestrial materials in the 8-14 μm atmospheric window," *Remote Sensing Environment* **42**, pp. 83–106, 1992.
21. J. W. Salisbury and D. M. D'Aria, "Emissivity of terrestrial materials in the 3-5 μm atmospheric window," *Remote Sensing Environment* **47**, pp. 345–361, 1994.
22. A. Berk, L. S. Bernstein, and D. C. Robertson, "MODTRAN: a moderate resolution model for LOWTRAN7," Tech. Report GL-TR-89-0122, Air Force Geophysical Laboratory, Hanscom AFB, MA, October 21 1989.
23. A. W. England, "Radiobrightness of diurnally heated, freezing soil," *IEEE Trans. Geosci. Remote Sensing* **28**(4), pp. 464–476, July 1990.
24. I. K. Sendur and B. A. Baertlein, "Simulation of thermal IR buried mine signatures," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, T. J. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **4038**, pp. 156–166, 2000.
25. I. K. Sendur and B. A. Baertlein, "Reference solution for thermal mine signature modeling," in *Detection and Remediation Technologies for Mines and Minelike Targets VI*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **4394-24**, 2001.

Analysis of Polarimetric IR Phenomena for Detection of Surface Mines

İbrahim Kürşat Şendur, Joel T. Johnson, and Brian A. Baertlein

The Ohio State University, ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

It has long been recognized that surface-laid land mines and other man-made objects tend to have different polarization characteristics than natural materials. This fact has been used to advantage in a number of mine detecting sensors developed over the last two decades. In this work we present the theoretical basis for this polarization dependence. The theory of scattering from randomly rough surfaces is employed to develop a model for scattering and emission from mines and natural surfaces. The emissivity seen by both polarized and unpolarized sensors is studied for smooth and rough surfaces. The polarized and unpolarized emissivities of rough surfaces are modeled using the solution of the reciprocal active scattering problem via the second order small perturbation method/small slope approximation (SPM/SSA). The theory is used to determine the most suitable angle for passive polarimetric IR detection of surface mines.

Keywords: Thermal infrared imagery, land mines, polarimetry, small perturbation method, small slope approximation, modeling, emissivity

1. INTRODUCTION

Infrared (IR) sensors have several attractive properties for detection of surface-laid mines, including an ability to detect mines at long range and their independence of the mine's metal content. The latter is particularly attractive in detecting non-metallic mines and in using IR sensors as part of a fused sensor suite, in which they complement GPR and EMI sensors.

A number of previously fielded IR systems exploit polarization to detect surface-laid and flush-buried mines. Man-made objects such as mines have smooth surfaces and both reflect and emit radiation depending on their electrical and thermal properties and the imaging geometry. Conversely, natural objects such as soil, have rougher surfaces and the reflected and emitted radiation depends on the statistical surface characteristics. As a result, natural backgrounds tend to have a different polarization signature than man-made objects and, hence, polarization can be a good basis on which to discriminate man-made and natural objects.

Several groups have previously explored polarization for mine detection. DiMarzio et al.[1] investigated the use of polarimetric measurements with IR spectral imagery. They have found that polarimetric signatures in the 8-12 μm band were promising and complementary to spectral imaging for flush-buried land mines. Furthermore, they confirmed the earlier findings of Johnson et al.[2,3] that disturbed soil tended to show a polarimetric signal. They suggested that this emissivity variations may not detect mines alone, but it could be used in conjunction with other sensors to reduce the false alarm rate. Larive et al.[4] made an experimental study of mine detection using polarimetric imagers. Their findings suggest that polarimetric IR imagery increases the contrast between the mines and the background soil, and that high spatial resolution offers additional benefits. In addition, they reported that the 8-12 μm band is preferable to the 3-5 μm band for polarimetric imagery. In a more recent study, Larive et al.[5] presented an automatic detection algorithm for surface-laid and flush-buried mines in polarimetric imagery. The algorithm depends on the extraction of two features, namely axis tilt and ellipticity. They reported success in locating surface mines.

Barbour et al.[6] experimentally explored the effectiveness of polarimetric IR imagery for mine detection. They reported a benefit for polarimetric IR imagery, especially under heavy background clutter. They also pointed out the high signal-to-clutter ratio and the natural dual mode capability of thermography and polarization in the same

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

Detection and Remediation Technologies for Mines and Minelike Targets VI, Abinash C. Dubey, James F. Harvey, J. Thomas Broach, Vivian George, Editors, Proceedings of SPIE Vol. 4394 (2001) © 2001 SPIE · 0277-786X/01/$15.00

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

image frame. Barnes et al.[7] described a 3-5 μm polarimetric system. For detection of surface mines, the proposed system uses polarimetric features such as degree of polarization (DoP), degree of linear polarization (DoLP), degree of circular polarization, ellipticity and orientation of major axis. Their experimental studies suggest that DoP and DoLP tend to highlight man-made objects, ellipticity tends to highlight all the objects within the scene equally, and the polarization vector orientation contains information regarding the relative orientation of surfaces within the scene.

The detection performance of a 3-5 μm camera and a visible (VIS) camera with and without polarization filter was investigated by de Jong et al.[8] They reported higher detection rates for classification based on polarization features than based on intensity only. They further state that use of a polarization filter enhances IR performance, especially when there is low thermal contrast. Based on their experimental studies in homogeneous sand, the VIS camera performed better than the MWIR sensor. Conversely, in a forest background where there is extensive natural clutter, the performance of the MWIR was reported to be superior. They reported further improvement when intensity and polarimetric results were fused.

Active sensors can be used for polarimetric detection independent of the diurnal cycle. Previously, active polarimetric systems have been used for various purposes. Miles et al.[9] reported a polarization-based active/passive sensor system which has been developed for the U.S. Army's STandoff MInefield Detection System (STAMIDS). They used two channels of near IR polarization reflectance information in addition to conventional thermal information. Although they reported good detection capability for both night and day operation, the detection rates were low for buried mines. Based on their studies, they concluded that the supplementary active polarization and reflectance data were superior to thermal data alone.

Although the existing experimental and empirical studies suggest that polarization information can improve IR mine detection, effective use of polarimetric IR sensors requires that we understand the emission characteristics of the mine and soil. To address this need, in this work a model based on second order small perturbation method/small slope approximation (SPM/SSA) is developed to study the effects of material composition, geometry, and statistical surface properties. This work begins with a description of the problem and the radiometric quantities in Sect. 2. The polarimetric emissivity vector and the unpolarized emissivity are introduced in Sect. 3. Evaluation of the emissivity vector via the reciprocal active scattering problem using bistatic scattering coefficients is also discussed in that section. In Sect. 4 we introduce the polarized and unpolarized emissivity of a perfectly flat surface, which will be used in modeling mines. A second order SPM/SSA solution of scattering from a rough surface is presented in Sect. 5. Using this solution, the polarimetric emissivity vector and the unpolarized emissivity are also formulated. Modeling results for polarimetric signatures of mines and soil surfaces are presented in Sect. 6. A summary and concluding remarks appear in Sect. 7.

2. PROBLEM DESCRIPTION

The spectral radiance of a blackbody is given by Planck's radiation law

$$I_{BB}(\lambda, T) = \frac{2c^2 h}{\lambda^5} \frac{1}{e^{h\nu/kT} - 1} \qquad [\mathrm{W\ m^{-2} sr^{-1} \mu m^{-1}}], \tag{1}$$

where $h = 6.63 \times 10^{-34}$ [J s] is Planck's constant, ν [Hz] is the frequency of the optical radiation, λ [m] is the wavelength of that radiation, $k = 1.38 \times 10^{-23}$ [J K^{-1}] is Boltzmann's constant, c is the speed of light [m s^{-1}] and T [K] is the temperature. Spectral distribution of the blackbody emission given by Planck's law is a function of both wavelength and temperature.

Real materials emit less than a blackbody and the emitted spectral radiance generally depends on the direction and polarization in addition to the temperature of the object and the frequency. The IR camera receives spectral radiance I_β emitted by the object, where β denotes the polarization. This can be expressed as

$$I_\beta(\lambda, T, \theta, \phi) = \mathcal{E}_\beta(\lambda, \theta, \phi) I_{BB}(\lambda, T) \tag{2}$$

Furthermore, the total spectral radiance emitted by an opaque object also involves the reflected radiance. The reflectance for an opaque object can be determined via Kirchoff's law. Taking the reflected radiation into account, the total spectral radiance from soil and mine can be expressed as

$$I_\beta^{mine}(\lambda, T_{mine}, \theta, \phi) = \mathcal{E}_\beta^{mine}(\lambda, \theta, \phi) I_{BB}(\lambda, T_{mine}) + \left(1 - \mathcal{E}_\beta^{mine}(\lambda, \theta, \phi)\right)\left(I_{sky}(\lambda) + I_{sun}(\lambda)\right) \tag{3}$$

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

$$I_\beta^{soil}(\lambda, T_{soil}, \theta, \phi) = \mathcal{E}_\beta^{soil}(\lambda, \theta, \phi) I_{BB}(\lambda, T_{soil}) + \left(1 - \mathcal{E}_\beta^{soil}(\lambda, \theta, \phi)\right)\left(I_{sky}(\lambda) + I_{sun}(\lambda)\right) \quad (4)$$

The components in Eqs. (3) and (4) are illustrated in Fig. 1. In Eqs. (3) and (4) the emissivity expressions depend

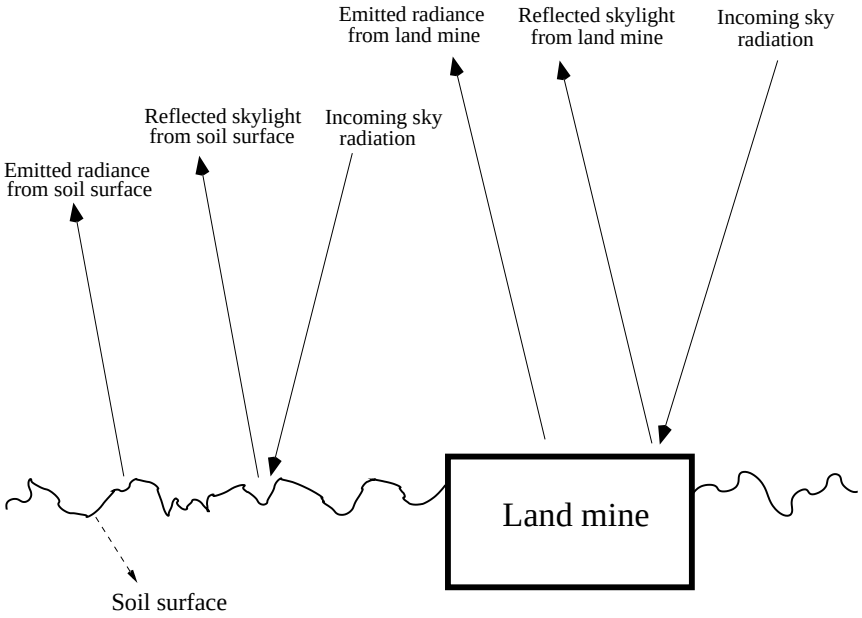


Figure 1. The reflected and emitted radiance components from the rough soil surface and the smooth land mine surface.

on the composition (optical properties) of the soil and mine, statistical description of the surface, sensor geometry, wavelength and polarization. It should be noted that, generally, the soil and mine have different optical properties. Throughout this work the mine surface is assumed to be a smooth surface, while the soil is assumed rough. This dissimilarity in material properties and surface roughness are the main sources of polarimetric signatures.

3. EMISSIVITY FOR POLARIZED AND UNPOLARIZED SENSORS

In this section we focus on the polarized and unpolarized emissivity variations. The polarization states of the electric field are characterized by four quantities known as Stokes parameters, which we can write in the following vector form:

$$\mathbf{I} = \begin{bmatrix} I_h \\ I_v \\ U \\ V \end{bmatrix} = \begin{bmatrix} \langle E_h E_h^* \rangle \\ \langle E_v E_v^* \rangle \\ 2\mathrm{Re}\langle E_h E_v^* \rangle \\ 2\mathrm{Im}\langle E_h E_v^* \rangle \end{bmatrix} \quad (5)$$

where E_h and E_v are the horizontal and vertical polarizations of the fields sensed by the receiver. Yueh and Kwok[10] state that the emission vector $\mathcal{E}$ can be related to the Stokes parameters as

$$\mathcal{E} = \begin{bmatrix} \mathcal{E}_h \\ \mathcal{E}_v \\ \mathcal{E}_U \\ \mathcal{E}_V \end{bmatrix} = c' \begin{bmatrix} \langle E_h E_h^* \rangle \\ \langle E_v E_v^* \rangle \\ 2\mathrm{Re}\langle E_h E_v^* \rangle \\ 2\mathrm{Im}\langle E_h E_v^* \rangle \end{bmatrix} \quad (6)$$

where c' is a constant for a fixed frequency.

Kirchoff's law relates the emissivity of a body to its absorptivity. For an object in thermal equilibrium, the energies emitted and absorbed by the object are equal. Therefore, the emissivity of the object can be obtained by solving for the absorptivity of the object using the reciprocal active scattering problem.[11] Consider a plane wave E_α^i incident on a medium with area A. The power intercepted by the surface area A is

$$P_i = \frac{|E_\alpha^i|^2 A \cos\theta_i}{2\mu} \quad (7)$$

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

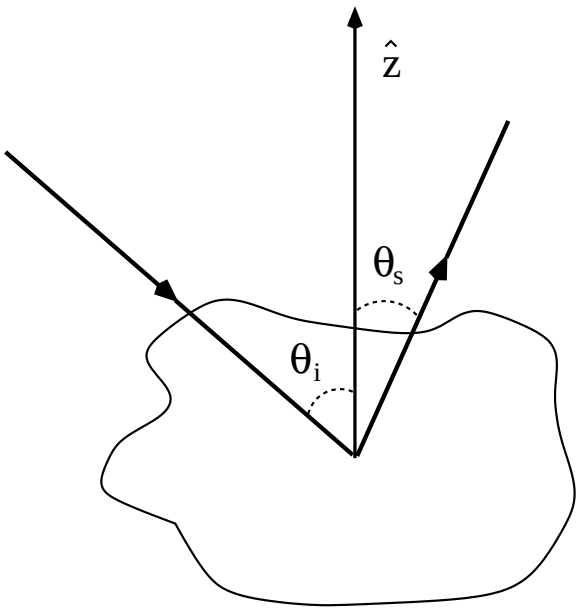


Figure 2. Reciprocal active scattering problem geometry.

where θ_i is the incidence angle with respect to the surface normal and μ is the free space impedance. The total scattered power at range R can be written as

$$P_s = \sum_{\beta=v,h} \int_0^{\pi/2} d\theta_s \int_0^{2\pi} d\phi_s R^2 \sin\theta_s \frac{|E_\beta^s|^2}{2\mu} \tag{8}$$

where θ_s and ϕ_s describe the scattering directions. Using Eqs. (7) and (8) we can write the absorptivity as

$$a_\alpha = \frac{P_i - P_s}{P_i} = 1 - \frac{1}{4\pi} \sum_{\beta=v,h} \int_0^{\pi/2} d\theta_s \sin\theta_s \int_0^{2\pi} d\phi \gamma_{\beta\alpha}(\theta_s, \phi_s; \theta_i, \phi_i) \tag{9}$$

where $\gamma_{\beta\alpha}(\theta_s, \phi_s; \theta_i, \phi_i)$ is the bistatic scattering coefficient defined as

$$\gamma_{\beta\alpha}(\theta_s, \phi_s; \theta_i, \phi_i) = \frac{4\pi R^2 |E_\beta^s|^2}{|E_\alpha^i|^2 A cos\theta_i}. \tag{10}$$

Kirchoff's law and the well known relation between emissivity and reflectivity r for an opaque body

$$\epsilon = 1 - r \tag{11}$$

allows us to write

$$\mathcal{E}_h = 1 - r_h = 1 - \frac{1}{4\pi} \sum_{\beta=v,h} \int_0^{\pi/2} d\theta_s \sin\theta_s \int_0^{2\pi} d\phi \gamma_{\beta h}(\theta_s, \phi_s; \theta_i, \phi_i) \tag{12}$$

$$\mathcal{E}_v = 1 - r_v = 1 - \frac{1}{4\pi} \sum_{\beta=v,h} \int_0^{\pi/2} d\theta_s \sin\theta_s \int_0^{2\pi} d\phi \gamma_{\beta v}(\theta_s, \phi_s; \theta_i, \phi_i) \tag{13}$$

in which r_h and r_v are horizontal and vertical reflectivities, which are obtained by integrating the bistatic scattering coefficient over the upper hemisphere. An unpolarized sensor will measure

$$\mathcal{E}_{up} = \frac{\mathcal{E}_h + \mathcal{E}_v}{2} \tag{14}$$

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

4. EMISSIVITY OF PERFECTLY FLAT SURFACE

In this section we review the emissivity seen by both polarized (vertical and horizontal) and unpolarized sensors. The IR detector is assumed to view a perfectly flat, dielectric surface, and the direction of observation makes an angle θ with the surface normal ($\theta = 0^\circ$ corresponds to nadir viewing). The reflectivities r_h and r_v of a flat surface are determined from the Fresnel reflection coefficients as

$$r_h = |R_h|^2 \tag{15}$$
$$r_v = |R_v|^2 \tag{16}$$

where the coefficients for horizontal and vertical polarizations are defined by

$$R_h = \frac{k_{zi} - k_{1zi}}{k_{zi} + k_{1zi}} \tag{17}$$
$$R_v = \frac{\epsilon_1 k_{zi} - \epsilon_0 k_{1zi}}{\epsilon_1 k_{zi} + \epsilon_0 k_{1zi}} \tag{18}$$

in which

$$k_{zi} = k\cos\theta_i \tag{19}$$
$$k_{1zi} = k\sqrt{(\epsilon_1/\epsilon_0) - \sin^2\theta_i} \tag{20}$$

$k = 2\pi f/c$ is the free space wavenumber, and ϵ_0 and ϵ_1 are the permittivities of the air and soil regions, respectively. In Eq. (18), k_{zi} and k_{1zi} represent the z component of the wave vectors in air and soil regions, respectively. Using these equations together with Eqs. (12), (13), and (14), the polarized and unpolarized emissivities can be obtained. In this work we consider mines with plastic surfaces. Refractive index values (ϵ_1/ϵ_0) of 1.4 to 1.6 are typical of a variety of plastics in the optical regime.[12–15]

Consider first the polarized and unpolarized emissivities as a function of observation angle. The state of polarization for the reflected radiation is defined as

$$p = \frac{r_h - r_v}{r_h + r_v} \tag{21}$$

In Fig. 3 the horizontally and vertically polarized emissivities, unpolarized emissivity, and the state of polarization of the reflected radiation are presented for dielectric surfaces with the previously noted refractive index values. As expected, the horizontally and vertically polarized emissivities are equal at $\theta = 0$ and vanish at $\theta = 90^\circ$. The vertically polarized emissivity attains a maximum of unity at the Brewster angle, at which there is no reflection for the vertical component. As the relative permittivity of the surface increases, the emissivity values decrease. Figure 3 (c) presents the emissivities seen by an unpolarized sensor. We noted above that for angles smaller than the Brewster angle, there is little variation in the unpolarized emissivity. This observation suggests that the diffuse surface approximation can be used if the observation angle is smaller than the Brewster angle. In Fig. 3 (d) the reflected polarization is plotted as a function of observation angle. This value is zero when the horizontal and vertical components are equal (nadir viewing) and unity when the vertical reflected component is zero (Brewster angle).

5. THE EFFECTS OF SURFACE ROUGHNESS ON IR EMISSION

The emissivities seen by polarized and unpolarized sensors are now investigated for the rough surface case. The rough surface scattering problem will be solved to obtain the bistatic scattering coefficients, which can be used to determine the emissivity. Several approximate theories can be used to solve the rough surface scattering problem. In this work we present a second order SPM/SSA solution.

The SPM/SSA is essentially an iterative technique for rough surfaces with small slopes. A second-order SPM formulation decomposes the fields into coherent and incoherent parts. Zeroth-order SPM solutions are the reflected and transmitted fields of a perfectly smooth surface. These fields are characterized by the Fresnel reflection coefficients. A first-order SPM solution gives the lowest-order incoherent fields. A second-order solution gives the lowest-order correction to the coherent fields. SPM assumes that the surface variations are much smaller than the incident wavelength, but recent studies have shown that[16] SPM can be used for emissivity calculations even when the surface height is large, if the small slope condition is satisfied.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

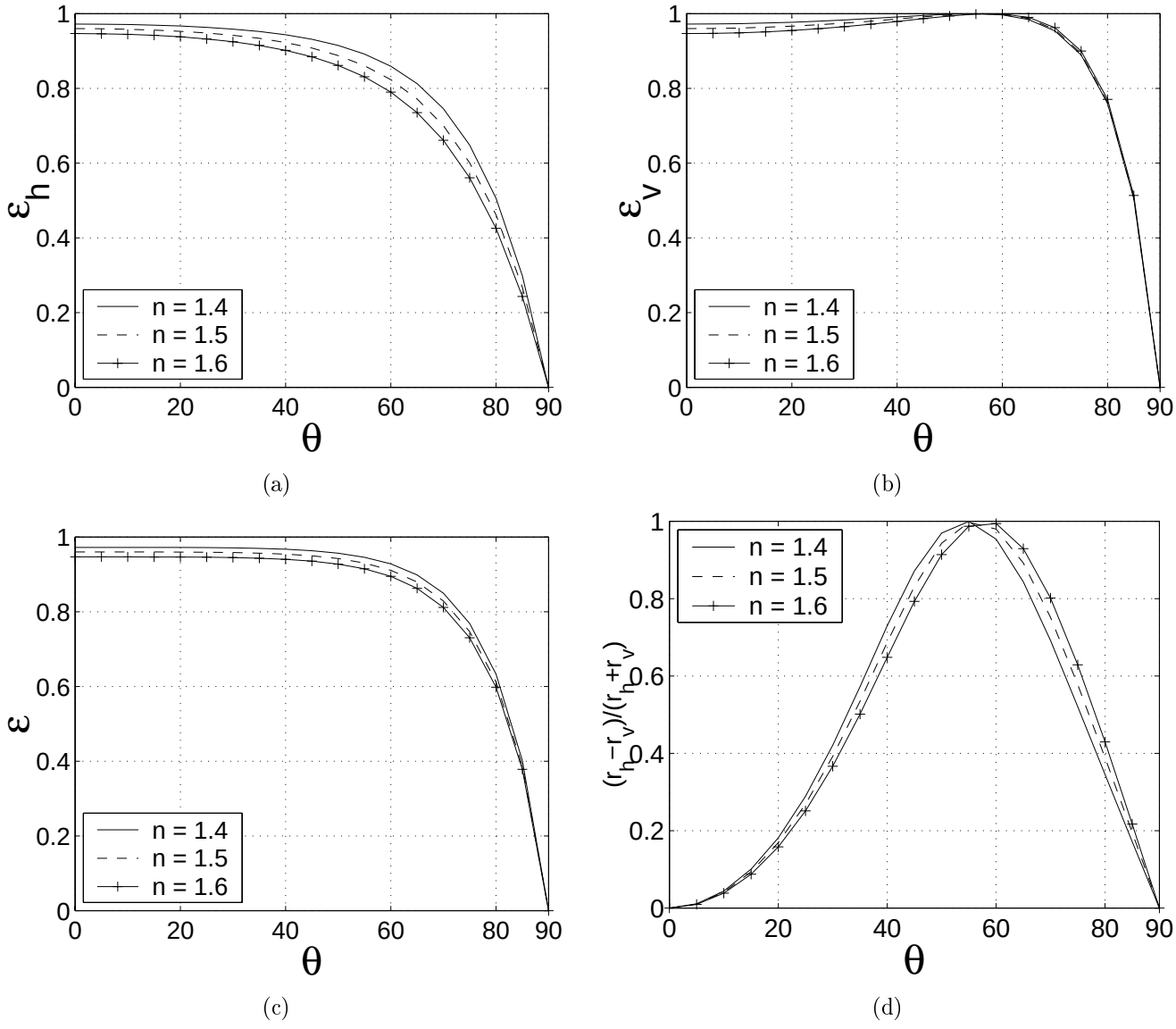


Figure 3. Emissivity and received polarization as a function of observation angle for plastic surfaces having different permittivity values. (a) Horizontally polarized emissivity. (b) Vertically polarized emissivity. (c) Emissivity seen by an unpolarized sensor.(d) Polarization of the reflected radiation.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

Yueh et al.[17] presented the second-order reflectivity matrix in a compact form. Using the formulations in Section 3 and Appendices 2 and 3 of the aforementioned paper, the emissivity calculations can be done, but there are several ambiguities in the formulation, and the approach to numerical evaluation of the integrals is unclear. Johnson and Zhang,[16] referred to hereafter as JZ, reported that separate evaluation of the coherent and incoherent terms requires very accurate evaluation of these integrals and may cause numerical problems due to the cancellation of large numbers. JZ addressed those problems by combining the coherent and incoherent terms in a compact form for the emissivity.

The emissivity expressions of JZ for a second order SPM/SAA are

$$\begin{bmatrix} \mathcal{E}_h \\ \mathcal{E}_v \\ \mathcal{E}_U \\ \mathcal{E}_V \end{bmatrix} = \begin{bmatrix} 1-|R_{hh}^{(0)}|^2 \\ 1-|R_{vv}^{(0)}|^2 \\ 0 \\ 0 \end{bmatrix} - \int_0^\infty dk'_\rho k'_\rho W(k'_\rho,\phi') \begin{bmatrix} g_h(f,\theta_i,\phi_i,\epsilon,k'_\rho,\phi') \\ g_v(f,\theta_i,\phi_i,\epsilon,k'_\rho,\phi') \\ g_U(f,\theta_i,\phi_i,\epsilon,k'_\rho,\phi') \\ g_V(f,\theta_i,\phi_i,\epsilon,k'_\rho,\phi') \end{bmatrix} \tag{22}$$

where R_{hh}^0 and R_{vv}^0 are the horizontally and vertically polarized flat surface Fresnel reflection coefficients, respectively, and $W(k'_\rho,\phi')$ is the surface spectrum. The quantities g_h, g_v, g_U, and g_V are weighting functions given by

$$g_h(f,\theta_i,\phi_i,\epsilon,k'_\rho,\phi') = 2\mathrm{Re}\{R_{hh}^{(0)*} f_{hh}^{(2)}\} + \frac{k_{zi}}{k_z}[|f_{hh}^{(1)}|^2 + |f_{hv}^{(1)}|^2]F \tag{23}$$

$$g_v(f,\theta_i,\phi_i,\epsilon,k'_\rho,\phi') = 2\mathrm{Re}\{R_{vv}^{(0)*} f_{vv}^{(2)}\} + \frac{k_{zi}}{k_z}[|f_{vv}^{(1)}|^2 + |f_{vh}^{(1)}|^2]F \tag{24}$$

$$g_U(f,\theta_i,\phi_i,\epsilon,k'_\rho,\phi') = 2\mathrm{Re}\{(R_{hh}^{(0)*} - R_{vv}^{(0)*}) f_{hv}^{(2)}\} + \frac{2k_{zi}}{k_z}\mathrm{Re}\{f_{vh}^{(1)} f_{hh}^{(1)*} + f_{vv}^{(1)} f_{hv}^{(1)*}\}F \tag{25}$$

$$g_V(f,\theta_i,\phi_i,\epsilon,k'_\rho,\phi') = 2\mathrm{Im}\{(R_{hh}^{(0)*} + R_{vv}^{(0)*}) f_{hv}^{(2)}\} + \frac{2k_{zi}}{k_z}\mathrm{Im}\{f_{vh}^{(1)} f_{hh}^{(1)*} + f_{vv}^{(1)} f_{hv}^{(1)*}\}F \tag{26}$$

The $f_{\alpha\beta}^1$ and $f_{\alpha\beta}^2$ functions are given in Appendices 2 and 3 of Yueh et al.[17] with modifications proposed by JZ. The function F was proposed by JZ to limit incoherent contributions to the upper hemisphere.

As noted by JZ, these results may be simplified extensively. In the absence of cultural or agricultural features (e.g., crop furrows), it is reasonable to assume that the soil surface spectrum is isotropic. JZ has shown that as a result of symmetry, $\mathcal{E}_U$ and $\mathcal{E}_V$ vanish in that case. The soil surface height variation are much larger than the sensor wavelengths, which permits a further approximation that leads to the result

$$\begin{bmatrix} \mathcal{E}_h \\ \mathcal{E}_v \end{bmatrix} = \begin{bmatrix} 1-|R_{hh}^{(0)}|^2 \\ 1-|R_{vv}^{(0)}|^2 \end{bmatrix} - \begin{bmatrix} S^2 h_{h,l}^{(0)}(\theta_i,\epsilon) \\ S^2 h_{v,l}^{(0)}(\theta_i,\epsilon) \end{bmatrix} \tag{27}$$

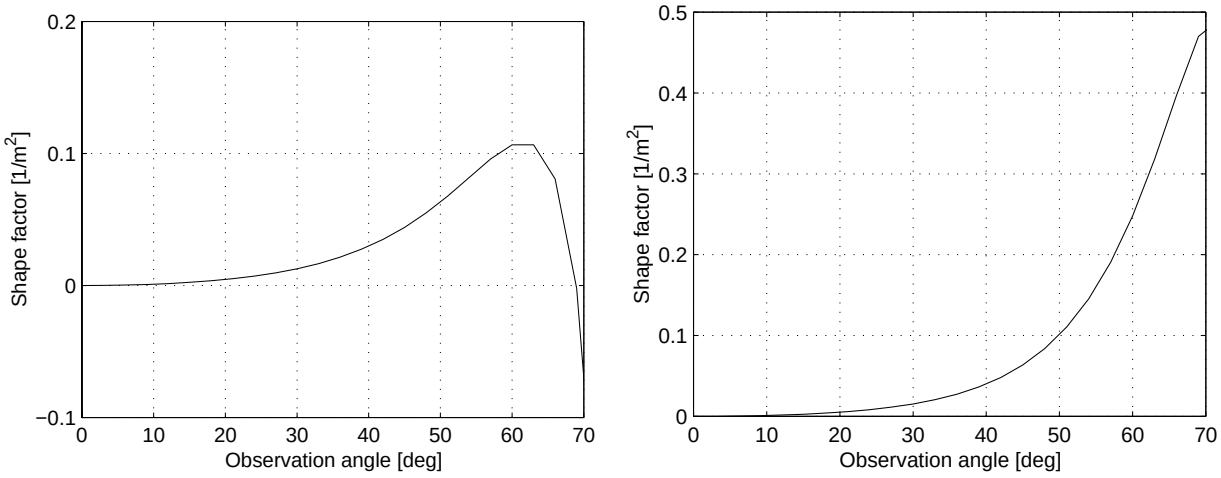
where S is the slope variance and $h^{(0)}$ are shape functions defined as an asymptotic limit of g functions (See JZ). Eq. (27) suggests that the emissivity vector for a rough soil surface is a function of observation angle, soil permittivity, and surface slope variance.

6. RESULTS

It is a straightforward matter to use Eq. (27) to compute soil emissivity. For this work we required the complex permittivity of soil in the IR spectral regime. The composition of soil and the thermal and optical properties of its components are important in many applications,[18–22] but surprisingly little data is available in the IR regime. Refractive index values in the visible and near-IR portions of the spectrum are available[23,24] and will be used here. Around 0.616 μm spectral region, the refractive index of soil with a mean particle radius of 10 μm is approximated as $1.6 + 0.002j$. Using this value the shape functions in Eq. (27) can be calculated, and they appear in Figure 4. From these data we compute the polarized and unpolarized emissivities and the state of polarization for a rough surface. The results are illustrated in Fig. 5 for different slope surface variances.

Predictions can be made from these data regarding the performance of polarimetric sensors of surface mines. A key issue is the polarization metric being used for detection. As noted in Section 1, different metrics have been used. The state of polarization p for the reflected radiation, which was defined with Eq. (21), can be used for this purpose, and it is plotted as a function of observation angle for mine-like surfaces and soil in Figs. 3 (d) and 5 (d), respectively.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms



(a) Horizontal (b) Vertical

Figure 4. Shape factors $h^{(0)}_{\gamma,l}(\theta_i, \epsilon)$ as a function of observation angle for soil at a wavelength of 0.616 μm.

The results suggest that for the very smooth soil case, the mine and plastic surfaces produce similar p values. As the soil roughens, however, the peak value of p moves to smaller observation angles. Operating the sensor at the angle of optimum p may lead to improved detection.

Another widely used metric for mine detection is the DoLP, which is defined as

$$\text{DoLP} = \frac{I_h(\lambda, T, \theta, \phi) - I_v(\lambda, T, \theta, \phi)}{I_h(\lambda, T, \theta, \phi) + I_v(\lambda, T, \theta, \phi)} \tag{28}$$

When the physical temperature of the soil surface and the mine surface are equal, i.e., when $T_{mine} = T_{soil}$, then DoLP is related to differences in the polarimetric emissivity terms $\mathcal{E}_\beta(\lambda, \theta, \phi)$. In this case Eq. (28) can be written as

$$\text{DoLP} = \frac{\big(\mathcal{E}_h - \mathcal{E}_v\big)}{\big(\mathcal{E}_h + \mathcal{E}_v\big) + \dfrac{2\big(I_{sky}(\lambda) + I_{sun}(\lambda)\big)}{\big(I_{BB}(\lambda, T_{mine}) - I_{sky}(\lambda) - I_{sun}(\lambda)\big)}} \tag{29}$$

In the visible and near-IR regime we have $\big(I_{sun}(\lambda) + I_{sky}(\lambda)\big) \gg I_{BB}(\lambda, T_{mine})$, and Eq. (29) can be approximated by

$$\text{DoLP} \approx \frac{\big(\mathcal{E}_h - \mathcal{E}_v\big)}{\big(\mathcal{E}_h + \mathcal{E}_v\big) - 2} \qquad \text{visible and near-IR} \tag{30}$$

Conversely, for the LWIR regime we have $\big(I_{sun}(\lambda) + I_{sky}(\lambda)\big) \ll I_{BB}(\lambda, T_{mine})$, which leads to

$$\text{DoLP} \approx \frac{\big(\mathcal{E}_h - \mathcal{E}_v\big)}{\big(\mathcal{E}_h + \mathcal{E}_v\big)} = p \qquad \text{LWIR} \tag{31}$$

We can compute the DoLP for a mine-like plastic surface and soil. To simulate the plastic mine we have chosen Nylon-66 (Polyhexamethylene-adipamide), which has a refractive index[15] of 1.53. Using the soil refractive index noted above for the visible band, yields the results presented in Figs. 6 and 6 for a range of surface variance values. For the smooth soil surface, the DoLP for the Nylon-66 surface is similar to those of the soil, but as the soil roughens, the contrast in mine and soil DoLP values increase. This finding is in qualitative agreement with earlier experimental studies.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

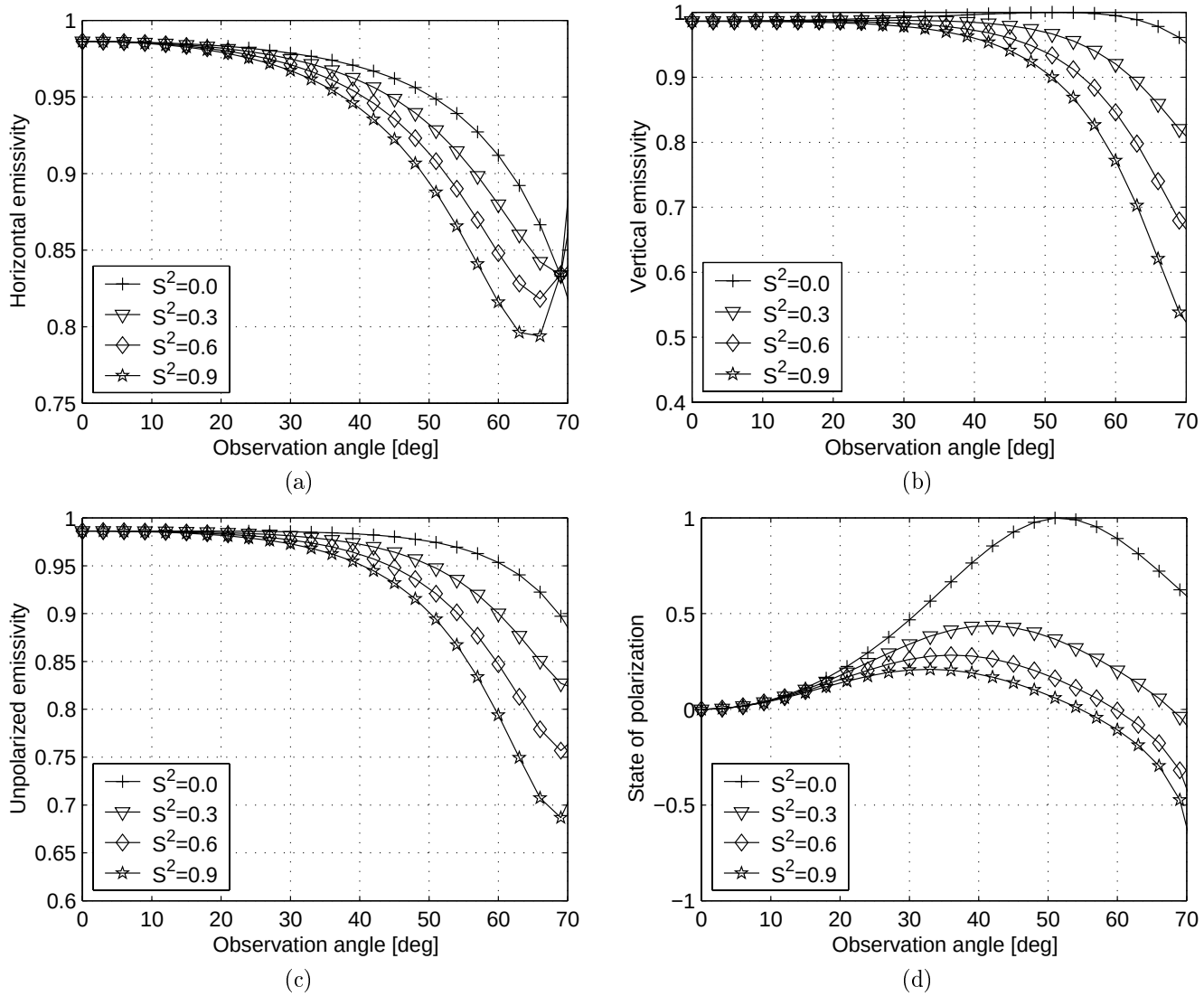
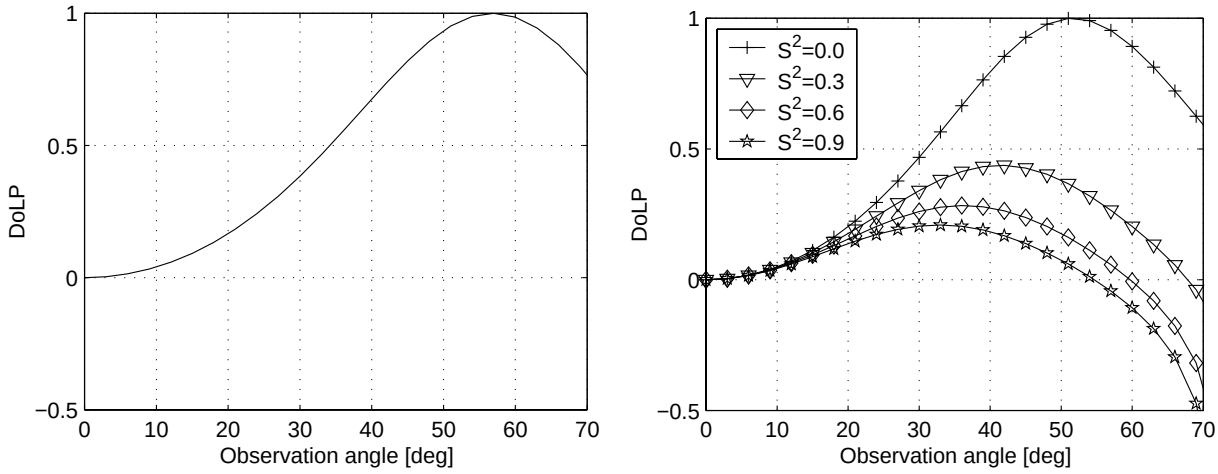


Figure 5. The emissivity and radiated polarization as functions of observation angle for rough soil having different slope variances (a) Horizontal polarized emissivity for different permittivity values. (b) Vertical polarized emissivity for different permittivity values. (c) Emissivity seen by an unpolarized sensor for different permittivity values. (d) Polarization of reflected radiation for different observation angles.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms



(a) Nylon

(b) A rough soil surface with different slope variances.

Figure 6. Degree of linear polarization (DoLP) as a function of observation angle for mine-like and soil surfaces.

7. SUMMARY AND CONCLUDING REMARKS

The theoretical basis for polarimetric IR phenomena was investigated for use in land mine detection. A model was developed for emission from mines and natural surfaces using a second order small perturbation method/small slope approximation (SPM/SSA) formulation for the rough-surface scattering problem. Emissivities seen by both polarized and unpolarized sensors were studied for smooth and rough surfaces. To model non-metallic mines, mine surfaces were represented by smooth plastics. Conditions of interest in mine detection make possible approximations to the theory that permit key polarimetric phenomena to be expressed in terms of the soil surface variance and complex permittivity. Results for polarimetric mine signatures were presented, which confirm earlier experimental findings that signature contrast improves for rougher soil surfaces.

ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed herein are those of the author and are not necessarily those of Duke University or the ARO.

REFERENCES

1. C. A. DiMarzio, T. Vo-Dinh, and H. E. Scott, "Some approaches to infrared spectroscopy for detection of buried objects," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *Proceedings of SPIE* **3392**, pp. 158–166, 1998.
2. J. R. Johnson, P. G. Lucey, K. A. Horton, and E. M. Winter, "Infrared measurements of pristine and disturbed soils 1. Spectral contrast differences between field and laboratory data," *Remote Sens. Environ.* **64**, pp. p. 34–46, 1998.
3. J. R. Johnson, P. G. Lucey, K. A. Horton, and E. M. Winter, "Infrared measurements of pristine and disturbed soils 2. Environmental effects and field data reduction," *Remote Sens. Environ.* **64**, pp. p. 47–52, 1998.
4. M. Larive, D. Spoliansky, and O. Trézières, "Pre-processing of 8-12 μm polarimetric features for laid and flush-buried mines detection," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **3710**, pp. 197–202, 1999.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

5. M. Larive, L. Collot, S. Beugnot, H. Botma, and P. Roos, "Laid and flush-buried mines detection using 8-12 μm polarimetric imager," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *Proceedings of SPIE* **3392**, pp. 115–120, 1998.
6. B. A. Barbour, M. W. Jones, H. B. Barnes, and C. P. Lewis, "Passive IR polarization sensors: A new technology for mine detection," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *Proceedings of SPIE* **3392**, pp. 96–103, 1998.
7. H. Barnes and M. Jones, "Infrared polarimetric camera system development," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **3710**, pp. 189–196, 1999.
8. W. de Jong, F. Cremer, K. Schutte, and J. Storm, "Usage of polarisation features of landmines for improved automatic detection," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **4038**, pp. 241–252, 2000.
9. B. H. Miles, E. R. Cespedes, and R. A. Goodson, "Polarization-based active/passive scanning system for minefield detection," in *Polarization and remote sensing*, W. G. Egan, ed., *Proceedings of SPIE* **1747**, pp. 239–252, 1992.
10. S. H. Yueh and R. Kwok, "Electromagnetic fluctuations for anisotropic media and generalized Kirchoff's law," *Radio Sci.* **28**, pp. 471–480, 1993.
11. L. Tsang and J. A. Kong and R. T. Shin, *Theory of Microwave Remote Sensing*, Wiley-Interscience, New York, NY, 1985.
12. D. E. Gray, ed., *American Institute of Physics Handbook*, McGraw-Hill, New York, NY, 1972.
13. J. Brandrup and E. H. Immergut, eds., *American Institute of Physics Handbook*, John Wiley & Sons, New York, NY, 1975.
14. *Proceedings of the London Conference on Optical Instruments*, John Wiley & Sons, New York, NY, 1950.
15. G. W. C. Kaye and T. H. Laby, *Tables of Physical and Chemical Constants*, Longman, New York, NY, 1973.
16. J. T. Johnson and M. Zhang, "Theoretical study of the small slope approximation for ocean polarimetric thermal emission," *IEEE Trans. Geosc. Rem. Sens.* **37**(5), pp. 2305–2316, 1999.
17. S. H. Yueh, R. Kwok, F. K. Li, V. Nghiem, W. J. Wilson, and J. A. Kong, "Polarimetric passive remote sensing of ocean wind vectors," *Radio Science* **29**(4), pp. 799–814, 1994.
18. J. M. Bigham and E. J. CiolKosz, eds., *Soil Color*, Soil Science Society of America, Inc., Madison, WI, 1993.
19. R. F. Keefer, *Handbook of Soils for Landscape Architects*, Oxford, New York, NY, 2000.
20. D. Kirkham and W. L. Powers, *Advanced Soil Physics*, John Wiley & Sons, New York, NY, 1972.
21. H. Kohnke, *Soil Physics*, McGraw-Hill, New York, NY, 1968.
22. *Remote Sensing with Special Reference to Agriculture and Forestry*, National Acedemy of Sciences, Washington, D.C., 1970.
23. T. Ishida, H. Ando, and M. Fukuhara, "Estimation of complex refractive index of soil particles and its dependence on soil chemical properties," *Remote Sens. Environ.* **38**, pp. 173–182, 1991.
24. T. Ishida and J. C. Price, "A procedure to infer complex refractive index and mean particle radius of soils from visible and near-infrared reflectance data," *Int. J. Remote Sensing* **17**(11), pp. 2145–2164, 1996.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 04/10/2013 Terms of Use: http://spiedl.org/terms

Predicting sensor fusion performance using theoretical models

B. A. Baertlein*, W.-J. Liao and D.-H. Chen

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

Simple theoretical models can be constructed to study the behavior of sensor-fused systems using idealized sensor suites. Models are available for feature-level and decision-level fusion, both of which are now being used with demining sensors. These models are attractive as design tools and for estimating the expected performance of new sensor suites, since their performance can be evaluated with relatively little effort. In this paper we review some simple idealized models and their predictions for fused system performance. The data produced by demining sensors are often correlated, and the effect of correlation is explored for both feature-level and decision-level fusion.

Keywords: mine detection, sensor fusion, independence, theoretical analysis, performance prediction

1. INTRODUCTION

Demanding performance requirements are placed on sensors that detect hazards such as buried land mines. A high probability of detection is obviously necessary and, when the hazards are spatially disperse, as is often the case for land mines, low false alarm rates are also required. Multi-sensor data fusion is often proposed as a means of meeting these requirements, but additional resources must be allocated to develop hardware and software for a multi-sensor system, and there is no a priori guarantee that the performance of the fused suite will meet its goals.

In this work we review some basic properties of fusion that can be predicted from theoretical work. Several relatively simple idealized models can be used to estimate the performance gains achievable through fusion given individual sensor performance data such as P_d and P_{fa}. We employ these theoretical results to examine the benefits achievable through fusion at the feature level and at the decision level. We compute the performance improvement (expressed as a reduction in error rate) for fusion. We also consider the important problem of correlated sensors. Many demining sensors tend to produce errors on a common set of targets, and theory allows us to predict the effect of this correlation on the performance of the fused sensor suite.

This work is organized in three major sections. In Section 2 we review some useful theoretical results for uncorrelated sensors. Expressions for the likelihood ratios of fused systems are presented for feature-level (FL) and decision-level (DL) fusion. The effects of correlation are discussed in Section 3. We present experimental data from demining sensors that suggest that correlation among demining sensors is common. The analysis shows that correlation strongly degrades performance. Concluding remarks appear in Section 4.

2. INDEPENDENT SENSORS

The assumption of independent sensors is relatively common in fusion studies, and it greatly simplifies the analysis. Both feature-level and decision-level fusion are addressed in what follows.

2.1. Feature-Level Fusion

In FL fusion target-related parameters are acquired by several sensors and merged into a feature vector. Detection is performed on the basis of the feature vector. In theory, the best features for discrimination of K classes are the K a posteriori probabilities.[1] In practice, those probabilities are unknown, and selecting good features is frequently a problem.

Some important aspects of FL fusion can be illustrated by considering the simple case of normally distributed features. The required analysis in this case is well known. Consider the case in which N features are fused into a single feature vector $\mathbf{X}$. Define hypotheses H_k, k=0,1 to be the no-mine and mine case, respectively. We assume

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

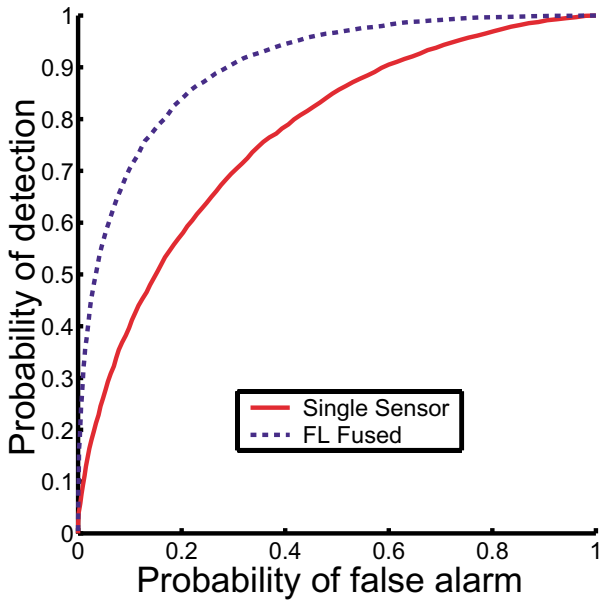


Figure 1. ROC curves for fusion of three uncorrelated identical sensors having $P_d = 0.7$ and $P_{fa} = 0.3$.

that under hypothesis H_k, $\mathbf{X}$ comprises independent multivariate normal deviates with mean μ_k and variance σ_k^2. Minimizing the Bayes cost leads to a likelihood ratio $\Lambda(\mathbf{X}) = \frac{\Pr(\mathbf{X}|H_1)}{\Pr(\mathbf{X}|H_0)}$, the logarithm of which has the well-known form

$$\ln(\Lambda(\mathbf{X})) = -(\mathbf{X}-\mu_1)^T(\mathbf{X}-\mu_1)/(2\sigma_1^2) + (\mathbf{X}-\mu_0)^T(\mathbf{X}-\mu_0)/(2\sigma_0^2) + N\ln(\sigma_0/\sigma_1) \tag{1}$$

It is possible to express the probability density of this likelihood ratio (for both independent and correlated features to be discussed later) in a semi-analytical form,[2] which yields values for P_d and P_{fa}. Analogous simplifications are not possible for the DL fusion results to be presented next, and in the interest of consistency we evaluated the performance of FL fusion using Monte Carlo methods. A sequence of normal random deviates was generated for each of the hypotheses. For each random deviate we computed the likelihood ratio. For a fixed decision threshold this yields a hard decision, which can be scored, ultimately producing a ROC curve. A total of 20,000 random deviates were used to generate each of the ROC curves shown in this work.

In this work we use as a performance metric the probability of error P_E (or the error rate) given by

$$P_E = \Pr(H_0)P_{fa} + \Pr(H_1)(1 - P_d) \tag{2}$$

where $\Pr(H_k)$ are the prior probabilities, P_{fa} is the false alarm probability, and P_d is the detection probability. Legitimate objections can be raised to the use of this metric for mine detection, since it implicitly assigns equal costs to a missed detection and a false alarm. It is, however, a common standard, and the extension to other metrics is straightforward. The value of P_E is also dependent on the decision threshold. Throughout this work we assume that any system would be operated at its minimum error point, and we use the minimum value of P_E in our analyses. In addition, we consider only the case $\Pr(H_0) = \Pr(H_1) = 0.5$.

In general, fusion produces a performance gain, which we quantify as follows:

$$\text{Fusion gain} = (minimum P_E \text{ for best single sensor}) - (minimum P_E \text{ for fused sensor suite}) \tag{3}$$

As an example, in Figure 1 we have plotted the performance of a FL system using three uncorrelated features each of which has a ROC curve that passes through the point $P_d = 0.7$, $P_{fa} = 0.3$. For this uncorrelated-feature case, fusion reduces P_E by 11.4%.

The performance improvement attainable with fusion depends on the performance of the individual sensors. This dependence is illustrated in Figure 2, where we again show the result of fusing three sensors. Two of the sensors have

a specified error rate, while the error rate of the third sensor is varied. To reduce the number of variables involved we have imposed the added constraint $P_d + P_{fa} = 1$ for the third sensor. As the error rate of the third sensor increases, the performance of the fused suite tends to that of a two-sensor fused system, and the third sensor provides little benefit. Conversely, as the error rate for the third sensor decreases, it tends to dominate the performance of the suite. In such cases, the contribution of the two more poorly performing sensors is reduced. This tendency is clear from Figure 2(b) where the performance gain achievable under fusion is shown. It is evident that the maximum gain is achieved when the sensors have equal error rates, although the best performance is obtained when one sensor dominates the others.

2.2. Decision-Level Fusion

In the simplest form of DL fusion, a vector of sensor decisions is produced by applying individual thresholds to the test statistics output by each sensor. In the interest of brevity we restrict our attention to binary decisions. The decision vector is processed to form a new fused test statistic for the sensor suite, and a (possibly) different threshold is applied to the fused test statistic. The relation between thresholds for different sensors and for the fused suite is not readily addressed theoretically, but some work in this area has appeared previously.[3]

Consider the case of N independent sensors each of which produces a random, binary decision u_n, $n = 1, \ldots, N$. Writing the decisions as the vector $\mathbf{U}$, the likelihood ratio for this formulation is given by

$$\Lambda_I(\mathbf{U}) = \frac{\Pr(\mathbf{U}|H_1)}{\Pr(\mathbf{U}|H_0)} = \prod_{n=1}^{N} \frac{\Pr(u_n|H_1)}{\Pr(u_n|H_0)} \tag{4}$$

An optimum decision rule for this case was first described by Chair and Varshney[4] and is developed by decomposing the above product as follows:

$$\Lambda_I(\mathbf{U}) = \prod_{S_1} \frac{\Pr(u_n = 1|H_1)}{\Pr(u_n = 1|H_0)} \prod_{S_0} \frac{\Pr(u_n = 0|H_1)}{\Pr(u_n = 0|H_0)} = \prod_{S_1} \frac{P_{d,n}}{P_{fa,n}} \prod_{S_0} \frac{1 - P_{d,n}}{1 - P_{fa,n}} \tag{5}$$

where S_1 (S_0) denotes the set of sensors that produce $u_n = 1$ ($u_n = 0$) for a given observation of the target, and $P_{d,n}$ and $P_{fa,n}$ are the respective detection and false alarm probabilities for sensor n.

Taking the logarithm of this result permits it to be expressed as

$$\ln(\Lambda_I(\mathbf{U})) = \sum_{n=1}^{N} \left[u_n w_n^{(1)} + (1 - u_n) w_n^{(0)} \right] \tag{6}$$

where

$$w_n^{(1)} = \ln\left(\frac{P_{d,n}}{P_{fa,n}}\right) \tag{7}$$

$$w_n^{(0)} = \ln\left(\frac{1 - P_{d,n}}{1 - P_{fa,n}}\right) \tag{8}$$

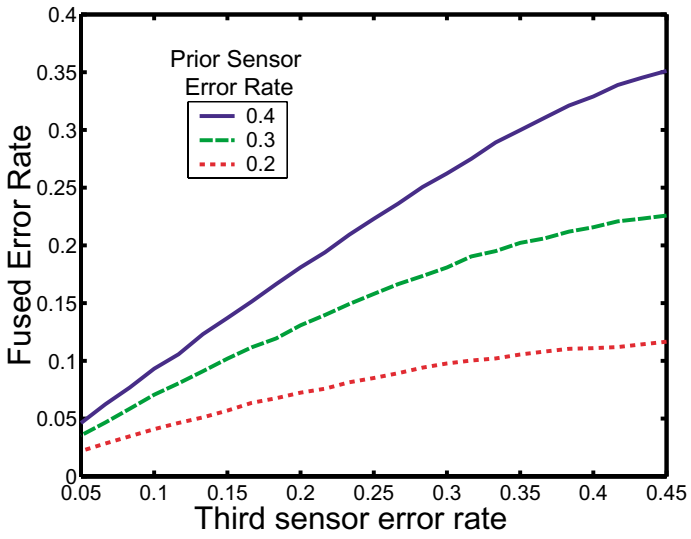
are weights are defined by sensor performance data. The foregoing summation comprises a weighted voting fusion, in which the fused decision is determined from the detections and the reliability of the sensors. Note that since only $2N$ distinct weights are involved, a finite number of values can be assumed by $\ln(\Lambda(\mathbf{U}))$. Some special cases are of interest. If the sensors have identical performance and if $P_{d,n} = 1 - P_{fa,n}$ then the weights satisfy

$$w_n^{(1)} = -w_n^{(0)} \equiv w \tag{9}$$

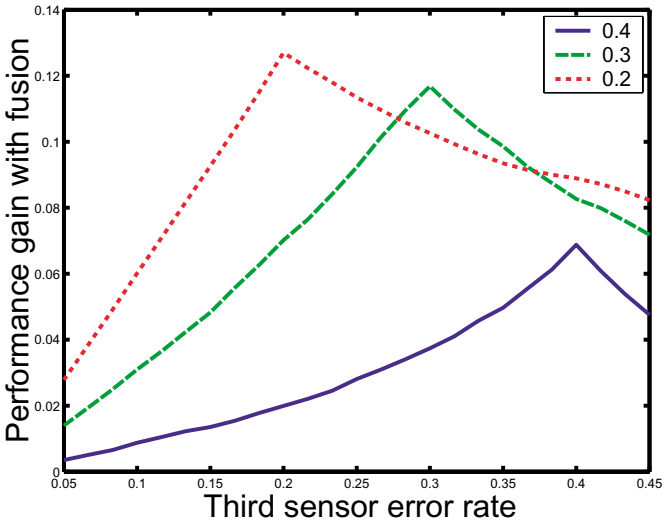
for every n. In this case we have

$$\ln(\Lambda_I(\mathbf{U})) = -w \left[N + \sum_{n=1}^{N} u_n \right] \tag{10}$$

If we apply the decision thresholds 0, -wN and $-wN/2$ to $\ln(\Lambda_I(\mathbf{U}))$, this expression leads to AND, OR, and majority voting fusion, respectively.



(a) Achievable performance level.



(b) Fusion-related performance gain.

Figure 2. The performance of two sensors having a specified error rate fused with a third sensor.

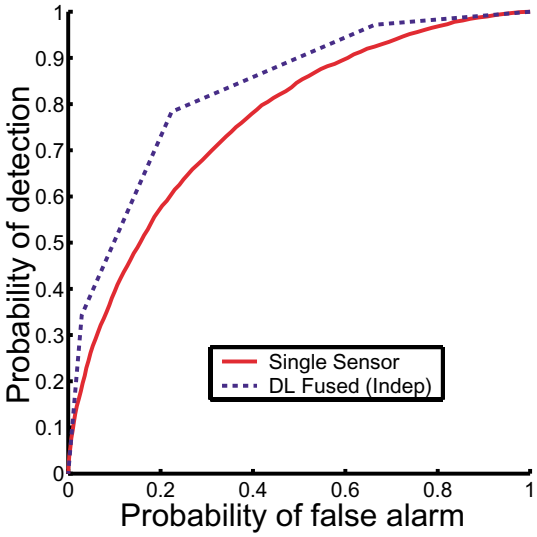


Figure 3. Performance of decision-level fusion for three sensors having $P_d = 0.7$ and $P_{fa} = 0.3$.

Simulation is again used to investigate the performance of DL fusion. Fixed thresholds can be applied to independent normal deviates yielding a sequence of independent binary random variables with specified P_d and P_{fa} values. These binary variables can be used in DL fusion performance assessments. A ROC curve for three independent sensors having $P_d = 0.7$ and $P_{fa} = 0.3$ is shown in Figure 3. For three identical sensors with the given performance values, $\ln(\Lambda_I)$ can take on only four discrete values corresponding to detection by none, one, two, or all sensors. The maximum gain for fusion in this case is 7.9%, less than that noted for feature-level fusion above.

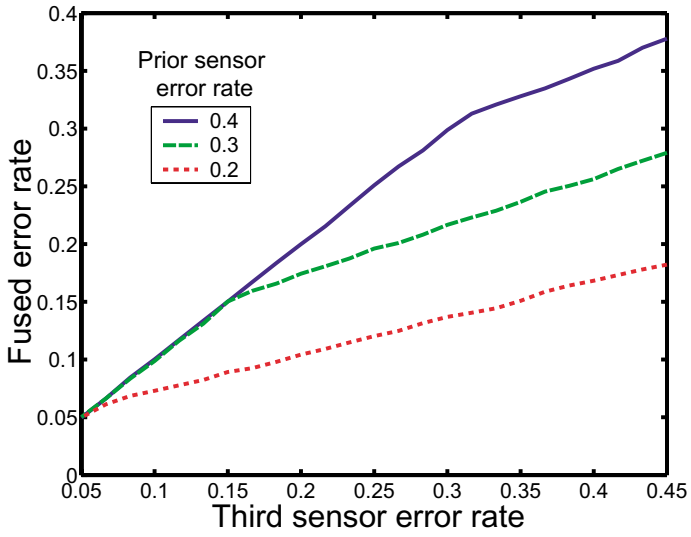
The gain of DL fusion is also dependent on the performance of the individual sensors. As done previously for FL fusion, we compute the performance of DL fusion algorithms for uncorrelated sensors as a function of P_E. The results are shown in Figure 4. We find that when the error rate of the third sensor falls below a specific value, there is no benefit to fusion since the best performance of the fused suite simply tracks that of the best sensor. (Recall that the decision thresholds are being selected to produce the best performance.) This finding is confirmed by the plot of fusion-related performance gain in Figure 4(b). As in the case of FLfusion, the largest gain is obtained when all three sensors have comparable error rates.

3. CORRELATED SENSORS

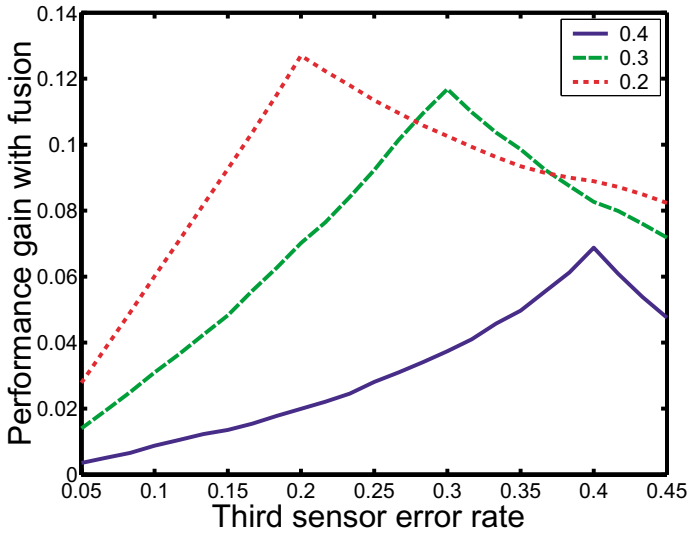
Data produced by demining sensors are frequently not independent. The outputs of these sensors are typically dominated by clutter, rather than internal noise, and a variety of clutter phenomena can produce detectable signals in multiple sensors. For example, a sufficiently large metal object will be detected by both an EMI sensor and a GPR. Similarly, a small depression in the soil will produce artifacts for IR, EMI and GPR sensors. In addition, small, non-metallic mines, are poorly detected by many sensors. In this section we use correlation as a metric for the lack of independence. In general, correlation is an imperfect measure of sensor dependence, but it is sufficient for the Gaussian features to be considered here.

3.1. Evidence for Sensor Correlation

A number of sensors have been used to acquire data at the JUXOCO Calibration Grid, Site 71A, Fort A. P. Hill, VA. This test site contains a grid of mines and clutter objects, the locations and identities of which are known. The site offers 125 sampling locations, of which 29 contain buried mines (both AP and AT) and 33 contain buried clutter artifacts (various metal, plastic, wood objects). The remaining sample locations are blanks (homogeneous soil with no emplaced clutter). The mines are diverse, with one to five examples of each mine. For that site we will consider data from three sensors: the GEM3 EMI sensor, a laser Doppler vibrometer, and a LWIR camera. These sensors and the processing used with them are described elsewhere,[3,5] and only a brief summary is presented here.



(a) Best achievable performance.



(b) Fusion-related performance gain.

Figure 4. The performance of two sensors having a specified error rate fused (at the decision level) with a third sensor.

The GEM3 EMI data were acquired in 1999 by Duke University. The results of that collection have been documented elsewhere.[6,7] In brief, those data comprise complex (I & Q), frequency-domain measurements using the GEM3 at ten points over each mine. The cells sampled in this manner were limited to those having known mines or emplaced clutter, i.e., no blank cells were sampled. Data from these scans were used to train a classifier, which detects mines on the basis of the spatial distribution of energy.[3]

During 1998 The Ohio State University acquired LWIR data at the JUXOCO calibration grid. The sensor was positioned on a tripod of a known height and at a known distance from the grid cells. Data were acquired at all 125 grid cells. The sensors and processing algorithms were described in a previous work.[3]

The University of Mississippi has demonstrated that buried mines can be detected by vibrating the soil with acoustic speakers and sensing the vibration with a laser Doppler vibrometer (LDV).[8,9] The resulting data comprise a multi-spectral image of the surface at a few tens of frequency bands and at a few tens of spatial positions. Preprocessing of the data to remove trends in the image and a CFAR algorithm used for detection have been described by Baertlein et al.[5]

Fusion of the EMI, LWIR, and LDV data from the JUXOCO site allow us to illustrate the concepts discussed in Sections 2.1 and 2.2. The size of the data set is small, which limits the confidence that can be placed in the results. In addition, other limitations in the data set preclude us from considering some sensor combinations: although each sensor sampled a large fraction of the 125 grid cells, each sensor also sampled a different set of locations and the intersection of these sets (i.e., the set of locations sampled by multiple sensors) is often small. In particular, only 13 of the 125 grid cells were sampled by both the GEM3 and LDV sensors. This makes it impossible to consider fusion of any sensor suite that includes both sensors. The situation is slightly better if we separately consider fusion of the LWIR data with the GEM3 and LDV data. The data available for these pairs of sensors comprises roughly 40 samples each, of which about half are mines.

In addition to the data from Fort A. P. Hill, data have been acquired by a forward-looking GPR and the LDV over a common set of targets at Yuma Proving Grounds, AZ. That data set produced a large common set of mine samples, but a smaller number of clutter samples. The details of that collection and its processing are described elsewhere.[5]

We define the feature-level correlation ρ_{ij} between features x_i and x_j in the usual manner

$$\rho_{ij} = \frac{\frac{1}{M}\sum_{m=1}^{M}(x_{i,m}-\mu_i)(x_{j,m}-\mu_j)}{\sqrt{\mathrm{var}(x_i)\mathrm{var}(x_j)}} \tag{11}$$

where μ_i (μ_j) is the mean of feature i (j), and M is the number of samples. A similar definition is used for the decision correlations. When a sensor produces multiple features, a classifier was used to reduce the features from each sensor to a single value.

Correlation coefficients among the sensors can be computed and are given in Tables 1 and 2 for features and decisions. We see that both features and decisions have significant correlation for some sensor combinations.

	Correlation ρ	
Sensors	H_1	H_0
LWIR-GEM3	-0.32	0.18
LWIR-LDV	0.55	-0.19
GPR-LDV	-0.23	-0.20

Table 1. Feature correlation coefficients.

3.2. Fusion of Correlated Features

For the case of correlated features, the log-likelihood ratio has the well-known form

$$\ln(\Lambda(\mathbf{X})) = -(\mathbf{X}-\mu_1)^T\boldsymbol{\Sigma}_1^{-1}(\mathbf{X}-\mu_1)/2 + (\mathbf{X}-\mu_0)^T\boldsymbol{\Sigma}_0^{-1}(\mathbf{X}-\mu_0)/2 + \ln(|\boldsymbol{\Sigma}_0|/|\boldsymbol{\Sigma}_1|)/2 \tag{12}$$

	Correlation ρ	
Sensors	H_1	H_0
LWIR-GEM3	-0.12	0.18
LWIR-LDV	0.26	-0.17
GPR-LDV	-0.09	-0.07

Table 2. Decision correlation coefficients.

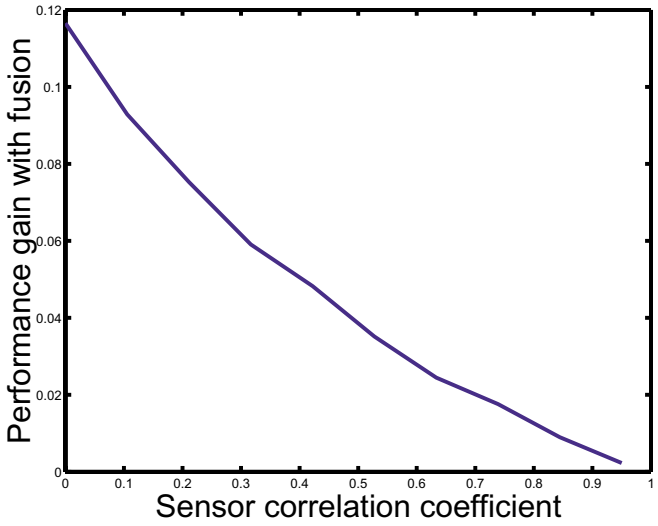


Figure 5. Performance gain of feature-level fusion for three sensors having $P_d = 0.7$ and $P_{fa} = 0.3$ and a common variable correlation level among all sensors.

where $\mathbf{\Sigma}_k$ is the covariance matrix under hypothesis H_k. Monte Carlo simulation of this case requires that we generate random deviates with the prescribed distributions, which can be done from independent, identically distributed, zero-mean, unit-variance, random normal deviates. We denote the feature for sensor n and trial i by $x_n^{(i)}$. If the covariance matrix $\mathbf{\Sigma}$ is factored using a Cholesky decomposition,

$$\mathbf{\Sigma}^{-1} = \mathbf{A}^T\mathbf{A} \tag{13}$$

then multivariate normal deviates $y_n^{(i)}$ with mean μ and covariance matrix $\mathbf{\Sigma}$ are generated by the linear transformation

$$\mathbf{y}^{(i)} = \mathbf{A}^T\mathbf{x}^{(i)} + \mu \tag{14}$$

Simulations show that sensor-to-sensor correlation rapidly eliminates the benefits of fusion. This effect is illustrated in Figure 5, where we show the result of fusing three sensors mutually correlated at a given level. A theoretical interpretation of this result is that forming the matrix product with $\mathbf{\Sigma}_k^{-1}$ projects out of the data the influence of the correlated sensors, leaving only the information from the uncorrelated residual data. Comparing Figure 5 with the feature correlations shown in Table 1 suggests that FL fusion of these sensors may achieve only 30% to 65% of the performance of independent sensors.

3.3. Fusion of Correlated Local Decisions

In this section we consider the problem of hard decision fusion using correlated sensors. The problem has received attention previously. Aalo and Viswanathan[10] performed numerical simulations of the problem and characterized how performance degrades with decreasing correlation for several different fusion strategies. Drakopoulos and Lee[11] derived an optimum fusion rule for the Neyman-Pearson criterion, and used simulation to study its performance for

a specific type of correlation matrix. Kam et al.[12] considered the case in which the class-conditioned sensor-to-sensor correlation coefficients are known, and expressed the result in a compact form. Blum and Kassam[13] studied the problem of locally most powerful detection for correlated local decisions.

The formulation of Kam et al. is attractive, since it permits the effects of correlation to be addressed explicitly and simply. Their formulation begins by reducing the decisions to a standard zero-mean, unit-variance form using the transformation

$$y_n = \frac{u_n - p_n}{\sqrt{p_n(1-p_n)}} \tag{15}$$

in which

$$p_n = \Pr(u_n = 1) = P_1 P_{d,n} + P_0 P_{fa,n} \tag{16}$$

One next uses the Bahadur-Lazarsfeld (BL) polynomials $\psi_i(\mathbf{U})$ described by Duda and Hart.[14] These polynomials comprise all possible products of the y_n taken none, one, two, and on to N at a time. Explicitly

$$\psi_i(\mathbf{U}) = \begin{cases} 1 & i = 0 \\ y_i & i = 1 \\ y_2 & i = 2 \\ \vdots & \vdots \\ y_N & i = N \\ y_1 y_2 & i = N+1 \\ y_1 y_3 & i = N+2 \\ \vdots & \vdots \\ y_{N-1} y_N & i = N + N(N-1)/2 \\ y_1 y_2 y_3 & i = 1 + N + N(N-1)/2 \\ \vdots & \vdots \\ y_1 y_2 \ldots y_N & i = 2^N - 1 \end{cases} \tag{17}$$

The BL polynomials can be used to expand the joint density $\Pr(\mathbf{U})$. If we define the density of independent variables as

$$P_1(\mathbf{U}) = \prod_{n=1}^{N} p_n^{u_n} (1-p_n)^{1-u_n} \tag{18}$$

then using P_1 as a weighting function, the ψ_i are orthonormal, viz:

$$\sum_{\mathbf{U}} \psi_i(\mathbf{U}) \psi_j(\mathbf{U}) P_1(\mathbf{U}) = \delta_{ij} \tag{19}$$

where $\mathbf{V}$ ranges over all possible values of $\mathbf{U}$ and δ_{ij} is the Kronecker delta function. Using these polynomials, any probability $\Pr(\mathbf{U})$ can be expanded in the form

$$\Pr(\mathbf{U}) = P_1(\mathbf{U}) \sum_{i=1}^{2^N - 1} a_i \psi_i(\mathbf{U}) \tag{20}$$

where the expansion coefficients are given by

$$a_i = \sum_{\mathbf{V}} \psi_i(\mathbf{V}) \Pr(\mathbf{V}) \tag{21}$$

Equation (20) shows that the density of correlated variables $\mathbf{U}$ can be represented by the density of the independent variables $P_1(\mathbf{U})$ times a correction factor, which is written as a sum of BL polynomials. It was noted by Kam et al.[12] that for an expansion using the normalized variables y_n, we have $a_0 = 1$, $a_i = 0$ for $i = 1, 2, \ldots, N$ (since the y_i are zero mean), and that for $i > N$ the a_i are correlation coefficients, which have the form

$$\rho_{ijk\ldots n} = \sum_{\mathbf{V}} y_i y_j \ldots y_n \Pr(\mathbf{V}) \tag{22}$$

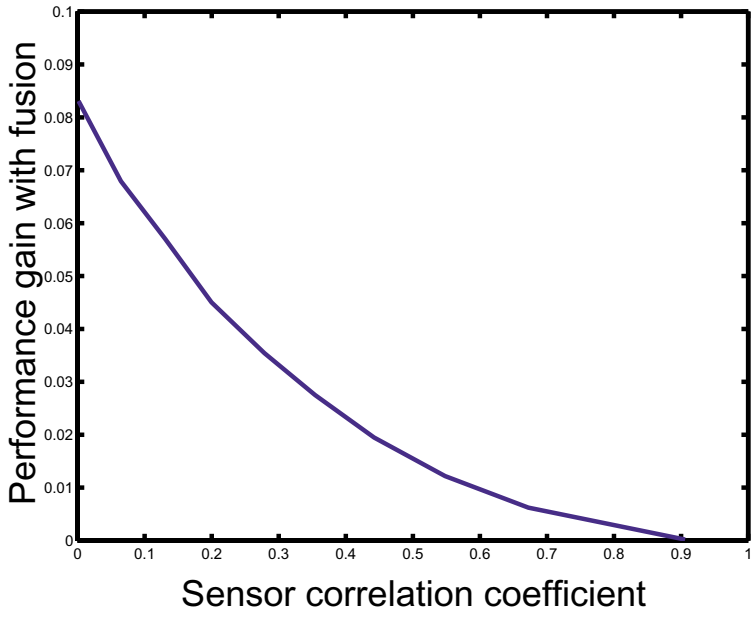


Figure 6. Performance gain for all sensors correlated at a common level.

Using these facts we can write

$$\Pr(\mathbf{U}) = P_1(\mathbf{U}) \left[1 + \sum_{i<j} \rho_{ij} y_i y_j + \sum_{i<j<k} \rho_{ijk} y_i y_j y_k + \cdots \rho_{123\ldots N} y_1 y_2 \ldots y_N \right] \tag{23}$$

This procedure can be used to formulate a likelihood ratio for correlated sensors. If we define normalized random variables $y_n^{(k)}$ conditioned on hypothesis H_k

$$y_n^{(k)} = \frac{u_n - \Pr(u_n = 1|H_k)}{\sqrt{\Pr(u_n = 1|H_k)[1 - \Pr(u_n = 1|H_k)]}} \tag{24}$$

then the densities $\Pr(\mathbf{U}|H_k)$ can be expanded as above to yield the following expression for the likelihood ratio $\Lambda_C(\mathbf{U})$ of correlated sensors

$$\Lambda_C(\mathbf{U}) = \Lambda_I(\mathbf{U}) \left[\frac{1 + \sum_{i<j} \rho_{ij}^{(1)} y_i^{(1)} y_j^{(1)} + \sum_{i<j<k} \rho_{ijk}^{(1)} y_i^{(1)} y_j^{(1)} y_k^{(1)} + \cdots \rho_{123\ldots N}^{(1)} y_1^{(1)} y_2^{(1)} \ldots y_N^{(1)}}{1 + \sum_{i<j} \rho_{ij}^{(0)} y_i^{(0)} y_j^{(0)} + \sum_{i<j<k} \rho_{ijk}^{(0)} y_i^{(0)} y_j^{(0)} y_k^{(0)} + \cdots \rho_{123\ldots N}^{(0)} y_1^{(0)} y_2^{(0)} \ldots y_N^{(0)}} \right] \tag{25}$$

in which

$$\rho_{ij\ldots n}^{(k)} = \sum_{\mathbf{V}} y_i^{(k)} y_j^{(k)} \ldots y_n^{(k)} \Pr(\mathbf{V}|H_k) \tag{26}$$

This result makes explicit the effect of the correlation. It is easy to see that for uncorrelated sensors we have $\Lambda_C(\mathbf{U}) \to \Lambda_I(\mathbf{U})$ as expected.

The effect of sensor correlation can be quantified using Monte Carlo methods as shown in Figure 6. Care is required since the correlation of random normal variables $\mathcal{E}[x_i x_j]$ is in general not equal to the desired correlation $\mathcal{E}[u_i u_j]$. As noted above for feature correlation, the fusion-derived performance gain decreases rapidly with increasing decision correlation. The correlation levels noted in Table 2 suggest that DL fusion of real sensors may experience only 50% to 80% of the performance benefit obtained for independent sensors.

4. SUMMARY AND CONCLUDING REMARKS

Simple theoretical models can be used to predict trends in sensor fusion. These models can be used to qualitatively estimate the performance of fused suites using minimal information including the single-sensor probabilities of detection and false alarm and the hypothesis-conditioned sensor-to-sensor correlations. The probability of error was used to quantify the performance benefit obtained through fusion. This benefit was calculated for both feature-level and decision-level fusion of uncorrelated sensors. We observed that the performance improvement is limited if one sensor in the suite dominates (or is dominated by) the remaining sensors. The peak benefit occurs when all sensors have comparable performance.

Because essentially all demining sensors are clutter limited, and because many sensors respond to similar clutter features, it is reasonable to expect correlation among the outputs of demining sensors. A survey of a few prior experimental works shows that a number of sensors are correlated at levels of 10% to 50%, with larger correlations observed among features than among decisions. Theoretical calculations of performance for correlated sensors were presented for both feature-level and decision-level fusion. It was observed that correlation greatly reduces the performance improvement attainable through fusion. A comparison of theoretical results and experimentally derived sensor-to-sensor correlations suggests that fusion-related performance gains for real sensors may be significantly smaller than what would be predicted for independent sensors.

5. ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

REFERENCES

1. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, Boston, MA, 1990.
2. S. M. Kay, *Fundamentals of Statistical Signal Processing: Volume II, Detection Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1998.
3. B. A. Baertlein and A. H. Gunatilaka, "Optimizing fusion architectures for limited training data sets," in *Detection Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *Proc. SPIE* **4038**, pp. 804–815, 2000.
4. Z. Chair and P. R. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerospace and Electronic Systems* **AES-22**, pp. 98–101, 1986.
5. B. A. Baertlein, W.-J. Liao, and D.-H. Chen, "Fusion of GPR and acoustic LDV sensors," in *Detection and Remediation Technologies for Mines and Minelike Targets VI*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **4394**, 2001.
6. L. M. Collins, "Statistical signal processing for demining: Experimental validation," tech. rep., Duke University, 1998.
7. P. Gao, L. Collins, J. Moulton, L. Makowsky, R. Weaver, D. Keiswetter, and I. J. Won, "Enhanced detection of landmines using broadband EMI data," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *SPIE* **3710**, pp. 2–13, April 5-9 1999.
8. J. M. Sabatier and N. Xiang, "Laser-doppler based acoustic-to-seismic detection of buried mines," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. Dugan, eds., *Proc. SPIE* **3710**, pp. 215–222, 1999.
9. N. Xiang and J. M. Sabatier, "Land mine detection measurements using acoustic-to-seismic coupling," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. Dugan, eds., *Proc. SPIE* **4038**, pp. 645–655, 2000.
10. V. Aalo and R. Viswanathan, "On distributed detection with correlated sensors: Two examples," *IEEE Trans. Aerospace and Electronic Systems* **AES-25**(3), pp. 414–421, 1989.
11. E. Drakopoulos and C. C. Lee, "On distributed detection with correlated sensors: Two examples," *IEEE Trans. Aerospace and Electronic Systems* **AES-27**(4), pp. 593–605, 1991.

12. M. Kam, Q. Zhu, and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerospace and Electronic Systems* **AES-28**(3), pp. 916–920, 1992.
13. R. S. Blum and S. A. Kassam, "Optimum distributed detection of weak signals in dependent sensors," *IEEE Trans. Information Theory* **IT-38**(3), pp. 1066–1079, 1992.
14. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley, New York, NY, 1973.

Using Physical Models to Improve Thermal IR Detection of Buried Mines

D.-H. Chen, I. K. Sendur, W.-J. Liao and B. A. Baertlein*

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

Many aspects of a buried mine's thermal IR signature can be predicted through physical models, and insight provided by such models can lead to better detection. Several techniques for exploiting this information are described. The first approach involves ML estimation of model parameters and followed by classification of those parameters. We show that this approach is related to an approximate evaluation of an integral over the parameters that arises in a Bayesian formulation. This technique is compared with a generalized likelihood ratio test (GLRT) and with computationally efficient, model-free approaches, in which soil temperature data are classified directly. The benefit of using the temporal information is also investigated. Algorithm performance is illustrated using broadband IR imagery of buried mines acquired over a 24 hour period. It is found that the detection performance at a suitably selected time is comparable to the performance achieved by processing all times. The performance of the GLRT, for which detection is based only on the residual error, is inferior to a classifier using the parameters.

Keywords: IR, thermal, buried mines, modeling, space-time processing

1. INTRODUCTION

The thermal IR signature of a buried land mine depends on a number of mine-related or environmental parameters. This dependence is poorly understood, which limits one's ability to predict the signature and to design effective mine detection algorithms.

Recent studies of the physics that govern thermal IR signatures of buried land mines have shed light on the phenomena at work in mine signatures. A simulation tool described by Sendur and Baertlein[1] has shown an ability to predict these signatures. Recent work in this area appears in a companion paper.[2] Physical signature models are complex, but an analysis of measured signature data and use of the aforementioned tool suggest that relatively simple models can be constructed for these signatures. We show herein that key model parameters can be estimated from imagery, and that such simple models can be used to improve buried mine detection.

Several methods can be used to exploit information derived through modeling. A Bayesian approach to detection leads to an integral over unknown model parameters. Alternatively, a generalized likelihood ratio test (GLRT) can be used, in which detection is based on the residual error between data and model having used ML estimates of the model parameters. We describe another approach, referred to herein as the "estimator-classifier" technique, which makes explicit use of the model parameters. We show that the estimator-classifier technique is consistent with an approximate evaluation of the Bayesian parameter integrals.

Mine signatures are time varying, and we also investigate the benefits of using both spatial and temporal information in detection. Temporal signature information is not available in many mine-detection operations, but it is attractive for humanitarian demining, in which the observer is able to acquire multiple looks at the region of interest. Although it has long been known that mine signatures are time dependent, this fact has seldom been exploited for detection. We have identified only one group[3] that has previously explored this concept.

The performance of several detection algorithms is evaluated using broadband IR data acquired at regular intervals over a 24 hour cycle. The estimator-classifier technique is compared with the GLRT, and the benefit of basing detection on single versus multiple time samples is determined. All of these methods are computationally expensive. For comparison, we also explore some computationally efficient model-free methods, including a classifier of temporal signature behavior and the well-known "RX" algorithm of Reed and Yu.[4]

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

This work is organized in four major sections. The problem of mine detection using a space-time model is formulated in Section 2. Our choice of model is motivated by some aspects of the data, which we describe in Section 3. An example multi-temporal data set of thermal IR signatures is also discussed in that section. A computationally efficient (albeit less effective) model-free algorithm is described in Section 4, which involves a classifier trained to recognize mines from their thermal history. Detection results in the form of ROC curves are reported in Section 5. Concluding remarks appear in Section 6.

2. SPACE-TIME DETECTION ALGORITHM

We can formulate a model-based detection algorithm for the spatial and temporal characteristics of the signature as follows: The measured image data $\mathbf{D}(t)$ are assumed to comprise clutter or the mine signature in additive noise $\mathbf{N}(t)$, for which we write

$$\mathbf{D}(t) = \begin{cases} \mathbf{G}_0(\mathbf{\Theta}, t) + \mathbf{N}(t), & H_0 \\ \mathbf{G}_1(\mathbf{\Theta}, t) + \mathbf{N}(t), & H_1 \end{cases} \quad t = t_1, t_2, \ldots \tag{1}$$

In what follows we suppress the time dependence to simplify the notation. Here $\mathbf{G}_k(\mathbf{\Theta})$ is the (assumed known) physical model (described below) for the signal under hypothesis H_k, $k = 0, 1$, given model parameters $\mathbf{\Theta}$. The noise $\mathbf{N}$ is assumed to be independent of the hypothesis and to have the distribution $\mathbf{N} \sim f_{\mathbf{N}}$. Adopting a Bayesian approach, the distribution of the parameters $\mathbf{\Theta} \sim f_{\mathbf{\Theta}|H_k}(\mathbf{\Theta})$ will depend on the hypothesis, and we have

$$f_{\mathbf{D}|H_k}(\mathbf{D}) \quad = \quad \int d\mathbf{\Theta}_k f_{\mathbf{\Theta}_k|H_k}(\mathbf{\Theta}_k) f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta}_k)) \tag{2}$$

which leads to the likelihood ratio

$$\Lambda(\mathbf{D}) = \frac{f_{\mathbf{D}|H_1}(\mathbf{D})}{f_{\mathbf{D}|H_0}(\mathbf{D})} = \frac{\int d\mathbf{\Theta} f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_1(\mathbf{\Theta})) f_{\mathbf{\Theta}|H_1}(\mathbf{\Theta})}{\int d\mathbf{\Theta} f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_0(\mathbf{\Theta})) f_{\mathbf{\Theta}|H_0}(\mathbf{\Theta})} \tag{3}$$

Direct evaluation of this expression is possible if the densities $f_{\mathbf{N}}$ and $f_{\mathbf{\Theta}|H_k}$ are available. This approach has been used previously by Gao and Collins[5] for the GEM-3 EMI sensor. EMI sensors are largely independent of the environment and, hence, only aspects of the target need be modeled. In their work, Gao and Collins modeled mine parameters such as tilt, depth, metal content and metal distribution. In principle, the same approach could be used here: measurements of relevant environmental parameters could be used with the signature model to predict the observed signature as a function of unknown variables including the mine depth and size. Our experience to date with the model suggests that such an approach would not be successful. Although the spatial shape of a mine signature can be predicted with some confidence, the temperature contrast over the mine is strongly dependent on environmental conditions including the soil emissivity, soil moisture distribution with depth and the initial temperature distribution of the soil. Measurements of those data are often not available, which makes it difficult to accurately predict contrast. In addition, the thermal signature model is computationally expensive to run, and it would be prohibitively time consuming to synthesize the signatures necessary to evaluate Eq. (3) directly.

A common alternative approach to similar problems is the GLRT,[6] in which a maximum likelihood estimate $\hat{\mathbf{\Theta}}$ of the model parameters is used as follows:

$$\Lambda_G(\mathbf{D}) = \frac{\max_{\mathbf{\Theta}} f_{\mathbf{D}|\mathbf{\Theta},H_1}}{\max_{\mathbf{\Theta}} f_{\mathbf{D}|\mathbf{\Theta},H_0}} = \frac{f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_1(\tilde{\mathbf{\Theta}}_1))}{f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_0(\tilde{\mathbf{\Theta}}_0))} \tag{4}$$

where the estimated parameters $\tilde{\mathbf{\Theta}}_k$ maximize the densities $f_{\mathbf{D}|\mathbf{\Theta},H_k}$. Thus, GLRT detection is based primarily on the density of the model error. The model parameters are considered indirectly, via their contribution to the error. The estimator-correlator method[6] uses a similar approach, in which the parameters are estimated and the resulting model signature is correlated with the data. For Gaussian statistics, detection is based on the correlation.

In this work we propose a somewhat different approach. We estimate parameters $\hat{\mathbf{\Theta}}_k$ which minimize the residual error $|\mathbf{D} - \mathbf{G}_k(\hat{\mathbf{\Theta}}_k)|$ and supply these values to a classifier. We refer to this approach as an estimator-classifier, by analogy with the estimator-correlator technique. The technique is motivated by an approximate evaluation of the

Bayesian representation in Eq. (3). Consider the case in which the model $\mathbf{G}_k$ is a good fit to the data $\mathbf{D}$ for the parameter value $\hat{\mathbf{\Theta}}_k$. In that case, $f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}(\mathbf{\Theta}_k))$ will have a well defined maximum for $\mathbf{\Theta}_k = \hat{\mathbf{\Theta}}_k$ and we can write

$$f_{\mathbf{D}|H_k}(\mathbf{D}) \approx f_{\mathbf{\Theta}_k|H_k}(\hat{\mathbf{\Theta}}_k) \int d\mathbf{\Theta}_k f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}_k(\mathbf{\Theta}_k)) \approx E_k f_{\mathbf{\Theta}_k|H_k}(\hat{\mathbf{\Theta}}_k) f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}_k(\hat{\mathbf{\Theta}}_k)) \tag{5}$$

where we have assumed that $f_{\mathbf{\Theta}_k|H_k}$ is smooth, and the constant factor

$$E_k = \int d\mathbf{\Theta}_k h_k(\mathbf{\Theta}_k) \tag{6}$$

$$h_k(\mathbf{\Theta}) = \frac{f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}_k(\mathbf{\Theta}))}{f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}_k(\hat{\mathbf{\Theta}}_k))} \tag{7}$$

is related to the residual error. The function h_k satisfies $h_k(\hat{\mathbf{\Theta}}_k) = 1$ and $(\partial/\partial\mathbf{\Theta})h_k(\hat{\mathbf{\Theta}}_k) = 0$. Defining the matrix

$$[\mathbf{\Omega}_k]_{ij} = -\frac{\partial}{\partial\Theta_i}\frac{\partial}{\partial\Theta_j} h_k(\mathbf{\Theta}_k)\Big|_{\hat{\mathbf{\Theta}}_k} \tag{8}$$

We must also have $(\mathbf{\Theta}-\hat{\mathbf{\Theta}}_k)^T\mathbf{\Omega}_k(\mathbf{\Theta}-\hat{\mathbf{\Theta}}_k) >> 1$ for a well defined maximum in $f_{\mathbf{N}}$, which implies that near $\hat{\mathbf{\Theta}}_k$ we may use the approximation

$$h_k(\mathbf{\Theta}) \approx \exp\left(-(\mathbf{\Theta}-\hat{\mathbf{\Theta}}_k)^T\mathbf{\Omega}_k(\mathbf{\Theta}-\hat{\mathbf{\Theta}}_k)/2\right) \tag{9}$$

This result is distribution free but immediately evident for normally distributed $\mathbf{N}$. We are led to the result

$$E_k \propto \det(\mathbf{\Omega}_k)^{1/2} \tag{10}$$

Thus,

$$\Lambda(\mathbf{D}) \approx \frac{f_{\mathbf{\Theta}|H_1}(\hat{\mathbf{\Theta}}_1)\int d\mathbf{\Theta} f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}_1(\mathbf{\Theta}))}{f_{\mathbf{\Theta}|H_0}(\hat{\mathbf{\Theta}}_0)\int d\mathbf{\Theta} f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}_0(\mathbf{\Theta}))} = \frac{f_{\mathbf{\Theta}|H_1}(\hat{\mathbf{\Theta}}_1)}{f_{\mathbf{\Theta}|H_0}(\hat{\mathbf{\Theta}}_1)}\frac{f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}_1(\hat{\mathbf{\Theta}}_1))}{f_{\mathbf{N}}(\mathbf{D}-\mathbf{G}_0(\hat{\mathbf{\Theta}}_0))}\frac{\det(\mathbf{\Omega}_1)^{1/2}}{\det(\mathbf{\Omega}_0)^{1/2}} \tag{11}$$

It is notable that a GLRT approach in which the covariance matrix of the imagery is estimated using an ML approach yields a likelihood ratio that involves only the ratio of covariance matrix determinants.[4] In the estimator-classifier formulation detection is performed on the basis of the class-conditioned densities of $\mathbf{\Theta}$ and the residual model error $|\mathbf{D}-\mathbf{G}_k(\hat{\mathbf{\Theta}}_k)|$. Since a classifier of $\hat{\mathbf{\Theta}}_k$ estimates quantities comparable to $f_{\mathbf{\Theta}|H_k}(\hat{\mathbf{\Theta}}_k)$, it is reasonable to use such classifiers, with additional information on the residual error, for detection. This is the basis for the estimator-classifier approach. In our formulation, a nonlinear optimization procedure is used to find $\hat{\mathbf{\Theta}}_k$, and the corresponding model error. We use a support vector machine (SVM)[7,8] as a classifier. Radial basis functions are used, and the basis function parameter is adjusted to produce the best performance in leave-one-out testing of the results.

It is also easy to show that an expression resembling the GLRT can be derived from Eq. (2) using this approach if one makes the approximation that the densities $f_{\mathbf{\Theta}|h_k}$ are strongly peaked about parameter values $\tilde{\mathbf{\Theta}}_k$. A lack of a priori information regarding the parameters motivated our investigation of the estimator-classifier technique.

3. SIGNATURE MODEL DEVELOPMENT

We now define signature models $\mathbf{G}_k(\mathbf{\Theta})$ to be used for targets and clutter. IR signatures of buried land mines are functions of both space and time. We describe the spatial and temporal characteristics of these signatures using a combination of experimental data and modeling.

3.1. A Measured Data Set

In this work we use signatures measured as part of the 1998 NVESD MWIR and LWIR diurnal collection study. During the summer of 1998 a staring IR camera was used to image surface-laid and buried targets at two sites, Fort A.P. Hill, VA and Fort Huachuca, AZ. The layout of the minefields were similar for these two collections, and a description of the Ft. A.P. Hill data set is shown in Fig. 1. Data were acquired at 15 minute intervals for a duration of roughly 24 hours. Both Agema 570 (8-12 μm) and Amber Radiance (3-5 μm) cameras were used. We used data from the Agema camera, one image of which is shown in Fig. 2. These data were collected 1 July 1998 at Ft. A. P. Hill (Collection 3). Meteorological data were collected with the imagery. For this study fifteen target chips (one for each buried mine) and an equal number of clutter chips were extracted from this imagery. Examples of those chips are shown in Fig. 3.

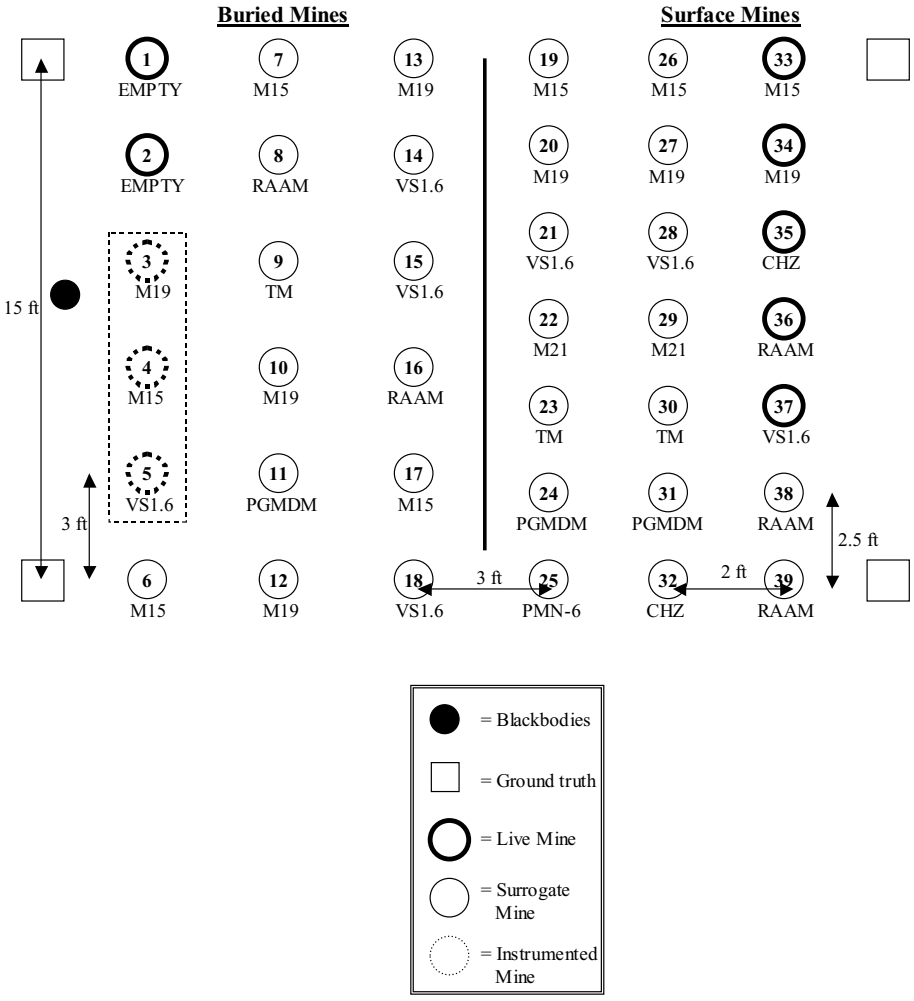


Figure 1. Mine layout for the AP Hill collection.

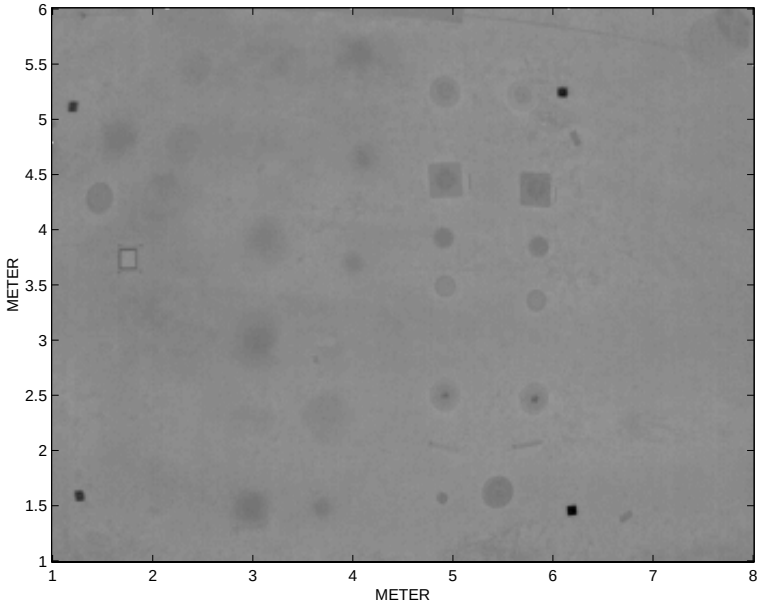


Figure 2. Example imagery from the Fort A. P. Hill collection.

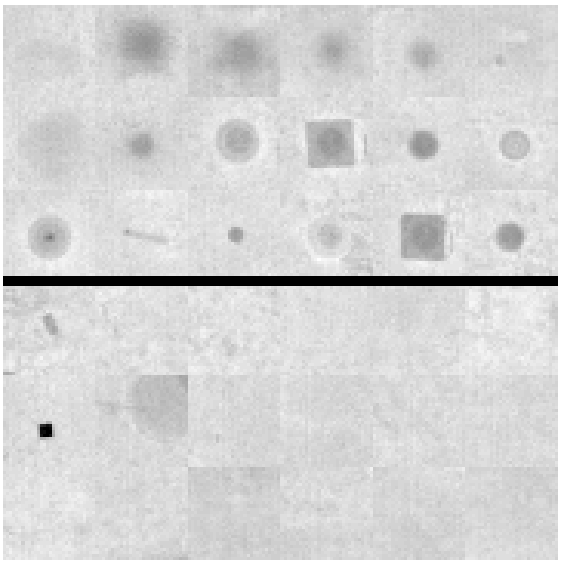
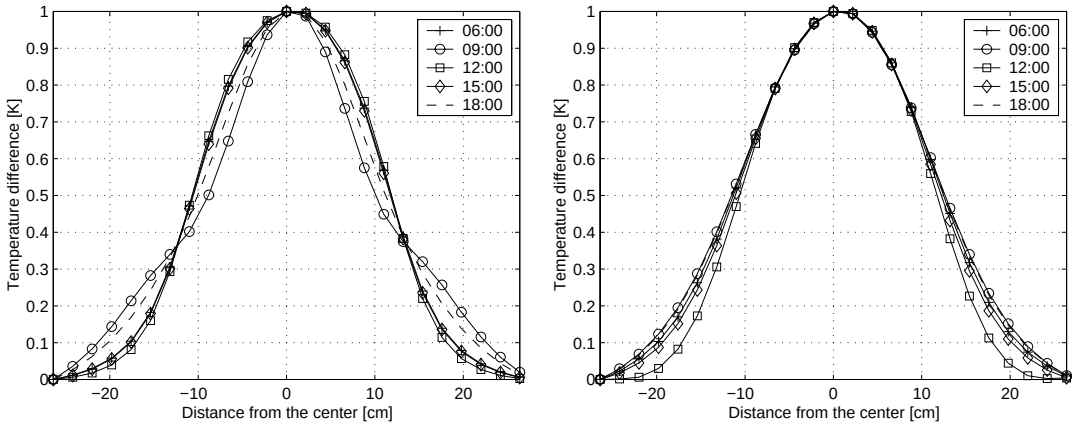


Figure 3. Chips from buried mines (above) and clutter (below).

3.2. Characteristics of Thermal IR Buried Mine Signatures

We can broadly characterize the signature model parameters as mine-related or environmentally related. The role of environmental variables is discussed in a companion paper.[2] In that work it is shown that the signature's shape is weakly dependent on the environmental variables. In addition, many details of the mine's construction have a small effect on soil surface temperature, since diffusion of the thermal signature through the overlying soil tends to smooth small-scale spatial variations. It remains to consider the signature's temporal behavior and the relation of mine-related variables to signature shape.

Consider first the dependence of the signature shape on mine-related variables such as the mine diameter and burial depth. Each buried mine comprises a single sample of a specific mine diameter and burial depth. Since the number of measured mine signatures is small (only 15), it is appropriate to examine the influence of these variables using modeling. Simulations of surface temperature contrast as a function of time and depth reveal that the spatial character of the signature is (after normalization to unit amplitude) only weakly dependent on mine depth. Although the peak amplitude of the contrast changes from positive to negative over the course of a diurnal cycle, the spatial distribution of the signature varies only moderately. Examples of this behavior are shown in Fig. 4 where we find that the shape of the signature is largely invariant with time. Fig. 5 shows that the depth of the mine also has little effect on the shape of the signature. The mine diameter has a more significant effect on the shape of the signature, as shown in Fig. 6. We see that for a fixed time and burial depth, the radius of the signature varies strongly with mine diameter. This suggests that the signature shape can be modeled with a small number of parameters, one of which should be related to diameter. These findings also suggest that the spatial dependence of the signature can be decoupled from the temporal dependence.



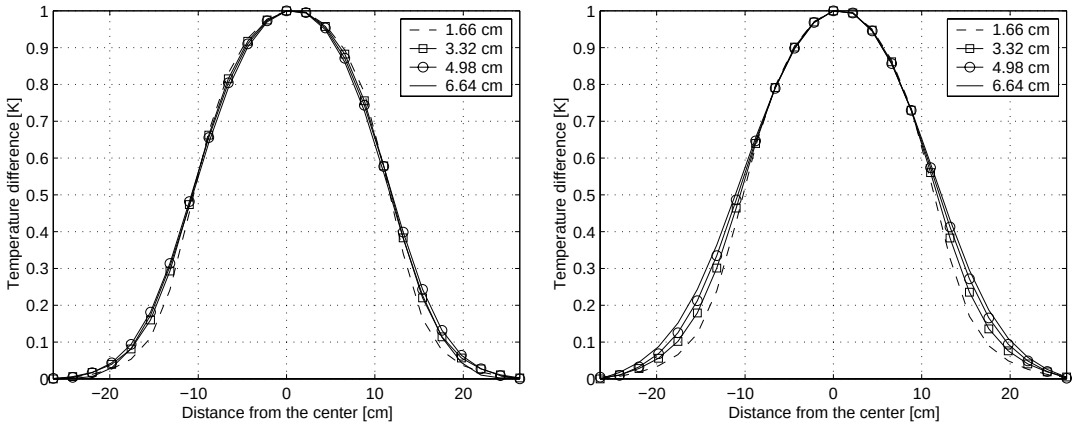
(a) 3.32 cm depth

(b) 6.64 cm depth

Figure 4. Spatial distribution of temperature contrast for mines buried at two depths and various times. The curves have been normalized to produce unit maximum.

Conversely, the signature's temporal behavior is strongly dependent on the environment and, in particular, on the solar insolation. The effect is most noticeable during daylight hours when broken clouds are present, since the temperature of the thin layer of soil at the surface can change rapidly in response to the incident sunlight. Figure 7 presents temperature histories for buried mines in relation to clutter. The temperature history during daylight hours is strongly dependent on the incident solar flux, which varies as a result of solar insolation.

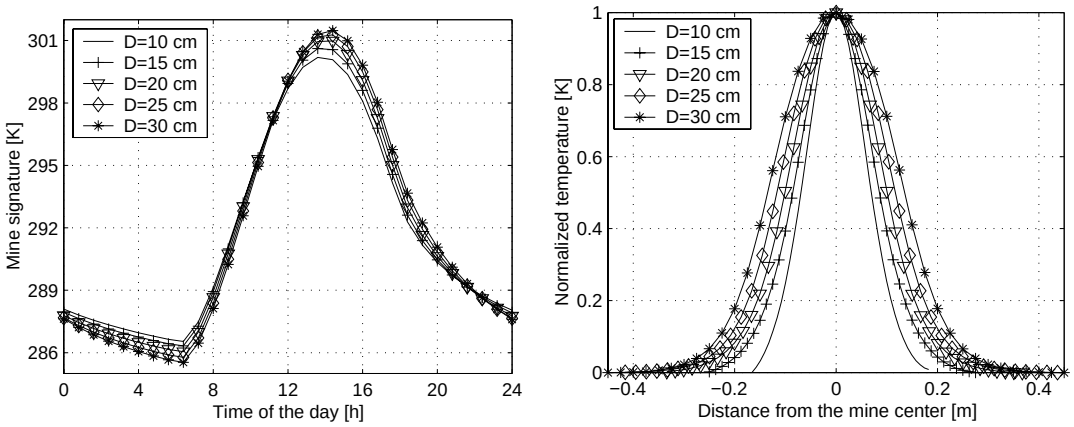
Although the temperature over a mine varies strongly with environmental conditions, detection is typically based on the thermal contrast — the difference between the temperature over a mine and the temperature of the surrounding soil. Plots the thermal contrast for buried mines are shown in Fig. 8. The curves pass through zero twice daily, and these times are referred to as "crossover times". For times in the early morning and late evening (i.e., between the thermal crossover times), a detectable contrast is present. The soil over a mine tends to be somewhat cooler during the evening hours, as predicted by the model.[2] The model also predicts that the soil over a mine will be somewhat



(a) 1200 Hours

(b) 1500 Hours

Figure 5. Spatial distribution of temperature contrast for two times and various mine depths. The curves have been normalized to produce unit maximum.



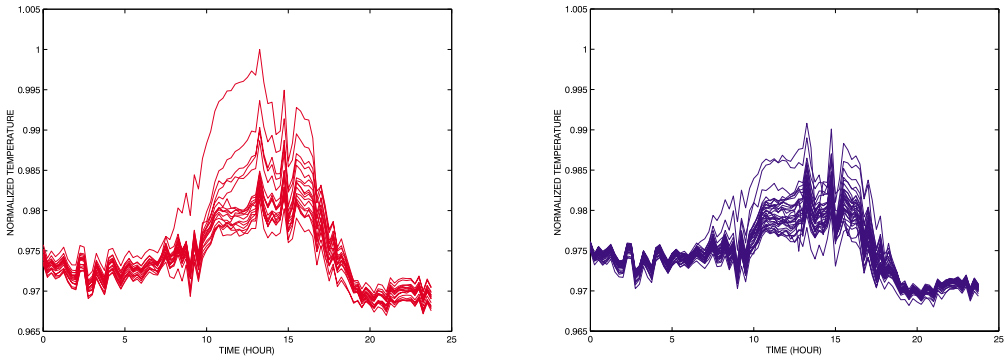
(a) Temporal variations in the temperature over the center of the mine.

(b) Spatial distribution of temperature contrast for a fixed time. The curves have been normalized to unit maximum.

Figure 6. Variations in the signature with mine radius.

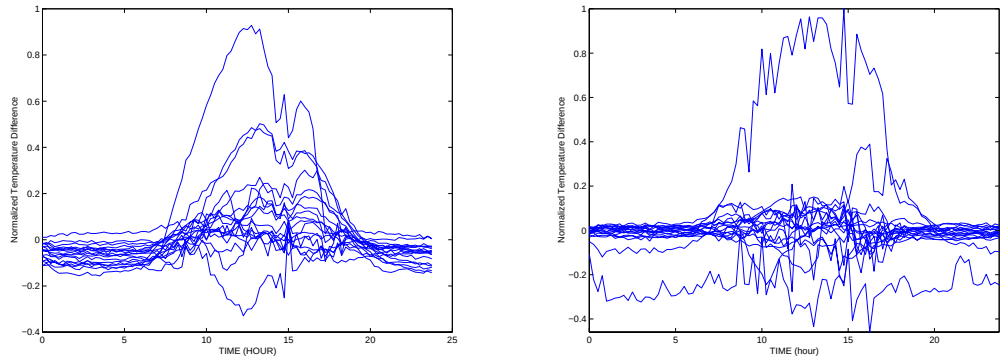
warmer during the day. The temperature difference between two points in the clutter chips are also shown in Fig. 8, where we see that those signatures are generally flat. One clutter chip contains a surface object, the temperature of which closely tracks the incident solar radiance.

A review of the results presented in this section suggests that spatial target signatures can be described by a model with a small number of physical parameters. Moreover, several of the physical parameters have a similar effect on the model, which implies that to first order the signature can be described by a smaller number of (possibly non-physical) parameters. Conversely, the temporal dependence of the signature is difficult to model, because it depends strongly on environmental factors.



(a) Mines (b) Clutter

Figure 7. Temperature histories over buried mines and clutter



(a) Mines (b) Clutter

Figure 8. Thermal contrast over buried mines and clutter.

3.3. Model Definition

Since the shape and time dependence of the signature are only weakly coupled, and since the shape depends on a small number of parameters, it is attractive to represent the IR signature by a relatively simple parameteric model. The spatial signature dependence is particularly amenable to this treatment.

A good fit to the aforementioned data set can be obtained with the following simple expressions:

$$\mathbf{G}(\mathbf{\Theta}_1) \quad = \quad b(t) + \frac{c(t)}{1 + (|\mathbf{r}|/\Delta)^p} \tag{12}$$

in which

$$\mathbf{\Theta}_1 \quad = \quad [b(t) \quad c(t) \quad \Delta \quad p] \tag{13}$$

where $b(t)$ and $c(t)$ are the time-dependent background and contrast temperatures, Δ is a spatial scale factor, and p controls the spatial decay rate of the signature. The same model is used for both clutter and mines. The quantities Δ and p control the shape of the signature, and they subsume variations with mine diameter and depth. The time

dependence is described by $b(t)$ and $c(t)$, of which only $c(t)$ is relevant to detection. The background temperature $b(t)$ can be obtained from modeling or measurements of the local environment.

As noted above, the mine contrast $c(t)$ and the background $b(t)$ depend strongly on a number of time-dependent quantities including the local solar insolation. A study of the experimental data shows that the background and contrast information are well approximated by a normalized reference signature $b_0(t)$ and $c_0(t)$ with random constants b and c, which leads to

$$\mathbf{G}(b, c, \Delta, p) = bb_0(t) + \frac{cc_0(t)}{1 + (|\mathbf{r}|/\Delta)^p} \tag{14}$$

where $b_0(t)$ and $c_0(t)$ are developed from models or from training data. An example of this model and it's fit to experimental data are shown in Fig. 9. Although the data are noisy, the model fits reasonably well.

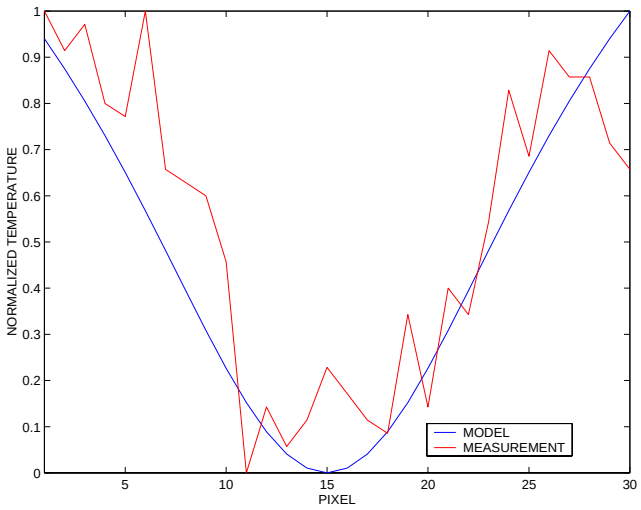


Figure 9. A comparison of the empirical model in Eq. (14) and measured data $\mathbf{D}$.

4. TEMPORAL DETECTION ALGORITHM

Because the model $\mathbf{G}(\boldsymbol{\Theta})$ depends nonlinearly on the parameters Δ and p, a nonlinear optimization procedure must be used to find the ML estimates $\hat{\boldsymbol{\Theta}}_k$. This process is computationally expensive, which motivates consideration of other more efficient detection methods. Since the nonlinearities arise in the spatial signature dependence, one approach is to base detection entirely on the temporal signatures. It is relatively easy to adapt the foregoing approach to temporal-only processing.

Mine detection in multi-temporal imagery was investigated using two techniques. In the first approach we searched for spatial temperature anomalies. The time histories were sampled at six points taken roughly four hours apart starting at approximately 4:00 AM. This set of data was used with the well-known RX algorithm of Reed and Yu.[4] That techniques comprises a CFAR implementation of a GLRT detector for a known target shape with unknown spectral amplitude and covariance in spatially white Gaussian noise. An ML technique is used to solve for the unknown amplitude c. The algorithm produces a likelihood ratio derived from local estimates of the residual model error and clutter variance. Because RX attempts to match to a known target shape, some spatial signature information is used. Like the GLRT approach described above, the unknown amplitude c is only used indirectly in detection, i.e., via its influence on the covariance matrix.

The second approach was model-free. We used a classifier to detect mines based on their temperature history. A support vector machine was trained on the image chips and used in classification. No spatial information was used in this detector. Each pixel in a chip was presented to the classifier, resulting in a binary distribution of single-pixel

detections. That binary image was processed to reject small pixel groupings, and detection was based on the presence of a "blob" at the location of the target.

5. RESULTS

Figure 10 summarizes the performance of the estimator-classifier detector and the GLRT detector, both of which exploit the full space-time signatures. The figure also shows the performance of the model-free temporal-only detector using the SVM. The results used leave-one-out testing of the 30 image chips described in Section 3.1.

The improved performance of the estimator-classifier using space-time information suggests that there is significant detection information in the model parameters. The GLRT, which ignores that information, shows markedly worse performance. The performance of the temporal-only SVM classifier is comparable to that of the GLRT. The explicit use of spatial information, conveyed using the model parameters, appears to be correlated with good performance.

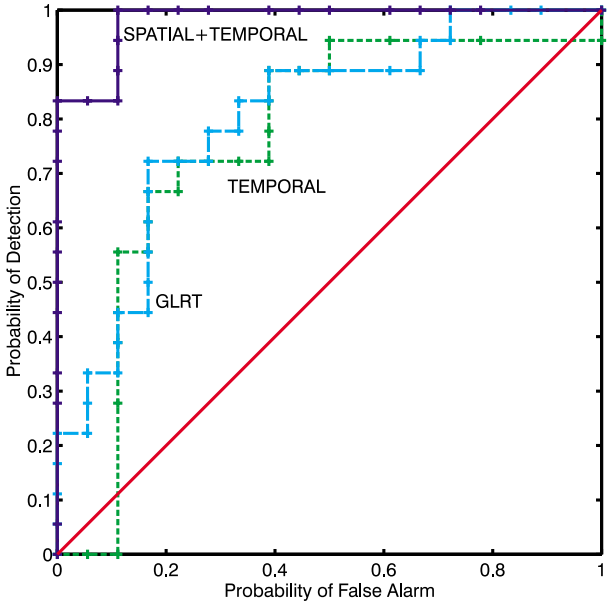


Figure 10. ROC curves for classifiers using (1) the estimator-classifier for the full space-time signature ("Spatial+Temporal"), (2) the GLRT for the space-time signature ("GLRT"), and (3) the SVM classifier using only temporal information ("Temporal").

Several of the algorithms discussed above use both spatial and temporal information in detection. Since the temporal information is often inconvenient or impossible to acquire, we next investigate its contribution to detection performance. This issues is addressed in Fig. 11 where we show detection of buried mines using a single "snap-shot" image and the entire time history. Two snap-shots are considered in Fig. 11. In the first, which occurred between crossover times, we see that the single-time and full-time performance are similar. The performance for the second snap-shot is dramatically worse. In that case, the sample time corresponds to a thermal crossover, when the mine signature is weak. We conclude that the performance of a detector that uses a single look can be comparable to that of a detector using the full time history but, as expected, performance is strongly dependent on the sample time and its relation to thermal crossover.

Finally we compare the performance of the relatively efficient detectors considered in Section 4. The ROC curves generated by RX (with spatial and temporal information) and by leave-one-out testing of the SVM using temporal data only are shown in Fig. 12 for buried mines. We see that the classifier approach performs significantly better than the RX algorithm for these data. The results suggest that assuming an incorrect signature shape can be worse than ignoring the spatial information entirely.

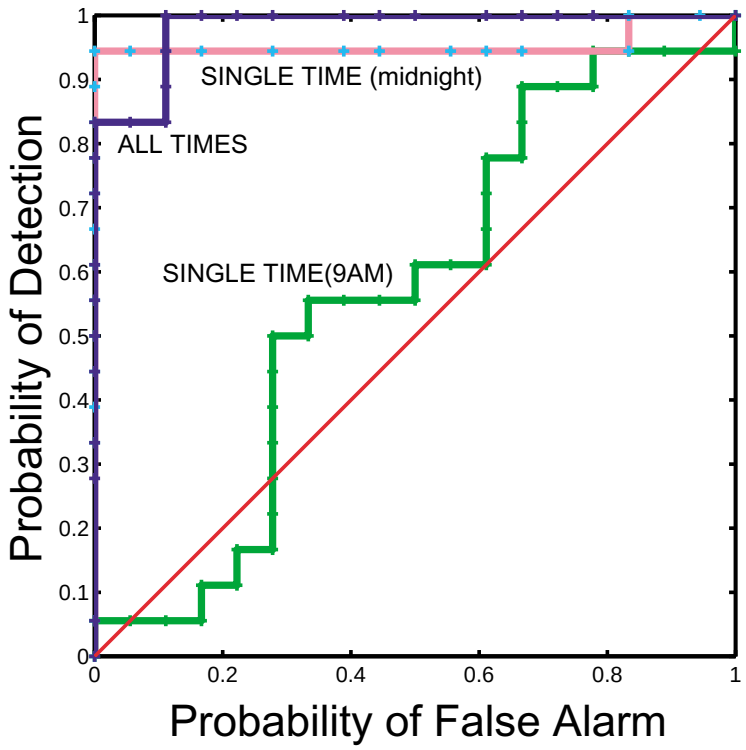


Figure 11. ROC curves comparing detection for a single time and for the full space-time history.

6. CONCLUDING REMARKS

We have described a number of model-based techniques for detecting buried land mines. Results from a signature simulation tool and an analysis of experimental data were used to devise a relatively simple signature model having a small number of parameters. The techniques examined include (1) the estimator-classifier technique developed by the authors, which uses model parameters and residual error via a classifier; (2) the GLRT, which uses only the residual error; (3) the RX algorithm of Reed and Yu,[4] which assumes a signature shape; and (4) a model-free classifier of temporal samples, which ignores spatial information present in the data.

These algorithms were used to detect buried mines in imagery acquired during July 1998 at Fort A.P. Hill, VA. That data set comprises imagery from 15 mines collected at several times over a 24 hour cycle. The detection performance of these algorithms was compared. The results suggest the following:

First, model-based detection algorithms can out-perform model-free algorithms, even when the model parameters must be estimated from the data. A direct comparison of the estimator-classifier method and the GLRT, both of which use the same model, suggests that the model parameters have information that significantly improves detection.

Second, detection performance using only spatial information from a single well-chosen time, can be comparable to the performance achieved using the complete time history of the signature. The choice of sample time, however, is critical to success.

Finally, the spatial shape of mine signatures varies considerably, and it is not sufficient to use a single universal shape for all mines. The performance of the RX algorithm, which assumes a constant mine shape, was significantly inferior to even that of the temporal classifier.

7. ACKNOWLEDGMENTS

This project was supported in part by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

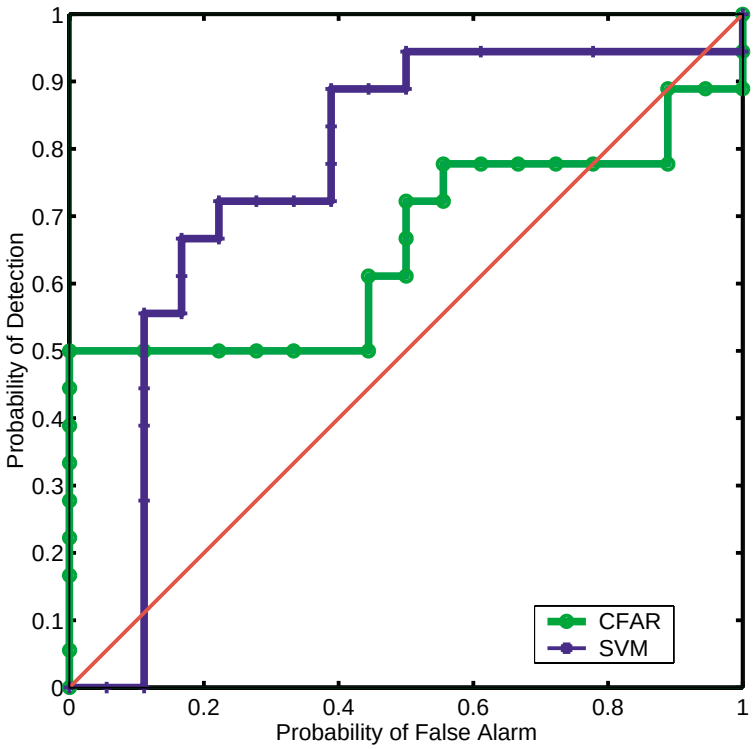


Figure 12. ROC curves for (1) a CFAR algorithm that uses spatial and temporal information with an assumed constant mine signature, and (2) a temporal-only model-free SVM classifier.

REFERENCES

1. I. K. Sendur and B. A. Baertlein, "Simulation of thermal IR buried mine signatures," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, T. J. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **4038**, pp. 156–166, 2000.
2. B. A. Baertlein and I. K. Sendur, "The role of environmental factors and mine composition in thermal IR mine signatures," in *Detection and Remediation Technologies for Mines and Minelike Targets VI*, A. C. Dubey, J. F. Harvey, T. J. Broach, and V. George, eds., *Proceedings of SPIE* **4394**, p. this issue, 2001.
3. M. Lundberg and I. Y. H. Gu, "A 3-D matched filter for detection of land mines using spatio-temporal thermal modeling," in *Detection Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *Proc. SPIE* **4038**, pp. 179–188, 2000.
4. I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoustics, Speech and Signal Proc.* **38**, pp. 1760–1770, October 1990.
5. P. Gao and L. Collins, "Improved signal processing approaches for landmine detection," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *SPIE* **3392**, pp. 1034–1043, 1998.
6. H. V. Poor, *An Introduction ot Signal Detection and Estimation*, Springer-Verlag, New York, NY, 2nd ed., 1994.
7. V. Vapnik, *The Nature of the Statistical Learning Theory*, Springer Verlag, New York, NY, second ed., 2000.
8. B. Schölkopf, C. J. C. Burges, and A. J. Smola, *Advances in Kernel Methods: Support Vector Learning*, MIT Press, Cambridge, MA, 1999.

A Reference Solution for Thermal Mine Signature Modeling

İbrahim Kürşat Şendur and Brian A. Baertlein

The Ohio State University, ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

Predicting the thermal signature of a buried land mine requires modeling the complicated inhomogeneous environment and the structurally complex mine. It is useful, both in checking such models and in making rough calculations of expected signatures, to have an accurate, easily computed solution for a relatively simple geometry. In this paper, a reference solution is presented for the integral equation that governs the temperature distribution. Our solution procedure uses the method of weighted residuals. The problem comprises a homogeneous cylindrical body (the mine model) buried in an infinite homogeneous half space (the soil model) with a planar interface. Using periodic boundary conditions in time at the planar interface, the temperature distribution in the lower-half space is expanded in a Fourier series. A volume integral equation for the Fourier series coefficients is obtained via Green's second identity. The Green's function for the Fourier coefficients is derived and reduced to a computationally efficient form. The integral equation is reduced to a matrix equation, which is then solved for the unknown temperature distribution. The integral equation solution is compared with a finite element model.

Keywords: Thermal infrared imagery, land mines, heat transfer, integral equations, thermal model, numerical simulation

1. INTRODUCTION

Research in infrared (IR) detection of land mines has been ongoing for several decades, and those sensors (especially multi-spectral instruments) are considered promising. Since IR sensors provide information independent of the electrical properties of the target, they complement metal detectors and ground-penetrating radar in a sensor-fused system.

Although many experimental and empirical studies have been performed, the physics that define IR signatures are poorly understood. These signatures are dependent on several factors, including solar insolation, cloud cover, vegetation, surface irregularities, and past meteorological conditions. In order to improve detection performance and to better utilize IR sensors, there is a critical need to understand the effects of these environmental factors.

Recently, combined thermal and radiometric models of IR mine signatures have been proposed, which offer insight into the underlying physics. Sendur and Baertlein[1] used a one-dimensional analytical thermal model to predict the surface temperature distribution over a mine. The land mine is assumed to be buried under a layered-earth heated by natural diurnal sources. The analytical formulation is based on a Fourier analysis of the periodic phenomenon. Closed form expressions were derived for the temperature distribution on the soil surface and at depth. That simple analysis confirmed a number of results that are seen in experimental studies, including the presence of two thermal "cross-over" times, a lag between the solar illumination and the soil temperature, the attenuation of the signature strength with depth, and a marked difference in the signatures of thermally conducting and insulating buried objects. Pregowski and Swiderski[2] developed a two-dimensional model to predict the thermal IR signatures of buried land mines. Those numerical simulations were based on a finite-difference time-domain (FDTD) formulation and they included water transport in the soil medium. Pregowski and Swiderski concluded that a single IR image taken at a single spectral range may not be sufficient for IR detection because of clutter. In a more recent study Sendur and Baertlein[3] developed a three-dimensional thermal and radiometric model to predict IR mine signatures. The thermal model was based on the finite element method (FEM), which is capable of modeling realistically shaped mines and inhomogeneous soil. The radiometric model accounted for both the spectral and spatial character of the surface, and the camera geometry and field of view (FOV) are included.

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

Although full three-dimensional solutions to the problem have now appeared, the accuracy and efficiency of such model must be addressed. For this reason, an accurate and easily computed solution for a relatively simple geometry is desired for checking the more detailed models and for making rough calculations. In this paper, we present a reference solution based on an integral-equation formulation. The problem geometry comprises a homogeneous cylindrical body (the mine model) buried in an infinite homogeneous half-space (the soil model) with a planar interface. The solution of the volume integral equation for the Fourier series coefficients is obtained using the method of weighted residuals, a technique better known to researchers in electromagnetics as the "Method of Moments" (MoM).

This work is organized as follows: Sect. 2 provides a description of the three-dimensional heat flow equation, including the convective and radiative boundary conditions at the soil-air interface. In addition, the Fourier series representation of the temperature distribution with appropriate boundary conditions is presented in that section. In Sect. 3 integral representations for the temperature distribution in soil and mine are obtained. The Green's function for this problem is derived in Sect. 4 and subsequently simplified and converted into a computationally efficient form. The numerical solution procedure for the Fourier series coefficients of the temperature distribution is formulated in Sect. 5. Example results are presented in Sect. 6, and concluding remarks appear in Sect. 7

2. PHYSICAL DESCRIPTION OF THE PROCESS

The temperature distribution in the soil and mine is described by the three-dimensional heat flow equation which, for piecewise constant media can be written

$$\nabla^2 T(\mathbf{r},t) - \frac{1}{\kappa_i(\mathbf{r})}\frac{\partial T(\mathbf{r},t)}{\partial t} = 0, \qquad i = s, m \tag{1}$$

in which $T(\mathbf{r},t)$ [K] is the temperature and κ is the thermal diffusivity, defined as the ratio of the thermal conductivity to the volumetric heat capacity. In this result the subscripts s and m refer to the soil and mine, respectively. To complete the thermal problem description, boundary condition at the soil-air interface must be specified. At that boundary convective and radiative heat transfer mechanisms are present. Numerous models have been developed in the literature to represent the heat transfer at an soil-air interface. Those models have been extensively studied in one-dimensional thermal analyses of homogeneous soil over a diurnal cycle,[4–9] and they are used here. The net heat flux into the ground, $\mathcal{F}_{net}$, can be written as

$$\mathcal{F}_{net}(t) = \mathcal{K}\frac{\partial T(\mathbf{r},t)}{\partial z}\bigg|_{z=0} = \mathcal{F}_{sun}(t) + \mathcal{F}_{sky}(t) - \mathcal{F}_{sh}(t) - \mathcal{F}_{gr}(t), \tag{2}$$

where $\mathcal{F}_{sun}$ is the incident solar radiation reduced by cloud extinction, atmospheric absorption, soil albedo and the cosine of the zenith angle; $\mathcal{F}_{sky}$ is the sky brightness with a correction for cloud cover; $\mathcal{F}_{sh}$ is the sensible heat transfer from land to atmosphere due to convection; and $\mathcal{F}_{gr}$ is the gray body emission from the soil surface.* Expressions for these terms have appear in prior works,[3,4] and here we summarize those findings. For the solar flux we write

$$\mathcal{F}_{sun}(t) = S_0 C_{sun}\mathcal{H}(t) \tag{3}$$

in which $S_0 = 1353$ [W/m^2] is the solar constant, C_{sun} is a factor that accounts for the reduction in absorbed energy due cloud cover and soil reflectivity, and $\mathcal{H}(t)$ is the local insolation function, assumed to be periodic over a diurnal cycle. The long-wavelength radiation from the atmosphere, $\mathcal{F}_{sky}$, is given as

$$\mathcal{F}_{sky}(t) = \sigma T_{sky}^4(t), \tag{4}$$

where $\sigma = 5.67 \times 10^{-8}$ [W m^{-2} K^{-4}] is the Stephen-Boltzmann constant and T_{sky} [K] is an effective sky radiance temperature. The heat loss due to ground radiation is given by Stefan's Law

$$\mathcal{F}_{gr}(t) = \mathcal{E}\sigma T^4(t, z=0), \tag{5}$$

where $\mathcal{E}$ [unitless] is the mean emissivity of the surface and $T(t, z=0)$ is the soil temperature at the soil-air interface. The sensible heat transfer between the surface and atmosphere is approximated by

$$\mathcal{F}_{sh}(t) = h(t)(T_{air}(t) - T(t, z=0)). \tag{6}$$

*In formulating this expression we have ignored evaporation of soil moisture and changes in the state in the medium.

where $h(t)$ is a convection coefficient which depends strongly on the local wind speed. Empirical models for the air and sky temperatures are

$$T_{air}(t) = T_{0,air} - T_{del}\cos(2\pi(t-2)/24) \quad (7)$$
$$T_{sky}(t) = T_{air}(t)C_{sky} \quad (8)$$

where t is the local time as above, and $T_{0,air}$ and T_{del} are estimated from local meteorological data available from the National Weather Service, and C_{sky} is a constant[7] that depends on the local meteorological conditions.

Using Eqs. (3) and (4)-(6) in Eq. (2) results in the boundary condition at the soil-air boundary. This condition is a nonlinear function of the surface temperature distribution, but over the limited temperature range of interest, the result can be linearized using the technique described by Watson[4] to obtain

$$\frac{\partial T(\mathbf{r},t)}{\partial z} \approx T(\mathbf{r},t)\frac{1}{\mathcal{K}_s}(h(t) + 4\epsilon\sigma T_{sky}^3(t)) - \frac{1}{\mathcal{K}_s}(\mathcal{F}_{sun}(t) + h(t)T_{air}(t) + 4\epsilon\sigma T_{sky}^4(t)). \quad (9)$$

The boundary condition at the soil-air interface given by Eq. (2) is time-varying. If the convection coefficient, $h(t)$, is assumed to be time invariant over the diurnal cycle and approximated by its mean value $\overline{h}$, then the boundary condition becomes a periodic function at the diurnal rate. This approximation suggests that the temperature distribution, $T(\mathbf{r},t)$, can be written as

$$T(\mathbf{r},t) = \sum_{n=-\infty}^{\infty} T_n(\mathbf{r})e^{i\omega nt} \quad (10)$$

where n denotes the nth component of the Fourier series expansion, $\omega = 2\pi/(24\times 60\times 60)$ is the angular frequency and $T_n(\mathbf{r})$ is the nth (complex) coefficient of the Fourier series expansion. Since $T(\mathbf{r},t)$ is a real quantity. the coefficients are conjugate symmetric, i.e., $T_n(\mathbf{r}) = T_{-n}^*(\mathbf{r})$ where * denotes the complex conjugate operator. Substituting Eq. (10) in Eq. (1), the Fourier coefficients of the three-dimensional heat-flux equation can be written as

$$\nabla^2 T_n(\mathbf{r}) - \frac{i\omega n}{\kappa_i(\mathbf{r})}T_n(\mathbf{r}) = 0, \qquad i = s, m \quad (11)$$

Solar heating is the dominant heat source in this problem, and this fact permits some additional approximations. Replacing $T_{air}(t)$ with its mean value $\overline{T}_{air} = T_{0,air}$, and simplifying T_{sky} in a similar manner yields

$$\frac{\partial T(\mathbf{r},t)}{\partial z} \approx T(\mathbf{r},t)\frac{1}{\mathcal{K}_s}\left(\overline{h} + 4\epsilon\sigma\overline{T}_{sky}^3\right) - \frac{1}{\mathcal{K}_s}\left(\mathcal{F}_{sun}(t) + \overline{h}T_{0,air} + 4\epsilon\sigma\overline{T}_{sky}^4\right) \quad (12)$$

By using the Fourier series expansion of the solar insolation function

$$\mathcal{F}_{sun}(t) = \sum_{n=-\infty}^{\infty} \mathcal{F}_n e^{i\omega nt} \quad (13)$$

the boundary condition for the nth Fourier coefficient can be written as

$$\frac{\partial T_n(\mathbf{r})}{\partial z} = \alpha T_n(\mathbf{r}) + \beta_n \quad (14)$$

where

$$\alpha = \overline{h} + 4\epsilon\sigma\overline{T}_{sky}^3 \quad (15)$$
$$\beta_n = -\mathcal{F}_n - \overline{h}T_{0,air} - 4\epsilon\sigma\overline{T}_{sky}^4 \quad (16)$$

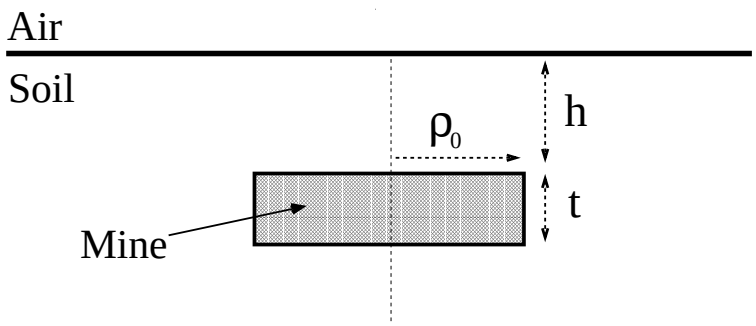


Figure 1. A cylindrical mine with radius ρ_0 and thickness t, buried at depth h under a smooth soil surface.

3. INTEGRAL REPRESENTATION FOR THE TEMPERATURE DISTRIBUTION

An integral representation of the temperature distribution for both soil and mine can be obtained using Green's second identity with an appropriate Green's function. In this section such an integral representation is obtained for a cylindrical mine with radius ρ_0 and thickness t, buried at depth h under a smooth soil surface as shown in Fig. 1. Homogeneous soil with thermal diffusivity $\kappa_s(\mathbf{r}) = \kappa_s$ and thermal conductivity $\mathcal{K}_s(\mathbf{r}) = \mathcal{K}_s$ is assumed. The thermal properties of the mine need not be uniform in the following formulation and the position ($\mathbf{r}$) dependence on the mine's thermal properties will be retained in this formulation.

For a cylindrical mine, the problem has rotational symmetry. Therefore, Eq. (11) can be written as

$$\begin{array}{rcl} \nabla^2_{\rho z} T^s_n(\mathbf{r}) - k_n^2 T^s_n(\mathbf{r}) & = & 0 \quad ; \quad \mathbf{r} \in \text{soil} \\ \nabla^2_{\rho z} T^m_n(\mathbf{r}) - k_1^2(\mathbf{r}) T^m_n(\mathbf{r}) & = & 0 \quad ; \quad \mathbf{r} \in \text{mine} \end{array} \tag{17}$$

where

$$\nabla^2_{\rho z} = \frac{1}{\rho}\frac{\partial}{\partial \rho}\Big(\rho \frac{\partial}{\partial \rho}\Big) + \frac{\partial^2}{\partial z^2} \tag{18}$$

$$k_n^2 = \frac{i\omega n}{\kappa_s} \tag{19}$$

$$k_1^2(\mathbf{r}) = \frac{i\omega n}{\kappa_m(\mathbf{r})} \tag{20}$$

In Eq. (17) $T^m_n(\mathbf{r})$ and $T^s_n(\mathbf{r})$ represent the Fourier coefficients in the mine and soil regions, respectively.

We employ a standard method[10] of solving the differential operator equation defined in Eq. (11). Forming the inner product of that equation and the Green's function G yields

$$\langle \nabla^2_{\rho z} T_n(\rho', z'), G(\rho, z; \rho', z') \rangle = \langle \nabla^2_{\rho z} G(\rho, z; \rho', z'), T_n(\rho', z') \rangle + \mathcal{R}(\rho, z) \tag{21}$$

where $\langle \cdot \rangle$ is the inner product, which consists in integration over the entire lower half space. The Green's function satisfies the equation

$$\big(\nabla^2_{\rho z} - k_n^2\big) G(\rho, z; \rho', z') = -\frac{\delta(\rho - \rho')}{\rho} \delta(z - z') \tag{22}$$

with boundary conditions defined subsequently. The conjunct[10] $\mathcal{R}$ is an integral over the problem boundaries and is also defined below. Using the above property of G and the differential equation satisfied by T_n we immediately obtain

$$T_n(\rho, z) = \int_{z_1}^{z_2} dz' \int_0^{\rho_0} d\rho' \rho' c(\rho', z') T^m_n(\rho', z') G(\rho, z; \rho', z') + \mathcal{R} \tag{23}$$

In this result $z_1 = h$ and $z_2 = h + t$, we have exchanged primed and unprimed coordinates, and we have defined

$$c(\mathbf{r}) = k_1^2(\mathbf{r}) - k_n^2 = i\omega n \frac{\kappa_s - \kappa_m(\mathbf{r})}{\kappa_s \kappa_m(\mathbf{r})} \tag{24}$$

Integration by parts is used to evaluate $\mathcal{R}$ from its definition in Eq. (21). In doing so, we exploit the thermal boundary conditions, namely that the temperature is continuous

$$T_n^m(\mathbf{r}=\mathbf{r}_{ms}) = T_n^s(\mathbf{r}=\mathbf{r}_{ms}) \tag{25}$$

and the thermal fluxes are continuous

$$\mathcal{K}_m \nabla T_n^m(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{r}_{ms}} = \mathcal{K}_s \nabla T_n^s(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{r}_{ms}} \tag{26}$$

where $\mathbf{r}_{ms}$ denotes the mine-soil boundary. The following result is obtained

$$\begin{aligned}\mathcal{R}(\rho,z) &= -\int_0^\infty d\rho'\rho'\mathcal{Z}^s(\rho,z;\rho',0) + \left(1-\frac{\mathcal{K}_m}{\mathcal{K}_s}\right)\left(\rho_0\int_{z_1}^{z_2} dz'\mathcal{F}^m(\rho,z;\rho_0,z')\right.\\ &\left. + \int_0^{\rho_0} d\rho'\rho'\Big(\mathcal{S}^m(\rho,z;\rho',z_2) - \mathcal{S}^m(\rho,z;\rho',z_1)\Big)\right)\end{aligned} \tag{27}$$

in which

$$\mathcal{Z}^i(\rho,z;\rho',z') = \mathcal{S}^i(\rho,z;\rho',z') - T_n^i(\rho',z')\frac{\partial G(\rho,z;\rho',z'')}{\partial z''}\Big|_{z''=z'} \tag{28}$$

$$\mathcal{F}^i(\rho,z;\rho',z') = G(\rho,z;\rho',z')\frac{\partial T_n^i(\rho'',z')}{\partial \rho''}\Big|_{\rho''=\rho'} \tag{29}$$

$$\mathcal{S}^i(\rho,z;\rho',z') = G(\rho,z;\rho',z')\frac{\partial T_n^i(\rho',z'')}{\partial z''}\Big|_{z''=z'} \tag{30}$$

In Eqs. (28)-(30) the subscript i can be s and m which refer to the soil and mine, respectively. By selecting the boundary condition for the Green's function at the soil-air interface as

$$\frac{\partial G(\rho,z;\rho',z')}{\partial z} = \alpha G(\rho,z;\rho',z') \tag{31}$$

the integral representation for the temperature distribution $T_n(\rho,z)$ in the lower half space can be obtained as

$$\begin{aligned}T_n(\rho,z) &= -\int_{z_1}^{z_2} dz'\int_0^{\rho_0} d\rho'\rho' c(\rho',z')T_n^m(\rho',z')G(\rho,z;\rho',z')\\ &- \beta_n\int_0^\infty d\rho'\rho' G(\rho,z;\rho',z'=0) + \left(1-\frac{\mathcal{K}_m}{\mathcal{K}_s}\right)\left(\rho_0\int_{z_1}^{z_2} dz'\mathcal{F}^m(\rho,z;\rho_0,z')\right.\\ &\left. + \int_0^{\rho_0} d\rho'\rho'\Big(\mathcal{S}^m(\rho,z;\rho',z_2) - \mathcal{S}^m(\rho,z;\rho',z_1)\Big)\right)\end{aligned} \tag{32}$$

Equation (32) is an integral relation from which one can determine the temperature distribution anywhere in the lower half space by integrating the temperature distribution over the mine. If the point (ρ,z) is within the mine, this relation becomes a Fredholm integral equation of the second kind[10] for the unknown mine temperature distribution, i.e., the unknown function appears both inside and outside of the integral. In the absence of the mine, i.e., for a homogeneous half space, the first and last terms in Eq. (32) vanish, and the temperature distribution can be found directly from

$$T_n(\rho,z) = -\beta_n\int_0^\infty d\rho'\rho' G(\rho,z;\rho',z'=0) \tag{33}$$

4. GREEN'S FUNCTION

The Green's function G is the solution of the heat transfer equation for an internal point source of heat. It must satisfy Eq. (22) with the boundary condition on the soil-air interface given by Eq. (31). In addition to this boundary condition, the Green's function must have a finite value for $\rho = 0$ and must vanish at $\rho \to \infty$ and $z \to \infty$. In Sect. 4.1 G is derived, and in Sect. 4.2 G is transformed into a form more appropriate for numerical evaluation.

4.1. Derivation of Green's Function

The Green's function is easily derived in the spectral domain. Taking the Hankel transform[10] of both sides of the Eq. (22) we obtain a one-dimensional Green's function problem

$$\left(\frac{d^2}{dz^2} - \left(k_\rho^2 + k_n^2\right)\right) G(k_\rho, z; z') = -J_0(k_\rho \rho')\delta(z - z') \tag{34}$$

in which k_ρ is the spectral variable and J_0 is the Bessel function of the first kind. The solution of this equation is readily obtained using standard methods,[10] and after inverting the Hankel transform we obtain

$$\begin{aligned} G(\rho, z; \rho', z') &= \int_0^\infty dk_\rho J_0(k_\rho \rho') J_0(k_\rho \rho) \frac{k_\rho}{2\sqrt{k_n^2 + k_\rho^2}} \exp\left(-\sqrt{k_n^2 + k_\rho^2} z_>\right) \\ &\times \left[\exp\left(\sqrt{k_n^2 + k_\rho^2} z_<\right) + r_0(k_n, k_\rho, \alpha) \exp\left(-\sqrt{k_n^2 + k_\rho^2} z_<\right)\right] \end{aligned} \tag{35}$$

where $z_>$ ($z_<$) is the greater (lesser) of z and z', and r_0 is a thermal "reflection coefficient" given by

$$r_0(k_n, k_\rho, \alpha) = \frac{\sqrt{k_n^2 + k_\rho^2} - \alpha}{\sqrt{k_n^2 + k_\rho^2} + \alpha} \tag{36}$$

4.2. Simplification of Green's Function

Efficient, accurate calculation of G is crucial to a numerical solution of the integral equation, but in general, numerical evaluation of the Green's function is computationally challenging. The major difficulties in the computation of this integral is that (1) for large ρ and ρ' values, the functions $J_0(k_\rho \rho)$ and $J_0(k_\rho \rho')$ oscillate rapidly, leading to slow convergence; and (2) for $|z - z'| \ll 1$ or $(z + z') \ll 1$ the exponential terms in the integrals decay slowly, again producing slow convergence.

The integral given by Eq. (35) is similar to the Sommerfeld integrals,[11] which have been studied extensively in the physics and electromagnetics literature. Straightforward manipulations ameliorate some of its problems. The reflection coefficient $r_0(k_n, k_\rho, \alpha)$ can be expressed as

$$r_0(k_n, k_\rho, \alpha) = 1 - \frac{2\alpha}{\sqrt{k_n^2 + k_\rho^2} + \alpha} \tag{37}$$

and, hence, the Green's function can be represented as

$$G(\rho, z; \rho', z') = G_1(\rho, z; \rho', z') + G_2(\rho, z; \rho', z') + G_3(\rho, z; \rho', z') \tag{38}$$

where

$$G_1(\rho, z; \rho', z') = \int_0^{2\pi} d\phi' \frac{\exp(ik_n R_1)}{4\pi R_1} \tag{39}$$

$$G_2(\rho, z; \rho', z') = \int_0^{2\pi} d\phi' \frac{\exp(ik_n R_2)}{4\pi R_2} \tag{40}$$

$$G_3(\rho, z; \rho', z') = -\alpha \int_0^\infty dk_\rho \frac{k_\rho J_0(k_\rho \rho') J_0(k_\rho \rho)}{\sqrt{k_n^2 + k_\rho^2}(\sqrt{k_n^2 + k_\rho^2} + \alpha)} \exp\left(-\sqrt{k_n^2 + k_\rho^2}(z + z')\right) \tag{41}$$

in which we have defined

$$R_1 = \sqrt{\rho'^2 + \rho^2 - 2\rho'\rho\cos\phi' + (z-z')^2} \tag{42}$$

$$R_2 = \sqrt{\rho'^2 + \rho^2 - 2\rho'\rho\cos\phi' + (z+z')^2} \tag{43}$$

and we have used the following identities[12,10]

$$J_0(k_\rho\rho')J_0(k_\rho\rho) = \frac{1}{2\pi}\int_0^{2\pi} J_0(k_\rho\rho_t)d\phi \tag{44}$$

$$\frac{\exp(-k\sqrt{r^2+z^2})}{\sqrt{r^2+z^2}} = \int_0^\infty \gamma J_0(\gamma r)\frac{\exp(-|z|\sqrt{k^2+\gamma^2})}{\sqrt{k^2+\gamma^2}}d\gamma \tag{45}$$

where

$$\rho_t = \sqrt{\rho^2 + \rho'^2 - 2\rho\rho'\cos\phi} \tag{46}$$

The form of G_3 can be simplified by using a device suggested by Kuo and Mei.[13] Omitting the details in the interest of brevity, we obtain

$$G_3(\rho, z; \rho', z') = -\frac{\alpha}{2\pi}\int_0^{2\pi} d\phi \exp(-\alpha(z+z'))\int_{z+z'}^\infty \frac{\exp(ik_nR_3)}{R_3}\exp(-\alpha\zeta)d\zeta \tag{47}$$

where

$$R_3 = \sqrt{\rho'^2 + \rho^2 - 2\rho'\rho\cos\phi' + \zeta^2} \tag{48}$$

The last integral in Eq. (47) has infinite integration limits, but the integrand is a rapidly decaying function for typical values of α, and it is well approximated by an integral over a finite domain. When the source and observation points approach one another, we have $R_1 \to 0$, and the Green's function has a well-known integrable singularity. The treatment of this case is is discussed in Sect. 5.2.

5. NUMERICAL SOLUTION OF THE INTEGRAL EQUATION

In this section we describe a numerical solution procedure for the integral equation in Eq. (32) the yield the Fourier coefficients $T_n^m(\mathbf{r})$. Our approach is to solve for the temperature T_n^m within the mine. Given the mine's temperature distribution, we then evaluate Eq. (32) to find the Fourier coefficients $T_n(\rho, z=0)$ for the temperature distribution at the surface. The time history of the temperature is then found by evaluating the Fourier series.

We employ the method of weighted residuals (MWR) in this work. The electromagnetics community has developed an extensive body of knowledge on MWR solution of integral equations in the temporal frequency domain under the guise of the MoM. The MoM discretizes the integral equation into a matrix equation, the solution of which is obtained using standard methods. Following the pioneering works of Richmond[14] and Harrington,[15] an extensive literature has developed on this procedure, and summary references are also available.[16–18]

5.1. Solution Procedure

The solution of the integral equation using the MoM begins with a discretization of the unknown temperature distribution over the buried mine with expansion functions. It is important to select expansion functions appropriate for the problem at hand. The integral representation of the temperature distribution given in Eq. (32) requires both the temperature and its derivatives with respect to ρ and z. In this work we employ subsectional expansion function $\Lambda(\rho, z)$ that are linear in ρ and z with the result

$$T_n^m(\rho, z) = \sum_{m=1}^{M}\sum_{n=1}^{N} A_{m,n}\Lambda_{nm}(\rho, z) \tag{49}$$

where M and N are the number of divisions in the z and ρ directions. This representation yields a continuous, piecewise-linear representation of T_n^m and a piecewise constant form for its derivatives.

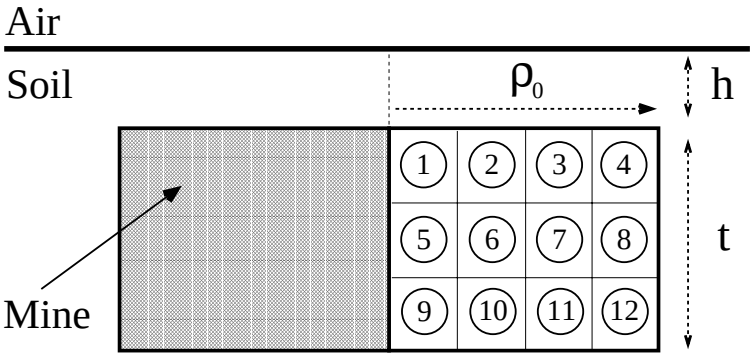


Figure 2. An example discretization of the buried mine with 4 subdivisions in the $\hat{\rho}$-direction and 3 subdivisions in the $\hat{z}$-direction.

A sample discretization of the buried mine with $N = 4$ subdivisions in the $\hat{\rho}$-direction and $M = 3$ subdivisions in the $\hat{z}$-direction is illustrated in Fig. 2. The thermal properties over each subsectional basis function are assumed constant and taken into account in the constant c_{mn} which can be defined as

$$c_{mn} = c\left((\rho_n + \rho_{n-1})/2, (z_m + z_{m-1})/2\right) \tag{50}$$

where c was previously defined with Eq. (24) and the indexed values of ρ and z are the subsection boundaries. By substituting the discretized temperature distribution given by Eq. (49) into the integral equation Eq. (32), we obtain

$$\begin{aligned}
\beta_n \int_0^\infty d\rho' \rho' G(\rho, z; \rho', z' = 0) \quad = \quad & -\sum_{m=1}^{M}\sum_{n=1}^{N} A_{mn}\left[\Lambda_{mn}(\rho, z) + c_{mn}\int\int dz' d\rho' \rho' G(\rho, z; \rho', z')\Lambda_{mn}(\rho', z')\right] \\
& + \left(1 - \frac{\mathcal{K}_m}{\mathcal{K}_s}\right)\left[\rho_0 \sum_{m=1}^{M} A_{mN}\int dz' G(\rho, z; \rho' = \rho_0, z')\frac{d}{d\rho}\Lambda_{nm}(\rho', z')\right. \\
& + \sum_{n=1}^{N} A_{Mn}\int d\rho' \rho' G(\rho, z; \rho', z' = z_2)\frac{d}{dz'}\Lambda_{nm}(\rho', z') \\
& \left. - \sum_{n=1}^{N} A_{1n}\int d\rho' \rho' G(\rho, z; \rho', z' = z_1)\frac{d}{dz'}\Lambda_{nm}(\rho', z')\right]
\end{aligned} \tag{51}$$

To complete the formulation, we multiply both sides of Eq. (51) by testing functions $w_p(\rho, z)$ and integrate over the mine. The result is a linear system of equations

$$\mathbf{V} = \overline{\mathbf{Z}}\,\mathbf{A} \tag{52}$$

for the constants A_{mn} which we represent by the matrix $\mathbf{A}$. Solving Eq. (52) using any convenient scheme yields the unknown coefficient vector $\mathbf{A}$ which can be used to compute the temperature distribution over the buried mine via Eq. (49). If desired, the temperature distribution everywhere in the lower half-space can be computed using Eq. (32). The details of computing the matrix elements Z_{pq} are presented in Sects. 5.2.

Various testing functions have been used in the literature. Galerkin methods are known to be optimum in the least-square sense, but they are also computationally expensive in terms of matrix-fill time. To reduce the computational cost, we have employed a point-matching technique, for which the testing functions are delta functions at appropriately chosen points in each subsection.

5.2. Computation of Z Matrix

In this section evaluation of the matrix elements Z_{pq} is discussed. To make the notation more concise, we use the index p to refer to a particular point in ρ, z, which we denote ρ_{n_p}, z_{m_p} with a similar convention for q. In general, Z_{pq} involves integrals of the form

$$\int_{z_{m_q-1}}^{z_{m_q}} dz' \int_{\rho_{n_q-1}}^{\rho_{n_q}} d\rho' G(\rho_{n_p}, z_{m_p}; \rho', z') f(\rho', z') \tag{53}$$

where $f(\rho, z)$ is polynomial function of the arguments. Consider first the off-diagonal matrix elements Z_{pq} with $p \neq q$. For this case, the testing point (ρ_{n_p}, z_{m_p}) is not in the integration domain and the Green's function is a smooth function throughout the integration domain. Therefore, Z_{pq} can be directly evaluated by numerical methods.

For the diagonal matrix elements Z_{pp}, however, the testing point (ρ_{n_p}, z_{m_p}) lies in the integration domain. This results in $R_1 \to 0$ which causes a singularity as R_1^{-1}. Consequently, an integrand regularization technique (e.g., a singularity extraction technique) is necessary in numerical evaluation of the integrals. The singularity extraction technique is widely used in MoM solutions in electromagnetics for both surface and volume formulations.[19–22] Singularity extraction techniques of the Green's functions for bodies of revolution have also been studied in the electromagnetics literature.[23,24]

In brief, calculation of the diagonal elements requires that the singular parts of the integrand are evaluated separately by analytical methods. A careful examination suggests that the singular part of the Green's function $G(\rho, z; \rho', z')$ is due to function $G_1(\rho, z; \rho', z')$, which leads us to write

$$G_1(\rho, z; \rho', z') = \int_0^{2\pi} d\phi' \frac{\exp(ik_n R_1) - 1}{4\pi R_1} f(\rho', z') + \int_0^{2\pi} d\phi' \frac{1}{4\pi R_1} f(\rho', z') \tag{54}$$

The first integral is evaluated numerically. The second term is integrated analytically in a small region about the singularity, and integration over the remainder of the domain is done numerically. Details may be found in the literature.

The numerical procedure described in Sect. 5.1 yields the Fourier coefficients of the temperature distribution within the mine. The temperature distribution over the soil surface $z = 0$ is found by using Eqs. (32) and (51) to obtain the Fourier coefficients at the surface. Some numerical problems arise in the surface integrals, since $z + z' = 0$, but an approximate, accurate, and efficient form for the Green's function is

$$\begin{aligned} G(\rho, z = 0; \rho', z' = 0) \quad &\approx \quad \frac{1}{2\pi} \int_0^{2\pi} d\phi' \frac{\exp(ik_n \rho_t)}{\rho_t} - \alpha K_0(-ik_n \rho_>) I_0(-ik_n \rho_<) \\ &- \quad \frac{\alpha^2}{2\pi} \int_0^{2\pi} d\phi' \frac{\exp(ik_n \rho_t)}{ik_n} - \frac{\alpha^3}{4\pi} \int_0^{2\pi} d\phi' \frac{\rho_t K_{-1}(-ik_n \rho_t)}{ik_n} \end{aligned} \tag{55}$$

The first term in Eq. (51) can be integrated numerically using the Green's function expression in Eq. (47). For the second term in Eq. (51), the Green's function representation given by Eq. (55) is used.

6. RESULTS

In this section we present example results that illustrate the accuracy of this solution procedure. Although the ultimate test of any model is its ability to predict measurements, uncertainty in environmental parameters makes it difficult to obtain a meaningful comparison. As a substitute, we compare our results with both analytical formulations and with FEM results. Other than the heat transfer equation and the assumed parameter values, these methods have little in common with one another, which makes the comparison informative.

We begin with a calculation of the temperature over a homogeneous half space performed using three different techniques, namely, an analytical model, an integral-equation based model, and an FEM based model. The problem is modeled using a half-space extending to infinity with constant thermal properties. The analytical solution is based the result of Watson,[4] which uses a Fourier series solution of the heat transfer equation.[25] In Fig. 3, the predicted soil surface temperature is given as a function of time. The integral equation solution and the FEM solution are in agreement with the analytical results, which lends confidence to the approximations used in this work for the Green's function. Note that the Green's function is independent of the thermal properties of the mine.

We next assess the ability of the model to describe the temperature over a simulated circular anti-tank mine buried in soil. The simulant mine has a diameter of 20 cm and a height of 7.5 cm. The mine is assumed to be a perfect insulator. Figure 4 shows the surface temperature over the center on the mine as a function time. Again, good agreement is noted, which provides a level of confidence that these disparate methods are both producing correct results.

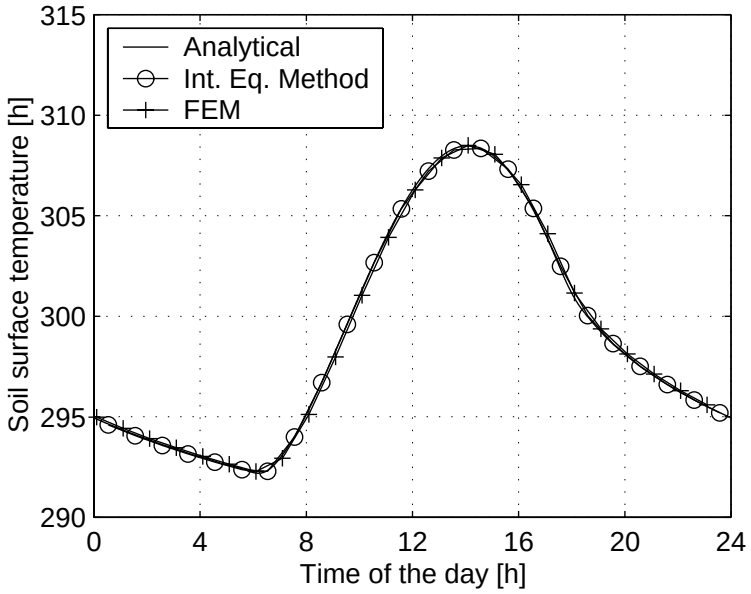


Figure 3. The surface temperature over homogeneous soil as a function of time using an analytical solution, the integral equation solution described herein, and a FEM solution.

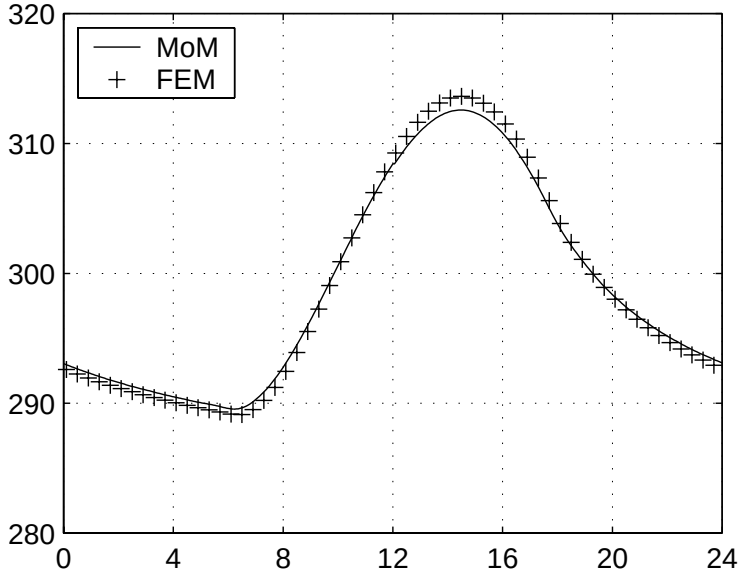


Figure 4. The surface temperature distribution over the center of a cylindrical time using the integral equation solution and a finite element solution.

7. SUMMARY AND CONCLUDING REMARKS

An integral equation is formulated and solved for the temperature over a buried land mine. Periodic (diurnal) solar heating is assumed, which makes possible a Fourier decomposition of the time variation. The integral equation for each Fourier harmonic of the temperature distribution is solved using the method of weighted residuals. The resulting model provides for a relatively simple geometry a reference solution, which can be used to check more sophisticated FEM-based models.

ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

REFERENCES

1. I. K. Sendur and B. A. Baertlein, "Thermal analysis of IR signatures for buried land mines," Progress Report to US Army Research Office on Contract 97-SC-ARO-1015, The Ohio State University ElectroScience Laboratory, Columbus, OH, March 1999.
2. P. Pregowski and W. Swiderski, "Using of (sic) comparison method in IR thermal detection of buried mines," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, eds., *Proceedings of SPIE* **3392**, pp. 1241–1248, 1996.
3. I. K. Sendur and B. A. Baertlein, "Numerical simulation of thermal signatures of buried mines over a diurnal cycle," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **4038**, pp. 156–167, 2000.
4. K. Watson, "Geologic application of thermal infrared images," *Proc. IEEE* **63**(1), pp. 128–137, Jan. 1975.
5. A. W. England, J. F. Galantowicz, and M. S. Schretter, "The radiobrightness thermal inertia measure of soil moisture," *IEEE Trans. Geosci. Remote Sensing* **30**(1), pp. 132–139, Jan. 1992.
6. A. B. Kahle , "A simple thermal model of the earth's surface for geologic mapping by remote sensing," *Journal of Geophysical Research* **82**, pp. 1673–1680, 1977.
7. A. W. England, "Radiobrightness of diurnally heated, freezing soil," *IEEE Trans. Geosci. Remote Sensing* **28**(4), pp. 464–476, July 1990.
8. Y.-A. Liou and A. W. England, "A land-surface process/radiobrightness model with coupled heat and moisture transport in soil," *IEEE Trans. Geosci. Remote Sensing* **36**(1), pp. 273–286, Jan. 1998.
9. Y.-A. Liou and A. W. England, "A land-surface process/radiobrightness model with coupled heat and moisture transport for freezing soils," *IEEE Trans. Geosci. Remote Sensing* **36**(2), pp. 669–677, March 1998.
10. I. Stakgold, *Green's Functions and Boundary Value Problems*, Academic Press, 1980.
11. A. Sommerfeld, *Partial Differential Equation in Physics*, Academic Press, 1949.
12. W. A. Johnson, "Analysis of a vertical, tubular cylinder which penetrates an air-dielectric interface and which is excited by an azimuthally symmetric source," *Radio Science* **18**(6), pp. p. 1273–1281, 1983.
13. W. C. Kuo and K. K. Mei, "Numerical approximations of the Sommerfeld integral for fast convergence," *Radio Science* **13**(3), pp. p. 407–415, 1978.
14. J. H. Richmond, "Digital computer solutions of the rigorous equations for scattering problems," *Proc. IEEE* **53**, pp. p. 796–804, 1965.
15. R. F. Harrington, "Matrix methods for field problems," *Proc. IEEE* **55**(2), pp. p. 136–149, 1967.
16. R. F. Harrington, *Field Computation by Moment Methods*, IEEE Press, 1993.
17. E. K. Miller, L. Medgyesi-Mitschang, and E. H. Newman, Eds., *Computational Electromagnetics*, IEEE Press, 1992.
18. R. C. Hansen, Ed., *Moment Methods in Antennas and Scattering*, Artech, 1990.
19. D. R. Wilton, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak and C. M. Butler, "Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains," *IEEE Trans. Antennas Propagat.* **32**(2), pp. p. 276–281, 1984.
20. D. E. Livesay, and K. M. Chen, "Electromagnetic fields induced inside arbitrarily shaped biological bodies," *IEEE Trans. Microwave Theory Tech.* **22**(12), pp. p. 1273–1280, 1974.
21. S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.* **30**, pp. p. 409–418, 1982.
22. D. H. Schaubert, D. R. Wilton, and A. W. Glisson, "A tetrahedral modeling method for electromagnetic scattering by arbitrarily shaped inhomogeneous dielectric bodies," *IEEE Trans. Antennas Propagat.* **32**(1), pp. p. 77–85, 1984.
23. T. K. Sarkar, "A study of the various methods for computing electromagnetic field utilizing thin wire integral equations," *Radio Science* **18**, pp. p. 29–38, 1983.

24. M. G. Andreasen, "Scattering from bodies of revolution," *IEEE Trans. Antennas Propagat.* **13**(2), pp. p. 303–310, 1965.
25. H. S. Carslaw and J. C. Yaeger, *Conduction of Heat in Soils*, Oxford Univ. Press, New York, NY, 1953.

Algorithms for detection of surface mines in multi-spectral IR and visible imagery

W.-J. Liao, D.-H. Chen and B. A. Baertlein*

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

Algorithms are presented for detecting surface mines using multi-spectral data. The algorithms are demonstrated using visible and MWIR imagery collected at Fort A.P. Hill, VA under a variety of conditions. For imagery with a resolution of a few centimeters there is significant correlation in the clutter. Using a first-order Gauss Markov random field model for the clutter, an efficient pre-whitening filter is proposed. A significant improvement in detection is demonstrated as a result of this whitening. Further improvement in the detection of specific mine types is demonstrated by using a random signal model with a known covariance matrix. That approach leads to an estimator-correlator formulation, in which the random signature estimate is the output of a Wiener filter. It is suggested that by fusing the output of a bank of such filters one could improve detection of all mine types.

Keywords: IR, multi-spectral, signal processing, surface mines

1. INTRODUCTION

Whereas many mine-detecting sensors must operate in close proximity to the target, electro-optical (EO) sensors can be used at significant standoff distances, which makes it possible to detect mines and mine fields from airborne platforms. For surface mines, there is considerable discriminative power in the spectral content of passive EO signatures. These facts have led to the development of several airborne multi-spectral and hyper-spectral detection systems.

In this work we present algorithms for surface mine detection using multi-spectral data. Detection of targets in multi-spectral imagery has been investigated by many groups, and several algorithms have been proposed for this task. Some current EO mine sensors have relatively high resolution (a few centimeters per pixel), which makes it possible to exploit spatial covariance information for targets and clutter. The approaches described herein use statistical detection theory for this purpose. We demonstrate that significant performance improvements are possible with respect to a standard detection algorithm.

This work is organized in five major sections. Prior work in multi-spectral sensors and detection algorithms is reviewed in Section 2. The algorithms in this paper are demonstrated on a particular multi-spectral data set, which we describe in Section 3. We have found that the performance of a standard algorithm can be improved by taking the spatial correlation of the clutter into account. This approach, which is related to the generalized matched filter concept, is described in Section 4. Allowing for a random target signature leads to a Wiener filter approach, which is described in Section 5. Detection results in the form of ROC curves appear in Section 6. Concluding remarks appear in Section 7.

2. PRIOR WORK

A number of systems have been developed to collect and process multi-spectral imagery for mine detection. Early multi-spectral work was reported by Witherspoon and Holloway[1] who fused six channels of imagery collected by a 400-900 nm camera with a spinning filter wheel. That sensor was later used on an airborne platform under the Coastal Battlefield Reconnaissance and Analysis (COBRA) program.[2,3] The REMIDS sensor[4] combined a passive thermal IR channel with two co-registered linearly polarized near-IR sensors of laser reflectance. That combined passive/active sensor concept later became part of the ASTAMIDS system.[5] There has also been extensive work in hyper-spectral imaging for demining. McFee et al. have developed a compact airborne spectrographic imager

Corresponding author: B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

(*casi*),[6–9] which employs up to 288 spectral bands over the range 400-1000 nm. An extensive experimental study of hyper-spectral phenomenology has recently been presented by Smith et al.[10]

Many researchers have explored target detection in multi- and hyper-spectral imagery. A widely referenced approach to this problem is the RX algorithm, named for its authors, Reed and (Xiaoli) Yu.[11] This algorithm, which we describe in more detail in Section 4, has been extended to include spectrally correlated (but spatially uncorrelated) clutter.[12] The spectral covariance matrix is estimated with an adaptive algorithm and used to design a spectral pre-whitening filter. It was found that a priori knowledge of the target spectral distribution often makes little difference in performance.[12] Those authors and their co-workers have proposed several forms of the algorithm,[13–15] and RX has previously been used in the COBRA system.[16] An application of RX to hyper-spectral detection of military and cultural features was described by Stellman et al.[17]

Other approaches to multi-spectral detection have been examined. Schweizer and Moura[18] developed a CFAR algorithm based on a first-order Gauss-Markov random field (GMRF) model for the clutter. In that algorithm a maximum likelihood (ML) technique was used to estimate the clutter parameters, which were then used in a GLRT detector. Several authors have attempted to exploit target spectral characteristics. Ashton[19] used clustering algorithms to find sub-pixel anomalies in multi-spectral IR terrain imagery. Ashton and Schaum[20] used RX to search for anomalies in background-suppressed spectral signatures. A very different approach was explored by Banerji and Goutsias,[21] who used mathematical morphology to detect mines in individual bands followed by fusion of the band information. Correlation among the bands was addressed by using a maximum noise fraction transform to generate independent bands.

3. A MULTI-SPECTRAL DATA COLLECTION

Multi-spectral data was collected by the US Army Night Vision Electronic Systems Directorate (NVESD) at Fort A.P. Hill, VA in May 2000. The data comprise images of clutter and anti-tank mines including TM62, VS1.6 and RAAM laid on the surface and buried. Targets arranged in lanes with several ground covers were imaged including sand, dirt, short grass and tall grass. Registered, multi-spectral images were collected using a visible broadband CCD camera and an MWIR camera. Filters were used to extract specific spectral bands. The visible camera produces 512×768 pixels per frame with a $20^\circ \times 30^\circ$ FOV and a pixel dimension of approximately 1.6 cm. The MWIR camera provides 256×256 pixels per frame with a $22^\circ \times 22^\circ$ FOV. Both cameras were mounted on a lift elevated 60 feet above the ground.

To efficiently explore the response of our algorithms to a variety of ground covers and environmental conditions while keeping the computational burden manageable, we extracted a set of 102 mine and clutter image chips from three background types (sand, short grass and tall grass). Figure 1 shows sample target chips for different environments. The target chips comprise surface-laid VS1.6 and TM62 mines only. Buried mines and RAAM have been ignored in this study. The clutter chips, samples of which are shown in Figure 2, contain homogeneous backgrounds, inhomogeneous backgrounds (without emplaced objects), and emplaced clutter objects that have a mine-like appearance. Those objects included crushed paint cans, pizza boxes, and wooden sticks.

Each image chip comprises six channels of data: five visible/IR bands (blue, green, yellow, red, and NIR) and one MWIR band. The chip size is 46×46 pixels. Because different cameras were involved, the visible/NIR and MWIR chips have different resolutions, which is evident in Figure 1.

It is difficult to gauge the true false alarm rate of an algorithm from chips and, hence, we also processed an extensive set of complete images. Because of concerns about image registration, only the visible/NIR data were used in those tests. The raw imagery contain fiducial markers, which were cropped from the processed data, resulting in images of 260×768 pixels or 4.2 m $\times$ 12.3 m = 51 m^2 area per image. Because the content of the target lanes varied, different numbers of images were collected for each background type. We processed 16 images with sandy soil backgrounds (Lane 7), 10 images with tall grass backgrounds (Lane 5) and 15 images of short grass backgrounds (Lane 3), for a total area of roughly 2050 m^2, which includes a small amount of overlap between adjacent images.

4. CLUTTER WHITENING FILTER

For a target with a known signature in spatially white Gaussian noise, a likelihood ratio test can be formulated. When the target signature is known with the exception of its amplitude, a generalized likelihood ratio test may be

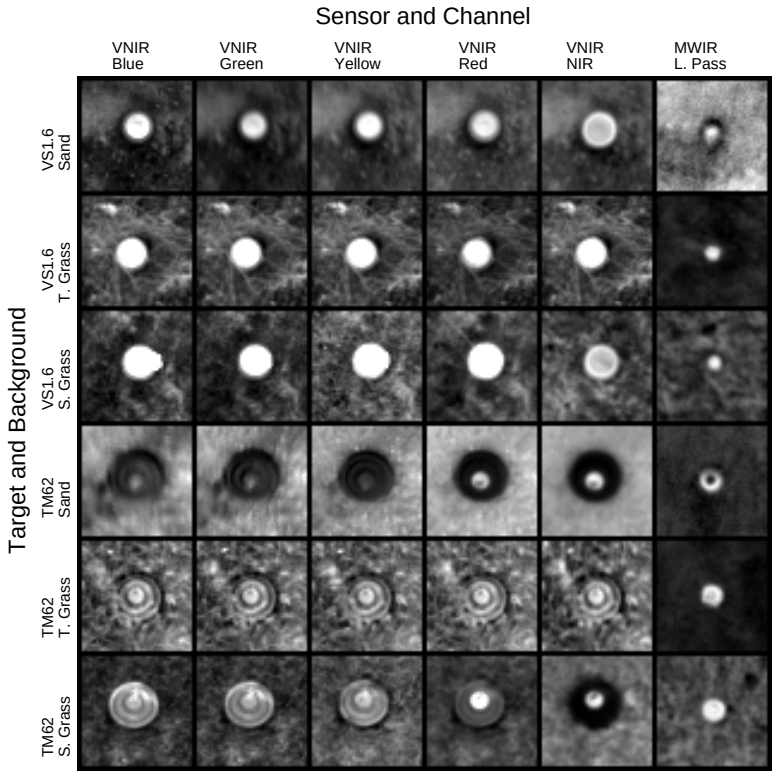


Figure 1. Sample multi-spectral target chips.

constructed in which a maximum likelihood (ML) estimator is used to determine the signature amplitude. For the special case of Gaussian noise, the amplitude estimator is linear.

The RX algorithm embodies this approach, extending the unknown variables to the covariance matrices. A detailed description of the algorithm has appeared,[11] which the reader may consult for more details. That algorithm is a locally adaptive CFAR detector, which assumes spatially uncorrelated Gaussian clutter and a known target spatial signature with unknown spectral distribution and covariance matrices. At each position within the image an estimate of the target amplitude is made. Spectral covariance matrices are then estimated for the clutter and for the residual noise (after the estimated target signature has been subtracted). A local likelihood ratio is estimated from these covariance matrices.

In the original presentation of the RX algorithm[11] it is assumed that the clutter is spatially uncorrelated. In an earlier work it was suggested that the assumption of spatially uncorrelated clutter is valid for low-resolution imagery,[12] but the images described in Section 3 contain significant correlation. Figure 3 shows imagery and power spectra for mines and clutter. It is evident that the clutter is not white and that the spectra overlap. Both of these facts have a strong effect on algorithm performance.

In this paper we apply a clutter whitening filter to the data to reduce spectral overlap between the target and clutter. This work involves a filter that appears in the context of the generalized matched filter, in which both the measured data $\mathbf{X}$ and the known template (signature) $\mathbf{S}$ are pre-whitened. The covariance matrix of the residual (signal-subtracted) clutter depends on the correlation between the data and reference signal, weighted by the inverse covariance matrix $\mathbf{\Sigma}^{-1}$. This quantity becomes

$$\mathbf{S}^T\mathbf{\Sigma}^{-1}\mathbf{X} = (\mathbf{V}\cdot\mathbf{S})^T\cdot(\mathbf{V}\cdot\mathbf{X}) \tag{1}$$

where $\mathbf{V}$ is a factorization of Σ^{-1}

$$\mathbf{\Sigma}^{-1} = \mathbf{V}^T\mathbf{V} \tag{2}$$

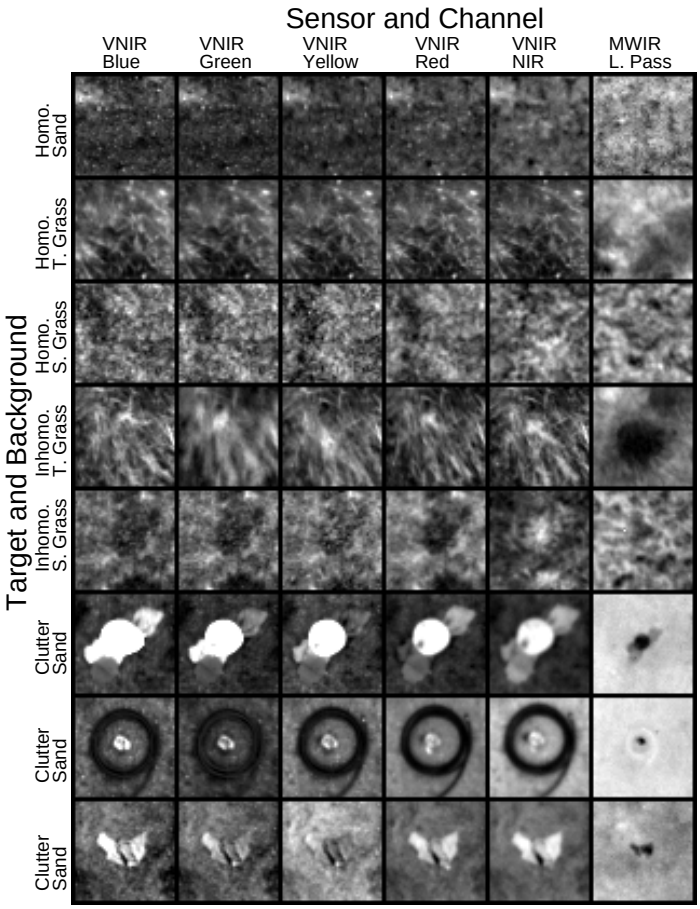


Figure 2. Sample multi-spectral clutter chips

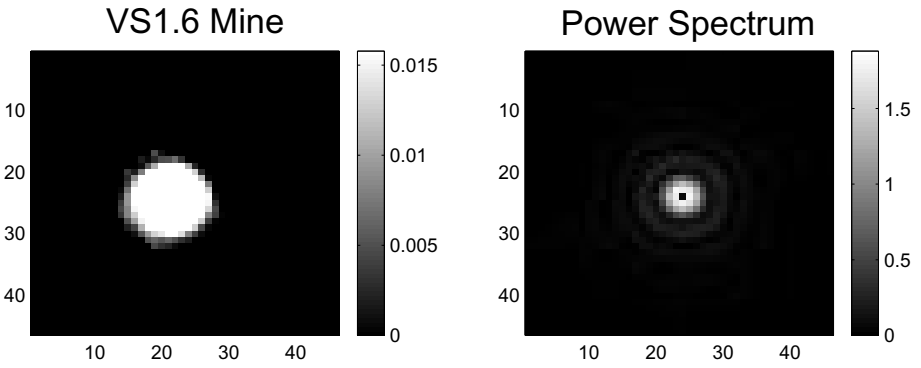
The factorization matrix $\mathbf{V}$ is found from an eigenvector decomposition or a Cholesky factorization of the positive definite matrix $\mathbf{\Sigma}$.

It is difficult to implement this filter because of both computational and practical limitations. For an image chip of size $J \times K$ the dimensions of $\mathbf{\Sigma}$ are $(JK) \times (JK)$. The local covariance matrix must be estimated from a set of clutter chips, the number of which is nominally much larger than JK. For chips of reasonable size, this computation becomes impractical. Furthermore, clutter statistics vary spatially. As a result, it may be impossible to find the number of chips required to estimate $\mathbf{\Sigma}$ within an image region having consistent clutter characteristics.

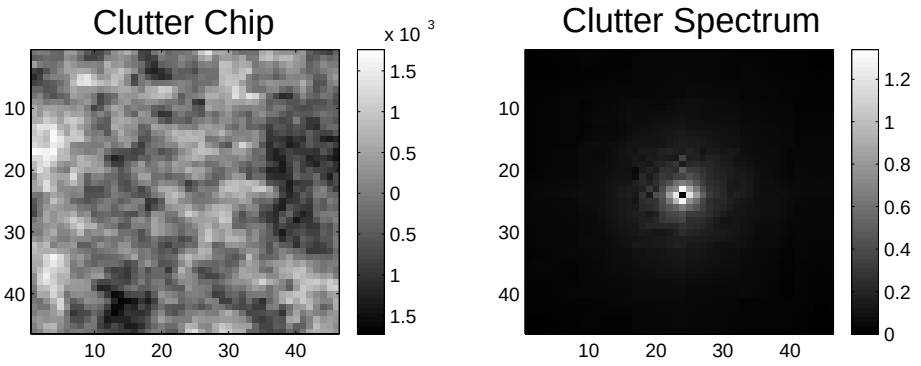
In this work we make the assumption that over the spatial scales of interest in mine detection the clutter is Gaussian and approximately first-order Markov. This leads to an exponential correlation function, the covariance matrix for which can be decomposed analytically. For isotropic clutter having correlation coefficient ρ, one can express the whitening operation as a convolution with a simple 3×3 kernel[22]

$$\begin{bmatrix} \rho^2 & -\rho(1+\rho^2) & \rho^2 \\ -\rho(1+\rho^2) & (1+\rho^2)^2 & -\rho(1+\rho^2) \\ \rho^2 & -\rho(1+\rho^2) & \rho^2 \end{bmatrix} \tag{3}$$

Using a locally derived estimate of ρ, this operator can be used to produce spatially whitened imagery at a modest computational cost. The RX algorithm can then be applied to that data using a whitened reference signal.



(a) Surface mine image and power spectral density.



(b) Clutter image and power spectral density.

Figure 3. Some mine chips and their power spectra.

5. WIENER FILTER PROCESSING

A key limitation of the foregoing algorithm is that the target signature must be known. Real mine signatures vary as a result of changes in viewing geometry, environmental conditions and illumination. To account for these effects we introduce a model in which the signal is also random.

A classical approach to this problem is the estimator-correlator detector,[23] in which a filter is designed to estimate the unknown random signal and then correlated with the data. When the noise and signal are Gaussian distributed, the underlying filter is a Wiener filter, which is most conveniently implemented in the space-spectral domain (k_x, k_y).[24] The measured image in the space-spectral domain $Y(k_x, k_y)$ is modeled by the sum of random noise $N(k_x, k_y)$ and a known target $H(k_x, k_y)$ distorted by a random function $T(k_x, k_y)$

$$Y(k_x, k_y) = H(k_x, k_y)T(k_x, k_y) + N(k_x, k_y) \tag{4}$$

The Wiener filter $W(k_x, k_y)$ is given by

$$W(k_x, k_y) = \frac{\mathcal{E}[T^*(k_x, k_y)]H^*(k_x, k_y)}{|H(k_x, k_y)|^2 S_{TT}(k_x, k_y) + S_{NN}(k_x, k_y)} \tag{5}$$

where S_{NN} and S_{TT} are the power spectral densities of the additive noise and target. Forming the filtered signature WY, it is clear that the Weiner filter rejects noise in frequency bands where the signal-to-noise ratio $|H|^2 S_{TT}/S_{NN}$ is high.

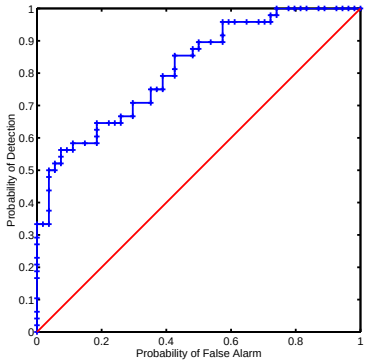
The principal limitation of the Wiener filter is that it requires knowledge of the target shape, captured by the function H and the target distortion spectrum S_{TT}. In real demining imagery, target shapes vary and it is difficult to make a compelling estimate of S_{TT}. Moreover, the amount of data available to estimate S_{TT} is limited. In our implementation we assume that the true target distortion has an isotropic Gaussian shape at the (unknown) spatial location of the mine. The correlation length of T is determined empirically.

Use of the Wiener filter is also motivated by another of its characteristics. For low signal-to-noise ratios, the Wiener filter degenerates to the generalized matched filter described above. For high signal-to-noise ratios, it is easy to see that $WY \approx 1$ and, hence, the Wiener filter is essentially performing deconvolution. In that environment, it whitens the signal, rather than the noise. For high S/N, the filter output will tend to an impulse, which leads to high contrast in the output image.

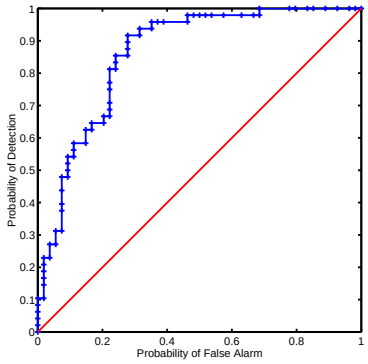
A convenient implementation of this approach in mine detection is to use the Wiener filter as a preprocessor for RX. An estimate of the clutter spectrum is derived using the first-order Markov model described above. The RX reference signal is taken to be an impulsive function with narrow width, consistent with the model for S_{TT}. A small amount of white noise is added to S_{NN} for regularization.

6. EXAMPLE RESULTS

The foregoing algorithms were applied to the multi-spectral image chips described above. We begin by applying the basic RX algorithm and RX with a pre-whitening filter to a single spectral band (blue). The results, shown in Figure 4 suggest that even for the single channel case, pre-whitening offers a performance gain, primarily as a reduction in P_{fa} for fixed P_d.



(a) The baseline RX algorithm.



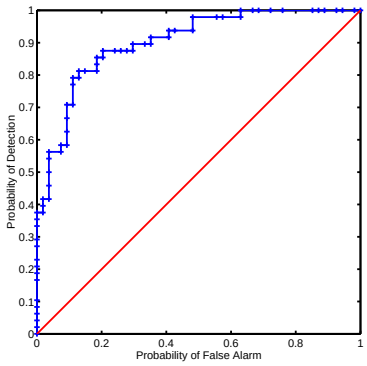
(b) The RX algorithm with clutter whitening.

Figure 4. ROC curves for the RX algorithm on single-band imagery.

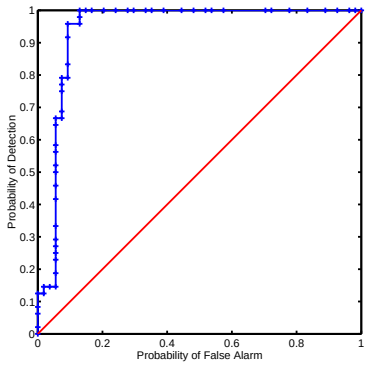
ROC curves for multi-spectral imagery are shown in Figure 5. Prewhitening significantly increases P_d for all but the smallest false alarm rates. At the lowest P_{fa} levels, the small loss in signal intensity caused by whitening is sufficient to cause some weak targets to be missed. In general, however, there is an improvement in false alarm rate for a fixed level of P_d. The improvement is most pronounced at the highest levels. For $P_d = 0.9$ prewhitening decreases P_{fa} roughly four-fold. Finally, a comparison of Figures 4 and 5 makes clear the performance benefit offered by the additional spectral channels.

To better gauge the false alarm rate of these algorithms we computed their performance over image sequences from the Fort A.P. Hill data set. We used the image sets described in Section 3, which comprise an area of roughly 2050 m^2. The performance of the algorithms over these data are shown in the ROC curves of Figure 6. We see that clutter whitening improves performance, which confirms the results obtained earlier for image chips.

The performance of the Wiener filter was first assessed using the image chips. As the ideal signature H we used the mean of several mine signatures. The target distortion spectrum estimate S_{TT} was approximated by a Gaussian



(a) The baseline RX algorithm.



(b) The RX algorithm with clutter whitening.

Figure 5. ROC curves for the RX algorithm applied to multi-spectral imagery.

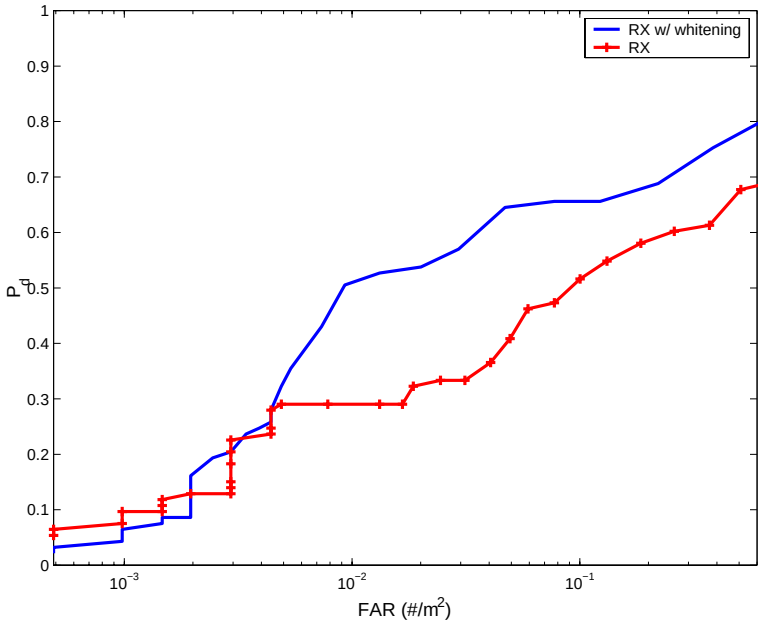
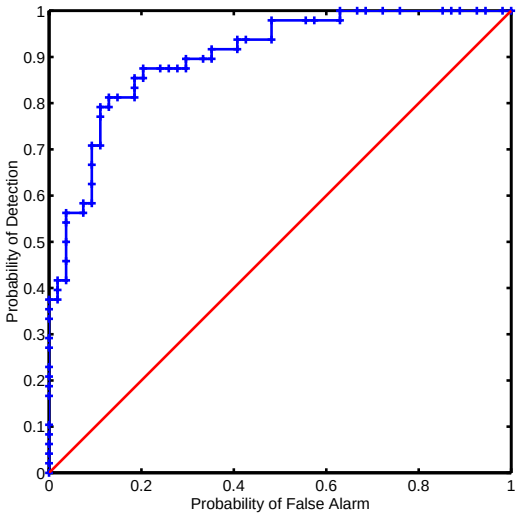


Figure 6. The performance of the RX algorithm without and with clutter whitening for 2050 m^2 of imagery over a variety of backgrounds. Both VS1.6 and TM62 mines are present and detected by the algorithm.

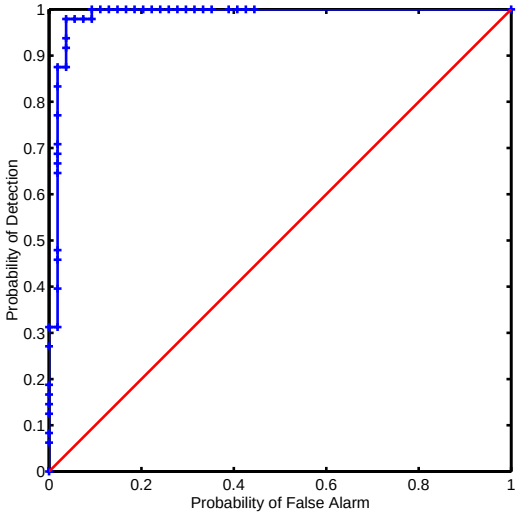
function with wide bandwidth. The clutter spectrum S_{NN} was estimated from measured clutter and approximated by an exponential with the same decorrelation length. The ROC curve for this algorithm is shown in Figure 7. The performance benefit from this approach is clear from this figure.

As seen in the chips in Figure 1, the signatures produced by VS1.6 and TM62 mines are significantly different. The VS1.6 mines tend to have a consistent positive contrast across different backgrounds, while the TM62 mine can show both positive and negative contrast. As a result, the reference signature used in the RX algorithm has a strong effect on performance. A similar sensitivity to target shape was noted in tests of the Wiener filter.

It is feasible to exploit this sensitivity to improve performance. A bank of pre-processing filters, each designed for a particular mine shape, can be applied individually and the results fused at the decision level using an "OR"



(a) The baseline RX algorithm.



(b) RX with Wiener filter preprocessing.

Figure 7. ROC curves for the RX algorithm applied to multi-spectral image chips containing both VS1.6 and TM62 mines.

algorithm. An indication of the performance achievable with that concept appears in Figure 8, which shows the ROC curve computed on the large image set for the VS1.6 mines. In that figure we used a reference signature that better matches the VS1.6 mines. For the Wiener filter, the ideal signature H was taken to be a VS1.6 signature, while the reference signal in the subsequent RX algorithm was taken to be impulsive. The ROC curves were computed for VS1.6 mines only — TM62 mines were not counted as detections or false targets. At FAR levels of 10^{-3}, the probability of detection is still quite large. Because that FAR level corresponds to only one false alarm in the entire 2050 m^2 of imagery, it was not possible to assess the performance at lower FAR levels.

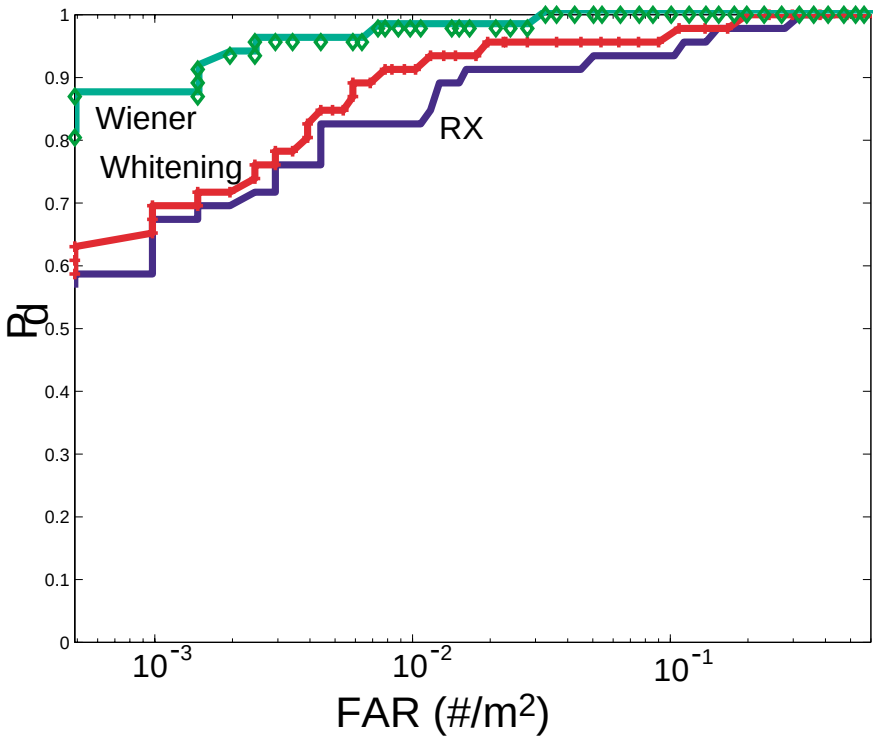


Figure 8. The performance of the RX algorithm with clutter whitening and Wiener filtering for 2050 m^2 of imagery having various backgrounds. The detector is optimized for VS1.6 mines, and only those mines were counted as detections. Other detected mines were not counted as false alarms.

Finally, in Figure 9 we show the result of using the filter bank concept for multiple mine types. Individual signatures H for VS1.6 and TM62 mines were used to separately process the aforementioned imagery, and the results were fused. In this work fusion was accomplished at the soft-decision level by scaling the likelihood ratios to a common range and using the $\max(\cdot)$ operation to produce a fused likelihood ratio. We see that in this case the Wiener filter approach provides superior performance, which suggests that a filter bank approach may be attractive.

7. CONCLUDING REMARKS

Algorithms for surface mine detection in multi-spectral imagery have been developed and demonstrated. Our algorithms exploit primarily the spatial information in the signature. We used a whitening filter to eliminate spatial correlation in the imagery. A first-order Gauss-Markov random field clutter assumption leads to a computationally efficient method for whitening prior to processing with other algorithms. Using a diverse set of image chips, which contain a variety of clutter types and backgrounds, it was shown that pre-whitening can greatly improve the performance of the widely used RX algorithm.

The aforementioned algorithms assume a known target signature. The potential benefits of assuming a random target signature were explored by using an estimator-correlator approach. The processing in that case is equivalent to a Wiener filter, which we implemented efficiently in the spatial-frequency domain. The Wiener filter is capable of additional performance gains, but it is sensitive to assumptions about the target signature.

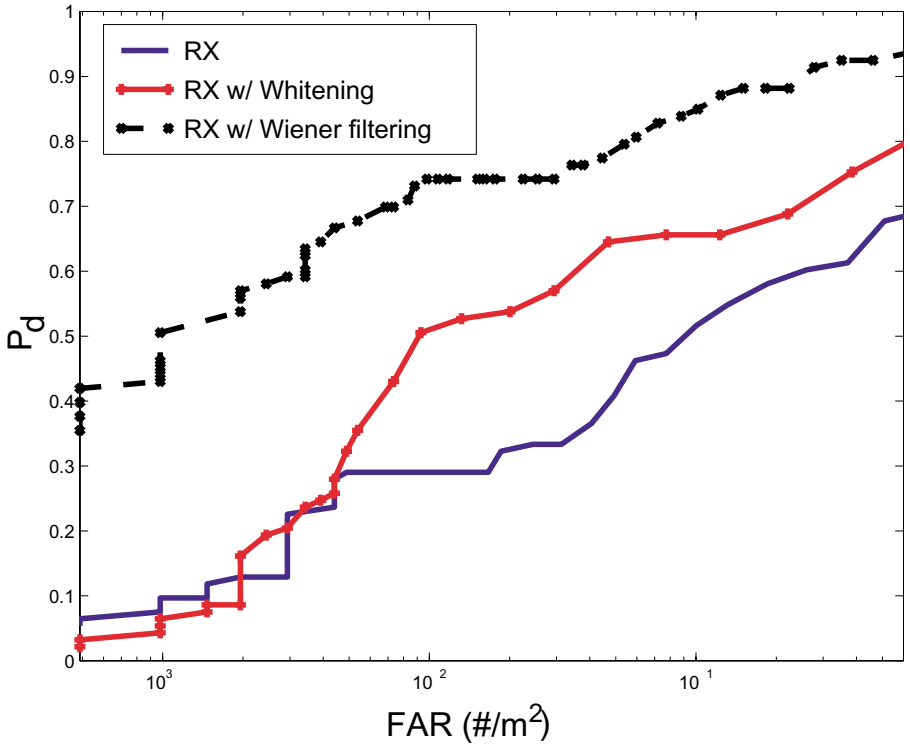


Figure 9. A comparison of the RX algorithm alone, with clutter whitening and with a bank of Wiener filters for 2050 m^2 of imagery having various backgrounds. For the Wiener-filtered result, detectors optimized for both VS1.6 and TM62 mines were fused.

To better assess the false alarm rate and robustness of the algorithms, they were also used to process a larger quantity of imagery (approximately 2050 m^2). The performance gains demonstrated for image chips were confirmed in processing the larger images, but additional gains were obtained when the algorithms were tuned for a specific mine type. This result suggests that the most effective means of detecting mines is to use a bank of filters tuned to specific mine shapes and to fuse the results in post-processing.

8. ACKNOWLEDGMENTS

This project was supported in part by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

REFERENCES

1. N. H. Witherspoon and J. J. H. Holloway, "Video based multispectral detection of land mines, a technology applicable for use in law enforcement," in *Surveillance Technologies II*, *Proceedings of SPIE* **2496**, pp. 185–194, 1992.
2. N. H. Witherspoon, J. J. H. Holloway, K. S. Davis, R. W. Miller, and A. C. Dubey, "The Coastal Battlefield Reconnaissance and Analysis (COBRA) program for minefield detection," in *Detection Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, eds., *Proceedings of SPIE* **2496**, pp. 500–508, April 1995.
3. R. R. Muise, J. A. Wright, and Q. A. Holmes, "Coastal mine detection using the COBRA multispectral sensor," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, R. L. Barnard, C. J. Lowe, and J. E. McFee, eds., *Proceedings of SPIE* **2765**, pp. 15–24, April 1996.

4. J. H. Ballard, R. M. Castellane, B. H. Miles, and K. G. Wesolowicz, "The Remote Minefield Detection System (REMIDS) II Major components and operation," Tech. Report EL-92-30, NTIS AD-B175 759, US Army, Waterways Experiment Station, Corps of Engineers, 1992.
5. G. Maksymonko and K. Breiter, "ASTAMIDS minefield detection performance at Aberdeen Proving Ground test site," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, A. C. Dubey and R. L. Barnard, eds., *Proceedings of SPIE* **3079**, pp. 726–737, April 1997.
6. S. B. Achal, J. E. McFee, and C. D. Anger, "Identification of surface-laid mines by classification of compact airborne spectrographic imager (CASI) reflectance spectra," in *Detection Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, eds., *Proceedings of SPIE* **2496**, pp. 324–335, April 1995.
7. J. E. McFee, H. T. Ripley, R. Buxton, and A. M. Thriscutt, "Preliminary study of detection of buried landmines using a programmable hyperspectral imager," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, R. L. Barnard, C. J. Lowe, and J. E. McFee, eds., *Proceedings of SPIE* **2765**, pp. 476–488, April 1996.
8. J. E. McFee and H. T. Ripley, "Detection of buried landmines using a *casi* hyperspectral imager," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, A. C. Dubey and R. L. Barnard, eds., *Proceedings of SPIE* **3079**, pp. 738–749, April 1997.
9. S. B. Achal, C. D. Anger, J. E. McFee, and R. W. Herring, "Detection of surface-laid mine fields in VNIR hyperspectral high spatial resolution data," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **3710**, pp. 808–818, April 1999.
10. A. M. Smith, A. C. Kenton, R. Horvath, L. S. Nooden, J. Michael, J. A. Wright, J. Mars, J. K. Crowley, M. Sullivan, S. Causey, D. Lee, M. Williams, and K. Montavon, "Hyperspectral mine detection phenomenology program," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., *Proceedings of SPIE* **3710**, pp. 819–829, April 1999.
11. I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoustics, Speech and Signal Proc.* **38**, pp. 1760–1770, October 1990.
12. X. Yu and I. S. Reed and A. D. Stocker, "Comparative performance analysis of adaptive multiplespectral detectors," *IEEE Trans. Acoustics, Speech and Signal Proc.* **41**, pp. 2639–2654, August 1993.
13. J. Y. Chen and I. S. Reed, "A detection algorithm for optical targets in clutter," *IEEE Trans. Aerosp. Electron. Syst.* **AES-23**, pp. 46–59, February 1987.
14. A. Margalit, I. S. Reed, and R. M. Gagliardi, "Adaptive optical target detection using correlated images," *IEEE Trans. Aerosp. Electron. Syst.* **AES-21**, pp. 394–405, June 1985.
15. X. Yu, L. E. Hoff, I. S. Reed, A. M. Chen, and L. B. Stotts, "Automatic target detection and recognition in multiband imagery: A unified ML detection and estimation approach," *IEEE Trans. Image Processing* **6**, pp. 143–156, January 1997.
16. Q. A. Holmes, C. R. Schwartz, J. H. Seldin, J. A. Wright, and L. W. Witter, "Adaptive multispectral CFAR detection of land mines," in *Detection Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, eds., *Proc. SPIE* **2496**, pp. 421–432, 1995.
17. C. M. Stellman, G. G. Hazel, F. Bucholtz, J. V. Michalowsicz, A. Stocker, and W. Schaaf, "Real-time hyperspectral detection and cuing," *Opt. Eng.* **39**, pp. 1928–1935, July 2000.
18. S. M. Schweizer and J. M. F. Moura, "Hyperspectral imagery: Clutter adaptation in anomaly detection," *IEEE Trans. Information Theory* **46**, pp. 1855–1871, August 2000.
19. E. A. Ashton, "Detection of subpixel anomalies in multispectral infrared imagery using an adaptive Bayesian classifier," *IEEE Trans. Geoscience and Remote Sensing* **36**, pp. 506–517, March 1998.
20. E. A. Ashton and A. Schaum, "Algorithms for the detection of sub-pixel targets in multispectral imagery," *Photogrammetric Engineering and Remote Sensing* **64**, pp. 723–731, July 1998.
21. A. Banerji and J. Goutsias, "A morphological approach to automatic mine detection problems," *IEEE Trans. Aerospace and Electronic Systems* **34**, pp. 1085–1096, October 1998.
22. A. Arcess, P. H. Mengaert, and E. W. Trombini, "Image detection through bipolar correlation," *IEEE Trans. Information Theory* **IT-16**, pp. 534–541, September 1970.
23. S. M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice Hall, Upper Saddle River, NJ, 1998.
24. W. K. Pratt, *Digital Image Processing*, Wiley-Interscience, New York, NY, 1978.

Improving detection of buried land mines through sensor fusion

Brian A. Baertlein and Ajith Gunatilaka

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

A sensor-fused system has been developed for detection of buried land mines. The system uses a ground-penetrating radar, an infrared camera, and an electromagnetic induction sensor. In the current implementation each sensor is used independently, and fusion is performed during post-processing. We briefly describe the sensors and a data collection involving buried mine surrogates. Algorithms for preprocessing and feature extraction are reviewed. To deal with non-coincident sampling we have developed a new feature-level fusion algorithm, which does not require detection and subsequent association of putative targets. Results are presented for fusion of simulated data and data measured in an outdoor surrogate minefield.

Keywords: Sensor fusion, land mines, ground-penetrating radar, infrared imaging, electromagnetic induction, metal detectors

1. INTRODUCTION

The potential benefits of sensor fusion for demining have prompted several groups to investigate this subject. Among the attributes claimed for sensor fusion are improved detection rates, decreased false alarm rates, and greater robustness in the face of instrument failure and diverse environmental conditions.

In this paper we describe the development of a sensor-fused system for detection of buried mines. Our system currently employs three sensors, which are operated independently on separate platforms. Fusion is performed during post-processing. A quantity of training data has been acquired using these sensors at our facility, where we have buried a number of mine surrogates.

Different sensor technologies tend to produce data which vary in their sampling of the area (e.g., point sensors versus imaging sensors), leading to non-coincident measurement locations. Differences in sampling locations are exacerbated when the sensors reside on different platforms. One method of dealing with non-coincident sampling is association, whereby putative detections from individual sensors are combined. In sensing buried mines, however, detection by individual sensors can be challenging and false alarm rates are typically high, making correct association difficult.

We have developed a novel technique for fusion which avoids the need for preliminary detection by individual sensors and is therefore better suited to irregularly sampled data. The algorithm is described herein, and results are presented for simulated sensor outputs and measured data.

This work is organized in five major sections. In Section 2 we describe the sensors used in the system. Data collection activities are reviewed in Section 3. Preprocessing algorithms required to prepare the data for fusion are given in Section 4. The fusion algorithm is described in Section 5. In Section 6 we present results of the fusion algorithm for a suite of three sensors.

2. SENSORS

To facilitate tests of our fusion concepts under a variety of conditions, a suite of sensors was acquired. Sensors used in the present system are a ground penetrating radar (GPR), a commercial infrared (IR) camera, and a Schiebel electromagnetic induction (EMI) sensor. Brief descriptions of these devices are presented below.

BAB (correspondence), voice (614) 292-0076, fax (614)-292-7297, E-mail: baertlein.1@osu.edu

An abbreviated version of this paper will appear in *Detection and Remediation Technologies for Mines and Mine-Like Objects III*, A. C. Dubey, J. F. Harvey, J. Broach (eds), SPIE Proceedings 3392, AeroSense98, Orlando, FL, April 1998.

2.1. Ground Penetrating Radar

The GPR used in this effort was developed at The Ohio State University (OSU) ElectroScience Laboratory (ESL). It employs a single offset-fed, down-looking focussed parabolic reflector mounted on a wheeled platform. The antenna beam, which has a radius of 5 to 28 inches at the ground depending on frequency, is scanned in an arc over a distance of 38 inches. Complex (in-phase and quadrature) data are collected at 48 points along this arc. The transmitted waveform is frequency-stepped CW over the range 1-6 GHz in 100 MHz steps. The antenna is positioned about four feet off the ground, which avoids direct antenna-ground coupling.

2.2. Electromagnetic Inductions Sensor

We use the Schiebel AN-19/2 metal detector, a pulsed induction sensor now deployed by the US military. This sensor consists of a search head with concentric circular transmit (inner) and receive (outer) coils. Approximate diameters of the coils are 18.5 cm and 28.5 cm respectively. The presence of a metal object is indicated by an audible tone of 1.3 kHz in the user's earphone. As the search head approaches the metal object the amplitude of the tone and the harmonic content increase. Performance data quoted by the manufacturer are: detection of mines with very small (0.15 g) metal content at 10 cm, and detection of a typical anti-tank mine at 50 cm. The unit is supplied with a 0.15 g test piece, a steel pin of diameter 1.5 mm by 10 mm long (comparable to a Chinese Type 72 mine), which can be used to adjust its sensitivity. In addition to the standard audio tone, we have configured the instrument to provide a time-domain waveform output directly from the sensor head. This waveform includes the decay curve of the target, which (in principle) can be related to the shape and metal composition of the target.

2.3. Infrared Camera

The IR camera is an IRRIS 160ST produced by Cincinnati Electronics. This sensor has a InSb focal plane array of 160 by 120 12-bit pixels. It operates in the regime 2.2-4.6 μm, and it has a noise equivalent temperature difference (NEΔT) of 0.025 K. The instantaneous field of view (IFOV) with the current optics is 1 mrad. The camera can be interfaced directly to a personal computer, with the computer performing camera control and image acquisition for long-term studies or for sampling rapid transient events at rates up to 160 frames per second.

3. DATA COLLECTIONS

Sensor data are required for training and testing a sensor fusion system. Using the sensors described above, data sets have been collected and are now being analyzed.

3.1. Facility

To better understand data artifacts and to explore "what-if" sensor questions, it is convenient to have ready access to a test mine field. For this reason a number of mine surrogates were buried in a level grassy field near the ESL facility. The buried surrogates are arranged in a 4-by-10 target grid as shown in Figure 1. Descriptions of the objects corresponding to the abbreviations in Figure 1 are given in Table 1. The objects include aluminum soft-drink cans at the corners, which are fiducial markers for the GPR and EMI sensors. White, reflective surface fiducials are also used when IR data is being collected.

In addition to the fiducials, the grid includes one mine surrogate of high metal content, four surrogates of low metal content, ten non-metal surrogates, a few man-made clutter features (crushed soda cans and "pop-tops"), and natural inhomogeneities (stones and voids).

3.2. Data Acquisition

Data were acquired over the mine grid for each of the three sensors. Sampling for each sensor is different, as is the format of the output data. Figure 2 illustrates the relation of the grid and the acquired samples.

The GPR platform was moved through the mine grid from East to West along each row of the grid. Data sets (arcs of 48 samples) were acquired every three inches starting from a point some distance to the east of the grid and ending about 15 inches to the west of the grid. Calibration data from a sphere placed above a microwave absorber background were also acquired for compensation of the system impulse response, but such compensation has typically not been required. Because of problems in precisely positioning a wheeled platform on uneven terrain, there is some error in the location of these points. In addition, a number of scans were lost when a data transfer failed without warning.

Figure 1. The layout of the mine test grid. The key marked near each object is interpreted as the Abbreviation/Depth in inches.

Table 1. Description of the objects buried in the mine test grid

Abbreviation	**Object**	**Diameter** (inches)	**Height** (inches)
CC	Soda can		
CX	Crushed soda can		
PT	Pop top		
M35	Aluminum disk	3.5	5/8
P15	Plexiglass disk	1.5	2
R4	Rock	~ 4	~ 2
V4	Void(styrofoam)	~ 4	~ 2
H4	Refilled hole	~ 6	~ 2
N35	Nylon disk	3.5	5/8
N3D	Nylon disk with .196 dia .078 thick copper disk	3	1
T3P	Teflon disk with 0.078 dia 1 long stainless steel pin	3	1
T7	Teflon disk	7	1
N7	Nylon disk	7	1

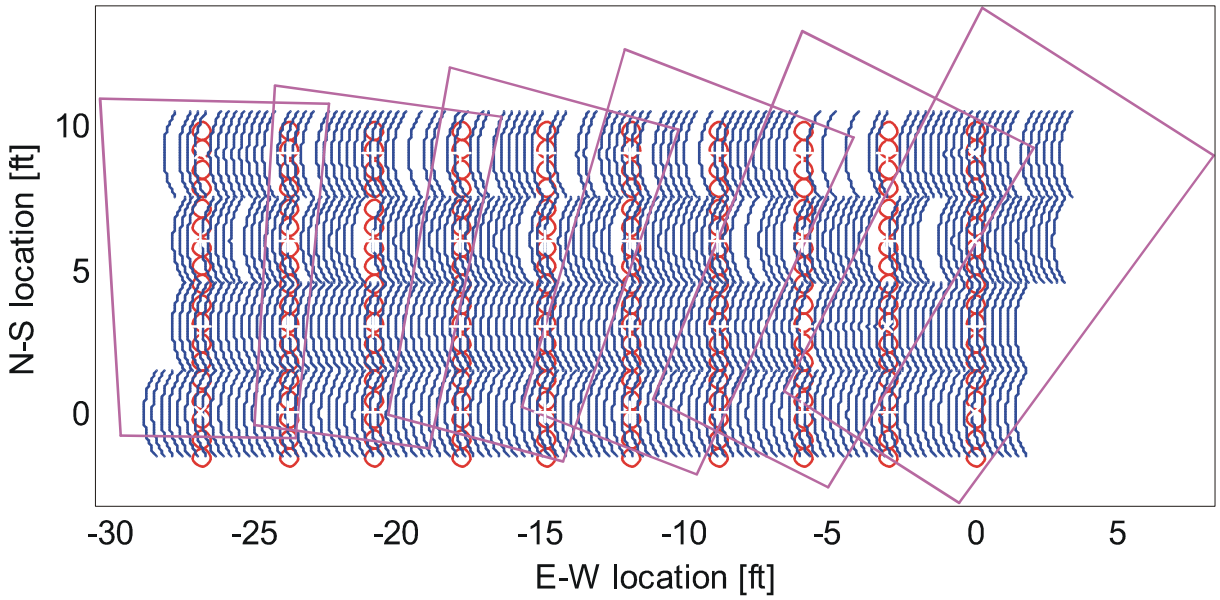


Figure 2. Sample locations for OSU data collection. The vertical columns of circles are EMI sample points (decimated by a factor of four for clarity), the rows of arcs indicate GPR scan locations, and the overlapping quadrilaterals define the boundaries of IR images. Mine-like and clutter-like targets are buried on a uniform grid of size 3 feet in each dimension (see Figure 1).

For the EMI sensor, measurements were performed by scanning South to North over each column of the grid. The metal detector was mounted on a nonmetallic cart with the search head placed about 1.5 inches above the ground. The sensitivity of the metal detector was adjusted to give no tone when the search head was located over a non target region but to produce an audible tone when the test piece (described in Section 2) was about 2 inches from the loop. The cart was placed on a pair of wooden rails. Each scan was started with the search head positioned at 18 inches prior to the start of the grid. The cart was advanced in 2 inch steps and at each location the time waveform and audio tone were recorded. To improve the signal to noise ratio, a digital oscilloscope was used to average 100 waveforms at each position. Data was collected until the search head was 16 inches north of the northern-most edge of the grid.

IR data collection was performed with the camera located atop the two-story ESL building. From this location the mine grid is larger than the field of view of the camera. A sequence of six photos were acquired across the grid to obtain complete coverage and image redundancy. Sequences of photos were acquired at 15 minute intervals throughout the day.

4. PRE-PROCESSING AND FEATURE EXTRACTION ALGORITHMS

To prepare the sensor data for fusion, a substantial amount of processing is required. For most sensors, the computational resources required for this processing exceeds that required for fusion.

4.1. GPR

As noted previously, stepped-frequency scans are acquired at 48 points located along arcs of 38 inch length. These data, which comprise complex samples at equally spaced frequencies, are first processed to subtract the system background response. The data are then windowed and inverse Fourier transformed to the time domain. A software range gate is applied to isolate the time interval containing the desired signal.

A major problem in most GPR data is the ground reflection, which typically dominates the received signal. We employ an iterative technique to eliminate this response. The onset time and duration of the ground reflection are estimated from low-pass filtered down-range (depth) profiles. The time-domain impulse response of the system is

estimated and then iteratively subtracted from the data at points within the ground reflection window. The process is effective, but care must be taken that near-surface targets are not also removed in the process.

Features used for the GPR include the cumulative energy (after suppression of the ground reflection) and the late-time spectra both as a function of sample position. The presence of multiple reflections between the mine and the ground surface or between the top and bottom surface of dielectric mines is evident in some data and appears in the late-time spectra.

4.2. EMI

The EMI sensor can output both time-domain waveforms directly from the sensor head and an internally processed signal that is provided to the user as an audio tone. The time-domain data is potentially more informative, and we have concentrated our initial work on that output. The usefulness of that data, however, has not met our expectations, and we are reconsidering our approach.

The principal feature derived from the time-domain data is a pointwise estimate of the waveform decay rate from the background-subtracted waveform. Several direct and indirect estimates of this quantity can be formed, including (1) the mean and standard deviation of each trace, (2) an integral of the decay profile with respect to time, (3) the value of the last point in the sample window, and (4) a multiple window method, in which ratios of the mean signal in three temporal windows are computed. After examining the output of each of these methods we found them to be highly correlated, with little difference in classifier performance.

Because of slow naturally occurring variations in the sensor background, it is useful to process these point feature estimates along a sensor track. Such processing helps to remove the background trends and it increases signal detectability. A similar technique has been used[1] to improve the performance of another EMI sensor. The algorithm was implemented using a high-pass filter with a cutoff frequency selected to enhance detectability of the smallest objects present.

4.3. IR

The IR data requires fairly extensive processing. After converting the camera data from the vendor's proprietary data format to temperature values, we perform a perspective remapping to transform the data to ground coordinates. This transformation is required to locate ground positions for subsequent fusion. The images do not cover the entire mine grid, and we currently process them individually rather than form a single full-scene mosaic.

IR images of natural scenes contain both fine-scale emissivity variations and large scale variations in illumination and scene content. To mitigate these phenomena we have employed a wavelet-based algorithm comparable to a spatial bandpass filter. The procedure involves forming the 2-D discrete wavelet transform of the image and dropping coefficients at spatial scales outside of the target band. This type of nonlinear editing of wavelet coefficients for noise suppression or image compression has been extensively developed in the literature[2].[3] The use of wavelets for detecting changes in trends is a closely related function and has also received attention.[4] Recently, both techniques were used in a demining context.[5]

The final processing step is feature extraction. The process begins with the detection of mine-like regions via a suitable filter. The filter is formed from a uniform circular disk with a concentric negative outer ring. The inner ring approximates a matched filter, and the outer ring imposes a penalty for non-circular shapes. From the filter output we identify regions of high and low temperature extremes. Thresholding these data produces a binary (segmented) image, which is the basis for further feature extraction. The resulting binary image contains "blobs" which comprise both true and false targets. Some of these blobs are the result of edge effects in the filtering operations described above. We reduce the number of false alarms by eliminating small blobs and by rejecting blobs on the edges. Within each of the segmented regions we compute the following characteristics, which comprise its feature vector: (1) blob area, (2) blob perimeter pixel count (an approximation to perimeter length), (3) bounding rectangle dimensions, (4) centroid location, (5) mean temperature offset in the original image, and (6) variance (a simple texture estimator).

5. FUSION ALGORITHM

Sensor position data are important for sensor fusion, since to combine data one must be reasonably confident that they represent views of the same physical location. As a result, position-related issues have a major role in determining the success of fusion. In particular, we note the following problems:

First, most countermine sensors produce ambiguous target position estimates in one or more dimensions. EO sensors sample points on the surface but are ambiguous in depth. Other sensors are sampled along tracks, typically along (reasonably) parallel lines at nearly equal intervals. GPR sensors typically sample depth and one along-track dimension. EMI sensors sample only in the along track-dimension.

Second, as noted previously, the locations of sample points are often not coincident. The sensors involved may be deployed on different platforms, or it may be impractical (or simply inconvenient) to sample all sensors at the same spatial location.

Finally, errors in positioning often arise when moving platforms are used. Such errors occur even in carefully conducted field experiments, and fusion algorithms that tolerate these errors are required.

In this section we present a new technique for performing sensor fusion in the presence of these position-related problems. The technique uses probabilistic models of sensor signatures to estimate the probability of various target hypotheses at locations away from the measured data. A modest capability for data extrapolation to points beyond the measurements is also inherent.

5.1. Fusion of Non-Coincident Samples Without Detection

Suppose that N_S countermine sensors are used to acquire data in a region containing mines, and that sensor i acquires J_i data samples $\mathbf{d}_{ij}$ at points $\mathbf{R}_{ij}$. These samples may comprise scalar measurements, vector measurements or image data. In general two sensors i and i' will produce samples $\mathbf{d}_{ij}$ and $\mathbf{d}_{i'j}$ that are different in number, dimensionality (i.e., scalar versus vector measurements), and sample positions $\mathbf{R}_{ij}$ and $\mathbf{R}_{i'j}$. For each sensor i and sample position $\mathbf{R}_{ij}$ we extract feature vectors $\mathbf{X}_i(\mathbf{R}_{ij})$, $j = 1, 2, \ldots, J_i$. (In some cases the features may be derived from a sequence of data records around each point $\mathbf{R}_{ij}$.) Let the set of all features acquired in the area be given by

$$\Omega = \{\mathbf{X}_i(\mathbf{R}_{ij}), i = 1, \ldots, N_s; j = 1, \ldots, J_i\} \tag{1}$$

Let $\mathbf{R}$ be any point at which we wish to determine the presence of a mine. This point need not be a sample point or even within the sample region. We form K hypotheses $H_k(\mathbf{R})$ regarding the presence or absence of various types of mines and clutter at $\mathbf{R}$. As usual, the set of hypotheses $\{H_k(\mathbf{R}), k = 1, 2, \ldots, K\}$ must be complete, i.e., it must encompass all possible decisions regarding the presence of a mine.

Our goal is to evaluate the a posteriori probabilities $\Pr(H_k(\mathbf{R})|\Omega)$ when the features Ω were not necessarily acquired at the point $\mathbf{R}$. Without loss of generality we cam define our coordinate system such that $\mathbf{R} = 0$. To simplify the notation in what follows we define the composite feature vector at some point $\mathbf{R}'$

$$\mathcal{X}(\mathbf{R}') = \begin{bmatrix} \mathbf{X}_1(\mathbf{R}') \\ \mathbf{X}_2(\mathbf{R}') \\ \vdots \\ \mathbf{X}_{N_s}(\mathbf{R}') \end{bmatrix} \tag{2}$$

$$\mathcal{X}_0 = \mathcal{X}(0) \tag{3}$$

We will assume that a classifier has been trained using data acquired over known mines to produce the densities

$$f_{\mathcal{X}_0|H_k(0)}(\mathcal{X}_0|H_k(0)) = f_{\mathcal{X}_0|H_k(0)}(\mathbf{X}_1(0), \mathbf{X}_2(0), \ldots, \mathbf{X}_{N_s}(0)|H_k(0)) \tag{4}$$

where we use the notation $f_{\mathcal{X}}(\mathcal{X}')$ for the probability density function of random variable $\mathcal{X}$ evaluated at $\mathcal{X}'$.

We can relate $f_{\mathcal{X}_0|H_k(0)}$ to the available data Ω as follows: Elementary properties of conditional densities permit us to write

$$\Pr(H_k(0)|\Omega) = \int d\mathcal{X}_0 \Pr(H_k(0)|\mathcal{X}_0, \Omega) f_{\mathcal{X}_0|\Omega}(\mathcal{X}_0|\Omega) \tag{5}$$

For each sensor i there exists a distance Δ_i such that for all mines of interest the sensor response is insignificant outside of this radius. For each point $\mathbf{R}$ we can then define a local collection of feature vectors $\mathcal{Y}$ that are significant in a neighborhood of $\mathbf{R}$. We write

$$\mathcal{Y} = \{\mathbf{X}_{ij_m}, i = 1, 2, \ldots, N_s;\ m = 1, 2, \ldots, M\} \tag{6}$$

where the subsequence $j_m,\ m = 1, 2, \ldots, M$ defines the points $\mathbf{R}_{ij_m}$ that lie close to $\mathbf{R}$. Since only these elements of Ω contribute to our estimate of $\Pr(H_k(0)|\Omega)$ we have

$$\Pr(H_k(0)|\Omega) = \Pr(H_k(0)|\mathcal{Y}) = \int d\mathcal{X}_0 \Pr(H_k(0)|\mathcal{Y}, \mathcal{X}_0) f_{\mathcal{X}_0|\mathcal{Y}}(\mathcal{X}_0|\mathcal{Y}) \tag{7}$$

Consider next the integrand in equation (7). The data $\mathcal{Y}$ are most consistent with a value of $\mathcal{X}_0$ that (hopefully) corresponds to the true feature vector at $\mathbf{R} = 0$. For this value the conditional density $f_{\mathcal{X}_0|\mathcal{Y}}$ has a maximum, which is the maximum a posteriori (MAP) estimate of $\mathcal{X}_0$. At such points we have

$$\Pr(H_k(0)|\mathcal{Y}, \mathcal{X}_0) f_{\mathcal{X}_0|\mathcal{Y}}(\mathcal{X}_0|\mathcal{Y}) \approx \Pr(H_k(0)|\mathcal{X}_0) f_{\mathcal{X}_0|\mathcal{Y}}(\mathcal{X}_0|\mathcal{Y}) \tag{8}$$

When the data strongly support this estimate, $f_{\mathcal{X}_0|\mathcal{Y}}$ will be peaked and the accuracy of the approximation is improved. This approximation is employed in what follows, and it leads to

$$\Pr(H_k(0)|\mathcal{Y}) \approx \int d\mathcal{X}_0 \Pr(H_k(0)|\mathcal{X}_0) f_{\mathcal{X}_0|\mathcal{Y}}(\mathcal{X}_0|\mathcal{Y}) \tag{9}$$

Thus, we obtain the intuitive result that the probability of a mine (or the absence of a mine) at position $\mathbf{R}$ given data $\mathcal{Y}$ from adjacent locations is a weighted integral of probabilities given all features $\mathcal{X}_0$ that could be acquired over the mine. The weighting factor is the probability of obtaining features $\mathcal{X}_0$ given features $\mathcal{Y}$ measured nearby.

Equation (9) has several important limiting cases. If for each sensor i and some sample j the distances $|\mathbf{R} - \mathbf{R}_{ij}|$ are small compared to the mine's signature radius Δ_i, then the measurements made at the nearby sample locations will be highly correlated. When the noise levels are small we have $f(\mathbf{X}_i(0)|\mathbf{X}_i(\mathbf{R}_{ij})) \to \delta(\mathbf{X}_i(0) - \mathbf{X}_i(\mathbf{R}_{ij}))$ and

$$\Pr(H_k(\mathbf{R})|\Omega) \approx \Pr(H_k(\mathbf{R})|\mathbf{X}_1(\mathbf{R}_{1j_1}), \mathbf{X}_1(\mathbf{R}_{1j_2}), \ldots, \mathbf{X}_{N_s}(\mathbf{R}_{N_s j_M})) \tag{10}$$

which implies that a "nearest-neighbor" approach to non-coincident sampling is effective under these conditions. If, however, the distances $|\mathbf{R} - \mathbf{R}_i|$ are large compared to the mine's sphere of influence, then the measurements made at the remote location will not be related to those over the mine. If all sensors satisfy this condition, then $f_{\mathcal{X}_0|\mathcal{Y}}(\mathcal{X}_0|\mathcal{Y})$ tends to $f_{\mathcal{X}_0}(\mathcal{X}_0)$ and we have $\Pr(H_k(\mathbf{R})|\mathcal{Y}) \sim \Pr(H_k(\mathbf{R}))$. In this case the sensors add no information, and fusion will be ineffective.

An important component of equation (9) is the conditional probability density $f_{\mathcal{X}_0|\mathcal{Y}}(\mathcal{X}_0|\mathcal{Y})$, which we can view as describing the consistency of measured data $\mathcal{Y}$ with the data extrapolation $\mathcal{X}_0$. Since the expected sensor output is dependent on the hypotheses employed, it is convenient to introduce conditioning on $H_k(0)$. It is straightforward to show

$$f_{\mathcal{X}_0|\mathcal{Y}}(\mathcal{X}_0|\mathcal{Y}) = \frac{\sum_{k'=1}^{K} \Pr(H_{k'}(0)) f_{\mathcal{X}_0|\mathcal{Y},H_{k'}}(\mathcal{X}_0|\mathcal{Y}, H_{k'}(0)) f_{\mathcal{Y}|H_{k'}}(\mathcal{Y}|H_{k'}(0))}{f_{\mathcal{Y}}(\mathcal{Y})} \tag{11}$$

where the density of $\mathcal{Y}$ is

$$f_{\mathcal{Y}}(\mathcal{Y}) = \sum_{k'=1}^{K} \Pr(H_{k'}(0)) f_{\mathcal{Y}|H_{k'}}(\mathcal{Y}|H_{k'}(0)) \tag{12}$$

We can interpret the components of this result as follows: If $\mathcal{X}_0$ represents an interpolation of the data $\mathcal{Y}$ to position $\mathbf{R} = 0$, then $f_{\mathcal{X}_0|\mathcal{Y},H_{k'}}(\mathcal{X}_0|\mathcal{Y}, H_{k'}(0))$ is related to the "goodness of fit" of our interpolated value to the data $\mathcal{Y}$, and $f_{\mathcal{Y}|H_{k'}}(\mathcal{Y}|H_{k'}(0))$ is the consistency of our data with hypothesis $H_{k'}$.

Using this result the required conditional probability $\Pr(H_k(0)|\mathcal{Y})$ in equation (9) can be written as a quotient of weighted sums

$$\Pr(H_k(0)|\mathcal{Y}) \quad \approx \quad \frac{1}{f_{\mathcal{Y}}(\mathcal{Y})} \sum_{k'=1}^{K} \Pr(H_{k'}(0)) f_{\mathcal{Y}|H_{k'}}(\mathcal{Y}|H_{k'}(0)) I_{kk'}(\mathcal{Y}) \tag{13}$$

where

$$I_{kk'}(\mathcal{Y}) \approx \int d\mathcal{X}_0 \Pr(H_k(0)|\mathcal{X}_0) f_{\mathcal{X}_0|\mathcal{Y},H_{k'}}(\mathcal{X}_0|\mathcal{Y}, H_{k'}(0)) \tag{14}$$

which is an integral over feature space involving a density function and a classifier.* We can express the classifier in terms of the available density function, which leads to

$$I_{kk'}(\mathcal{Y}) = \Pr(H_k(0)) \int \frac{d\mathcal{X}_0}{f_{\mathcal{X}_0}(\mathcal{X}_0)} f_{\mathcal{X}_0|H_k}(\mathcal{X}_0|H_k(0)) f_{\mathcal{X}_0|\mathcal{Y},H_{k'}}(\mathcal{X}_0|\mathcal{Y}, H_{k'}(0)) \tag{15}$$

where

$$f_{\mathcal{X}_0}(\mathcal{X}_0) = \sum_{k'=1}^{K} \Pr(H_{k'}(0)) f_{\mathcal{X}_0|H_{k'}}(\mathcal{X}_0|H_{k'}(0)) \tag{16}$$

To evaluate equations (13) and (15) we require the conditional densities $f_{\mathcal{Y}|H_k}$ and $f_{\mathcal{X}_0|\mathcal{Y},H_k}$. Directly over the mine we model the random measured features $\mathbf{X}(0)$ for a given hypothesis H_k as a random mine signature $\mathbf{S}_{ik}$ contaminated by zero-mean, additive noise $\mathbf{N}_i$, viz:

$$\mathbf{X}_i(0)|H_k(0) = \mathbf{S}_{ik} + \mathbf{N}_i \tag{17}$$

We take $\mathbf{N}_i$ to be independent of the measurement position and of $\mathbf{S}_{ik}$. At points away from the mine, we introduce a multiplicative factor $G_{ik}(\mathbf{R})$ as follows:

$$\mathbf{X}_i(\mathbf{R})|H_k(0) = \mathbf{S}_{ik} G_{ik}(\mathbf{R}) + \mathbf{N}_i \tag{18}$$

where G_{ik} is a monotonically decreasing† scalar function of $R = |\mathbf{R}|$ with $G_{ik}(0) = 1$. The form of the functions G_{ik} will depend on the sensor i and the hypothesis H_k, but they can be determined from modeling or measurements. Define the mine signature over all sensors for hypothesis k as

$$\mathbf{S}_k = \begin{bmatrix} \mathbf{S}_{1k} \\ \mathbf{S}_{2k} \\ \vdots \\ \mathbf{S}_{N_s k} \end{bmatrix} \tag{19}$$

Since $\mathbf{N}_i$ is assumed to be zero mean, the mean of $\mathbf{S}_k$ is equal to the mean of the feature vector $\mathcal{X}_0|H_k(0)$. An approach to estimating the required densities is to first estimate $\mathbf{S}_k$ from the data $\mathcal{Y}$, and then to estimate the density of $\mathcal{Y}$ from the density of $\mathbf{S}_k$ using the transformation in equation (18) and the independence of $\mathbf{S}_k$ and $\mathbf{N}$.

5.2. The Fusion Process

The fusion algorithm defined above involves a number of steps, which we summarize here. The preparatory steps are as follows:

1. Define K hypotheses H_k regarding the presence or absence of different types of mines at a point. The hypotheses should be chosen to minimize within-class variations in features.
2. Collect sensor data directly over known mine positions for each of these hypotheses.
3. Use these data to develop estimates of $f_{\mathcal{X}_0|H_k(0)}(\mathcal{X}_0|H_k(0))$.
4. Develop models $G_{ik}(\mathbf{R})$ for the features acquired by these sensors under each hypothesis.

Use of the algorithm begins with the collection of measurements

1. Acquire data over an unknown area of interest and form the feature set Ω.
2. Select the point of interrogation $\mathbf{R}$ at which to assess the presence of a mine The vector $\mathbf{R}$ might sequentially sample a grid over the entire area of interest at a spacing comparable to that of the smallest expected mine signature.

*Equation (14) becomes an exact result of equation (5) if we write $\Pr(H_k(0)|\mathcal{X}_0) \to \Pr(H_k(0)|\mathcal{Y}, \mathcal{X}_0)$

†It is implicitly assumed here that the features change monotonically when we move away from the mine. Features with non-monotonic behavior (e.g., raw magnetometer signatures) require a more complicated treatment.

3. Identify the points $\mathbf{R}_{ij_m}$, $m = 1, 2, \ldots, M$, close to $\mathbf{R}$, and form the local feature set $\mathcal{Y}(\mathbf{R}) \subset \Omega$.

4. Estimate $\Pr(H_k(\mathbf{R})|\mathcal{Y}(\mathbf{R}))$ for all $k = 1, 2, \ldots, K$. In evaluating these expressions, employ the feature models G_{ik} to obtain the density functions away from the mine.

5. Report these probabilities at each point $\mathbf{R}$.

Details of the process are described in the following sections.

5.3. Feature Models

Models for the feature distance dependence function $G_{ik}(\mathbf{R})$ can be developed from physical insight or experimental measurements. We assume that the features are linearly related to the data, which makes their amplitude dependence comparable to that of the sensor signature. On the basis of such measurements, we propose the simple model

$$G_{ik}(\mathbf{R}) = \frac{1}{1 + (|\mathbf{R}|/\Delta_{ik})^{\beta_{ik}}} \tag{20}$$

where Δ_{ik} and β_{ik} are constants determined by the sensor and the hypothesis. We take $G_{ik} = 0$ for the no-mine hypothesis. The quantity Δ_i determines the spatial scale of the signature while β_i determines its rate of decrease away from the mine. For a metal detector, values for Δ_i are on the order of the loop radius, and β_i=6. For a focussed beam GPR, we find Δ_i comparable to the beam radius and β_i=2.

5.4. Calculation of Probability Density Functions

Determining the probability density functions $f_{\mathcal{X}_0|H_k}$ is a crucial part of this work. From these functions and the feature models above we estimate $f_{\mathcal{X}(\mathbf{R})|H_k}$ as well. In addition, because $f_{\mathcal{X}_0|H_k}$ is integrated over the features space, it is important that its evaluation be reasonably efficient.

We have used a radial basis function (RBF) neural network to estimate $f_{\mathcal{X}_0|H_k}$. The PDF approximation produced by a RBF NN is of the form

$$f_{\mathcal{X}}(\mathcal{X}) = \sum_{q=1}^{Q} w_q f_q(|\mathcal{X} - \mu_q|) \tag{21}$$

where μ_q are cluster points derived from the data, Q is the number of such points, w_q are a set of scalar weighting coefficients, and f_q are a family of functions that depend only on the radial distance between the input argument and the cluster point. We have used multi-variate Gaussian functions for the f_q, where each has a different covariance matrix and the mean of the Gaussian is the cluster centroid.

Training such a RBF comprises finding the cluster points, the weights, and the covariance matrices of the clusters. A variety of algorithms have been suggested for this purpose, and we have employed a learning vector quantization (LVQ) neural network. In our technique the LVQ network is used to determine the classes of the training data. From the class membership we estimate mean and covariance matrices. The weights are determined by the number of points assigned to the cluster. To improve the continuity of the resulting density function, one can slightly expand the covariance matrix for each cluster.[6] The result is an expression of the form

$$f_{\mathcal{X}|H_k}(\mathcal{X}|H_k) = \sum_{q=1}^{Q} w_q \frac{1}{(2\pi)^{L/2}\det(\mathbf{\Sigma}_{qk})^{1/2}} e^{-(\mathcal{X}-\mu_{qk})^T \cdot \mathbf{\Sigma}_{qk}^{-1} \cdot (\mathcal{X}-\mu_{qk})/2} \tag{22}$$

where L is the length of the feature vector $\mathcal{X}$. It is easy to see that if each cluster point μ_{qk} is shifted by some amount μ_0, then the mean of $f_{\mathcal{X}|H_k}$ shifts by the same amount.

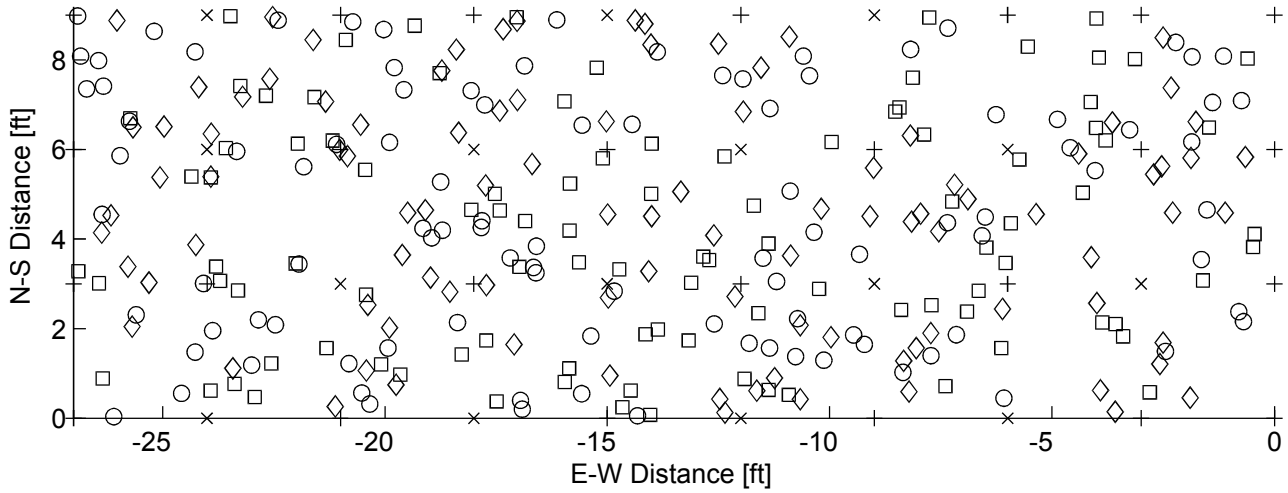


Figure 3. An example of randomly located measurements. The locations of 100 sensor measurements are indicated by the square, circle and diamond markers. Mine and non-mine grid points are indicated by the "x" and "+" symbols respectively.

6. RESULTS

We have developed software to implement the above-described algorithms. Tests of the algorithms were performed using simulated data. Using simulated data in development effectively decouples algorithm problems from problems with ineffective sensors and uninformative features.

To demonstrate the algorithm's insensitivity to measurement locations, we will use randomly placed sensor measurements, an example of which is shown in Figure 3. In that figure 100 measurements are shown for each of three sensors. The sensors are defined to be sensitive out to distances $\Delta_i = 12$ inches. Two classes, mine and no-mine, are considered. One feature is defined for each of the three sensors, and zero-mean Gaussian noise is added to the simulated features. As a test of the system's ability to deal with uninformative sensors, one of the three features is constrained to provide no useful information.

We interrogate the data set at points $\mathbf{R}$ located on a 9 inch grid. At each point, the nearest $M = 3$ adjacent measurements for each sensor were identified and used to form $\mathcal{Y}$. The probabilities $\Pr(H_k(\mathbf{R})|\mathcal{Y})$ were computed using equation (13) and the log-likelihood ratio was formed. The results are shown in Figure 4 for the case when the distance between features (in feature space) is 4 times the noise standard deviation. In the interest of brevity we will refer to this as a "feature-to-noise" ratio of 4. All the mines are clearly detected in this result. Because the algorithm outputs the probabilities $\Pr(H_k|\Omega)$ at each desired point $\mathbf{R}$, we are able to determine the confidence that a mine does or does not exist at every point. This is in contrast with a conventional association-based feature-fusion, which would only provide that information at discrete points where the sensors have made a detection.

Receiver operating characteristic (ROC) curves for several sample densities are shown in Figure 5. Feature-to-noise ratios of 4 and 2 are used in parts (A) and (B) respectively. Performance improves with increasing sample density as shown in the figures. The case of 100 samples comprises an average sample density of 2.4 ft^2/sample, which is only marginally better than the sensor's region of sensitivity (3.1 ft^2). As expected, performance also improves with decreased noise. Essentially perfect performance is achieved at high sample densities and high feature-to-noise ratios.

The algorithm has also been tested using experimentally measured data. The data collection depicted in Figure 2 was processed to extract samples near each of the grid positions. The results are shown in Figure 6. These data were partitioned into training and testing sets, and the fusion algorithm described above was applied. Because of poor performance by the IR sensor,[‡] the IR data was excluded from this tests. Because of the limitations of the resulting

[‡]In our first round of data collection the mine field was covered with grass which tends to obscure thermal signatures of buried targets. This grass cover has now been removed.

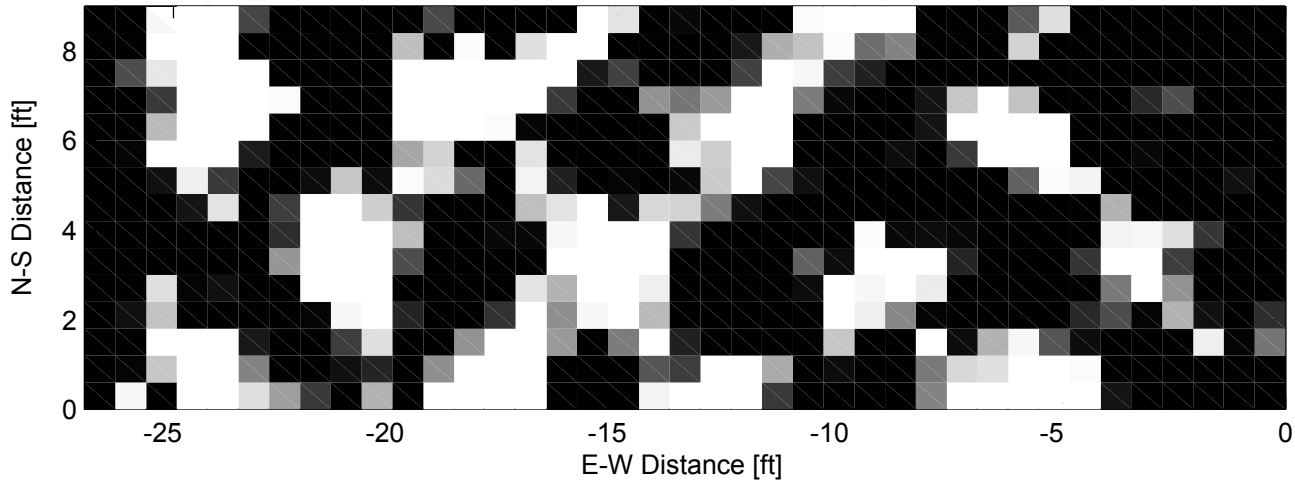


Figure 4. Log-likelihood ratio for fusion of 400 randomly located sensor measurements.

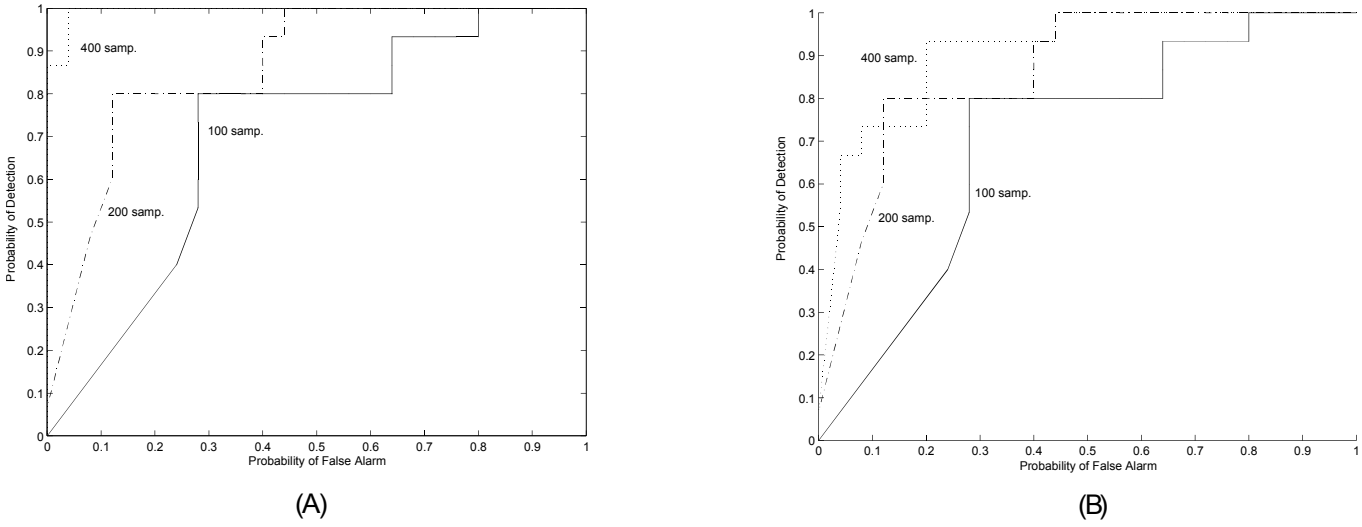


Figure 5. ROC curves for fusion of randomly located sensor measurements with different densities. The ratio of inter-class distance (in feature space) and noise standard deviation is 4:1 and 2:1 for Figures (A) and (B) respectively.

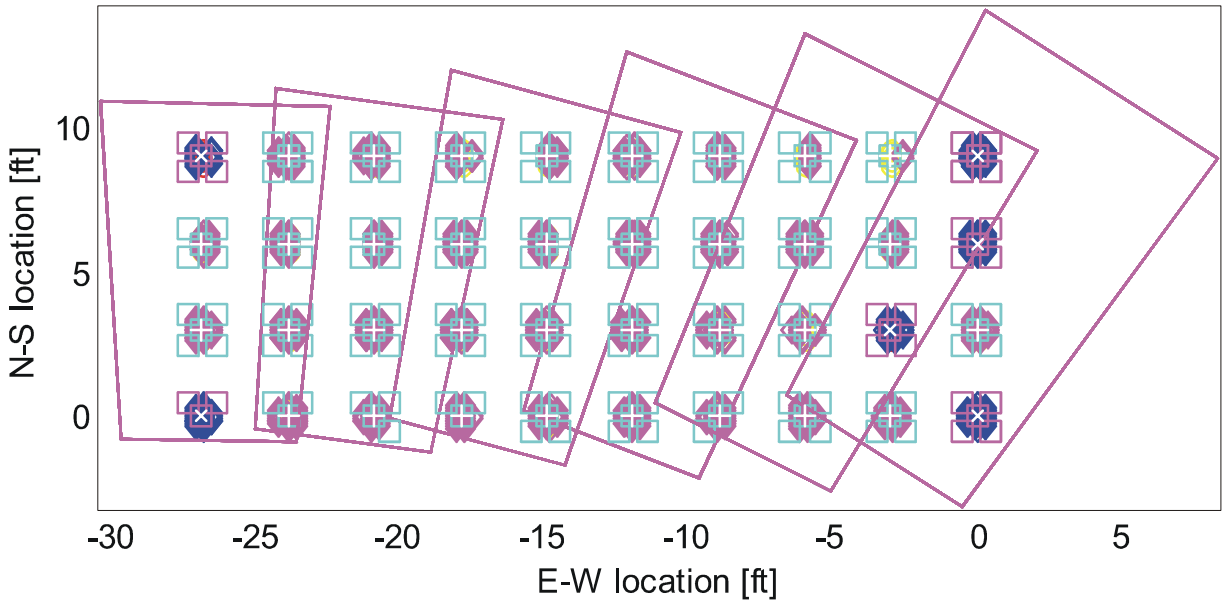


Figure 6. Training and testing data set. The dark masses on the grid points indicate the locations of relevant GPR and EMI data. The small square patches represent IR image chips. Because of overlapping IR imagery, multiple chips exist for some locations.

sensor suite, a two-class problem was defined, comprising metallic mines and non-metallic objects.

The performance of the fused EMI-GPR sensor suite is excellent. The ROC curve for this suite is shown in Figure 7. The benefits of fusion are evident from Figures 8 and 9, where we display the results of EMI-only and GPR-only detection. These results suggest that fusion can be highly effective, but because our data set is relatively small additional testing must be performed to verify this conclusion.

7. CONCLUDING REMARKS

We have developed a novel algorithm for fusion of sensor data. The algorithm is insensitive to non-coincident sampling and it does not require individual sensors to perform detection prior to fusion. To support the development of a fusion system based on this and other fusion techniques, pre-processing algorithms have been developed and tested. Measured data have been collected to support testing. Future efforts will focus on improvements and refinements of the algorithms.

ACKNOWLEDGMENTS

This work was funded by the US Army Research Office under the Multi-University Research Initiative (MURI) on Humanitarian Demining, contracts DAAH04-96-1-0448 and 97-SC-ARO-1015.

REFERENCES

1. L. Collins, V. George, T. Altshuler, L. Nolte, and L. Carin, "An improved decision-theoretic approach to the optimum detection of mines," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, A. C. Dubey and R. L. Barnard, eds., *SPIE* **3079**, pp. 716–723.
2. R. R. Coifman and M. V. Wickerhauser, "Wavelets and adapted waveform analysis," in *Wavelets: Mathematics and Applications*, J. Benedetto and M. Frazier, eds., pp. 399–424, CRC Press, Boca Raton, FL, 1993.
3. D. L. Dohono, I. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: Asymptopia?," *J. Royal Stat. Soc., series B* **57(2)**, pp. 301–369, 1995.

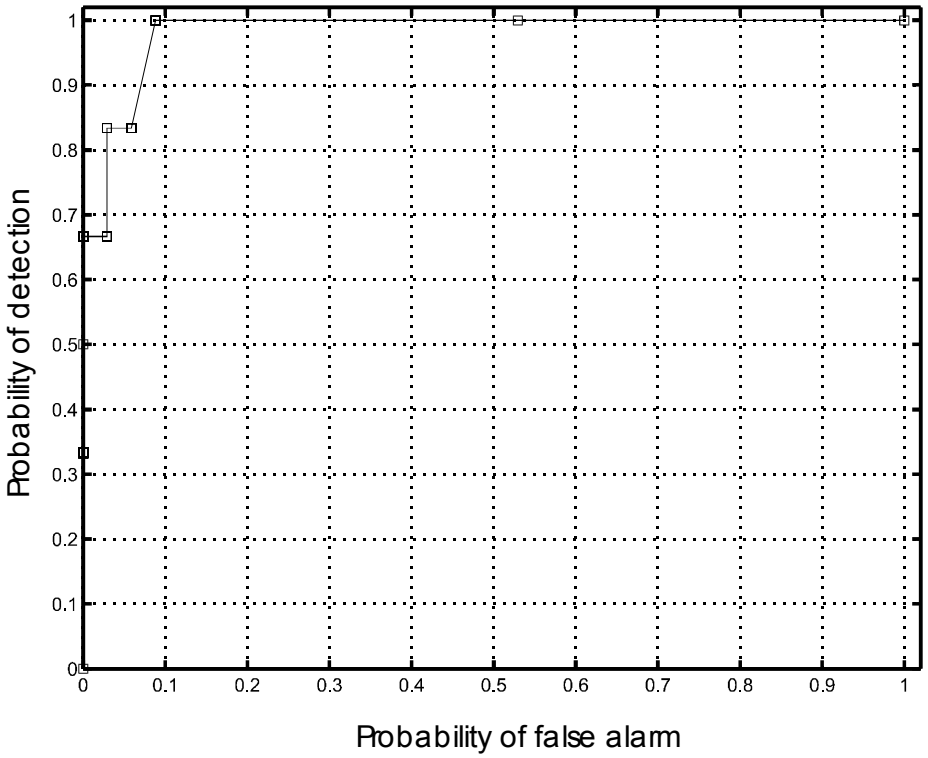


Figure 7. ROC curve for fused EMI-GPR sensor suite.

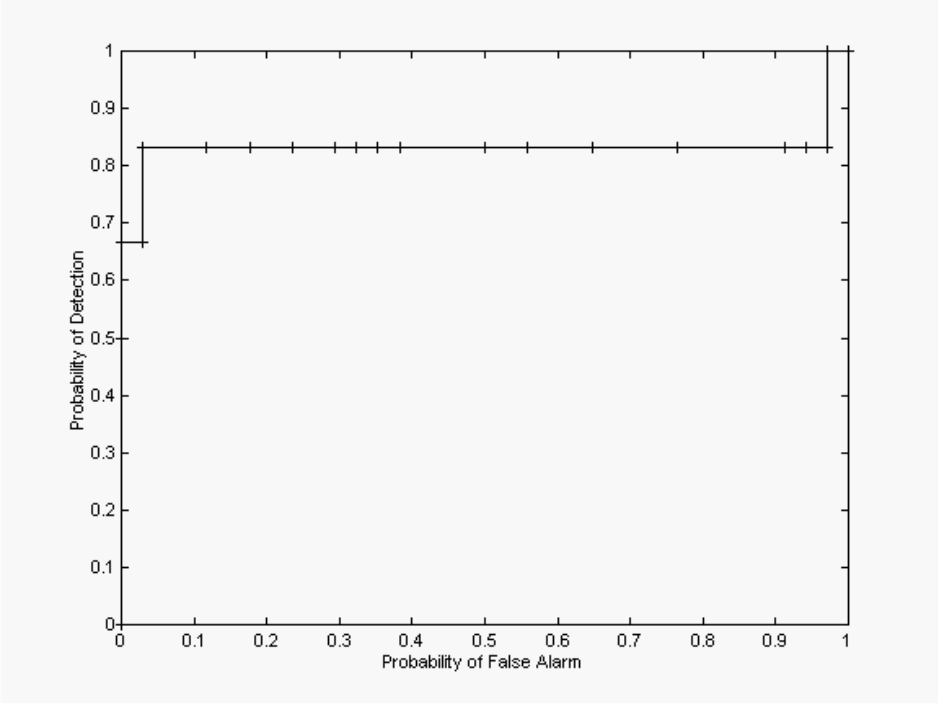


Figure 8. ROC curve for the EMI sensor only.

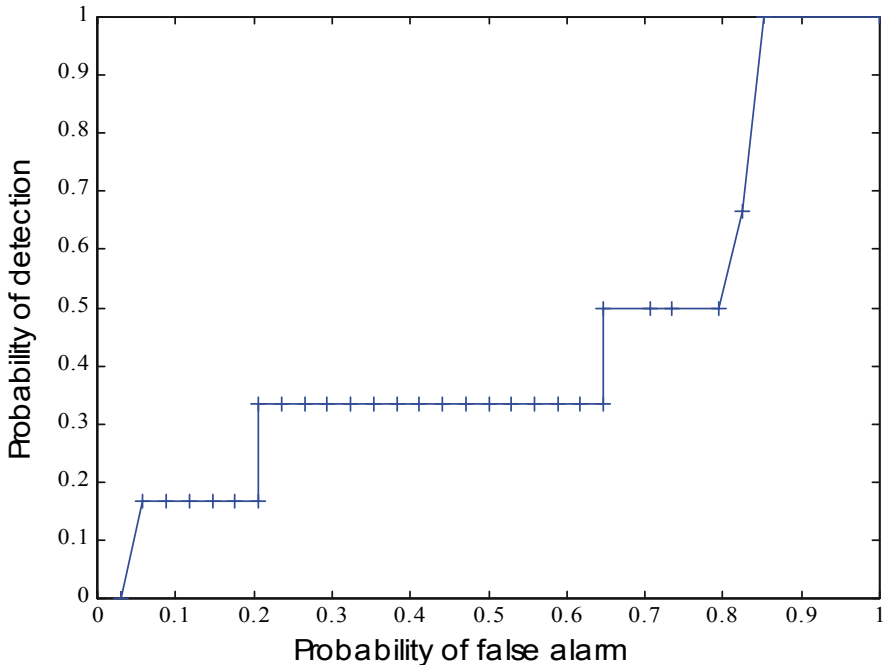


Figure 9. ROC curve for the GPR sensor alone.

4. U. Wiklund, M. Akay, and U. Niklasson, "Short-term analysis of heart-rate variability by adapted wavelet transforms," *IEEE Engineering in Medicine and Biology Mag.* **16(5)**, pp. 113–118, 1997.
5. T.-H. Chao, B. Lau, A. Yacoubian, and G. Henderson, "Mine detection using wavelet processing of electro-optic active sensor data," in *Detection Technologies for Mines and Minelike Targets*, A. Dubey, I. Cindrich, J. Ralston, and K. Rigano, eds., *SPIE* **2496**, pp. 17–21.
6. J. Schurmann, *Pattern Classification*, John Wiley, New York, 1996.

A Clutter Reduction Technique for GPR Data from Mine Like Targets

A. van der Merwe, I.J. Gupta and L. Peters, Jr.

ElectroScience Lab., Ohio State University, 1320 Kinnear Rd., Columbus, OH 43212

ABSTRACT

In this paper a signal processing technique is developed to reduce clutter due to ground bounce in ground penetrating radar (GPR) measurements. This technique is especially useful when a GPR is used to detect subsurface anti personnel mines. The GPR clutter is modeled using a simple parametric model. Buried mine and clutter contributions are separated through a set of iterative subspace projections. The algorithm outperforms existing clutter reduction approaches and also yields target features which are useful for detection and identification of these mines. It is shown that the proposed technique increases the probability of detection and at the same time significantly decreases the false alarm rates when combined with a matched filter detector.

Keywords: GPR, clutter reduction, AP mines.

1. INTRODUCTION

The global land mine crisis is creating immense social and economic problems worldwide. Safe and cost effective methods for clearing these mines are therefore needed. A promising sensor for detecting anti personnel mines is a ground penetrating radar (GPR). The GPR data, however, is distorted due to very high levels of interfering signals or clutter which leads to an unacceptably large number of false alarms. The dominant clutter contributors being the strong returns received from the rough ground surface. The clutter varies statistically with soil conditions and/or surface roughness and leads to uncertainty in the measurements. A new signal processing technique is discussed to reduce this clutter term in the GPR data.

The following physical limitations are taken into account in developing the new technique. First, anti-personnel mines are buried at shallow depths. Thus the returns from the mines and that from the rough ground interface overlap in time. It is very difficult to discriminate these coincident responses. Mines are made out of plastic materials and buried in lossy soil. The relative returns from the mines are therefore very low in energy. Commonly used complex natural resonance frequencies are very highly damped due to the lack of dielectric contrast between the soil and the mine and thus not suitable for detection. The GPR also receives returns from other subsurface inhomogeneities, for example rocks, tree roots or small pieces of metal in the ground which leads to high levels of false alarms.

As opposed to most current clutter reduction techniques, where the clutter is treated statistically, the new technique incorporates the clutter as a deterministic, but unknown component. Simple parametric models are used to represent the mine and the clutter in a GPR measurement. An iterative signal processing algorithm is developed whereby the mine and clutter contributions are separated through a set of iterative subspace projections. To evaluate the performance of the detector, the receiver operating characteristic (ROC) curves obtained from Monte Carlo simulations of a surrogate mine field are compared to various state-of-the-art detectors. The proposed processing technique also provides useful target features which can be used for identification of these mines. It is shown that the proposed technique increases the probability of detection and at the same time significantly decreases the false alarm rates.

To demonstrate the proposed technique, a two-dimensional finite-difference-time-domain (FDTD) algorithm is implemented to simulate surrogate land mines buried in a lossy clay medium under a rough ground surface. The FDTD technique is chosen here because it is ideally suited for modeling scatterers surrounded by inhomogeneous media, and does not require complicated Green's functions which is typically used in the method of moments. Another advantage of using the FDTD technique is that the wide band nature of the GPR sensor can easily be incorporated into the formulation.

Other author information– Gupta Email: gupta.11@osu.edu, van der Merwe Email: avm@avmpc.eng.ohio-state.edu.

The rest of this document is organized as follows. The numerical simulation of the data is discussed in Section 2. Some existing clutter reduction techniques are discussed in Section 3. The proposed iterative clutter reduction technique is developed in Section 4. The detection of subsurface targets is discussed in Section 5.

2. DATA SIMULATION

The two dimensional FDTD algorithm[1,2] incorporates a lossy half space, rough surface and a buried dielectric object or a buried randomly shaped inhomogeneity. The algorithm is implemented in the time domain, starting from Maxwell's equations and employing the classical Yee algorithm. The geometry and dimensions of the FDTD grid used here is shown in Figure 1.

To ensure stability of the FDTD computations, the time step needs to satisfy the Courant Stability condition which relates the time step Δt to the spatial increments as follows:

$$\Delta t \leq \frac{1}{v_g(1/\Delta^2 x + 1/\Delta^2 y)} \tag{1}$$

where Δx and Δy of the spatial increments on the two dimensional grid and v_g is the speed of wave propagation in the soil. Numerical dispersion causes the wave to propagate in the FDTD mesh at a speed less than the physical speed of the wave, giving rise to a non physical distortion. To combat this effect, the spatial discretization is chosen to be more than 10 samples per wavelength, at the highest frequency of interest.[3]

The FDTD grid is surrounded by an absorbing boundary condition to absorb outwardly propagating waves at the edges of the finite computational domain.[4–7] The perfectly matched layer (PML) is implemented here. Note that the original PML formulation is adjusted when applied to the lossy soil medium. The parameters for each PML layer is chosen to yield an impedance match with the adjacent medium.

Plane wave incidence is studied here. Due to the finite size of the FDTD computational domain, the amount of energy in the reflected waves is finite and will be determined by the size of the finite illuminated spot on the ground surface which is in turn a function of the size of the grid along the ground surface. The plane wave has a Gaussian pulse shape. This parameters of the Gaussian pulse is chosen such that the incident field has most of its energy in the frequency band between 1 GHz and 5 GHz. Results are presented for normal incidence and TE_z polarization. All scattered fields are observed in backscatter.

In our simulation, the interface between the soil and the free space is modeled as a rough surface. The surface roughness is assumed to have a Gaussian spectrum.[8] The spatial correlation function for the rough surface is defined as,

$$\rho(x) = h^2 e^{-x^2/l^2} \tag{2}$$

where h is the rms surface height and l the correlation length. The interface (surface profile) varies randomly between surface generations.

Three different deterministic targets are considered in these simulations. All targets are composed of a homogeneous, lossless dielectric medium. The relative dielectric constants of all three are fixed at $\epsilon_r = 3.5$, which is a typical value for plastic land mines.[9] The cross sections are chosen to be representative of commercial plastic land mines. Schematics of the dimensions and geometries of these targets are shown in Figure2. The depth of the target is defined as the distance, d, between the top of the target and the mean position of the rough surface as shown in Figure 1. The depth is varied between 2.54 cm and 10.16 cm here.

Scattering from randomly shaped subsurface inhomogeneities is also studied here. Soil inhomogeneities for example refilled trench holes, rocks, tree roots, and in particular disturbed clay introduces a distortion of the received waveform which depends on the location and characteristics of the inhomogeneity. The intensity of the scattering depends on the soil type and depth of the anomaly.[10]

A random shape is generated using the following perturbed circle formulation:

$$\sqrt{((x-x_c)^2 + (y-y_c)^2)} \leq R_o + \Delta r(x,y) \tag{3}$$

where (x_c, y_c) is the coordinates of the center of the circle with a radius R_o. The perturbation to the circle at each (x,y) location is defined as,

$$\Delta r(x,y) = \sum_{i=1}^{5} a_i \cos(\frac{2\pi}{T_o} i\theta) \tag{4}$$

where $\theta = \arctan(\frac{y}{x})$ and $T_o = 2\pi R_o$. The coefficients $a_i, i = 1, 2, \ldots, 5$ are random variables chosen from a uniform distribution between -0.5 and 0.5. A schematic representation of the geometry of the inhomogeneity is also shown in Figure 2.

3. DATA PROCESSING TECHNIQUES

The goal of this study is the detection of small, dielectric land mines buried close to a rough ground surface in a lossy soil. Since there are a relatively small number of land mines, the scattering characteristics (signatures) of which can easily be measured, a simple matched filter is proposed for detection.

The scattered fields typically received by a ground penetration radar contains contributions from the subsurface targets as well as from the ground surface and interaction terms. The latter two contributions are referred to as clutter. The GPR data needs to be preprocessed before detection is attempted to reduce the high levels of clutter typical for such measurements. Some state-of-the-art clutter reduction techniques will be briefly discussed next.

Assuming that the unwanted contributions due to scattering from the rough surface can be approximated by a correlated random process with a known mean and correlation matrix, a whitening filter can be used to transform the correlated component of the clutter into white Gaussian noise.[11,12] To determine the clutter statistics, many realizations of the rough surface, characterized by the same statistics are needed.[12] The whitening filter is implemented as a linear prediction error filter.[13,14] The GPR data is preprocessed by the whitening filter to reduce the effects of non-coherent clutter component. The coherent component is reduced by complex averaging. A limitation of this technique is the fact that in reality the soil is an inhomogeneous medium and therefore its statistical properties vary with position along the surface. Any mismatch in the clutter statistics used to design the filter and that of the GPR data degrades the performance of the whitening prefilter.

Since the dominant clutter contribution originates at the ground surface, early time gating is also a good choice for reducing this component of the clutter. Various techniques proposed in this regard perform only marginally.[15,16] An improved time gating approach has been suggested where the early time contributions are approximated using a superposition of damped exponentials.[17] The TLS-Prony technique[18] is used to estimate the necessary exponential parameters *. The estimated clutter is then subtracted from the data to reduce its contribution. However, both the whitening filter as well as the time gating approach do not perform well for anti personnel mines. Therefore, it is necessary to develop more efficient clutter reduction approaches.

4. PROPOSED CLUTTER REDUCTION TECHNIQUE

In the proposed approach, the clutter contained in a GPR measurement is treated as a deterministic, but unknown component. A reasonable assumption for high frequency ultra-wideband ground penetrating radar system measurements is that the total field is a superposition of individual contributions which originates at localized surface and subsurface scatterers. To account for the variability in a typical subsurface scenario, a parametric model is used to represent the clutter contributions. Assume that the total frequency domain scattered field received by the GPR can be represented by the following superposition,

$$S(\omega) = H_c(\omega) + T(\omega) + n(\omega) \tag{5}$$

where $H_c(\omega)$ represents the clutter contributions, $T(\omega)$ represents the contribution from the desired target, $n(\omega)$ represents additive Gaussian noise and $\omega = 2\pi f$. It is important to note that $H_c(\omega)$ represents all the unwanted contributions, for example scattering from the rough surface as well as interaction terms and contributions from subsurface soil inhomogeneities. These unwanted contributions are collectively referred to as *clutter*. An iterative method is proposed to remove clutter due to surface scattering and interaction terms. Note that the contributions

*Note that the early time gating approach correspond to a single iteration of the iterative technique discussed in Section 4

from the subsurface inhomogeneities do not obscure the return from the target. Therefore these will not necessarily be removed using the iterative approach. The subsurface inhomogeneities are treated as false alarms and are separated from the target contributions using the matched filter detector.

Assume that the characteristics of the desired target is known. It is represented by $T_r(\omega)$ and referred to as the reference signature. The part of the model containing the buried target return in (5) differ from this reference signature in two ways. The depth of the subsurface target is unknown and will result in a complex phase change in the scattered field. Secondly, the incident field which is transmitted through the rough surface is no longer a plane wave anymore and is also dispersed as it propagates through the soil due to the frequency dependence of the index of refraction.[19,12] For this reason a frequency dependent damping factor is included in the model to account for this effect. Incorporating these factors, the estimated target signature can be represented by,

$$\hat{T}^n(\omega) = A_r^n \, e^{-\omega\gamma_r^n} e^{-j\omega t_r^n} T_r(\omega) \tag{6}$$

where A_r^n represents a complex amplitude, γ_r^n represents a damping factor and t_r^n a time delay at the n^{th} iteration.

The estimated target defined in (6) contains the desired component in the received GPR data set. Define the following data set,

$$\tilde{S}^{n+1}(\omega) = S(\omega) - \hat{T}^n(\omega) \tag{7}$$

which contains the unwanted contributions of the GPR data set. It is assumed that the target return is zero for $n = 0$.

A damped exponential model is suitable for representing most contributions.[17] Accordingly, represent the data set $\tilde{S}^n(\omega)$ using damped exponentials as follows,

$$\tilde{S}^n(\omega) = \sum_{k=1}^{I_t} A_k \, e^{-\omega\alpha_k} e^{-j\omega t_k} \tag{8}$$

where I_t is the total number of damped exponential terms needed to model the data set, α_k is the damping factor and t_k the time delay of the k^{th} term for $k = 1, \ldots, I_t$. The TLS-Prony technique[18] is used to estimate these parameters. Some of the estimated complex damped exponentials represent the ground return and interaction terms while others represent subsurface inhomogeneities and noise. It is desired at this point to estimate the dominant early time clutter contributions, namely the surface returns and the interaction terms. Therefore, the poles which represent these terms are carefully identified. A way of doing this is to look at the estimated time locations and select only the poles within a certain time window. The fact that the dominant clutter resides in the earlier time contributions can be used in choosing an appropriate time window. The size of the window is varied between iterations. A very narrow window is selected in the beginning to ensure that the contribution from the shallow target is not also removed. The window size is increased at later iterations. The selected complex exponential parameters are then used to reconstruct the clutter contributions defined as,

$$\hat{C}^n(\omega) = \sum_{k=1}^{I_c} A_k^c \, e^{-\omega\alpha_k^c} e^{-j\omega t_k^c} \tag{9}$$

where $I_c \leq I_t$ and A_k^c, α_k^c and t_k^c represent the complex exponential parameters of the k^{th} early time pole. Although the estimated parameters can not be associated with any physical scattering mechanisms, these effectively model the early time clutter contributions.

The iterative signal processing technique is used to decompose a frequency domain measurement into the desired target signature and clutter contributions using the models defined above. A schematic representation of the procedure is shown in Figure 3. The subscript n indicates the iteration number and $S(\omega)$ is the measured frequency domain data containing both the clutter and the subsurface target contributions.

Subtracting the estimated clutter from the GPR data yields the residual data set defined as,

$$R^n(\omega) = S(\omega) - \hat{C}^n(\omega) \tag{10}$$

which contains the left-over clutter contributions and/or the target contribution. The residual data is used to obtain an estimate of the target. The parameters needed to estimate the target at iteration n are obtained by optimizing the following least squares error,

$$\text{Cost}^n(\theta) = \arg\min_{\theta^n}[|\sum_{\omega_1}^{\omega_N} R^n(\omega) - A_r^n e^{-\omega\gamma_r^n} e^{-j\omega t_r^n} T_r(\omega)|^2] \tag{11}$$

where ω_1 and ω_N are the end band frequencies and $\bar{\theta}$ the parameter vector defined as $\bar{\theta^n} = [\gamma_r^n, t_r^n]$. Since the target model is linear in amplitude, A_r^n can be removed from the cost function by linear least-squares.[13] Therefore no optimization is needed for this parameter. In the procedure outlined in Figure 3, $A_r^n = 0$ if $n = 0$.

The iterations are terminated when the change in the estimated target † is less than a predefined threshold. The output of the iterative procedure yields the two decomposed data sets, namely the estimated clutter contribution and a residual data set. The residual data is dominated by the buried target signature which can be further processed to yield target identification or detection. Note that the estimated target is not used in further processing since it may lead to unwanted detections for measurements without the target of interest. In some cases the residual data may also contain the contributions from a subsurface soil inhomogeneity but this will be classified as a false alarm by the match filter.

The algorithm also provides an estimate of the depth of the buried target. This estimate provides valuable information regarding the placement of the mines in the ground. Estimates of the amplitude of the scattering from the target as well as an indication of the dispersion of the wave through the soil is also obtained. These additional estimates may be used as features in target identification.

Next the iterative technique is applied to a simulated data set generated using the time domain FDTD algorithm. The data set is transformed into the frequency domain to simulate a wide band GPR measurement between 1 GHz and 5 GHz. The frequency sampling is Δf=50 MHz. To obtain the reference signature for the target, simple time gating is used to separate the target response from the interaction terms and the surface clutter. This signature is extracted from a simulation of the desired target buried deep underneath a flat surface.

The target used in this example has a rectangular cross section which resembles Target 1 of Figure 2. It is composed of a homogeneous, lossless dielectric with a relative dielectric constant of ϵ_r=3.5. The surrounding soil resembles a dry clay medium with a relative dielectric constant of ϵ_r=9 and a conductivity of $\sigma_g = 0.05$ S/m. The target is buried at a depth of 2.54 cm beneath the mean position of a rough surface in the lossy, dielectric soil. The Gaussian surface roughness parameters are h=0.5 cm and l=18.7 cm. Note that these parameters represents a fairly smooth surface. In the first set of examples, only a single realization of a surface with the latter roughness characteristics is considered.

First consider the responses at some crucial points after a single iteration of the algorithm. The residual data is shown in the broken lines in Figures 5 (a) and (b) in the frequency and time domains respectively. It is clear that there is a discrepancy between the residual data shown in the broken lines and the reference signatures. This can be attributed to the fact that there is still some clutter left. This is clear from Figure 6 (a) and (b) which show the estimated clutter after a single iteration in the broken lines. Only the early time clutter is removed as expected.

Proceeding, consider the same set of responses, but after multiple iterations of the algorithm has been completed. In this case, 5 iterations were performed to achieve convergence of the algorithm. The residual data is compared to the reference target response in Figures 5 (a) and (b) in the frequency and time domains respectively. It is clear that the residual data corresponds very well to that of the reference target. Thus, the proposed algorithm is very effective for clutter reduction.

The estimated clutter is compared to the total data in the time domain and frequency domain in Figure 6 (a) and (b) respectively. It is clear that the estimated clutter contains not only the rough surface response, but also some later time contributions which may possibly be due to interaction between the shallow target and the rough surface. Despite the fact that the some of the interaction terms are very close in time to the target returns, it is still possible to remove them from the data. This example demonstrates the advantages of widening time window as the number of iterations increases.

† $\frac{1}{E^n}|\sum_{\omega_1}^{\omega_N} T^n(\omega) - T^{n-1}(\omega)|^2 < \delta$, where E^n represents the energy in the estimated target at the current iteration.

5. DETECTION OF SUBSURFACE TARGETS

The performance of the detectors when applied to a surrogate mine field are evaluated using the receiver operating characteristic (ROC) curves. The preprocessors compared here are the whitening filter, the time gating approach and the proposed iterative technique. The performance of the matched filter with no clutter reduction is also shown.

The procedure used for detection is formulated using a simple binary hypothesis test.[20] The hypothesis H_o occurs in the absence of the desired target when there is only clutter and noise while the hypothesis H_1 occurs when the desired target is present. A schematic of the detector is given in Figure 4.

The matched filter used here is formulated as,

$$\rho(\tau) = \frac{1}{2\pi\,\Delta}\left|\int_{-\infty}^{\infty} S(\omega)H^{\star}(\omega)e^{-j\omega\tau}\mathrm{d}\omega\right| \tag{12}$$

where $H(\omega)$ is the reference signature and $S(\omega)$ is the total GPR data in the frequency domain. The normalization factor is defined as $\Delta = \sqrt{E_S E_H}$ where E_S and E_H represent the energy in the respective signatures. Note however that the dependence of the buried target return on depth is not taken into account in this definition. A more accurate formulation includes the factor $H(\omega, t_d)$, where t_d, related to the depth d of the buried target is given by $t_d = \frac{2d}{v_g}$ and where v_g is the speed of propagation in the soil. However, the approximation used in (12) is justified because at the frequency band and soil losses considered the soil is not dispersive and the speed of wave propagation can be approximated as $v_g \approx \frac{1}{\sqrt{(\mu\epsilon)}}$.

Since it is a controlled experiment, the target depth d, is known and it can be used to calculate the time delay τ_d. That is, $\tau_d = \frac{2d}{v_g}$. The output of the matched filter, defined in (12) at the delay $\tau = \tau_d$ is compared to the threshold as follows,

$$\begin{aligned} \rho\left(\tau = \tau_d\right) \quad &> \beta, \quad H_1 \\ &< \beta, \quad H_o \end{aligned} \tag{13}$$

where β denotes the decision threshold. Note that the value of the threshold is different for the various preprocessing techniques.

The performance measure used here are the detector ROC curves. Monte Carlo simulations are performed to yield the necessary realizations. Each Monte Carlo realization contains either one of the following,

- a rough ground surface without any buried objects
- a rough ground surface with a buried dielectric target at various depths
- a rough ground surface with a randomly, shaped, dielectric inhomogeneity at various depths with varying shape and relative dielectric constant

White Gaussian noise is added to each realization [‡]. The interface profile, subsurface inhomogeneities as well as the additive noise is changed from one realization to the next. The statistics of the rough surface are the same for all the realizations and is chosen as h=0.5 cm and l=18.7 cm. The surrounding soil resembles a dry clay medium with a relative dielectric constant of ϵ_r=9 and a conductivity of $\sigma_g = 0.05\,\mathrm{S/m}$.

The targets are deterministic and the geometry and material composition is kept fixed between realizations. The targets are composed of a homogeneous, loss less dielectric material with a relative dielectric constant of $\epsilon_r = 3.5$. The geometry and dimensions of the targets were defined in Section 2. To simulate the uncertainty in the placement of the subsurface target, the depth of each target is varied uniformly between 2.54 cm and 10.16 cm beneath the mean position of the rough surface.

The shape, depth and dielectric composition of the subsurface inhomogeneities are randomly varied between realizations. The relative dielectric constant is uniformly varied between 4.5 and 7.5. The conductivity is set equal to that of the surrounding medium which is $\sigma_g = 0.05\,\mathrm{S/m}$. The shape is varied using the formulation given in (2).

[‡] The signal-to-noise ratio is defined in the frequency domain as follows, SNR = 10 $log_{10}\frac{E_s}{E_n}$ dB where E_s and E_n denotes the energy in the simulated data and Gaussian noise respectively.

The value of the radius R_o used here is 25 mm. The depth of the inhomogeneities is also varied between 2.54 cm and 10.16 cm.

In all the examples considered next, the matched filter is designed to correspond to the target of interest. For example, in the first set of results, the target of interest is Target 1 and the reference signature corresponding to this target is used in the matched filter formulation. The reference signature is obtained from a simulation of the desired target buried deep under a flat surface and using time gating. In the first set of results, white Gaussian noise is added to each realization to yield a value of SNR = 40 dB.

Assume that Target 1 is the target of interest. The ROC curves for the various techniques are shown in Figure 7(a). The matched filter without any preprocessing does not yield satisfactory performance. This can be expected due to the high levels of the clutter which is typical for subsurface target measurements. Furthermore note that the performance improvement obtained with the whitening filter and with the removal of only the early time clutter terms are comparable. Note that in neither of the latter techniques any knowledge of the buried target is used. The time gating done here in the frequency domain is however preferred because it is applicable to any clutter environment while the performance of the whitening filter depends on the similarity of the measured clutter statistics and that used to design the whitening filter. From these curves it is also clear that the highest P_D for any P_F is obtained when the data has been preprocessed using the iterative algorithm. This may be attributed to the fact that this technique removes clutter from the rough surface as well as due to ground surface-target interaction terms. The clutter reduced data set therefore contains a very good signature of the subsurface target. Due to this reason it is very easy to distinguish between the desired target and a false alarm using a simple matched filter.

Results with Target 2 and Target 3 buried at different depths under the rough surface are shown in Figure 7 (b) and 7 (c) respectively. Note that new matched filters have been designed to correspond to the new targets of interest. As before, the proposed iterative algorithm outperforms the other techniques. The performance of the whitening filter and frequency domain time gating technique are comparable in these cases also.

Monte Carlo simulations similar to before are also performed at a SNR = 30 dB. The ROC curves for Target 1, Target 2 and Target 3 are shown in Figures 7 (d), (e) and (f) respectively. The performance of all the techniques degrade to some extent at the higher Gaussian noise level. Note that the performance degradation using the whitening prefilter is more than that of the early time gating approach. The performance of the detector with iterative clutter reduction still yields the highest P_D for all P_F.

6. SUMMARY AND CONCLUSIONS

A new signal processing technique to reduce clutter in GPR data is proposed here. It is demonstrated that the proposed algorithm reduces the clutter due to rough surface scattering as well as the surface-target interaction terms. This is true even for targets buried at very shallow depths under a rough, ground surface where the responses from the clutter and that from the target overlap in time.

The clutter reduction technique is combined with a simple matched filter to formulate a detector. The matched filter without any processing does not perform satisfactory, as can be expected due to the high levels of the clutter. Preprocessing to reduce the clutter improves the performance for all the techniques compared here. At the lower SNR the improvement of the whitening filter and that of the early time gating approach are comparable but at the lower SNR the latter performs slightly better. Both these techniques reduces the clutter to some extent but does not yield optimal performance. The new iterative clutter reduction algorithm yields the highest probability of detection for any probability of false alarm. This may be attributed to the fact that the algorithm decomposes the data into its clutter and target contributions. The target is the dominant contribution in the clutter reduced data set and therefore it can easily be distinguished from false alarms using a simple matched filter.

The new clutter reduction technique also yields a set of parameters for the target namely an estimate of the depth of the subsurface target and an estimate of the damping factor. The latter parameter is related to propagation through lossy soil. These features will be used to improve the matched filter formulation in future studies. A multi spectral approach to reduce the false alarms due to objects which are similar to the target of interest is also being developed. Finally, the performance of the proposed clutter reduction technique under rougher surface conditions is also being studied.

7. ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the ARO (The OSD MURI Program). The findings, opinions and recommendations expressed therein are those of the authors and are not necessarily those of Duke University or the ARO.

REFERENCES

1. A. Taflove, *Computational Electrodynamics: The Finite-Difference Time -Domain Method*, Artech House, Boston, MA, 1995.
2. K. L. Shlager and J. B. Schneider, "A selective survey of the finite-difference time domain literature," *IEEE Antennas and Propagat. Magazine* **37**, pp. 11–17, August 1995.
3. P. G. Petropoulos, "Stability and phase error analysis of FDTD in dispersive dielectrics," *IEEE Trans. Antennas and Propagat.* **42**, pp. 62–69, January 1994.
4. B. Engquist and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves," *Math. Comput.* **31**, pp. 629–651, July 1977.
5. G. Mur, "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagnetic Campat.* **23**, pp. 377–382, November 1981.
6. R. Higdon, "Numerical absorbing conditions for the wave equation," *Math. Comput.* **49**, pp. 65–90, July 1987.
7. J. P. Berenger, "Perfectly matched layer for the FDTD solution of wave-structure interaction problems," *IEEE Trans. Antennas and Propagat.* **44**, pp. 110–117, January 1996.
8. E. I. Thorsos, "The validity of the Kirchoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Am.* **83**, pp. 78–92, 1988.
9. D. of Defence, *Mine Facts*, CD-ROM by U.S Department of Defence, 1997.
10. Y. Miyazaki, Y. Jyonori, and J. Sonoda, "Statistical characteristics of reflections and scatterings of electromagnetic radar pulses by buried objects in random underground media," 6^{th} *International Conference on Ground Penetrating Radar* **1**, pp. 151–154, October, 1996.
11. A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 1984.
12. T. Dogaru and L. Carin, "Time-domain sensing of targets buried under a rough air-ground interface," *IEEE Trans. Antennas Propagation* **46**, pp. 360–372, 1998.
13. S. M. Kay, *Modern Spectral Estimation*, Prentice Hall, New Jersey, 1987.
14. S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 1986.
15. C. C. Chen, J. D. Young, L. Peters, and F. Paynter, "Radar signature measurements of buried, unexploded ordance targets," tech. rep., The Ohio State University, ElectroScience Laboratory, February 1994.
16. D. Carevic, M. Craig, and I. Chant, "Modelling GPR echos from land mines using linear combinations of exponentially damped sinusoids," *Proc. of SPIE* **3079**, pp. 568–580, 1997.
17. I. J. Gupta, A. van der Merwe, and C. C. Chen, "Extraction of complex resonances associated with buried targets," *Proc. of SPIE* **3079**, pp. 568–580, 1998.
18. M. P. Hurst and R. Mittra, "Scattering center analysis via Prony's method," *IEEE Trans. Antennas Propagation* **AP-35**, pp. 986–988, August 1987.
19. C. A. Balanis, *Advanced Engineering Electromagnetics*, Wiley, 1989.
20. H. L. van Trees, *Detection, Estimation and Modulation Theory*, Wiley, 1969.

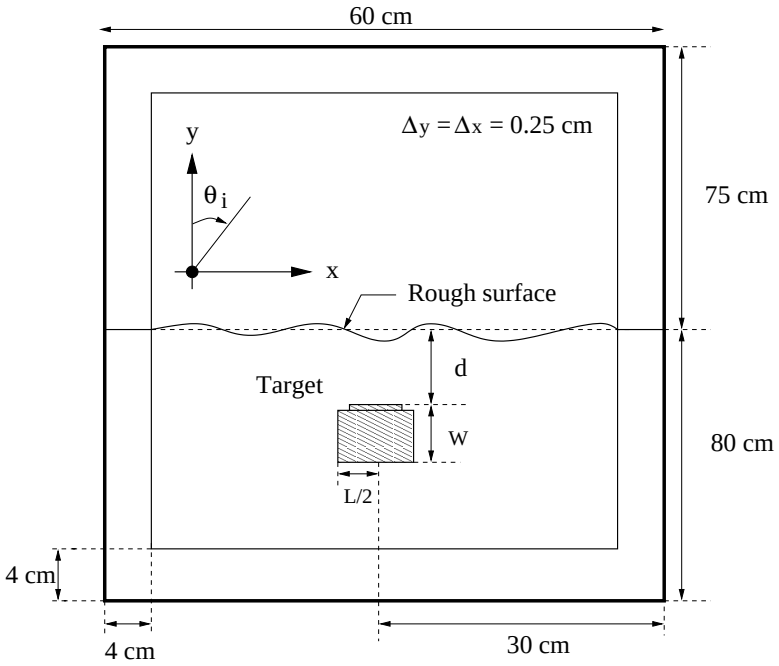


Figure 1. Schematic of the FDTD computational domain.

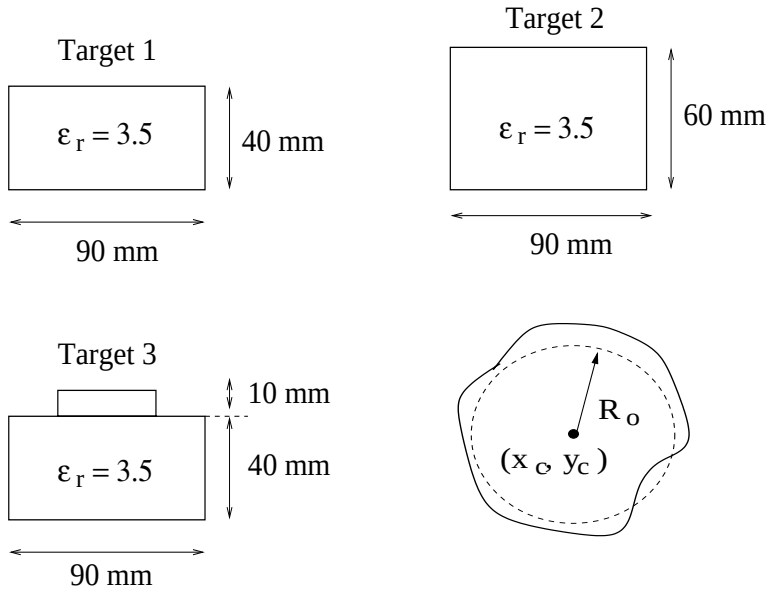


Figure 2. Cross sections of the three targets and of the subsurface inhomogeneity.

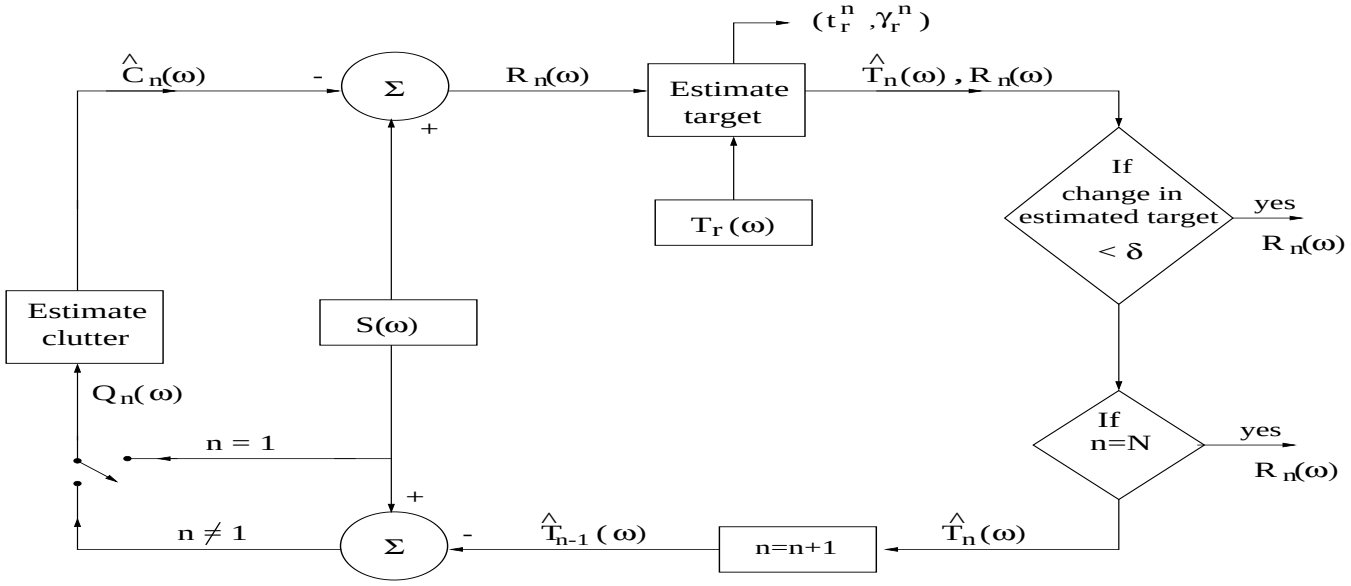


Figure 3. Schematic representation of the iterative clutter reduction procedure.

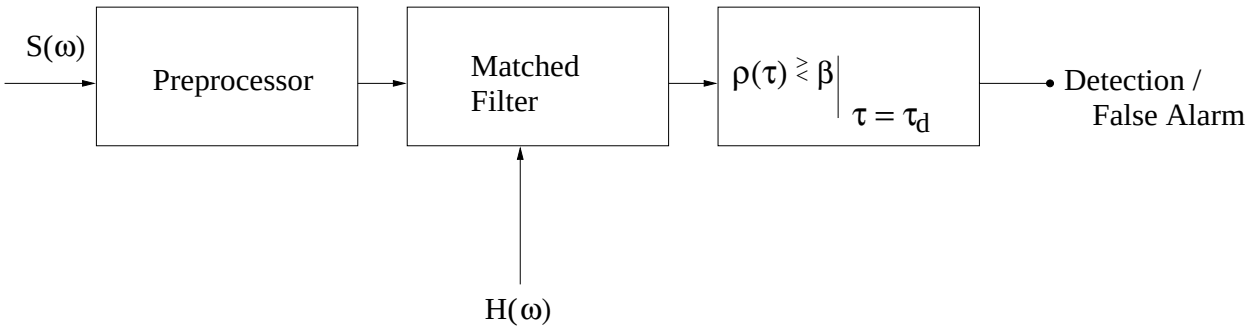
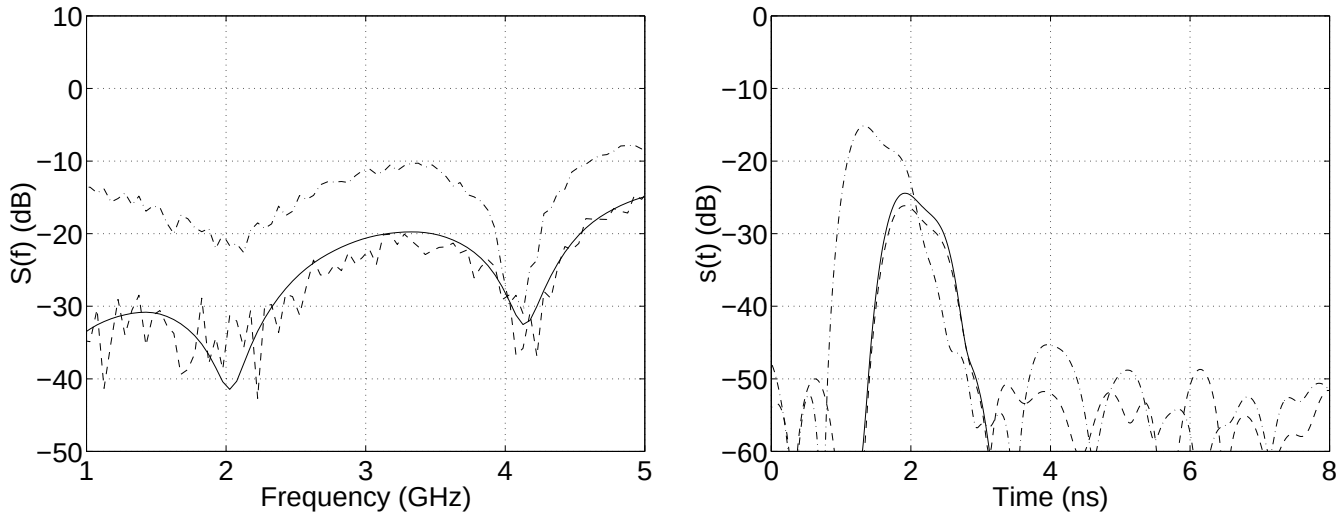
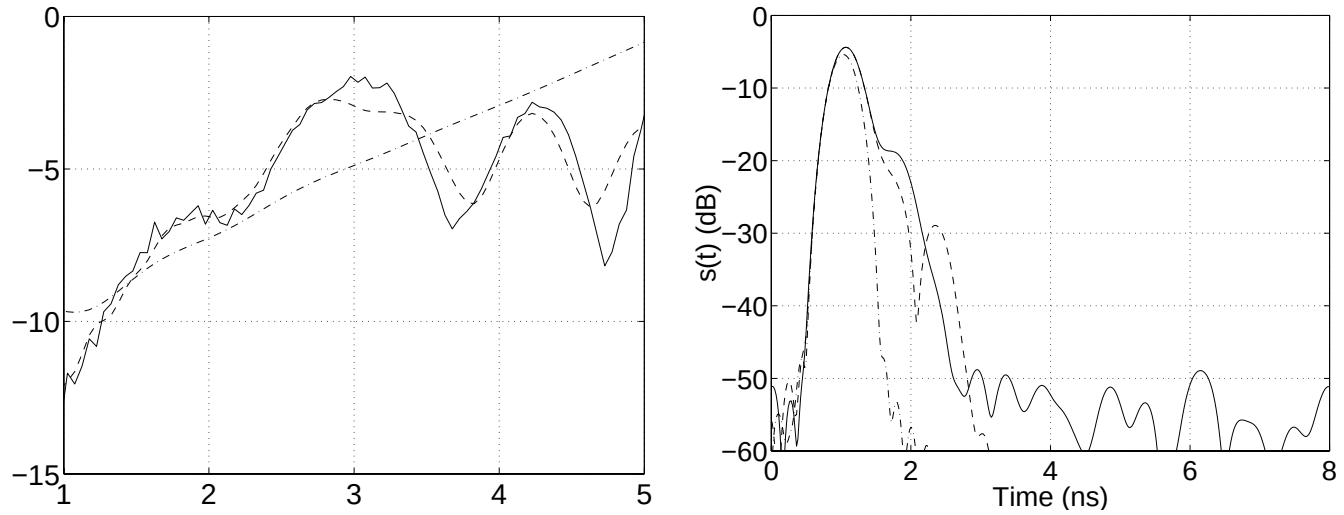


Figure 4. Block diagram to illustrate the formulation of the subsurface target detector.



(a) (b)

Figure 5. Comparison of the residual data to the reference signature in the (a) frequency and (b) time domains after a single iteration and after multiple iterations of the clutter reduction algorithm. The reference signature is shown using the solid line. The residual data after a single iteration and after multiple iterations are shown using the dash-dotted and the dashed lines respectively.



(a) (b)

Figure 6. Comparison of the estimated clutter to the total data in the (a) frequency and (b) time domains after a single iteration and after multiple iterations of the clutter reduction algorithm. The reference signature is shown using the solid line. The residual data after a single iteration and after multiple iterations are shown using the dash-dotted and the dashed lines respectively.

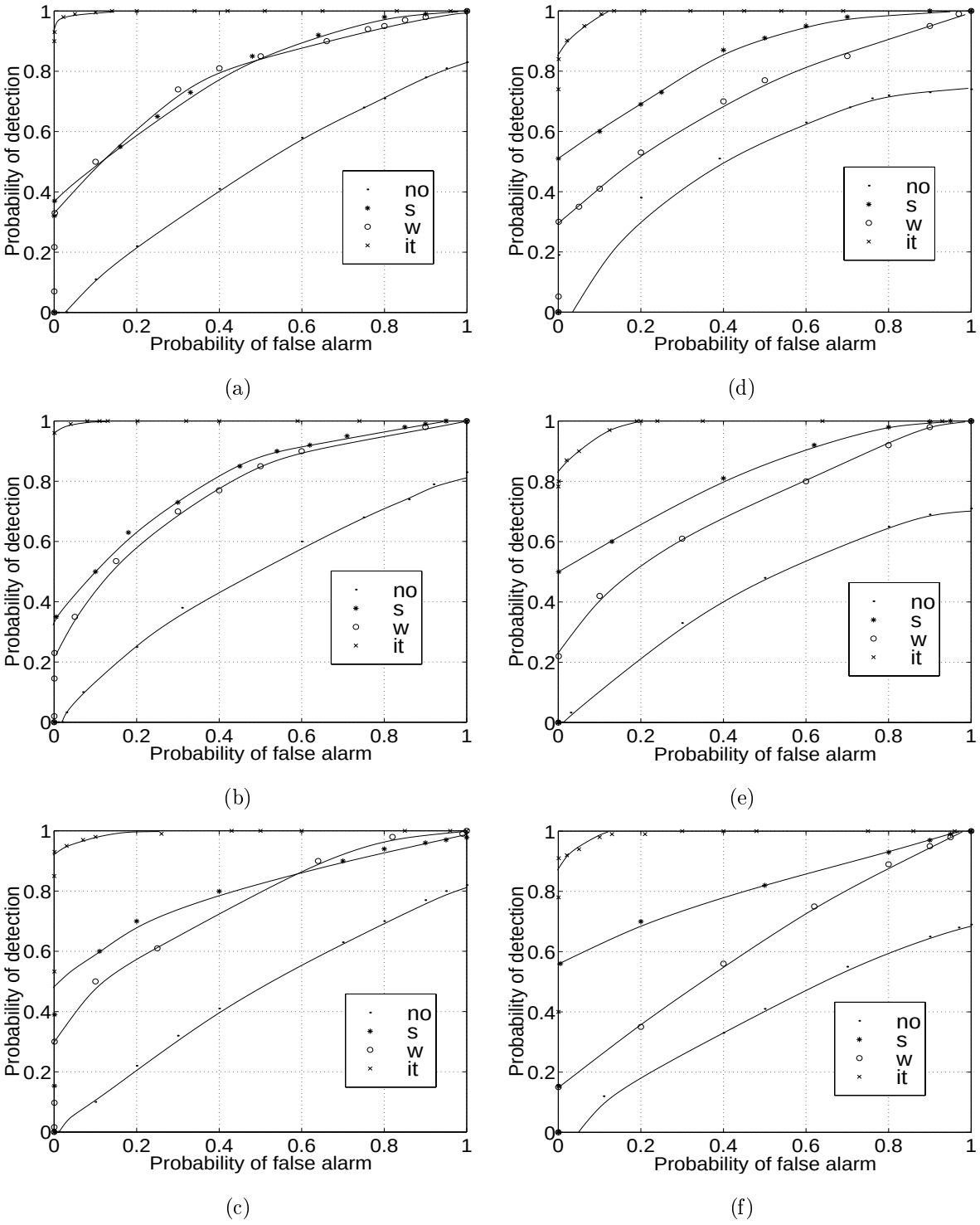


Figure 7. Comparison of the matched filter performance on the detection of (a)Target 1 (b)Target 2 and (c)Target 3 at an SNR = 40 dB and of (d)Target 1 (e)Target 2 and (f)Target 3 at an SNR = 30 dB. The various preprocessors are indicated by the following symbols, $\cdot$ indicates the no preprocessing cases, $\circ$ indicates the whitening filter, $*$ indicates time gating and $\times$ indicates the iterative algorithm.

Techniques for Improving Buried Mine Detection in Thermal IR Imagery

İbrahim Kürşat Şendur and Brian A. Baertlein

The Ohio State University, ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

We describe sensor-based and signal-processing-based techniques for improving the detection of buried land mines in thermal infrared imagery. Results of experimental studies using MWIR (2.2-4.6 μm) and LWIR (8-12 μm) imaging systems are reported. Thermal clutter due to surface reflected sunlight and skylight are investigated and shown to be the dominant clutter component for both MWIR and LWIR imagery collected during daylight hours. A sensor-based clutter reduction technique, spectral differencing, was considered and found to provide some benefit. The temporal evolution of thermal signatures was investigated. The imagery are found to have near-Gaussian statistics, and therefore the deflection coefficient is a valid measure of detectability. The deflection coefficient for some buried mines was found to improve with time after sunset. In addition, the LWIR band appears to offer some advantages in detection. Clutter mitigation via signal processing is also explored using an "estimator-classifier" technique in which target-related parameters (features) are estimated from the data and detected with a classifier. The theoretical basis of the method is discussed. MWIR and LWIR imagery are used to illustrate both the sensor-based and signal-processing-based techniques.

Keywords: Thermal infrared imagery, clutter, land mines, LWIR, MWIR

1. INTRODUCTION

In this work we examine methods for improving detection of buried mines in thermal infrared images. Naturally occurring clutter has long been recognized as a performance limiting factor in IR mine detection. A significant problem in dealing with this clutter is that it derives from several largely independent sources including variations in soil temperature, surface emissivity, and reflected light. These clutter sources are affected by diverse natural phenomena including variations in soil mineral composition and moisture content, inhomogeneous ground cover, the vitality of overlying vegetation, surface topography, and shadowing.

Several techniques for mitigating clutter are considered here. In Section 2 we describe clutter reduction techniques related to sensors and their use. Tests at our facility, reported here, indicate that during daylight hours surface-reflected sunlight and skylight are the dominant clutter sources both at MWIR and LWIR. In Section 2.1 we discuss the nature of reflected-light clutter. Supporting experimental data were acquired on a surrogate mine field constructed at our facility. In Section 2.2 we describe the time evolution of thermal imagery. The behavior of the mine signature and the surrounding soil are compared. In Section 2.3 we describe the concept of spectral differencing, in which surface reflectivity is estimated at one band and used to remove reflected-light clutter at another band. Measurements are presented to illustrate the effectiveness of the concept. Finally, in Section 3 we discuss a signal-processing approach to clutter rejection. The technique presented here is a two-stage process in which a parametric model of the mine signature is fit to the data. The parameters are then used in detection via a classifier trained on measurements. Concluding remarks appear in Section 4.

Send correspondence to B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

2. SENSOR-BASED TECHNIQUES

2.1. The Role of Reflected Light

In this section we consider the sources of surface-reflected IR clutter and we present some imagery to illustrate the phenomenon. Some important IR clutter mechanisms are shown in Figure 1. The received radiance L_R [W m^{-2} sr^{-1}] at an IR sensor can be written as

$$L_R(\lambda, x, y) = \rho(\lambda, x, y)L_{SUN}(\lambda) + \rho(\lambda, x, y)L_{SKY}(\lambda) + \epsilon(\lambda, x, y)L_T(\lambda, x, y), \tag{1}$$

where ρ is the surface reflectivity, ϵ is the surface emissivity, L_{SUN} is the radiance due to sunlight and L_{SKY} is the radiance due to skylight (sunlight scattered by particles and molecules in the earth's atmosphere and thermal radiation from the warm atmosphere). Both L_{SUN} and L_{SKY} are assumed constant over the surface, i.e., shadowing and surface topography are not considered.

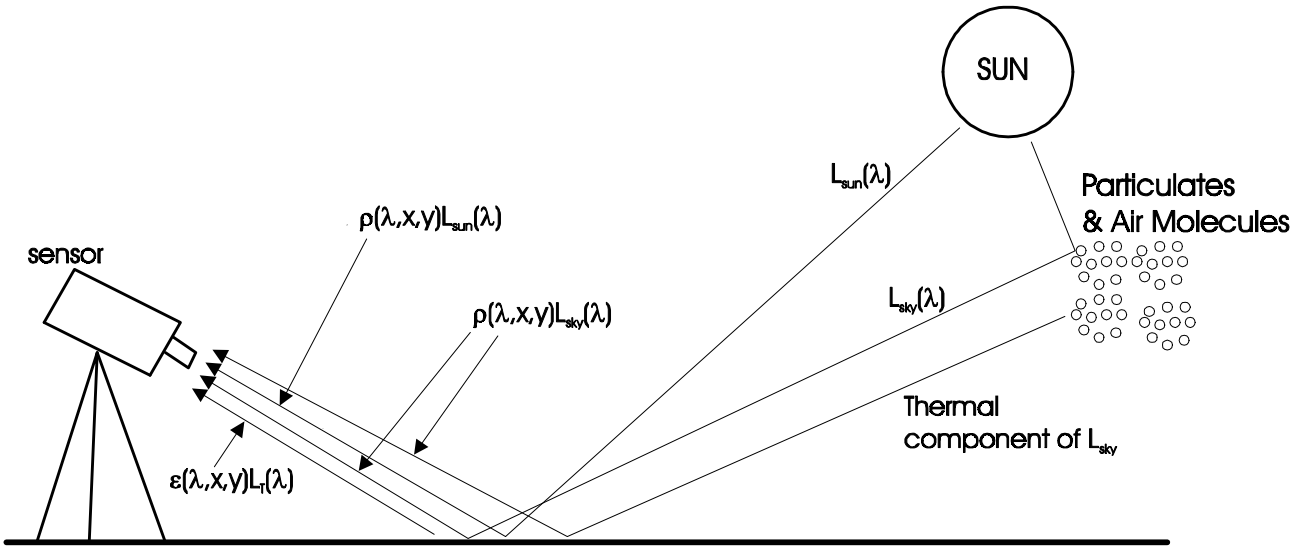


Figure 1. Contributors to a thermal image of the ground include direct sunlight, aerosol-scattered sunlight, thermal emissions from the air reflected by the ground, and thermal emissions from the soil.

For an opaque body that satisfies Kirchhoff's law, we have $\epsilon(\lambda, x, y) = 1 - \rho(\lambda, x, y)$, and defining

$$L_{SS}(\lambda) = L_{SUN}(\lambda) + L_{SKY}(\lambda) \tag{2}$$

leads to

$$L_R(\lambda, x, y) = [1 - \epsilon(\lambda, x, y)]L_{SS}(\lambda) + \epsilon(\lambda, x, y)L_T(\lambda, x, y) \tag{3}$$

Equation (3) illustrates a well known problem in IR detection of buried objects. Even if the incident radiance L_{SS} can be determined by other means, the surface emissivity ϵ and thermal radiance L_T are unknown and vary with surface position. Thus, it is essentially impossible to estimate the desired quantity L_T from a single measurement.

A study of the emissivity of natural materials has been presented by Salisbury and D'Aria.[1,2] The spectral responses of these materials are quite complicated, but in general over the 3-4 μm band rocks and soils exhibit reflectances of 5% to 30%, while vegetation has reflectances of 2% to 15%. For the 8-14 μm band they found rock reflectances of 1% to 10%. Vegetation reflectances were found to have approximately the same range.

Experiments were done to study the relative contributions of these illumination components. The experiments were performed at a surrogate minefield emplaced at the ElectroScience Laboratory (ESL) and further described elsewhere.[3] Some 40 mine-like and clutter-like targets were placed in the native Ohio topsoil and had been in situ for more that 18 months at the time these data were acquired. Surface vegetation on the mine field was completely removed. The test area was not specially conditioned for IR imaging prior to emplacing the targets. As a result,

Table 1. IR Camera Characteristics.

Parameter	IRRIS 160ST	Agema 1000
Spectral range	MWIR (2.2-4.6 μm)	LWIR (8.0-12.0μm)
Sensor type	InSb PV array	HgCdTe
Array size	160 (h) by 120(v) pixels	590 (h) by 401 (v) pixels
FOV	9.1° by 6.8°	NFOV: 5° by 3° WFOV: 20° by 13°
IFOV	1 mrad	NFOV: .15 mrad WFOV: 0.6 mrad
Dynamic Range	variable	100 K
NEΔT	0.025 K (typ.), 0.040 (max)	< 0.2 K
Pixel depth	12 bits	12 bits

there may exist buried objects or voids in this area. The surface of the mine field is planar to within an estimated variation of ± 2 inches.

As noted above, measurements were performed at different IR spectral bands to better judge the effects of wavelength on clutter. Since the spectrum of sunlight is strongest at visible wavelengths, sensors optimized for long wavelengths will be less affected by direct and reflected sunlight.* A commercial MWIR camera (Cincinnati Electronics IRRIS 160ST) and a LWIR camera (FLIR Sytems Agema 1000) were used in these tests. Sensor characteristics are shown in Table 1. The MWIR sensor offers a InSb array detector with low internally generated sensor noise, expressed as the noise equivalent temperature difference (NEΔT). This metric determines the weakest detectable signal. The LWIR camera was borrowed from the US ARMY Night Vision Laboratory for the duration of this test. This sensor is a second-generation scanning array. The spectral range of its HgCdTe detector is 8-12 μm and it produces 12 bit data with a dynamic range of 100 K. It also has the ability to switch between two fields of view.

We first illustrate the relative contributions of direct sunlight and skylight through shadowing for the MWIR sensor. We compared images of regions illuminated by direct sunlight and shadowed (i.e., without direct illumination). Figure 2 shows data collected at 11:55 on 25 July 1998. Figures 2(a) and (b) illustrate the shadowed and unshadowed IR images of the field, while Figure 2(c) shows the positions of fiducial markers, mines, and a flush-buried concrete piling. The mines in this scene are seven-inch plastic disks buried one inch in the soil. We find that it is nearly impossible to distinguish the fiducial markers and mines due to heavy clutter in the unshadowed image. There is less clutter in the shadowed image and the fiducial markers can be easily seen, but the mine signatures are not apparent. On the basis of these data we conclude that direct illumination of the scene by sunlight is the dominant MWIR clutter source, and skylight (both aerosol-scattered sunlight and thermal emission by the air) is also a significant clutter source.

The most dramatic illustration of the importance of surface-reflected light on MWIR imagery occurs when the sensor is operated at night. Figure 3 shows the results of an experiment performed on 9 June 1998 in which data were collected after sunset. Figure 3(a) shows the fiducial marker geometry, the mine positions and the concrete pilings. Figures 3(b)-(f) are the IR imagery taken at 19:10, 20:25, 21:40, 23:10 and 00:40, respectively. In these figures the fiducial marker positions are obvious, and one can also identify the mine signatures. It is evident that MWIR imagery is best captured at night to avoid surface-reflected light. Operating at night does not remove all reflected clutter, since the night sky continues to produce thermal radiation in the infrared region, which is reflected from the ground.

The role of surface-reflected light in LWIR sensors is explored in Figure 4. Figure 4(a) illustrates the region of interest including the mine positions, fiducial markers and the concrete pilings. Figure 4(b) shows the data collected on 13 October 1998 at 14:30. The direct sunlight contribution is present in this experiment. Although the fiducial markers can be distinguished, it is not possible to identify the mine positions due to heavy clutter. In Figure 4(c) data collected on 13 October 1998 at 18:00 are presented. This image is captured just after the whole mine-field is completely shadowed by an adjacent building. The image is less cluttered and it is now possible to identify the mine positions. In Figure 4(d) we present an image captured on 12 October 1998 after sunset at 21:00. In this figure the

*Using a 6000 K blackbody radiator as a model for solar radiation, we find that the solar radiance at 3 μm will be roughly two orders of magnitude stronger than at 10 μm. In addition, the reflectivity of soil is roughly an order of magnitude larger at 3 μm than at 10 μm.[2]

mines can be clearly identified. These results show that surface-reflected light is also a major source of clutter in LWIR imagery. Figure 5 shows a sequence of images taken after sunset.

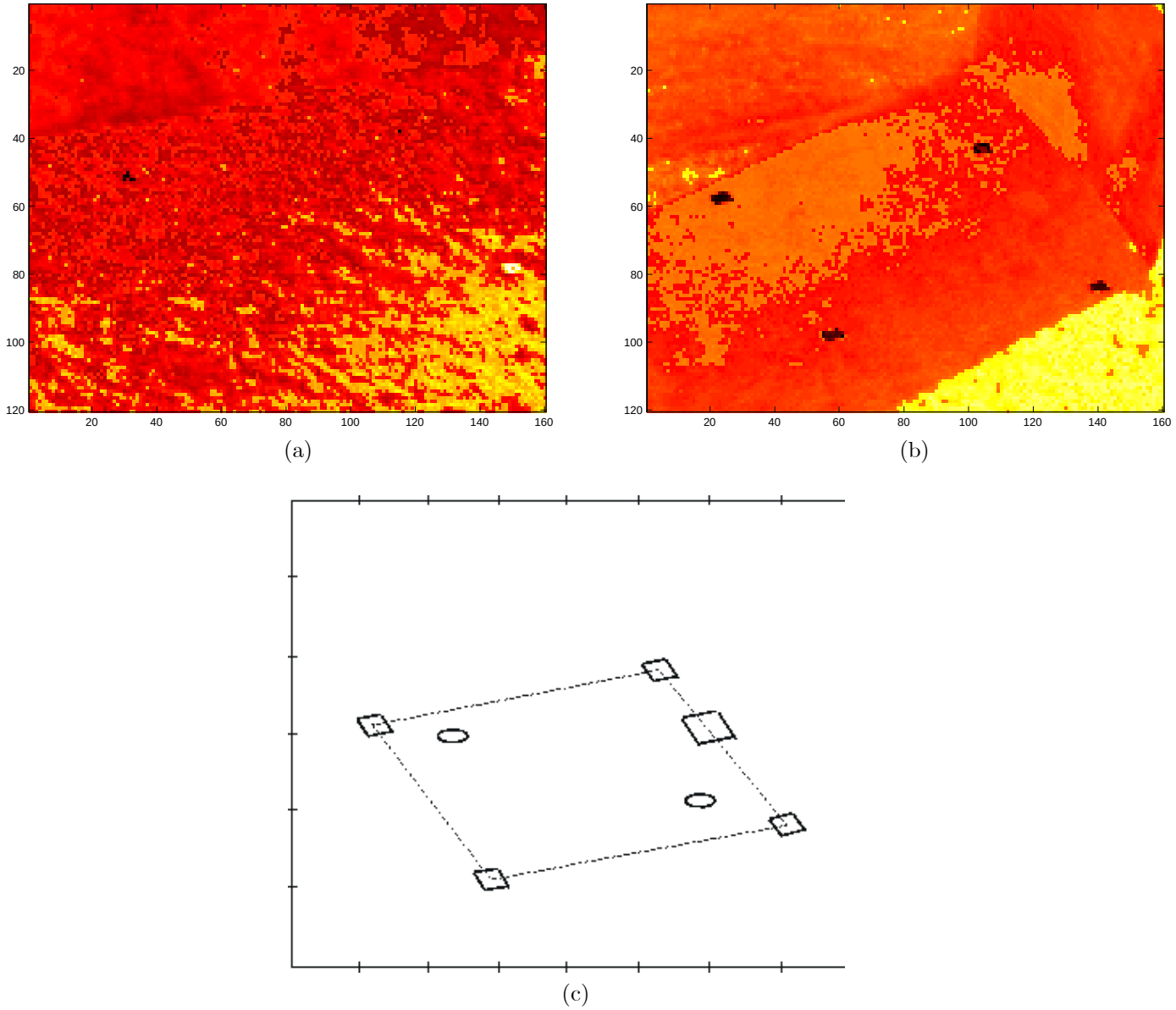


Figure 2. MWIR images collected on 25 July 1998 at 11:55 showing (a) unshadowed IR imagery, (b) shadowed IR imagery, and (c) a sketch of the region of interest including the mine positions (shaded circles), fiducial markers (small squares) and a flush-buried concrete piling (large square).

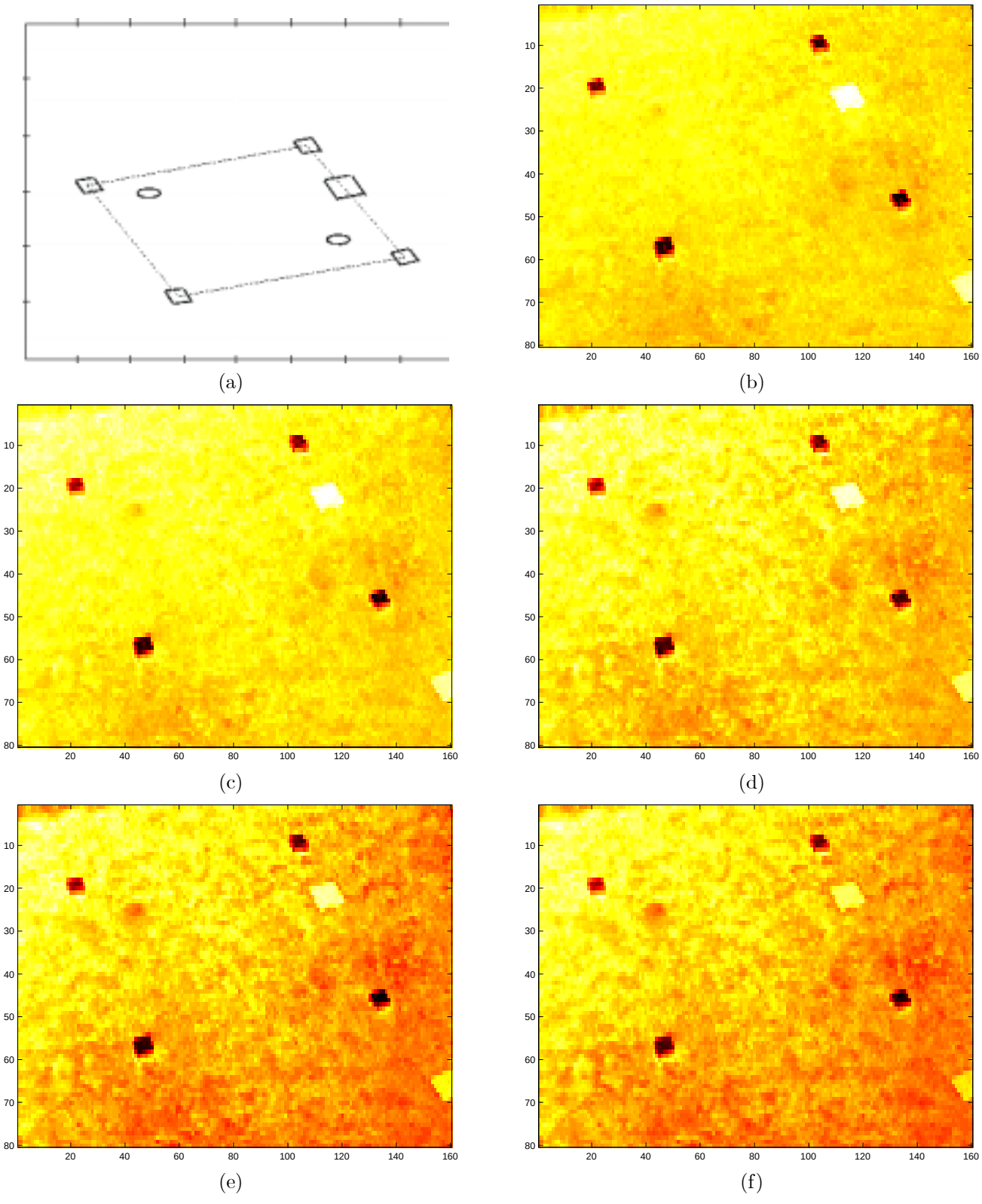


Figure 3. MWIR images collected on 9 June 1998. (a) Sketch of the region of interest including the mine positions, fiducial markers and the concrete pilings. Images were captured at (b) 19:10, (c) 19:25, (d) 20:10. (e) 20:40, and (f) 21:40.

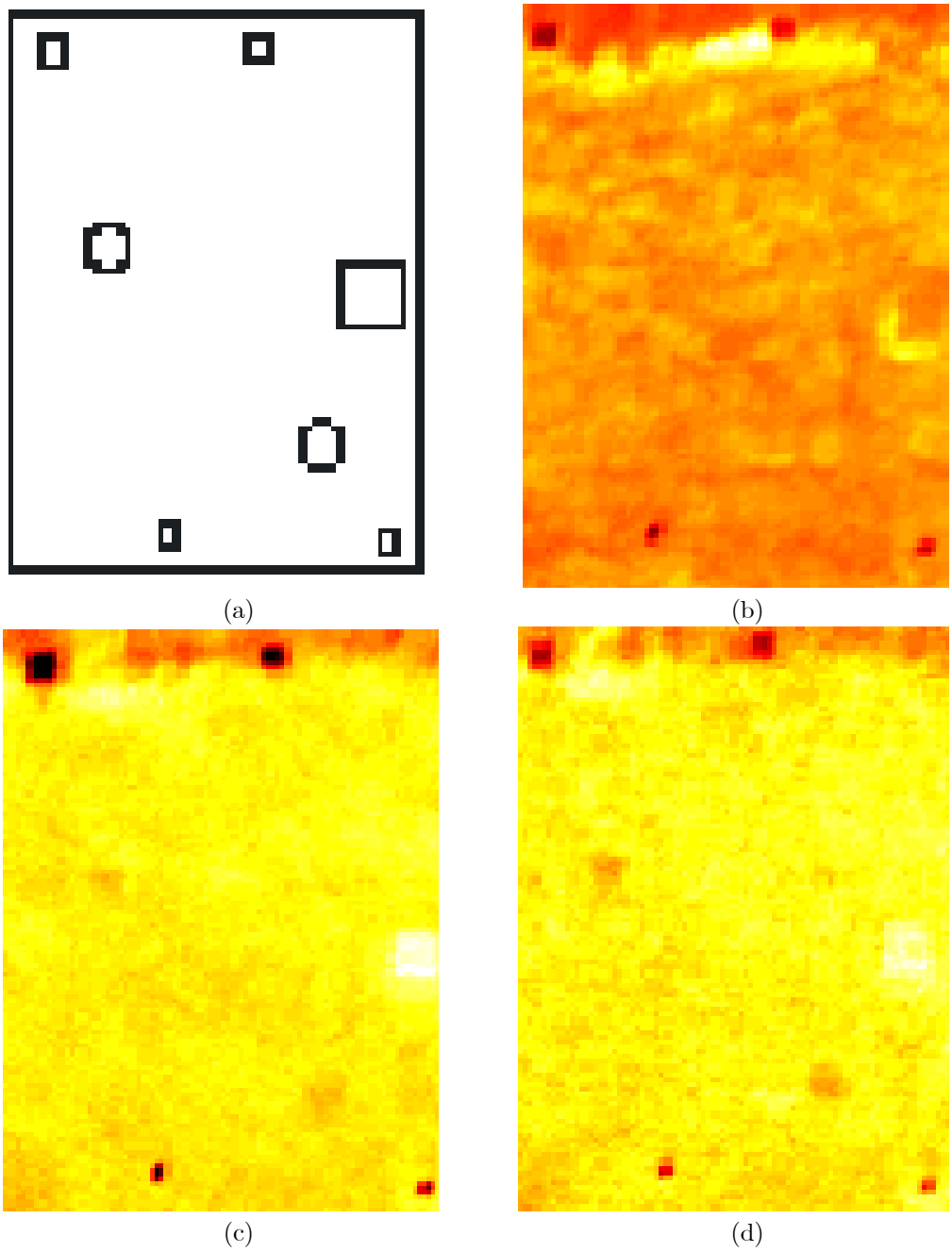


Figure 4. LWIR images captured 13 October 1998. (a) Sketch of the region of interest including the mine positions, fiducial markers and the concrete pilings. Images were captured at (b) 14:30, (c) 18:00, and (d) 21:00.

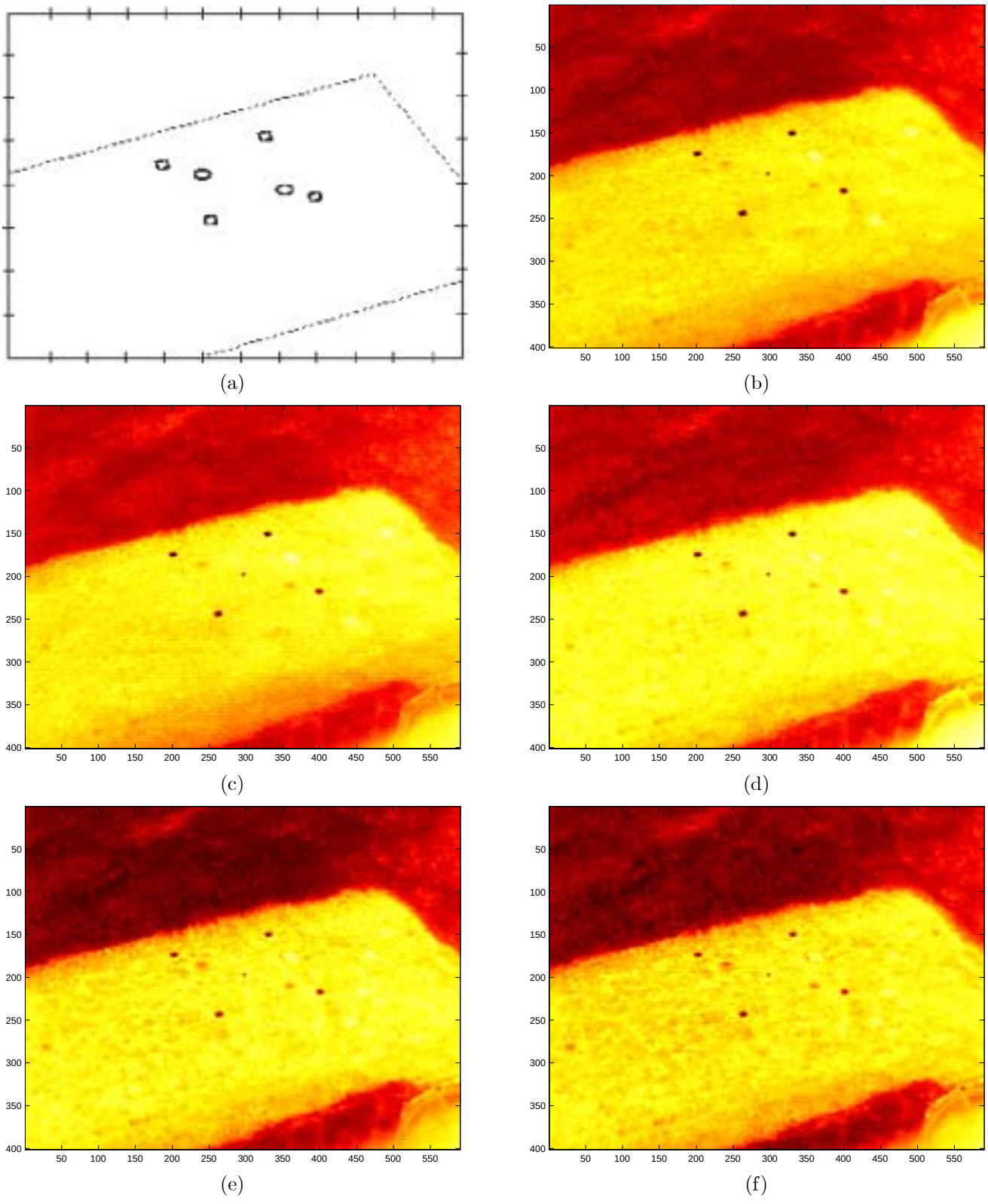


Figure 5. LWIR images captured 30 October 1998. (a) A sketch of the mine area showing the positions of mine surrogates and fiducial markers. Images were captured at (b) 18:10, (c) 19:10, (d) 20:10, (e) 21:40, and (f) 22:40.

2.2. Temporal Evolution Studies

The data displayed in Figures 3 and 5 were part of a study undertaken to identify time-dependent differences in the thermal behavior of soil over a mine with respect to the surrounding soil. Those images indicate that thermal mine signatures are dynamic phenomena and, in principle, their evolution is described by thermal models. In humanitarian demining we have the luxury of being able to acquire the time history of the minefield thermal imagery, which can provide additional information leading to improved detection.

The temperature of soil is constantly changing throughout the scene, and any mine-related temperature differences must be detected in the presence of this background change. To minimize the number of clutter phenomena present, the images are acquired after sunset. The timing of the measurements with respect to sunset is, however, an important factor as shown in the preceeding section. We collected image sequences at regularly spaced time intervals using both the MWIR and LWIR cameras. Software was developed to control the cameras and to automatically acquire imagery every 15 minute for several hours.

MWIR images were collected on 9 July 1998 starting at 19:10 (near sunset) and continuing for more than four hours. Representative samples of the data were shown in Figure 3. It can be seen that the clutter in the scene increases as time increases. The reason for this change has not been conclusively determined, but it is possible that small protruding surface features cool more rapidly than larger flat ground regions, because of the greater exposed surface area of the former.

Histograms of these images (not shown here) appear Gaussian. A test of this condition is possible by using the Kolmogorov-Smirnov (KS) statistic,[7] which permits us to determine the probability that data is drawn from a specified distribution. For these images typical levels of significance for the test are near unity, indicating a near-Gaussian distribution for the image pixels.

LWIR imagery was acquired 30 October 1998 starting at 19:10. Representative results, acquired between 18:10 and 22:40, were shown in Figure 5. The histograms of the LWIR imagery are strongly bimodal as a result of imaging both bare soil (over the mine field) and surrounding vegetation, but imagery with a single type of ground cover appears to have near-Gaussian statistics also.

It is well known that the detection and false alarm rates of targets in Gaussian noise are directly related to the "deflection coefficient" $d^2 = (\mu_1 - \mu_0)^2/\sigma^2$, which is analogous to the signal-to-noise ratio. In this expression μ_1 is the mean of the signal with the target present, μ_0 is the mean of the signal with the target absent, and σ_0^2 is the variance of the signal with the target absent. The deflection coefficient was computed for the two mines present in each of the MWIR and LWIR data shown in Figures 3 and 5, and the result is shown in Figure 6. We find that the deflection coefficient is generally increasing with time in both bands, which suggests that detection will be improved later in the evening. The MWIR data show inconsistent trends, and the LWIR data have markedly higher values. These limited data suggest that the LWIR band may offer some advantages for thermal mine detection.

2.3. Spectral Differencing

In deriving equation (3) we observed that reflected light is a significant source of clutter. Furthermore, even if L_{SS} could be determined, this equation cannot be used to estimate L_T, since the surface emissivity ϵ is unkown. In principle, two measurements can be used to eliminate the L_{SS} contribution in a straightforward manner. Another approach to dual-band operation was explored previously in a series of papers by Del Grande et al.[4,5,5] In that work it was shown that a *ratio* of measurements at two bands was directly related to soil temperature, although the effect of surface-reflected light was apparently ignored. We make the key assumption that although the emissivity may vary strongly with position, its wavelength variation is essentially independent of position.[4] Thus, the spatial and spectral variation of the emissivity can be factored as follows:

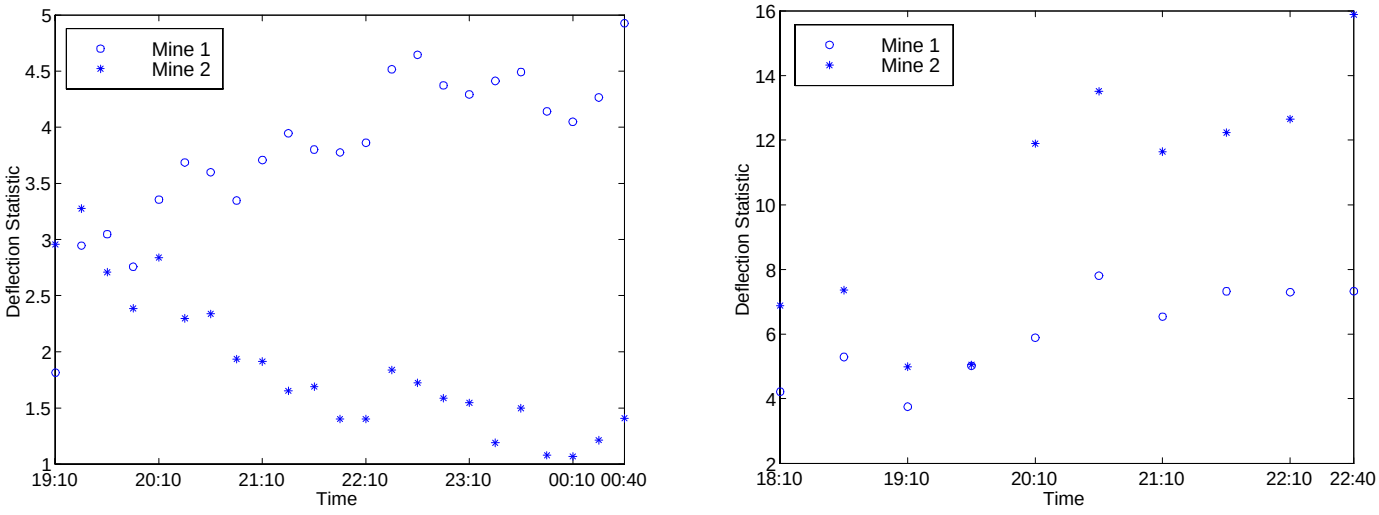
$$\rho(\lambda, x, y) \approx \rho(\lambda)A(x, y) = [1 - \epsilon(\lambda)]A(x, y) \tag{4}$$

where $A(x, y)$ is an unknown function that is independent of λ. Given measurements of a region at two wavelengths,

$$L_R(\lambda_j, x, y) = \rho(\lambda_j)A(x, y)L_{SS}(\lambda_j) + \epsilon(\lambda_j)A(x, y)L_T(\lambda_j, x, y) \qquad j = 1, 2 \tag{5}$$

we form the weighted difference image

$$\Delta L_R(x, y) \equiv L_R(\lambda_1, x, y) - C(\lambda_1, \lambda_2)L_R(\lambda_2, x, y) \tag{6}$$



(a) (b)

Figure 6. The deflection coefficient measured by (a) MWIR and (b) LWIR cameras as a function of time.

where C is a constant. If we choose

$$C(\lambda_1, \lambda_2) = \frac{\rho(\lambda_1) L_{SS}(\lambda_1)}{\rho(\lambda_2) L_{SS}(\lambda_2)} \tag{7}$$

then we have

$$\Delta L_R(x, y) = A(x, y)[\epsilon(\lambda_1) L_T(\lambda_1, x, y) - C(\lambda_1, \lambda_2)\epsilon(\lambda_2) L_T(\lambda_2, x, y)] \tag{8}$$

in which there is no contribution from the direct radiation. If one wavelength, say λ_2, is in the visible range, then this formulation simplifies slightly. We have $L_T(\lambda_2, x, y) \approx 0$ for $T \approx 300K$, and

$$\Delta L_R(x, y) = A(x, y)\epsilon(\lambda_1) L_T(\lambda_1, x, y) \tag{9}$$

Thus, we find that a weighted form of the thermal radiation ϵL_T can be determined by an appropriately weighted difference of visible and IR imagery. The appropriate value of C is seen to depend on environmental factors which are not easily measured. In this work we vary C to achieve a value that is most effective in suppressing clutter.

To explore the spectral differencing concept MWIR and LWIR images were combined. The imagery was acquired on 5 November 1998 after sunset to avoid surface-reflected clutter. The objective of differencing under these conditions is to remove the reflected radiance from the warm atmosphere. Figures 7 (a) and (b) show the original MWIR and LWIR imagery, respectively. Figures 7 (c)-(e) illustrate the compensated IR images for different C values. Increasing the value of the C increase the influence of the LWIR image. The results show that it is possible to reject a modest amount of clutter by choosing C apppropriately. Figures 7 (c)-(e) illustrate the trade-off between the maximum contrast and the minimum clutter. If we increase C further, we continue to decrease the clutter, but we also decrease the mine signature. Figure 7 (f) shows the ratio of these bands as suggested by LeSchack and Del Grande.[4] We find the performance of these methods to be almost indistinguishable.

3. ESTIMATOR-CLASSIFIER DETECTION

We also explored clutter reduction through image processing. Efforts described here show that image processing can provide an additional detection improvement beyond that achievable by sensor-based techniques.

IR signatures depend on a variety of unknown factors including target depth, target size, soil conditions, and thermal history. Thus, detection algorithms which require a known signature (e.g., matched filtering) are inappropriate. We will consider the more general case in which the target can be modeled by a known function $\mathbf{G}(\mathbf{R}; \mathbf{\Theta})$ with unknown parameters $\mathbf{\Theta}$. In this case we express the measured data $\mathbf{r}$ as

$$\begin{aligned} \mathcal{H}_0 : \mathbf{r}(\mathbf{R}) &= \mathbf{G}(\mathbf{R}; \mathbf{\Theta}) + \mathbf{n}(\mathbf{R}) \\ \mathcal{H}_1 : \mathbf{r}(\mathbf{R}) &= \mathbf{n}(\mathbf{R}) \end{aligned} \tag{10}$$

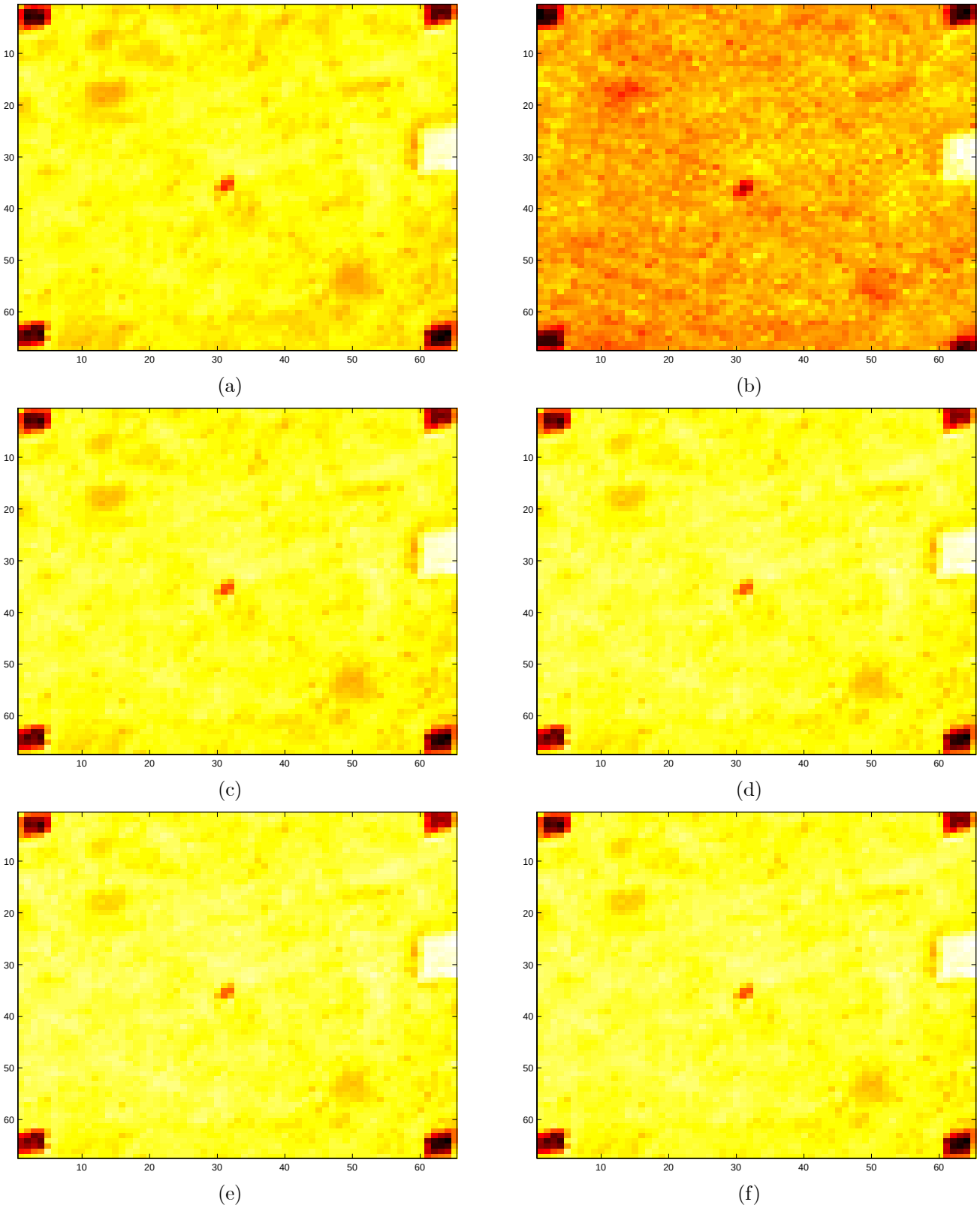


Figure 7. (a) The original MWIR image. (b) The original LWIR image. Compensated images for (c) $C = 0.03$, (d) $C = 0.05$, and (e) $C = 0.07$. Figure (f) shows the ratio of the MWIR and LWIR bands as suggested by LeSchack and Del Grande.

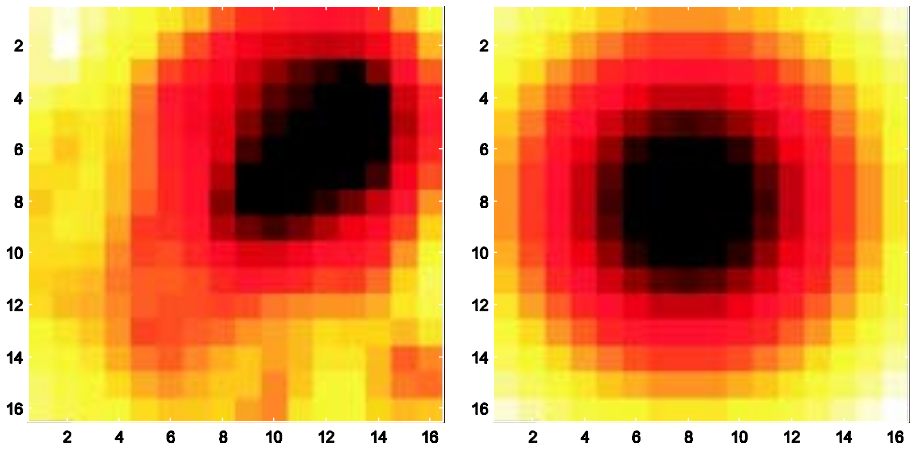


Figure 8. An example mine signature acquired with the MWIR camera (left) and the simulation $\mathbf{G}$ (right).

We assume that the components of $\mathbf{n}$ are jointly normal random variables with zero mean and covariance matrix $\mathbf{C}_n$.

For this composite hypothesis, we employ a generalized likelihood ratio test which involves $\hat{\mathbf{\Theta}}$, the maximum likelihood estimate of the unknown parameter $\mathbf{\Theta}$ given by $\hat{\mathbf{\Theta}} = \arg\min_{\mathbf{\Theta}} \mathcal{L}$. In this result the log-likelihood $\mathcal{L}$ is given by

$$\mathcal{E} = [\mathbf{r}(\mathbf{R}) - \mathbf{G}(\mathbf{R};\mathbf{\Theta})]^T \mathbf{C}_n^{-1} [\mathbf{r}(\mathbf{R}) - \mathbf{G}(\mathbf{R};\mathbf{\Theta})] \tag{11}$$

Selection of the model $\mathbf{G}$ is difficult in this case, since little is known about thermal IR mine signatures. Based on empirical data we propose the following function:

$$\mathbf{G}(\mathbf{R};\mathbf{\Theta}) = I_{back} + \frac{I_{mine}}{(1 + (\frac{|\mathbf{R}-\mathbf{R}_0|}{\alpha})^{\beta})} \tag{12}$$

where $\mathbf{R}$ is the position of the image pixel, $\mathbf{R}_0$ is the position of the mine, and the parameter vector $\mathbf{\Theta}$ is given by

$$\mathbf{\Theta} = [I_{back}\ I_{mine}\ \alpha\ \beta\ \mathbf{R}_0]^T \tag{13}$$

where I_{back} represents the background amplitude, I_{mine} represents the amplitude of the mine signal, α represents the size of the mine, and β represents the exponent of a power-law decay rate for the signature. With four parameters there is sufficient generality in this model to allow it to fit most targets. Not all of the parameters in $\mathbf{\Theta}$ are relevant to detection of the mine. We find that the parameters of greatest interest are $(I_{mine} - I_{back})$, α, β, and $\log(\mathcal{L})$. An example IR mine signature and the simulated signature $\mathbf{G}$ are shown in Figure 8.

We evaluated the performance of this algorithm on the surrogate mine field noted previously.[3] In Figures 9 the ROC curves obtained using the estimator-classifier approach are presented. In the left figure only the IR detectable targets (plastic mine surrogates and voids with diameter greater than a few inches) are declared to be valid detections. The classifier readily learns this class of target as shown by the excellent performance. Detection of all mine surrogates is a more challenging task as shown in the right figure. Many mine surrogates, because of their size or thermal properties, offer weak thermal signatures. This fact has motivated the use of sensor fusion, as we describe elsewhere.[3]

4. CONCLUSION

In this paper we have documented several investigations directed toward improving detection of buried mines in thermal IR imagery. We found that for both MWIR and LWIR sensors, surface-reflected radiation comprises the dominant source of daytime image clutter. Shadowing of the scene produced some benefit at MWIR and LWIR, but significant reflections are still seen in MWIR from aerosol-scattered sunlight and thermal radiation from the warm atmosphere. MWIR imagery acquired at night shows dramatically reduced clutter. For LWIR the benefits of night versus shadowed imagery are not as dramatic, but they are significant nonetheless. We showed that clutter can also

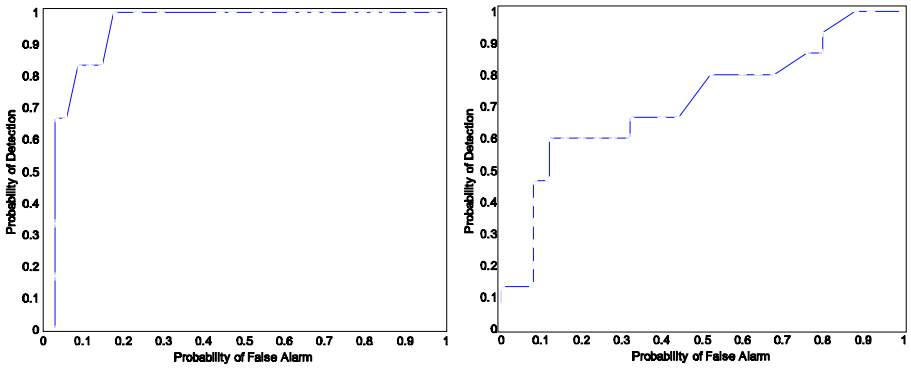


Figure 9. ROC curves obtained using the estimate-classify approach. For the curve on the left, only IR detectable objects (mine surrogates and voids larger than a few inches) were declared to be valide detections. For the curve on the right, only mine surrogates are declared to be valid detections.

be suppressed by using multi-spectral data. By forming a weighted difference of two images of the same scene it is possible to increase enhance the mine signature somewhat. Studies of the time evolution of thermal imagery suggest that the ability to detect buried mines in thermal imagery increases with time after sunset. The LWIR band appears to offer some advantages in detection.

Signal processing approaches to the mine detection problem were also considered. We developed a detection algorithm based on pattern recognition principles. In this approach signature parameters are computed and used in a classifier. The ROC curve for this approach was computed and found to be good for certain classes of mines, however, other mine types (i.e., those smaller than about three inches and those with significant metal content) were poorly detected in our tests.

ACKNOWLEDGMENTS

The authors would like to thank the US Army Night Vision Laboratory and John Moulton of E-OIR for the loan of their Agema camera. This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the authors and are not necessarily those of Duke University or the ARO.

REFERENCES

1. Salisbury, J.W. and D. M. D'Aria, "Emissivity of terrestrial materials in the 8-14 μm atmospheric window," *Remote Sensing Environment*, Vol. 42, pp. 83-106, 1992.
2. Salisbury, J.W. and D. M. D'Aria, "Emissivity of terrestrial materials in the 3-5 μm atmospheric window," *Remote Sensing Environment*, Vol. 47, pp. 345-361, 1994.
3. Gunatilaka, A. and B. A. Baertlein, "Comparison of pre-detection and post-detection fusion for mine detection," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, J. T. B. A. C. Dubey, J. F. Harvey and R. E. Dugan, eds., *SPIE* **3710-118**, 1999.
4. LeSchack, L. A., and N. K. Del Grande, "A dual-wavelength thermal infrared scanner as a potential airborne geophysical exploration tool," *Geophys.*, 41(6), p. 1318-1336, 1976.
5. Del Grande, N. K., G. A. Clark, P. F. Durbin, D. J. Fields, J. E. Hernandez, and R. J. Sherwood, "Buried object remote detection technology for law enforcement," *Surveillance Technol.*, Proc. SPIE 1479, pp. 335-351, 1991.
6. Del Grande, N. K., P. F. Durbian, M. R. Gorvad, D. E. Perkins, G. A. Clark, J. E. Hernandez, and R. J. Sherwood, "Dual-band infrared capabilities for imaging buried object sites," in *Underground and Obscured Object Imaging and Detection*, N. K. Del Grande et al. (eds.), Proc. SPIE 1942, pp. 166-177, 1993.
7. Kendall, M. G., *Kendall's Advanced Theory of Statistics: Volume 2, Classical Inference and Relationship*, 5th ed., Oxford University Press, New York, NY, 1991.

Comparison of Pre-Detection and Post-Detection Fusion for Mine Detection

A. Gunatilaka and B. A. Baertlein

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

ABSTRACT

We present and compare methods for pre-detection (feature-level) and post-detection (decision-level) fusion of multi-sensor data. This study emphasises methods suitable for data that are non-commensurate and sampled at non-coincident points. Decision-level fusion is most convenient for such data, but this approach is sub-optimal in principle, since targets not detected by all sensors will not achieve the maximum benefits of fusion. A novel feature-level fusion algorithm for these conditions is described. The optimal forms of both decision-level and feature-level fusion are described, and some approximations are reviewed. Preliminary results for these two fusion techniques are presented for experimental data acquired by a metal detector, a ground-penetrating radar, and an infrared camera.

Keywords: land mines, sensor fusion, infrared, ground penetrating radar, metal detectors

1. INTRODUCTION

Fusion of multi-sensor data for mine detection is complicated by several factors. The sensors used are diverse and produce non-commensurate data. The need for essentially perfect detection leads to high false alarm rates, and the clutter-rich environment in which these sensors operate leads to additional false alarms. Finally, the sensors may reside on different platforms, leading to problems with data registration.

One approach to these problems is to perform detection at the sensor level and to then combine the detections from individual sensors. This process, also known as "decision-level" fusion, is convenient since it reduces diverse sensor data to a common format (binary decisions or detection probabilities) that are readily combined using a variety of techniques developed for this purpose.[1,2] In principle, decision level fusion is suboptimal, since if a target is not detected by more than one sensor, it will not experience the benefits of fusion.

In this paper we discuss a feature-level fusion approach to non-coincident sensor sampling. The data are represented by a model with unknown parameters (features) and random additive clutter. Optimization techniques are used to determine the features from the available data, and classification is performed on the basis of the features. A technique for dealing with position uncertainty is also described. This approach is capable of detecting targets even when the features derived from any single sensor are insufficient for detection.

The work is organized in four major sections. In Section 2 we describe the theoretical basis for feature-level fusion. The basis for decision-level fusion is presented in Section 3. Descriptions of our sensors and techniques for sensor data processing and feature estimation are described in Section 4. Experimental data and examples of fusion are given in Section 5. Finally, concluding remarks appear in Section 6.

2. FEATURE-LEVEL FUSION OF RANDOMLY SAMPLED MULTI-SENSOR DATA

2.1. Problem Definition

Suppose that N_S countermine sensors are used to acquire data in a region. Sensor i acquires J_i data samples at locations $\mathbf{R}_{ij}$ which we denote by the row vector $\mathbf{d}_i(\mathbf{R}_{ij})$. Detection and/or classification is to be performed using the data set

$$\mathcal{D} = \{\mathbf{d}_i(\mathbf{R}_{ij}), i = 1, \ldots, N_s; j = 1, \ldots, J_i\} \tag{1}$$

We assume that different sensors produce samples that differ in number J_i, sample positions $\mathbf{R}_{ij}$, and the format and dimensionality of their data (i.e., they are non-commensurate).

Send correspondence to B.A.B. (614) 292-0076 (voice), (614) 292-7297 (fax), baertlein.1@osu.edu

Let $\mathbf{R}$ be any point at which we wish to determine the presence (or identify) of a mine. We refer to $\mathbf{R}$ as the "interrogation point." We form the K hypotheses $H_k(\mathbf{R})$, $k = 1, 2, \ldots, K$ regarding the presence or absence of various types of mines at $\mathbf{R}$. We may also include in the set $\{H_k\}$ discrete clutter objects, e.g., buried rocks, metallic clutter, etc., that might be false alarms.

At each point $\mathbf{R}$ we wish to determine the hypothesis $H_k(\mathbf{R})$ that minimizes the Bayes risk conditioned on the data $\mathcal{D}$

$$k = \arg \min_j \sum_{\ell=1}^{K} C_{j\ell} \Pr(H_\ell(\mathbf{R})|\mathcal{D}) \tag{2}$$

where $C_{j\ell}$ is the cost of deciding hypothesis H_j when H_ℓ is true. In the special case of $K = 2$ (binary detection) or when all the cost functions are equal, it is sufficient to compute likelihood ratios and to compare them to a common threshold value. In this work, however, we will use the more general criterion noted above, since multiple hypotheses may be involved and equal costs are inappropriate for the mine-detection problem.

We now suppose that data acquired by sensor i at a point $\mathbf{R}_{ij}$ over a target of type H_k at $\mathbf{R}$ can be modeled as a signal $\mathbf{G}_i(\mathbf{R}_{ij}; \mathbf{\Theta}_{ik})$ and additive noise $\mathbf{N}_i(\mathbf{R}_{ij})$

$$\mathbf{d}_i(\mathbf{R}_{ij})|H_k(\mathbf{R}) = \mathbf{G}_i(\mathbf{R}_{ij}; \mathbf{\Theta}_{ik}) + \mathbf{N}_i(\mathbf{R}_{ij}) \tag{3}$$

where $\mathbf{\Theta}_{ik}$ is a (column) vector parameter that describes the size, shape, amplitude, and position of targets of type k for sensor i. Since the discrete clutter objects have been assigned hypotheses H_k, the phenomena modeled by $\mathbf{N}_i(\mathbf{R})$ consist of instrument noise and large-scale clutter. We assume that the clutter $\mathbf{N}(\mathbf{R})$ is independent of the parameter $\mathbf{\Theta}$.

Since for each sensor a mine signature has a finite spatial extent, only a portion of the samples in $\mathcal{D}$ are relevant to interrogation in the region around $\mathbf{R}$. Let the relevant data be given by subsequences $j_1, j_2, \ldots, j_{M_i}$ of each sensor data set where M_i is the number of relevant samples for sensor i. It is convenient to assemble this reduced data set into the column vectors

$$\mathbf{D} = \left[\mathbf{d}_1(\mathbf{R}_{1j_1}) \cdots \mathbf{d}_1(\mathbf{R}_{1j_{M_1}}) \, \mathbf{d}_2(\mathbf{R}_{2j_1}) \cdots \mathbf{d}_{N_s}(\mathbf{R}_{N_s j_{M_{N_s}}})\right]^T \tag{4}$$

$$\mathbf{G}_k = \left[\mathbf{G}_{1k}(\mathbf{R}_{1j_1}; \mathbf{\Theta}_{1k}) \cdots \mathbf{G}_{1k}(\mathbf{R}_{1j_{M_1}}; \mathbf{\Theta}_{1k}) \, \mathbf{G}_{2k}(\mathbf{R}_{2j_1}; \mathbf{\Theta}_{2k}) \cdots, \mathbf{G}_{N_s k}(\mathbf{R}_{N_s j_{M_{N_s}}}; \mathbf{\Theta}_{N_s k})\right]^T \tag{5}$$

$$\mathbf{N} = \left[\mathbf{N}_1(\mathbf{R}_{1j_1}) \cdots \mathbf{N}_1(\mathbf{R}_{1j_{M_1}}) \, \mathbf{N}_2(\mathbf{R}_{2j_1}) \cdots \ldots \mathbf{N}_{N_s}(\mathbf{R}_{N_s j_{M_{N_s}}})\right]^T \tag{6}$$

If $\mathbf{d}_i(\mathbf{R})$ has length N_{Di}, then the length of these vectors is $N_D = \sum_i N_{Di} M_i$. Using this notation, we have for the vector signal model

$$\mathbf{D}|H_k(\mathbf{R}) = \mathbf{G}_k(\mathbf{\Theta}_k) + \mathbf{N} \tag{7}$$

To support decision-level fusion and feature-level fusion with uncertain sample locations we can include in $\mathbf{\Theta}_{ik}$ a position offset $\mathbf{R}_0$ that describes the nominal "center" of the mine signature with respect to $\mathbf{R}$. We write

$$\mathbf{\Theta}_{ik} = [\boldsymbol{\theta}_{ik} \; \mathbf{R}_{0,ik}]^T \tag{8}$$

In general, we will take $\boldsymbol{\theta}$ and $\mathbf{R}_0$ to be independent in what follows.

2.2. General Formulation

Feature-level fusion can be formulated in two ways. First consider the direct approach using the a posteriori densities. To minimize the risk given by equation (2) we are led to consider

$$\Pr(H_k(\mathbf{R})|\mathbf{D}) = \int d\mathbf{\Theta} \Pr(H_k(\mathbf{R})|\mathbf{\Theta}) f_{\mathbf{\Theta}|\mathbf{D}}(\mathbf{\Theta}) = \int d\boldsymbol{\theta} \int d\mathbf{R}_0 \Pr(H_k(\mathbf{R})|\boldsymbol{\theta}, \mathbf{R}_0) f_{\mathbf{\Theta}|\mathbf{D}}(\mathbf{\Theta}) \tag{9}$$

The expression in equation (9) is exact, but it requires the densities $\Pr(H_k|\mathbf{\Theta})$ (a classifier), $f_{\mathbf{\Theta}|\mathbf{D}}(\mathbf{\Theta})$ (related to an estimator of $\mathbf{\Theta}|\mathbf{D}$) and an integral over a feature space of possibly high dimensions. In many applications it is

attractive to employ approximations to this result. If the data $\mathbf{D}$ strongly imply the value $\hat{\mathbf{\Theta}}$, then $f_{\mathbf{\Theta}|\mathbf{D}}$ will be strongly peaked about $\hat{\mathbf{\Theta}}$ and we have a maximum a posteriori approximation

$$\Pr(H_k(\mathbf{R})|\mathbf{D}) \quad \approx \quad \Pr(H_k(\mathbf{R})|\hat{\mathbf{\Theta}}) \tag{10}$$

Conversely, if $\mathbf{D}$ provides no information about $\mathbf{\Theta}$ then $f_{\mathbf{\Theta}|\mathbf{D}}(\mathbf{\Theta}) = f_{\mathbf{\Theta}}(\mathbf{\Theta})$ and we recover the a priori probabilities.

A further approximation will permit us to separate the effects of features and position offsets. If we take the features $\boldsymbol{\theta}$ and the position $\mathbf{R}_0$ to be independent when conditioned on $H_k(\mathbf{R})$, then we obtain

$$\Pr(H_k(\mathbf{R})|\hat{\mathbf{\Theta}}) = \Pr(H_k(\mathbf{R})|\hat{\boldsymbol{\theta}}) \frac{f_{\mathbf{R}_0|H_k(\mathbf{R})}(\mathbf{R}_0)}{f_{\mathbf{R}_0}(\mathbf{R}_0)} \tag{11}$$

In the absence of any a priori information about mine position, we take $f_{\mathbf{R}_0}$ to be uniform over the region of interest.

Feature-level fusion can also be formulated using the class-conditioned densities. This approach has advantages when models for such densities exist. From Bayes' rule we can write

$$k = \arg\min_j \sum_{\ell=1}^{K} C_{j\ell} \Pr(H_\ell(\mathbf{R})) f_{\mathbf{D}|H_\ell(\mathbf{R})}(\mathbf{D}|H_\ell(\mathbf{R})) \tag{12}$$

Note that this formulation requires knowledge of the priors $\Pr(H_k(\mathbf{R}))$, which are sometimes difficult to obtain. Again using Bayes' rule, we find

$$f_{\mathbf{D}|H_k(\mathbf{R})}(\mathbf{D}) \quad = \quad \int d\mathbf{\Theta} f_{\mathbf{\Theta}|H_k(\mathbf{R})}(\mathbf{\Theta}) f_{\mathbf{D}|\mathbf{\Theta},H_k(\mathbf{R})}(\mathbf{D}) \tag{13}$$

Since $\mathbf{G}$ is presumed known, the data $\mathbf{D}$ conditioned on the parameter $\mathbf{\Theta}$ have the same density as the clutter $\mathbf{N}$, which leads to

$$f_{\mathbf{D}|\mathbf{\Theta},H_k(\mathbf{R})}(\mathbf{D}) = f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta})) \tag{14}$$

and

$$f_{\mathbf{D}|H_k(\mathbf{R})}(\mathbf{D}) \quad = \quad \int d\mathbf{\Theta} f_{\mathbf{\Theta}|H_k(\mathbf{R})}(\mathbf{\Theta}) f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta})) \tag{15}$$

Once again we have an exact result, but the cost of evaluating equation (15) can be high. As done above, we obtain an approximate expression by assuming that $f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta}))$ has a well defined maximum for $\mathbf{\Theta} = \hat{\mathbf{\Theta}}$. We find

$$f_{\mathbf{D}|H_k(\mathbf{R})}(\mathbf{D}) \approx f_{\mathbf{\Theta}|H_k(\mathbf{R})}(\hat{\mathbf{\Theta}}) \int d\mathbf{\Theta} f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta})) = (const) f_{\mathbf{\Theta}|H_k(\mathbf{R})}(\hat{\mathbf{\Theta}}) = (const) f_{\boldsymbol{\theta}|H_k(\mathbf{R})}(\hat{\boldsymbol{\theta}}) f_{\mathbf{R}_0|H_k(\mathbf{R})}(\hat{\mathbf{R}_0}) \tag{16}$$

where the constant term depends only on the data $\mathbf{D}$ and in the last step we have made the assumption that $\boldsymbol{\theta}$ and $\mathbf{R}_0$ are conditionally independent.

3. DECISION-LEVEL FUSION

The formulation of an optimal decision-level fusion algorithm is similar to that described in Section 2 for feature-level fusion. As before, each sensor i collects data $\{\mathbf{d}_i(\mathbf{R}_{i1}), \ldots, \mathbf{d}_i(\mathbf{R}_{iJ_i})\}$. On the basis of these data, it makes a declaration u_i regarding the truth of intermediate hypotheses $h_{iq}(\mathbf{R})$ for $q = 1, 2, ..., Q_i$. The intermediate hypotheses may be selected on the basis of the sensors involved. For example, for a sensor suite that comprises EMI, GPR and IR we might use h_{11} ="metallic object present," h_{12} ="metallic object not present," h_{21} ="dielectric discontinuity present," h_{22} ="dielectric discontinuity not present," h_{31} ="thermal discontinuity present," and h_{32} ="thermal discontinuity not present," rather than simply "mine present" and "mine not present." To simplify the discussion we will only treat the case in which all the Q_is are equal. The single-sensor detections u_i can be obtained by any process, including that described in Section 2, if in that discussion the data set is restricted to one sensor only. To simplify the notation we form the vector of local declarations given by

$$\mathbf{u} = [u_1 \; u_2 \; \ldots \; u_{N_s}]^T \tag{17}$$

We wish to combine these declarations into a fused declaration u_0 which describes the truth of hypotheses $H_k(\mathbf{R})$, $k = 1, 2, ..., K$. We will assume the following: (1) the fusion processor has no knowledge of the sensor data $\mathbf{D}$ and u_0 is based only on the declarations $\mathbf{u}$, (2) the association problem has been solved so that a putative detection $H_k(\mathbf{R})$ is not confused with another detection $H_k(\mathbf{R}')$, (3) for each sensor we know the detection confidences $\Pr(u_i|H_k(\mathbf{R}))$ for all k, and (4) the a priori probabilities $\Pr(H_k(\mathbf{R}))$ are known. For optimal fusion we must minimize the risk of the decision, which for a fixed value of $\mathbf{u}$ is equivalent to minimizing

$$\arg\min_j \sum_{\ell=1}^{K} C_{j\ell} \Pr(H_\ell|\mathbf{u}) = \arg\min_j \sum_{\ell=1}^{K} C_{j\ell} \Pr(\mathbf{u}|H_\ell) \Pr(H_\ell) \tag{18}$$

There are several approaches to decision fusion. The relation between u_0 and the individual declarations u_i, which is expressed by $\Pr(u_0 = j|\mathbf{u})$, is a mapping from the N_s-fold product of the space $\{1, 2, ..., Q\}$ to one of K output hypothesis values. There are $K^{Q^{N_s}}$ such functions, from which we must determine the most effective rule. Even for the case of binary decisions with $K = Q = 2$, this leads to a large number of possible mappings. Several ad hoc mappings from the u_i to u_0 have been used (some with considerable success) in performing fusion. These include logical AND and OR operations, as well as majority voting.

In some situations we can determine the optimal mapping.[1] Here we consider the simplest case, in which the local decision rules (i.e., the mappings from $\mathbf{D}$ to the local decisions u_i) are fixed. For $K = Q = 2$ the minimum in equation (18) reduces to a likelihood ratio test, given by

$$\frac{\Pr(u_1, u_2, \ldots, u_{N_s}|H_1)}{\Pr(u_1, u_2, \ldots, u_{N_s}|H_0)} \gtrless \eta \tag{19}$$

where η is a constant threshold and the cases $>$ and $<$ correspond to the decisions $u_0 = 1$ and $u_0 = 0$ respectively.

If the declarations u_i are conditionally independent for each sensor, then we have

$$\frac{\Pr(u_1, u_2, \ldots, u_{N_s}|H_1)}{\Pr(u_1, u_2, \ldots, u_{N_s}|H_0)} = \prod_{i=1}^{N_s} \frac{\Pr(u_i|H_1)}{\Pr(u_i|H_0)} \tag{20}$$

If we define miss and false-alarm probabilities for each sensor as follows:

$$P_{Mi} = \Pr(u_i = 0|H_1) \tag{21}$$

$$P_{Fi} = \Pr(u_i = 1|H_0) \tag{22}$$

then we can write the optimum test as

$$\sum_{i=1}^{N_s} \left[u_i \log\left(\frac{1 - P_{Mi}}{P_{Fi}}\right) + (1 - u_i) \log\left(\frac{P_{Mi}}{1 - P_{Fi}}\right) \right] \gtrless \log(\eta) \tag{23}$$

which is a weighted form of voting.

4. SENSORS AND SIGNAL PROCESSING

The foregoing algorithms are being used with a sensor suite that comprises a ground penetrating radar (GPR), a commercial infrared (IR) camera, and a Schiebel electromagnetic induction (EMI) sensor. Brief descriptions of these devices and their supporting signal preprocessing algorithms are presented below.

4.1. Electromagnetic Induction Sensor

4.1.1. Hardware

The EMI sensor used in this work was a Schiebel AN-19/2 pulsed-induction metal detector. The standard output signal, an audio tone, is not suitable for the modeling work described here. Instead, data were acquired with a digital oscilloscope at an internal signal within the sensor. The sensor was fitted to a linear scanner, which was programmed to acquire 56 samples over a 55 inch path. Scans were performed at six inch intervals along a track to obtain a grid of sample points over the region of interest.



Figure 1. The GPR dielectric rod antenna and scanning system.

4.1.2. Signal Processing

Signal processing for the EMI sensor is a two-step process. First, the sensor response over metal-free soil was acquired, and this background was subtracted from the waveforms collected over other locations. In principle, the response of a pulse-induction EMI sensor should be well approximated by a sum of decaying exponentials. The resulting difference signal, however, is not well approximated in this manner (especially for small mines), but its integral is a very effective detector of buried metal.

The second step involves estimation of the feature vector $\mathbf{\Theta}$. For targets not too near the sensor head and not too large compared to the loop radius, the integrated sensor response described above has a spatial variation that is approximated by the function

$$\mathbf{G}_1(\mathbf{R};\mathbf{\Theta}_1) = B + \frac{S}{1 + (||\mathbf{R} - \mathbf{R}_0||/a)^b} \tag{24}$$

where $\mathbf{R}_0$ is the location of the target's centroid, B is the response of the background, S is the amplitude of the target response and a and b are shape parameters. At each sample position $\mathbf{R}$ we fit all data within a window to this model. The size of the window is selected on the basis of the signal amplitude. A nonlinear optimization technique is then used to estimate the values of these parameters from samples within the window. The resulting feature vector is given by

$$\mathbf{\Theta}_1 = [a\ b\ S\ B\ E\ \mathbf{R}_0]^T \tag{25}$$

where $E = |\mathbf{D} - \mathbf{G}_1|^2/|\mathbf{D}|^2$ is the normalized residual error in the estimate.

4.2. Ground Penetrating Radar

4.2.1. Hardware

The GPR used in this effort was developed at The Ohio State University (OSU) ElectroScience Laboratory (ESL)[3] and is shown in Figure 1. It employs a novel dielectric rod antenna, which is scanned horizontally over the earth at a fixed height. The antenna and its supporting platform are moved along a linear track in two inch intervals, and a horizontal scan of 55-inch length and 101 samples is taken in the cross track dimensions at each position. A network analyzer is used to acquire complex reflection coefficient measurements at 51 frequencies between 1 and 6 GHz. The frequency regime from 5 to 6 GHz was found to be dominated by clutter and was not used in this analysis.

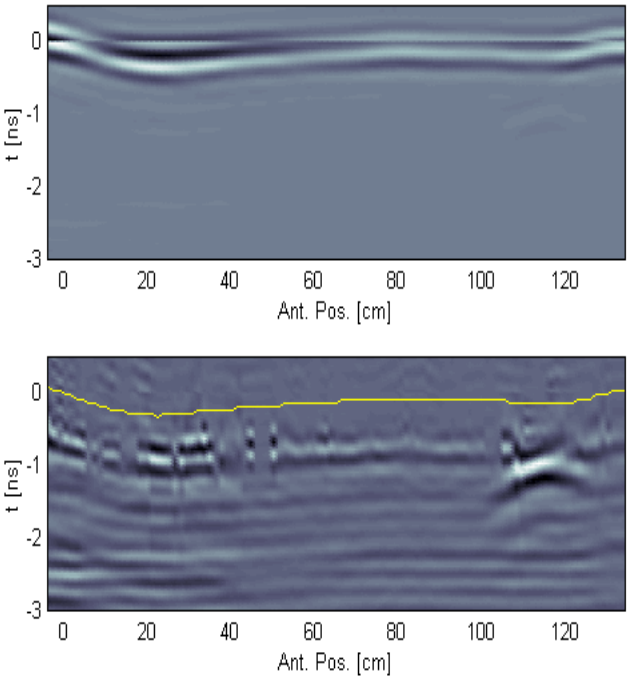


Figure 2. Example GPR data acquired over the 3.5 inch metal disk mine surrogate. The upper figure shows the time-domain data. The lower figure shows the same data with the surface reflection suppressed. The thin white line in the lower figure is the estimated position of the surface.

4.2.2. Signal Processing

A background signature (the response of the radar when the antenna points into free space) was acquired and subtracted from all data. The impulse response of the radar was determined by measuring scattering from a reference target (a short cylinder) in the absence of soil and comparing the result with a calculated scattered response. This impulse response was then deconvolved from the background-corrected measured data.

The resulting data are processed by first locating and removing the ground-reflected wave. This wave, a bandlimited approximation to an impulse response, is removed using an OSU-developed TLS-Prony technique.[4] An example of the process is shown in Figure 2 for a 3.5 inch metal mine surrogate buried at 1 inch depth. Mine signatures are detected by performing a three-dimensional convolution with a matched filter, which comprises $\mathbf{G}_2$. The filter is formed by simulating the response of a point scatterer in the soil medium at the depth of interest. GPR clutter tends to be highly correlated in the horizontal dimension (primarly because of the return from plane-stratified media), and whitening is performed prior to using this filter. The sensor feature vector $\mathbf{\Theta}_2$ used in classification is the position and maximum value of the matched filter within a window several samples wide at each position $\mathbf{R}$.

4.3. Infrared Camera

4.3.1. Hardware

We used a commercially available MWIR camera in this work. The sensor is a InSb photovoltaic array of 160 by 120 pixels. The optics provide a field of view of 9.1° by 6.8° for an IFOV of 1 mrad. The noise equivalent temperature difference for the sensor is 0.025 K. More information on the sensor and our IR measurements is available in a companion paper.[5]

The camera was positioned on the roof of the ESL building adjacent to the test site, and data were acquired from this vantage point. The camera's field of view does not permit us to sample the entire region simultaneously, and a sequence of images was necessary. Fiducial markers were placed over the region to permit registration.

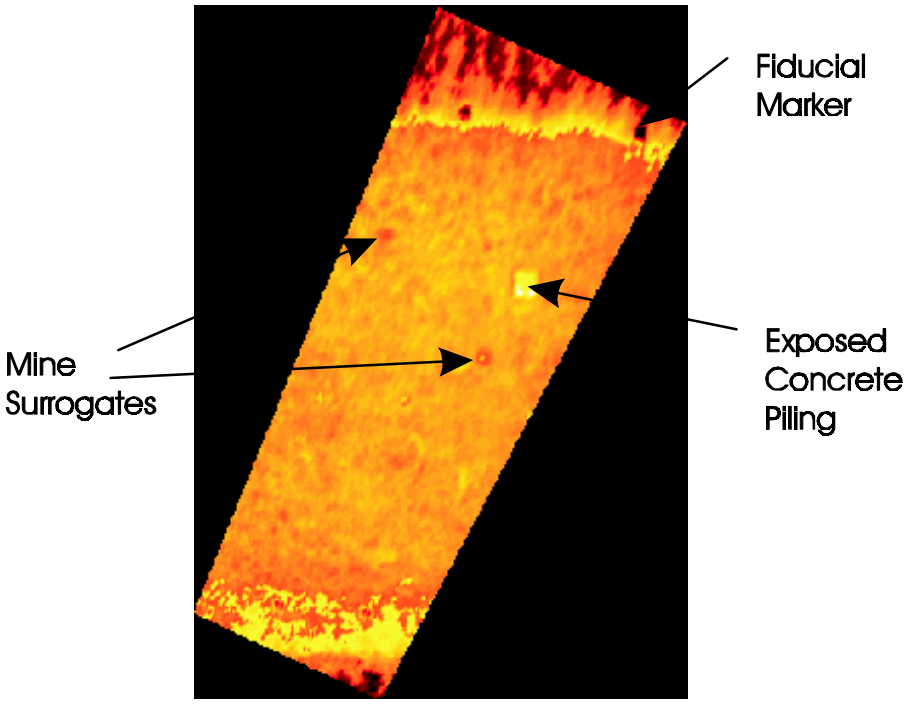


Figure 3. An example IR image after remapping to ground coordinates.

4.3.2. Signal Processing

Processing the IR data begins by remapping the imagery to ground coordinates using a standard perspective transformation. This step is followed by bilinear interpolation and resampling to a uniform pixel size. An example image is shown in Figure 3. An image chip is extracted at each interrogation point **R**. Based on empirical observations, we employ a signature model identical to that used for the EMI sensor. A nonlinear optimization process is used to determine parameter values within the chip.

5. RESULTS

The fusion procedures described above in Sections 2 and 3 have been tested on experimental data acquired by the sensors described in Section 4. These data are preliminary and require further analysis before a complete comparision of the methods is possible. The test procedures and results are described below.

5.1. Test Site

Our experimental work was performed using a simulated mine field of 40 mine-like and clutter-like surrogates. The layout of this mine field is shown in Figure 4 and the identities of the buried objects are given in Table 1. The objects in this field had been in situ for more than 18 months at the time these data were acquired.

Our objective in creating this mine field was to provide an environment for testing sensors, validating fusion algorithms, and identifying problems that arise in realistic conditions. The target set is quite challenging. Roughly half the mine surrogates contain no metal (in contrast to real mines, essentially all of which contain some metal). In addition, some of the targets are quite small (only 1.5 inches in diameter), which present detection problems for all of the sensors. Those targets (all P15 objects in Figure 4) were designated clutter in what follows.

Consistent with our initial goal of simulating realistic field conditions, the test area was not specially conditioned prior to emplacing the targets. As a result, there exists a variety of clutter and buried objects in this area. Since our tests began about two years ago, roughly 100 metallic fragments, and a number of stones have been located in (and subsequently removed from) the region. The four concrete pilings shown in the figure were concealed by a lush

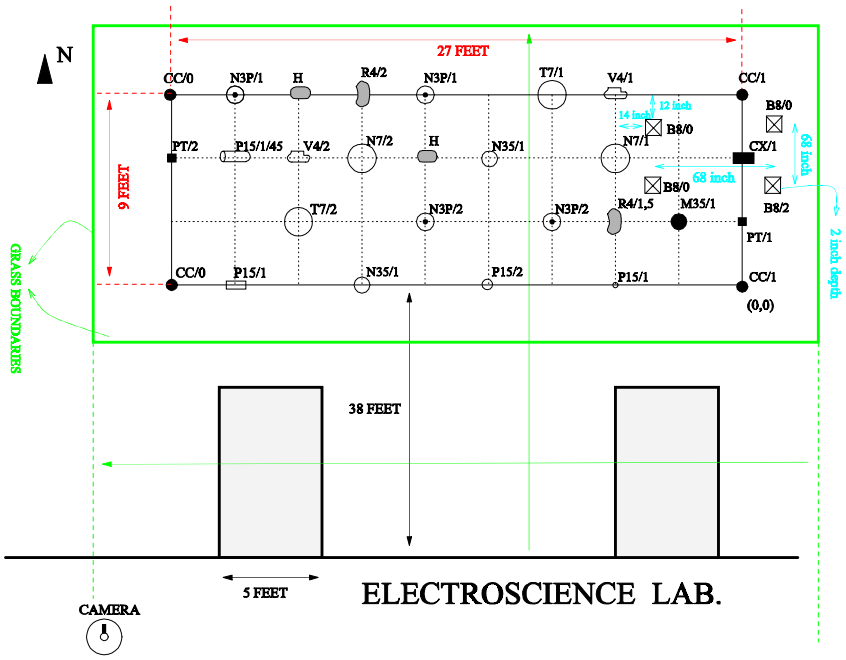


Figure 4. The layout of the ESL mine field.

grass cover that was present when the test objects were emplaced. Surface vegetation on the mine field has since been completely removed. In addition to these objects, the soil contains a large amount of clay, which affects GPR performance. The surface of the mine field contains a number of irregularities, but it is planar to within an estimated variation of ± 2 inches. As a result of this surface roughness, scanned sensors such as the EMI and GPR cannot maintain a constant height, and some undesirable variations in signal amplitude occur.

Position errors in the data arise from several sources. The positions of the targets shown in Figure 4 are accurate to within approximately one inch. Each sensor is operated on a different platform, and the platform position is known to within approximately one inch. As a result of roll, pitch, and yaw by the sensor platform, surface roughness produce errors in the sensor field of view on the order of one to two inches, for a total maximum position uncertainty of four inches. Processing of the individual sensor data sets has confirmed this estimate.

Data were acquired along the paths shown in Figure 5. The GPR samples were acquired at roughly 0.5 inch intervals along N-S lines spaced every two inches E-W, and the EMI sensor samples at a 1 inch interval along N-S lines spaced six inches E-W. The infrared camera acquired eight images as shown by the trapezoids. After remapping to ground coordinates, the pixel dimensions are approximately 0.5 inches. The small squares are IR fiducial markers, and the small circles are the target grid points shown in Figure 4.

5.2. Results for Feature-Level Fusion

Features were extracted from each of the sensors as described above, and those features were used in a classifier as indicated in equation (10). For feature-level fusion, in estimating the features $\hat{\mathbf{\Theta}}$ we did not permit the optimization process to use a variable shift $\mathbf{R}_0$. Features samples were computed on a two inch grid of interrogation points $\mathbf{R}$ over the entire mine field. The classifier $\Pr(H_k(\mathbf{R})|\hat{\mathbf{\Theta}})$ was approximated by a backpropagation neural network. We attempted only the detection problem, for which the output hypotheses were H_1 ="mine" and H_2 ="non-mine." The result of this process was compared to a threshold, leading to a binary map. Regions with small numbers of adjacent detections were discarded. The performance of the fused system is shown in Figure 6. Targets detectable by the fused sensor suite are indicated by circles, and we see that all of the mine surrogates have been detected with nine false alarms.

Table 1. Description of the objects buried in the MURI mine test grid

Abbreviation	**Object**	**Diameter** (inches)	**Height** (inches)
CC	Coke can		
CX	Crushed coke can		
PT	Pop top		
M35	Aluminum disk	3.5	5/8
P15	Plexiglass disk	1.5	2
R4	Rock	4	2
V4	Void(styrofoam)	4	2
H4	Refilled hole	6	2
N35	Nylon disk	3.5	5/8
N3P	Nylon disk with .078 diam 1 long steel pin	3	1
T7	Teflon disk	7	1
N7	Nylon disk	7	1
B8	Concrete pilings	8	varies

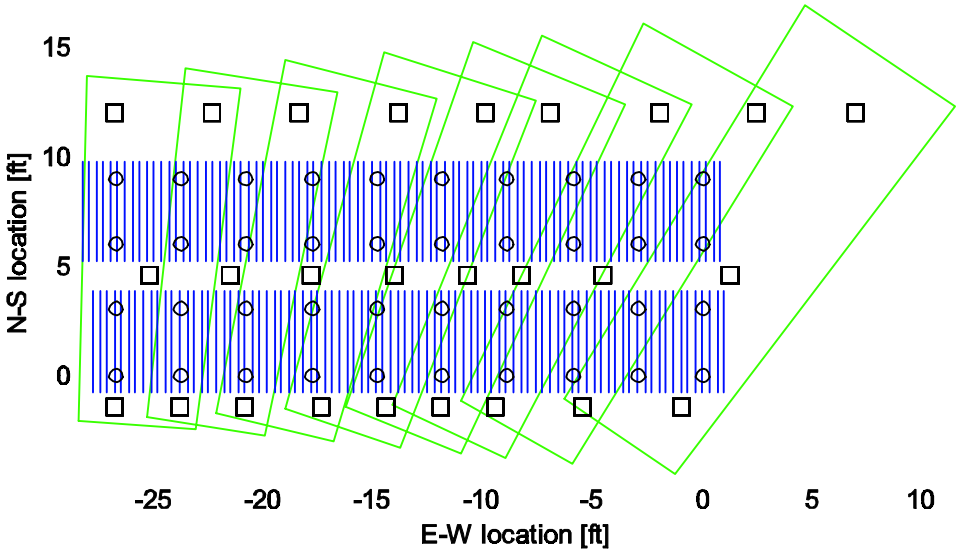


Figure 5. Locations of sample points for the sensor suite.

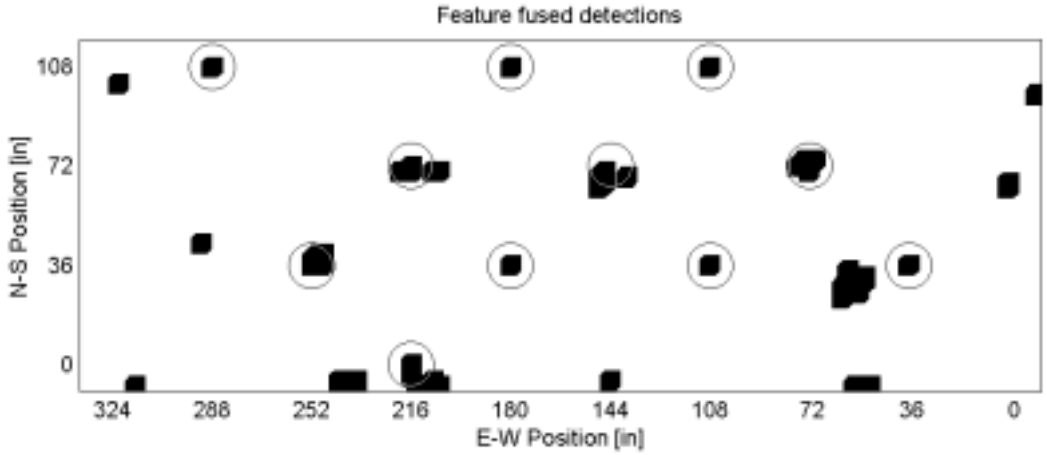


Figure 6. The results of feature-level fusion. The known targets are indicated by circles.

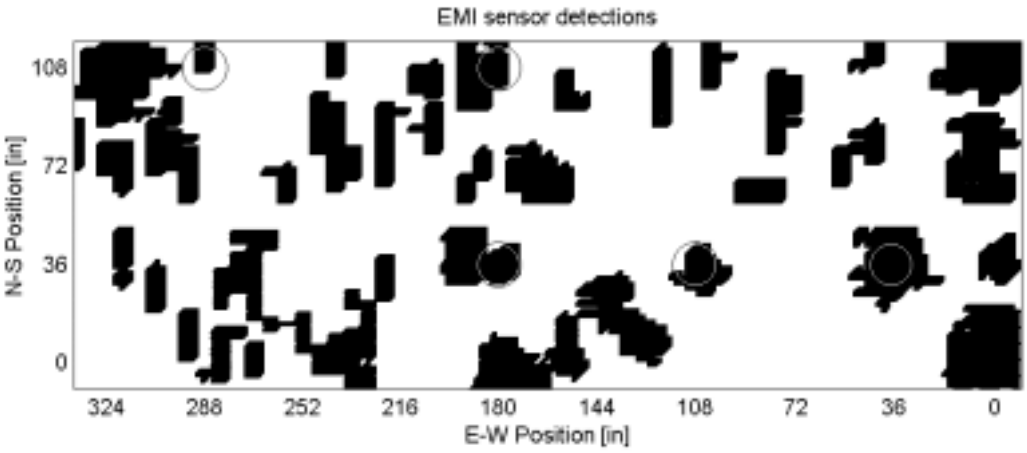


Figure 7. Detections by the EMI sensor.

5.3. Results for Decision-Level Fusion

For decision-level fusion we used the same grid of feature values described above. The resulting features $\boldsymbol{\theta}_i$ for each sensor i are applied to sensor-specific classifiers (backpropagation networks) $\Pr(h_k(\mathbf{R})|\boldsymbol{\theta}_i)$, and a target detection u_i is produced at each point. The target position information $\mathbf{R}_{0,i}$ was used to weight the classification results u_i as given in equation (16). The density $f_{\mathbf{R}_0|H_k(\mathbf{R})}$ was taken to be uniform over $|\mathbf{R}-\mathbf{R}_0|$ <4 inches and zero otherwise. This process leads to the three detection maps shown in Figure 7, 8, and 9.

Because our mine-field contains roughly equal numbers of metallic and metal-free mines, we used the EMI and GPR sensors to recognize metal-bearing objects and the IR and GPR sensors to recognize metal-free objects. A two-level fusion approach was employed. For the EMI-GPR sensor suite the intermediate hypotheses h_{11}= "metallic mine" and h_{12} ="metallic clutter" were used, while for the IR-GPR suite we used h_{21}= "non-metallic mine" and h_{22} ="non-metallic clutter." The second-level fusion involved the trivial mapping $H_1 = h_{11} \cup h_{21}$ and $H_2 = h_{12} \cup h_{22}$. The set of all mine detections is shown in Figure 10. We find that the decision-level fusion system has produced a detection rate equal to that of the feature-level fused system (all targets were detected), and its false-alarm rate is comparable. The reduction in false alarms compared with the individual sensors is significant.

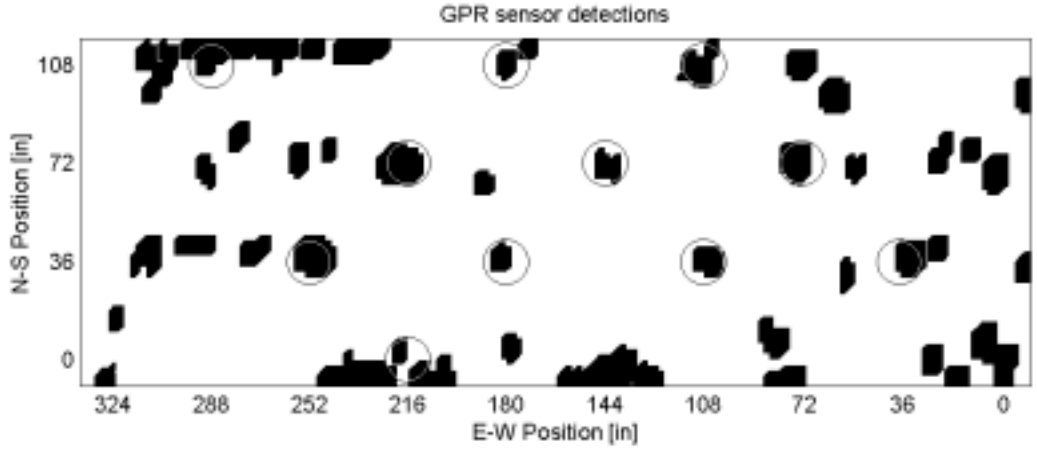


Figure 8. Detections by the GPR sensor.

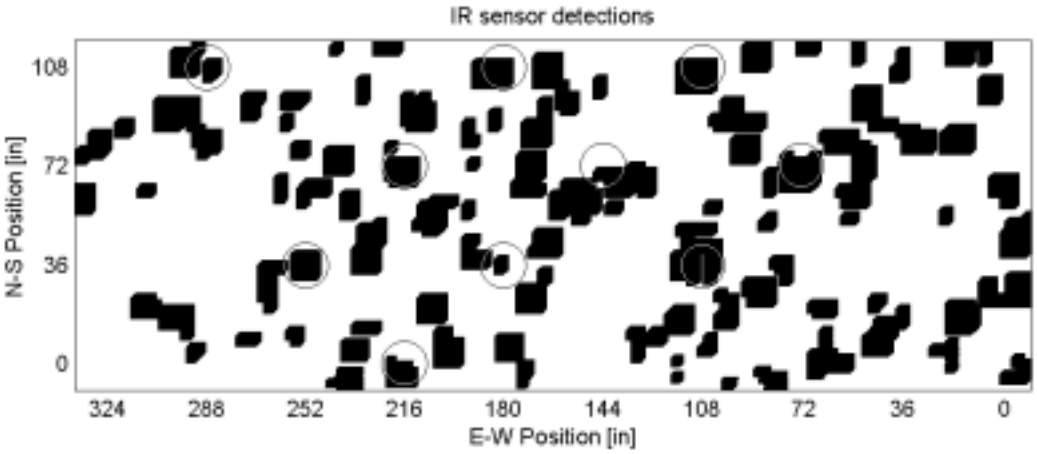


Figure 9. Detections by the IR sensor.

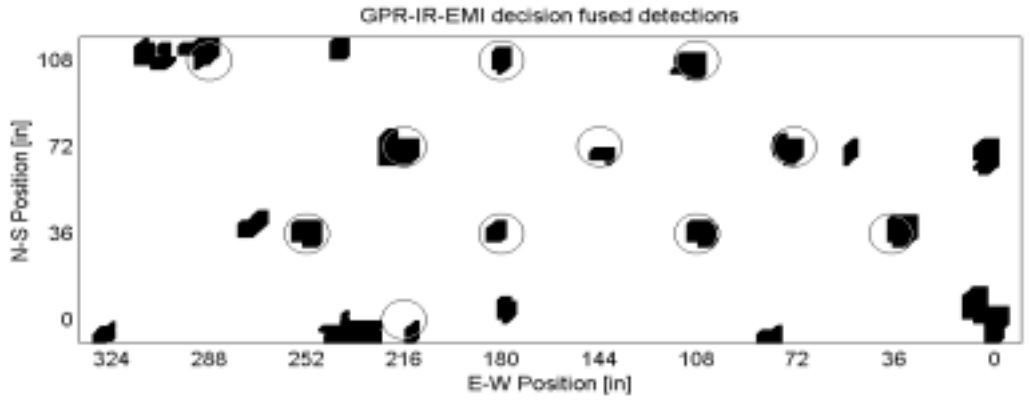


Figure 10. Detection of all mine types obtained by decision-level fusion of the EMI, GPR and IR sensors.

6. CONCLUDING REMARKS

Mine detecting sensors tend to produce data that are non-commensurate and sampled at non-coincident positions. We have developed and implemented a feature-level fusion technique for such data. The method has a rigorous basis, and with some approximations it can be reduced to form that is implemented efficiently and with a modest amount of information. The signal-processing tools required to use EMI, GPR and IR sensors for fusion have been developed.

We have conducted some preliminary tests of the algorithm and compared it to a decision-level fusion algorithm. Comparable performance was found for both forms of fusion. The theoretical work suggests that feature-level fusion may have some advantages, but further study is necessary to confirm these findings. To be more convincing, future tests of the algorithm must involve larger data sets. Development of more effective feature sets is also desirable to get maximum information from the available sensors.

ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the authors and are not necessarily those of Duke University or the ARO.

REFERENCES

1. P. K. Varshney, *Distributed Detection and Data Fusion*, Springer-Verlag, New York, NY, 1997.
2. B. V. Dasarathy, *Decision Fusion*, IEEE Computer Society Press, Los Alamitos, CA, 1994.
3. S. Nag, L. Peters, Jr., I. J. Gupta, and C.-C. Chen, "Ramp response for the detection of anti-personnel mines," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, J. T. B. A. C. Dubey, J. F. Harvey and R. E. Dugan, eds., *SPIE* **3710-137**.
4. I. J. Gupta, A. van der Merwe, and C.-C. Chen, "Extraction of complex resonances associated with buried targets," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, J. T. B. A. C. Dubey, J. F. Harvey, ed., *SPIE* **3392**, pp. 1022–1032, 1998.
5. I. K. Sendur and B. A. Baertlein, "Techniques for improving buried mine detection in thermal ir imagery," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, J. T. B. A. C. Dubey, J. F. Harvey and R. E. Dugan, eds., *SPIE* **3710-130**.

Radar Images of Penetrable Targets Generated from Ramp Profile Functions

Soumya Nag, *Member, IEEE*, and Leon Peters, Jr., *Life Fellow, IEEE*

***Abstract*—Images can be generated for penetrable targets from their scattered fields when the time dependence of the incident electromagnetic (EM) wave takes the form of a ramp function. Previous researchers have developed these concepts for conducting targets. This paper focusses attention on penetrable targets. Ramp response signatures of the targets for cases where the dielectric constant of the target is greater than and also less than that of the ambient medium are included. The latter case can be applied as a signature of antipersonnel mines. The results contained herein are based on: 1) scattering measurements in The Ohio State University ElectroScience Laboratory Compact Range; 2) scattering computation using an eigenfunction solution and a method of moments solution; and 3) a very limited set of measurements generated from a buried land mine using the ElectroScience Laboratory Ground Penetrating Radar. The targets presented in this paper include metallic and dielectric spheres and actual land mines.**

***Index Terms*—Dielectric scatterers, ground-penetrating radar (GPR), land mines.**

I. Introduction

THE ramp response of a radar target is defined herein as the far zone backscattered time-domain waveform resulting from a plane wave illumination whose time dependence is that of a ramp function. This ramp incident wave is a function of time and distance. The backscattered field from a ramp incident wave is proportional to the transverse cross-sectional area corresponding to the time and position on the target such that the argument of the incident ramp function is zero. This property has been demonstrated for conducting targets by use of physical optics (PO) and has been used to create images of such targets from only three orthogonal incidence angles. These concepts have been recently expanded to treat dielectric bodies.

Consequently, it represents a potential valuable identification tool for buried mines. It has several advantages over other ground-penetrating radar (GPR) incident waves. It is basically a low-frequency concept, and consequently the clutter for a typical case is decreased. In most earth media, the conductivity and the attenuation factor are increasing functions of frequency over GPR frequency bands. The most distinguishable characteristic of a mine as well as similar false targets is observed for frequencies just above the Rayleigh region, thus enhancing identification capabilities.

Applications of electromagnetic (EM) field waveforms, produced by scattering of transient plane waves from finite objects, to target identification was first discussed by Kennaugh and Moffatt [1]. Using physical optics, they showed that the scattered field from a finite metallic target due to a ramp driving function traces out the cross-sectional area profile of the target as a function of a distance along the line-of-sight (known as the *profile function* of the target). Young [2], [3] compared the measured ramp response signatures for different metallic targets of simple shapes with the corresponding profile functions. Using the ramp waveforms, he also generated three-dimensional images of simple conducting targets in free-space from orthogonal look angle radar data.

Nag *et al.* [4] have adapted the original solution to include dielectric (or penetrable) targets and have obtained the ramp profiles and images for such targets.

The ramp response of a target has the following characteristics. It is derived from the specular or early time component of the target backscattered impulse response. The ramp response is a low-frequency concept, since it is obtained by taking the inverse Fourier transform (IFT) of the target backscattered frequency response spectrum weighted by the inverse of the frequency squared. Other types of imaging techniques require target radar echo measurements at multiple antenna positions to form a usable image of the target. However, generation of an image of a target of simple shape via ramp waveforms requires the ramp response obtained typically from three antenna positions.

The ramp response has some inherent limitations. First, since the ramp response is a low-frequency phenomenon, it is essential for the lowest of the interrogating frequency components to lie in the Rayleigh region of the target's backscattered radar echo so that a valid ramp waveform can be generated. Since plane wave incidence is used, a large antenna could be required. The multiple reflections of waves inside a inhomogeneous penetrable target can perturb the ramp waveform from the desired transverse cross-sectional area of the target along line-of-sight. There are signal-processing techniques available to reduce such multiple reflections. Three-dimensional images of dielectric targets in free-space are presented using the processed ramp waveform obtained from three orthogonal look angles. For GPR applications, means of generating images of buried targets are also suggested.

The organization of this paper is as follows. In Section II, we discuss the theory of the ramp response for dielectric (penetrable) bodies. In Section III, we present the ramp response

Manuscript received September 1, 1999; revised May 3, 2000. This work was supported by Duke University under an award from ARO (OSD MURI Program).

S. Nag is with Time Domain Corporation, Huntsville, AL 35806 USA (e-mail: soumya.nag@timedomain.com).

L. Peters, Jr., is with the ElectroScience Laboratory, Department of Electrical Engineering, The Ohio State University, Columbus, OH 43212 USA (e-mail: peters.6@osu.edu).

Publisher Item Identifier S 0018-926X(01)02290-6.

0018–9268/01$10.00 © 2001 IEEE

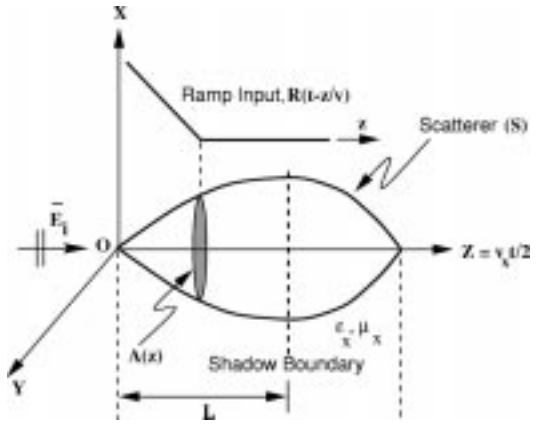


Fig. 1. Target geometry and ramp incident wave.

signatures of dielectric and conducting bodies immersed in both free-space and more dense dielectric media. In Section IV, complete three dimensional (3-D) images of land mines, obtained using the ramp response for three orthogonal look angles, are presented. Section V discusses the potential measurement of ramp response of buried targets using a dielectric rod antenna, developed at The Ohio State University ElectroScience Laboratory (OSU-ESL) [5]. In Section VI, an image of a buried land mine is shown, which is obtained from the measured GPR data. Finally, in Section VII, some concluding remarks are presented.

II. Ramp Response for Penetrable Targets

The ramp incident wave $R(z,t)$ is a time-spatial function given by

$$R(z,t) = R\left(t - \frac{z}{v_x}\right) \tag{1}$$

where t is time and v_x is the velocity of the wave in the ambient medium. The target geometry, as a function of coordinate z, is illustrated in Fig. 1 for a given time instant. It has been shown in [2] and [3] that the ramp response for a conducting target is given by

$$f_r^M(t) = -\frac{1}{\pi v_x^2} A\left(\frac{v_x t}{2}\right) \tag{2}$$

where $A(v_x t/2)$ is the transverse cross-sectional area of the target and the argument $(v_x t/2)$ accounts for the two-way radar path $2z = v_x t$.

The ramp response given in (2) for conducting targets is extended to dielectric targets by modifying the reflection coefficients that are used in the PO solution [1], [6]. In this case, the reflection coefficients become distance (or time) dependent because of their angle of incidence dependence for dielectric media. The ramp response for penetrable targets via PO takes the form

$$f_r^D(t) = \frac{1}{\pi v_x^2} \int_{0^-}^{t} \bar{R}_{12}\left(\frac{v_x t'}{2}\right) \frac{d}{dt'} A\left(\frac{v_x t'}{2}\right) dt' \tag{3}$$

where $\bar{R}_{12}(v_x t'/2)$ is the average reflection of the dielectric surface. For a rotationally symmetric target

$$\bar{R}_{12} = \frac{1}{2}[R_{12}^{\perp} + R_{12}^{\parallel}] \tag{4}$$

where $R_{12}^{\perp}$ and $R_{12}^{\parallel}$ are the Fresnel reflection coefficients of a dielectric surface perpendicular and parallel to the plane of incidence, respectively. Because of the functional form of $R_{12}^{\perp}, R_{12}^{\parallel}, \bar{R}_{12}$ can be approximated by their values at normal incidence. Then $\bar{R}_{12}(0)$ can be taken outside the integral of (3), and we obtain

$$f_r^D(t) = \frac{\bar{R}_{12}(0)}{\pi v_x^2} A\left(\frac{v_x t}{2}\right). \tag{5}$$

$A(v_x t/2)$ is maximum at the shadow boundary of the target. Also, the duration of the ramp function from zero to its maximum value is approximately

$$t_d = L\sqrt{\frac{\epsilon_r^{\text{rel}}}{15.24}}\ \text{ns} \tag{6}$$

where L is the distance from the onset of the target to the shadow boundary, as shown in Fig. 1 in centimeters, and ϵ_r^{rel} is the relative permittivity of the external medium. While this is valid over the lit region, it has also been found to represent a useful approximation in the shadow region when L is replaced by the length along the surface of the target from the shadow boundary to the furthest point on the target.

The ramp response of a scatterer can be obtained from its backscattered frequency response in two ways.

1) By double integrating the impulse response of the scatterer, i.e.,

$$f_r(t) = \int_{t'=0^-}^{t} \int_{\tau=0^-}^{t'} f_I(\tau)\, d\tau\, dt'. \tag{7}$$

The impulse response $f_I(t)$ is obtained via IFT of the backscattered frequency response.

2) By performing IFT of a target's backscattered frequency response $F_I(j\omega)$ weighted by $1/(j\omega)^2$, i.e.,

$$f_r(t) = \text{IFT}[F_R(j\omega)] = \text{IFT}\left[\frac{1}{(j\omega)^2} F_I(j\omega)\right]. \tag{8}$$

Note that (7) would also contain the additive terms of the form $c_0 t + c_1$, where c_0 and c_1 are constants. c_0 and c_1 are the dc biases of the step and ramp responses, respectively. c_0 is simply subtracted from the step response (single integral of the impulse response) before it is integrated once, and c_1 can be subtracted from the ramp response. Also in (8), $F_R(j\omega)$ contains additional additive terms: $(1/(j\omega)^2)\int_{-\infty}^{0-} f_i(\tau)\, d\tau + (1/j\omega)\int_{-\infty}^{0-}\{\int_{-\infty}^{\tau} f_i(t')\, dt'\}\, d\tau$. But these terms drop out since we are interested in zero initial conditions, i.e., $\int_{-\infty}^{0-} f_i(\tau)\, d\tau = 0$ and $\int_{-\infty}^{0-} f_u(\tau)\, d\tau = 0$, where $f_u(t)$ is the step response.

Fig. 2 is a plot of the physical optics ramp responses, as obtained from (2), for a perfectly conducting sphere of 7.6 cm diameter. Also shown in Fig. 2 is the result obtained using (3)

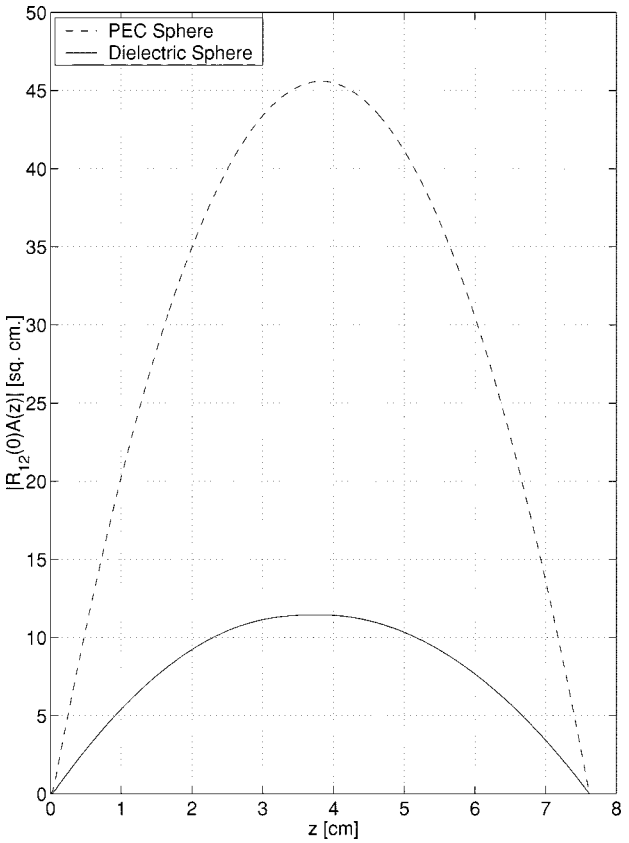


Fig. 2. PO ramp response functions of 7.6 cm diameter conducting and of dielectric spheres ($\epsilon_i = 3$) in free-space.

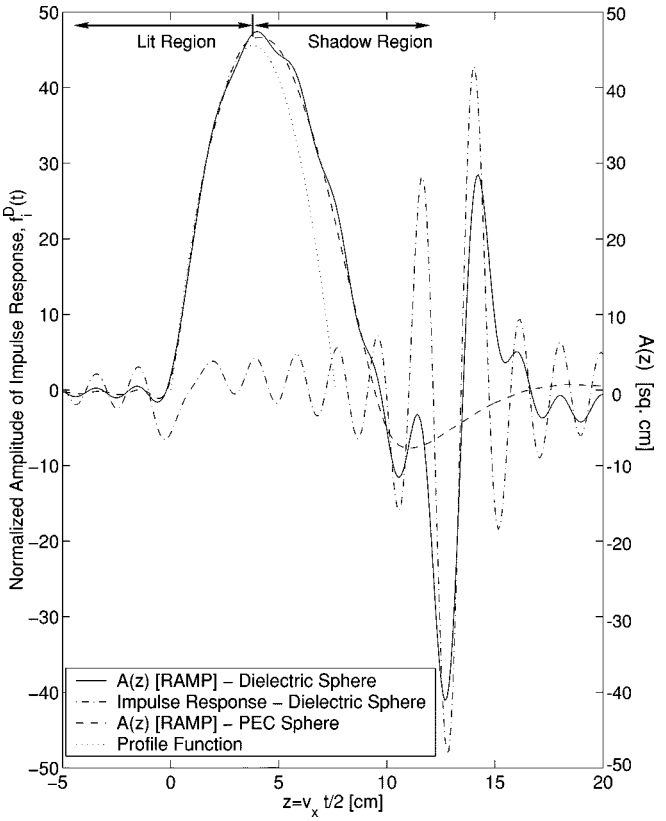


Fig. 3. Normalized impulse response and the transverse cross-sectional area $A(z)$, obtained from the ramp response, for a 7.6-cm diameter PEC sphere and dielectric ($\epsilon_i^{\rm rel} = 3$) sphere in free-space.

for a dielectric sphere ($\epsilon_i^{\rm rel} = 3$) of the same size. The difference in the amplitude of $f_r(t)$ for the two spheres is due to the different values of the reflection coefficient. In this case, the function $(1/2)(R^{\perp} + R^{\parallel}) \approx R^{\perp}(0) = R^{\parallel}(0)$. Note that there would be only small differences between these two cases if the curve for the dielectric sphere had been normalized to that of the conducting sphere. For this case, variations of $R^{\perp}(v_x t/2)$ and $R^{\parallel}(v_x t/2)$ tend to cancel each other. This result, of course, is valid for symmetric geometries. However, if the dielectric target is a cylindrical body, at endfire incidence the reflection coefficient must be retained inside the integral of (3).

III. Ramp Profile Function

The interest in the ramp response of targets is dictated by its close relationship to the physical geometry of the target. This is illustrated in Fig. 3, in which the transverse cross-sectional area obtained from the ramp response (also referred to in the paper as the ramp profile function) of both a conducting sphere and a dielectric ($\epsilon_i^{\rm rel} = 3$) in free-space, are shown. The profile function is simply its physical transverse cross sectional area as a function of range (z of Fig. 1). The ramp response is obtained from the exact solution of the scattered fields of each sphere for 0.02–8.02 GHz frequency band. The agreement between the profile function and the ramp response in the lit region is excellent. The normalized impulse response, obtained from the exact solution, for the dielectric sphere is also shown, and it is much more complicated than the ramp response.

If the operational frequency band does not contain sufficiently low-frequency components, the ramp waveform contains an unwanted low-frequency ripple corresponding to the minimum frequency ($f_{\rm min}$) used in the Rayleigh region. These sidelobes can be reduced by simply extrapolating the magnitude and phase of the $1/(j\omega)^2$ weighted frequency response to $f_{\rm min} \rightarrow 0$, respectively, as independent of frequency and linear function of frequency. Note that in the Rayleigh region of the backscattered frequency response, the magnitude is approximately proportional to ω^2 and the phase is a linear function of ω. However, this extrapolation is valid only if the minimum radar frequency lies in the Rayleigh region of the target in a given medium.

Ramp waveforms for a number of targets have been obtained from scattered field data for dielectric targets in free-space, dielectric targets in penetrable medium, and dielectric targets immersed in a lossy dispersive medium [7]. Additional processing is required for these latter cases. When the dielectric constant of the medium in greater than that of the target, internal reflections occur at times when the ramp is being generated, and these can be removed from the impulse response before forming the ramp response. This will also be the case for mines whose internal structure is more complex. However, even if these reflections are not removed, the ramp response still provides usable data. A more complex solution is required if the medium is lossy and dispersive. A means of treating this case is discussed in a recent paper [8].

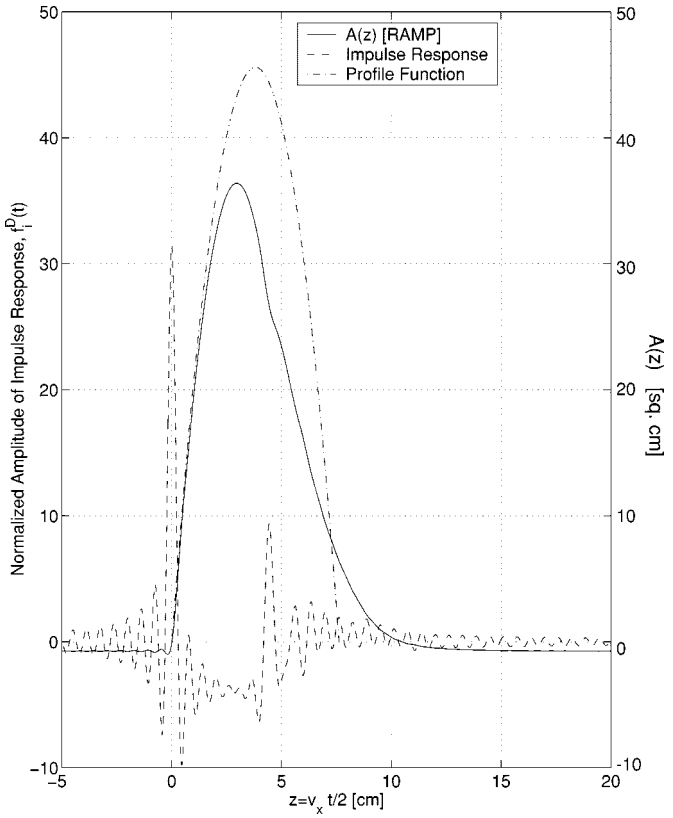


Fig. 4. Impulse and ramp responses of a 7.6-cm diameter lossless dielectric sphere with $\epsilon_i^{\rm rel} = 3$ embedded in lossless dielectric medium of $\epsilon_x^{\rm rel} = 9$.

The results obtained using an eigenfunction solution for a dielectric sphere $(\epsilon_i^{\rm rel} = 3)$ immersed in a dielectric medium $(\epsilon_x^{\rm rel} = 9)$ are shown in Fig. 4. There are some mechanisms that can distort the ramp profile functions. The impulse response of Fig. 4 contains a number of reflections that can cause perturbations of the ramp response. These can be reduced to improve the shape of the ramp response using signal processing developed to remove the ground-free-space interface reflections [9]. However, the major perturbation appears to occur when the angle of incidence approaches the critical angle, and the approximation implicit in (5) is no longer valid. Further analysis of this case is beyond the scope of this paper. While the ramp response for this case is disturbed, it can be used to provide an image whose radius would be slightly smaller than the actual radius of the original sphere.

The ramp response for normal incidence on a 2.54-cm dielectric cube measured in free-space represents an extreme case. It would require a very large bandwidth to obtain the sharp rise time expected for such a cube. Consequently, the frequency band was increased from 6 to 12 GHz for the cube [7]. In this case, the ramp waveform is slightly delayed about 0.51 cm from the profile function but otherwise gives good agreement for the greater bandwidth.

The ramp response of a pair of dielectric blocks $(\epsilon_i^{\rm rel} = 3)$, immersed in a dielectric medium $(\epsilon_x^{\rm rel} = 9)$ where $\epsilon_x^{\rm rel} > \epsilon_i^{\rm rel}$, has also been reported previously. Ramp responses for a dielectric block attached to a larger one have also been generated using a method of moment (MoM) solution [10]. Both blocks are clearly recognizable when the small block is directly illuminated. The presence of the smaller block is detectable, but its response is distorted when it is in the shadow of the larger block [7].

Measured free-space scattered fields (0.1–6.0 GHz bandwidth) are used to generate the ramp response for an SB-33 anti-personnel (AP) land mine for broadside and endfire incidence, shown in Fig. 5(a) and (b). The broadside and endfire dimensions for this mine are 8.8 and 3.2 cm, respectively. Note that the duration of the ramp response is not seriously perturbed by the internal reflections.

Finally, ramp response, generated using computed scattered fields, have been presented elsewhere [5] for a perfectly conducting sphere and cube immersed in a lossy dispersive medium ($\epsilon_x^{\rm rel} = 9, \sigma_x = 0.8$ S/m). An eigenfunction solution for 0.3–3.1 GHz frequency band, and an MoM solution for 0.3–4.3 GHz frequency band, were used to generate the backscattered fields for the sphere and the cube, respectively. In the lower part of the frequency band of interest, the medium is a conductor; in the upper part, it is a lossy dielectric; and in intermediate region, it is neither. It is emphasized that this medium would be far too lossy for any ground. Nevertheless, after substantial additional processing, the ramp responses in this lossy dispersive medium approximate the profile function particularly in the lit region of the target [8].

IV. Images of Dielectric Bodies in Free-Space Obtained from the Ramp Response

A substantial amount of information is generated from a single ramp response of a target concerning its length, width, and electrical properties. Much more can be generated in the form of images created by combining results from different incidence angles. Young [2], [3] developed this concept using three orthogonal observation angles, and this approach is used to obtain some of the images to be presented. The reader is referred to the above references for the details of this imaging process. Young considered only metallic targets where the reflection coefficient is independent of incidence angle. For dielectric targets, the reflection coefficient is as yet an unknown variable depending on the incidence angle and the electrical properties of the ambient medium. Hence, it becomes difficult to evaluate $A(x), A(y)$, and $A(z)$ from the measured ramp waveforms. Images of several land mines have been generated, and two of these are shown in in Figs. 6 and 7, which were obtained using Young's imaging technique and three orthogonal incidence angle ramp responses for the 0.1–6.0 GHz bandwidth. The duration and the shape of the three orthogonal ramp waveforms have been used to normalize the peak of each of these waveforms to the maximum value of the transverse cross-sectional $(A_{\max})$ area in each orthogonal direction. For targets with a flat surface, the target width obtained in the image can be too large. This is because the rise (and fall) times of the ramp waveforms for the front surfaces of these targets are stretched due to the limited bandwidth. This problem can be alleviated by extending the bandwidth. Otherwise, the images are reasonable pictures of the respective targets.

Most mines have rotational symmetry. If such mines, buried in the ground, are oriented with the axis of rotation almost

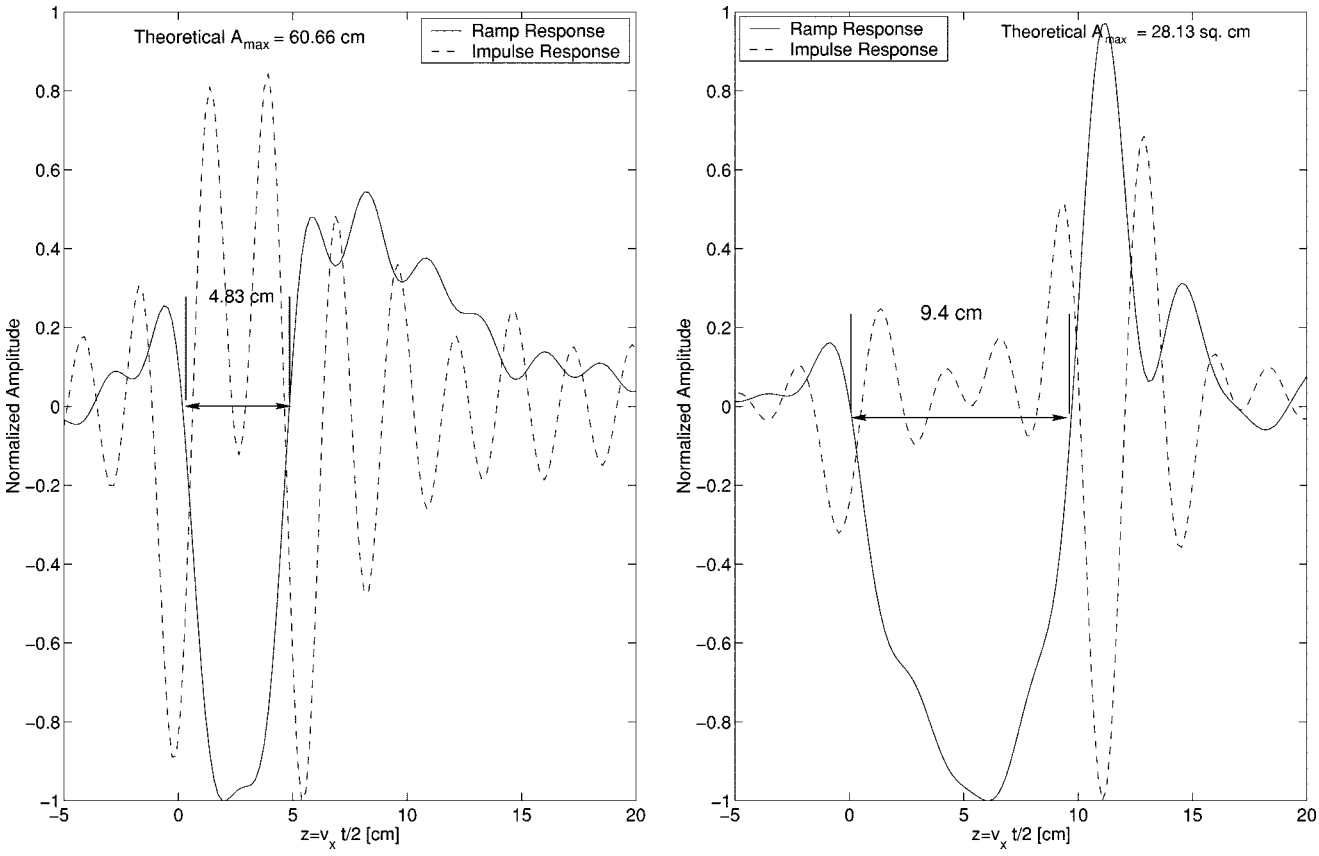


Fig. 5. Normalized ramp response and impulse response for the land mine SB-33 measured at OSU-ESL Compact range for (a) broadside and (b) endfire incidence.

perpendicular to the ground-air interface, only a single ramp response is required to generate an image. The rotational symmetry of a buried target can be demonstrated by measuring the target response as a function of antenna positions from different directions. The transverse cross-sectional function $A(z) = \pi r^2(z)$, where the radius $r(z)$ completely approximates the geometry. The maximum value (peak of the ramp profile function) $A_{\max}$ for a buried target could be generated in two ways: 1) by using the usual GPR measured data as the antenna is moved over the mine to estimate its lateral extent and 2) from (4) by using the maximum measured value of the ramp profile function, which implies that accurate, calibrated data is available. This last method is used to generate an image of an SB-33 mine in free-space using only the ramp waveform, given in Fig. 5(a), and assuming the rotational symmetry of the body. The resulting image is shown in Fig. 8.

V. Potential Measurement of Ramp Response Using GPR with a Dielectric Rod Antenna

In the preceding discussion, results have been generated under ideal conditions either from measured data taken in the OSU-ESL compact range or computed data. The task now is to generate these data using a GPR capable of acquiring the requisite data under field conditions. As has been observed, the fields of the focused reflector [11] do approximate a plane wave in the focused spot region, but this antenna would be far too large to include the required scattered fields in the Rayleigh region of A-P mines.

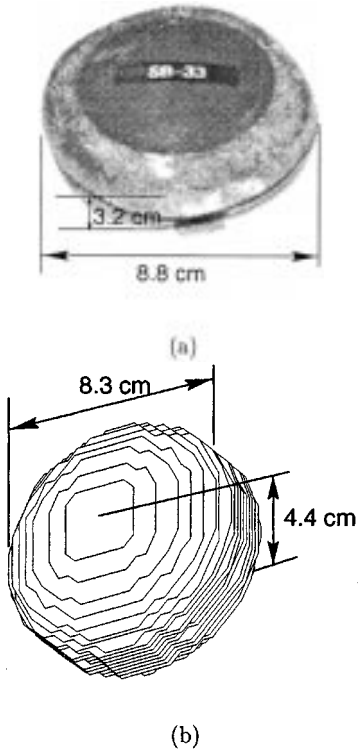
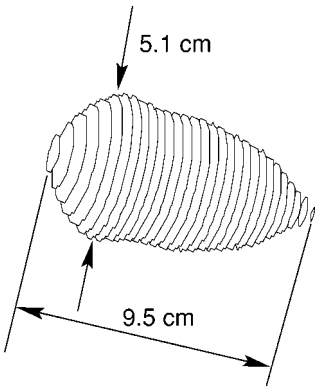


Fig. 6. (a) A-P mine SB-33 in free-space and (b) reconstructed image of the land mine.



(a)



(b)

Fig. 7. (a) A-P mine C4A1 in free-space, and (b) reconstructed image of the mine.

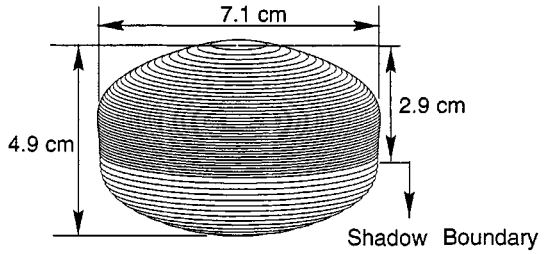


Fig. 8. An image of SB-33 measured in free-space, using only the ramp response given in Fig. 5.

The second type of antenna used in GPR for mine detection is a wide-band dielectric rod antenna (near-field probe) that also creates a highly localized illumination [5]. This dielectric rod antenna, however, could be hand held or vehicle-mounted for convenience and speed. It is interesting to note that this antenna is in the far field of the mine at a range of 7.6 cm up to 2 GHz. This is obtained using the range $R > (2D^2/\lambda)$ criterion and assuming $D = 7.6$ cm for a typical A-P mine. If the range is increased to 22.8 cm, the frequency increases to 6 GHz. This does not even include the refraction at the ground–air interface, which would further reduce this phase error.

It is suggested that plane wave representation could be improved by taking a series of measurements as the antenna is moved over the target. These time-domain waveforms could simply be averaged to further reduce the phase error inherent in the dielectric rod antenna pattern, which is simply a point source radiator. A radar can be used even though certain mechanisms such as the creeping wave on a sphere would not be observed.

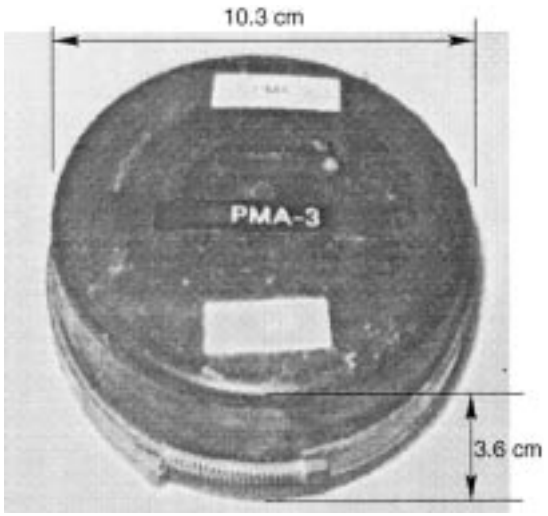


Fig. 9. PMA-3 A-P mine.

However, these mechanisms would only corrupt the data to be used in generating these ramp images.

The current dielectric rod antenna operates from 1 to 6 GHz and does not include the Rayleigh region of most of the A-P mines buried in clay. A simple redesign should produce such an antenna that operates as low as 300 MHz up to the Rayleigh region.

A more serious problem is that of launching a wave parallel to the ground–air interface to achieve the required orthogonal incidence on the target. The antenna would be another version of the dielectric rod antenna. This antenna would have a dielectric constant ϵ_a greater than that of the earth ϵ_x. The antenna would require contact with the ground and would be at an angle θ greater than the critical angle measured from the normal to the ground, or $\sin\theta > \sqrt{\epsilon_x/\epsilon_a}$, so that the lateral wave is launched below the ground.

VI. Ramp Image of a Mine Buried in Dry Clay

The goal of this section is to demonstrate that a measured ramp response scan be obtained for an actual mine in a real medium with clutter present. It further demonstrates that a GPR using the dielectric rod antenna can be calibrated with sufficient accuracy to generate a realistic image. The PMA-3 A-P mine (see Fig. 9) buried approximately 2.5 cm deep in a dry clay medium at a JUXOCO site, Fort A.P. Hill, Virginia, was selected for this purpose. A network analyzer and dielectric rod was used as the GPR. The measured strength of the impulse response is shown in Fig. 10 as a function of time (ns) and antenna position (cm). The stripes across the top of this figure (at around 0.2 ns) are the response due to the ground reflections. They can be removed from the data using a signal-processing scheme developed by van der Merwe [12]. The scattering from the mines occurs at about -0.54 ns to -1.35 ns and 7.6–20.3 cm antenna position range, as indicated by the hyperbolic arc in the figure. This time-domain plot is obtained via IFT on 1.2–6 GHz range frequency-domain response. The measured data at 12.7-cm position is time gated between the -0.54 to -1.35 ns range and is then transformed to the frequency domain (1.2–6

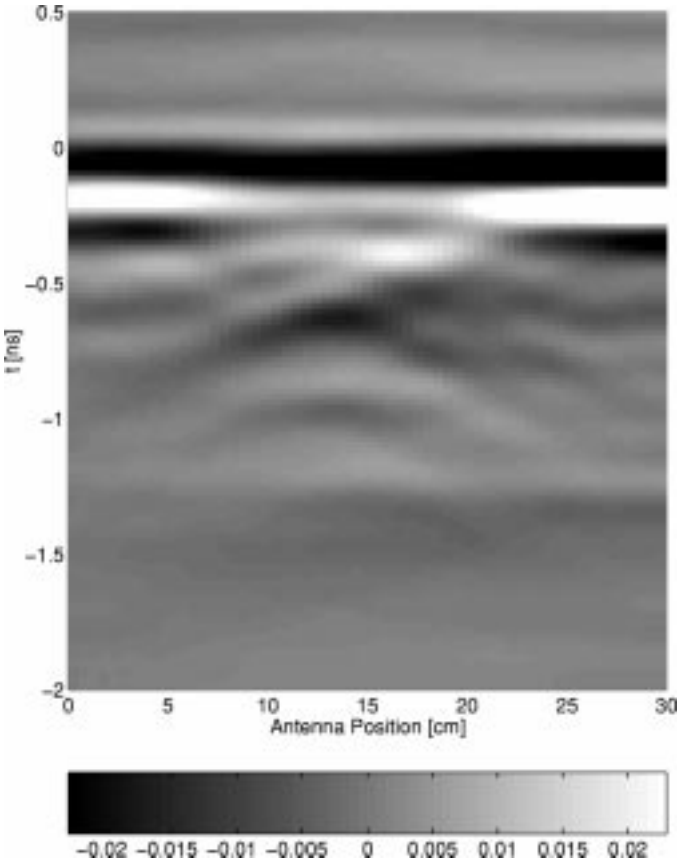


Fig. 10. Amplitude of the impulse response (for 1.2–6 GHz data) for the PMA-3 mine plotted as functions of antenna positions (cm) and time (ns).

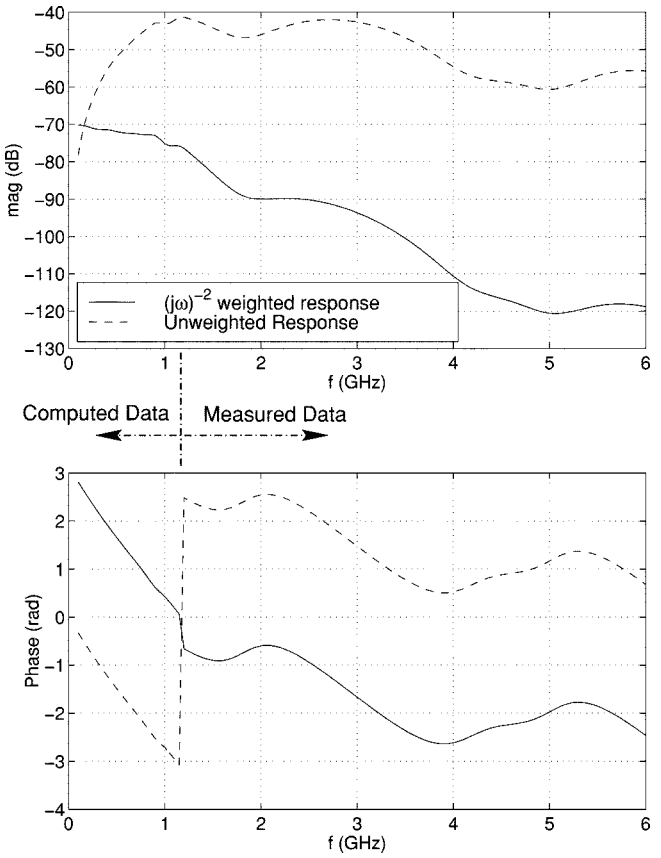


Fig. 11. The unweighted and $1/(j\omega)^2$ weighted spectrum for the buried PMA-3 A-P mine, shown with the numerical data supplemented for frequencies below 1.2 GHz with the measured data.

GHz band). This particular antenna operates for 1–6 GHz frequency range, and it does not operate in the Rayleigh region of most of the A-P mines buried in clay. The PMA-3 mine was selected since the lower frequency limit of the radar system was just above the Rayleigh region of this target. The measured fields for this are shown in Fig. 11 for the region 1.2–6.0 GHz. Note that the lowest frequency (1.2 GHz) is just above the Rayleigh region. The data in Fig. 11 were supplemented as will be discussed by computed values for 0.1– 1.2 GHz. To supplement these data in the Rayleigh region and remove this deficiency, the PMA-3 mine was approximated by a homogeneous dielectric body of the same overall shape, embedded in a lossless dielectric medium having relative permittivity the same as that of the clay ($\epsilon_x^{\rm rel} \approx 4$ for the observational bandwidth). The backscattered fields were then computed using an MoM solution for 0.1–1.2 GHz range. The computed field data, after appropriate amplitude scaling and phase shifting, was cascaded with the measured data. The spectrum of PMA-3 (0.1–6 GHz frequency band), after supplementing the measured data with the numerical data, is shown in Fig. 11. $1/(j\omega)^2$ weighting is applied on this composite response, and the ramp response is finally obtained using IFT on this weighted frequency response. It is observed that the curve below 1.2 GHz in Fig. 10 simply behaves as would be expected in the Rayleigh region and could have been estimated from the region of the curve above 1.2 GHz.

In practice, there are still several steps that would be required to create an image of the buried target. First, it would be necessary to demonstrate that an unknown target has rotational symmetry. This could be be achieved by moving the GPR in different directions across the center of the target and repeating the measurements of Fig. 10 to demonstrate that the target response does not change significantly as a function of position. Next, it is necessary to determine the lateral extent of the target. This can be achieved also using Fig. 10. It is visually estimated to be about 11.4 cm. However, because the dielectric rod antenna is essentially a point source, the target is illuminated by the antenna even after it is no longer directly over the target. Since a target may not have a flat upper surface, there will be geometrically reflected rays from this surface, even though the antenna is not directly over the target. This results in an estimation of a bigger lateral extent of the target from a plot like Fig. 10. It is noted that a diameter, much closer to that of an actual target, has been obtained for a target with a flat upper surface.

Alternatively, the maximum transverse cross-sectional area $A_{\max}$ can also be estimated from the ramp waveform. This requires the scaling of the ramp response by a factor $[T_{12}(0)T_{21}(0)R_{23}(0)]^{-1}$, where $T_{12}(0)$ is the transmission coefficient of the wave at the free-space–ground interface for normal incidence, $T_{21}(0)$ is the transmission coefficient of the wave at the ground–free space interface for normal incidence, and $R_{23}(0)$ is the reflection coefficient of the wave (in the clay medium) illuminating the target at normal incidence. R_{23} is assumed constant over the lit surface of the target. The lateral extent l_e of the target, assuming that it is

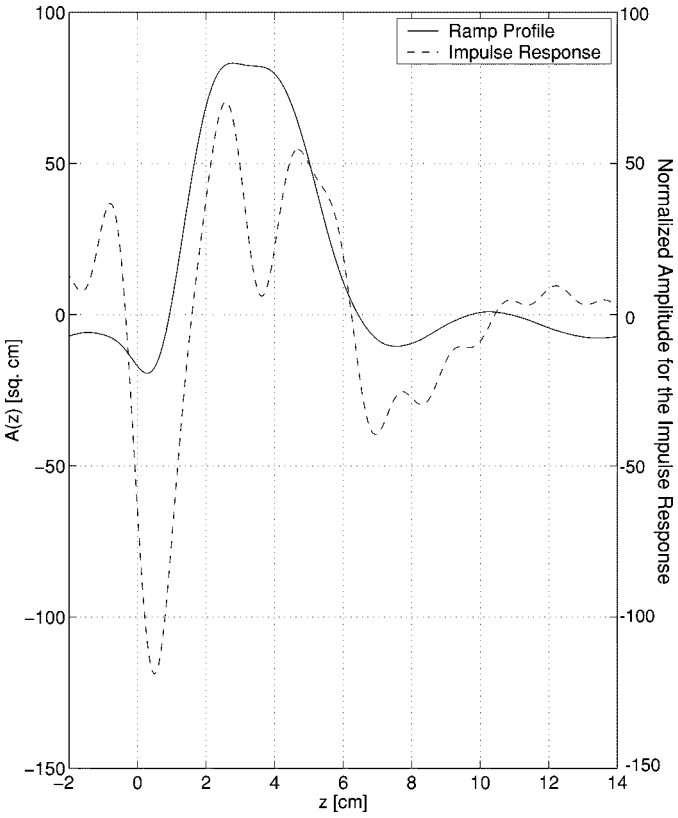


Fig. 12. The ramp profile function and the normalized impulse response obtained after supplementing 1.2–6 GHz measured data withthe 0.1–1.1 GHz MoM data.

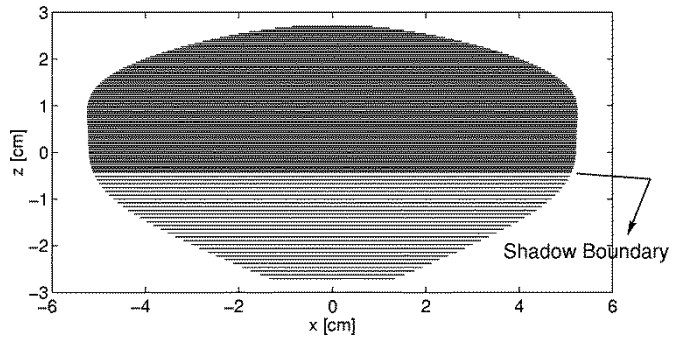


Fig. 13. Side-view image of PMA-3 A-P mine generated using the ramp profile function in Fig. 12.

rotationally symmetric, is obtained simply using the relation $l_e = 2\sqrt{A_{\max}/\pi}$. This process of estimation of l_e of a buried land mine is reasonably good provided a) the incident wave entirely illuminates the target and b) the dielectric constant of the target is known. If a given land mine, embedded in a given ambient medium, is to be detected, its dielectric constant can be determined from the measured ramp waveform of that mine buried in the same ground condition. For most dielectric mines, reasonable values of the relative permittivity lie in the range 3–4 for our observational frequency band. For this particular case, the relative permittivities of free-space, clay, and the mine are, respectively, $\epsilon_{\text{air}} = 1$, $\epsilon_{\text{clay}} = 4$ and $\epsilon_{\text{mine}} = 3$. $A(z)$, obtained from the ramp response, and the normalized impulse response are shown in Fig. 12. The diameter of the PMA-3 mine, estimated from the ramp profile function, is 10.4 cm, which is in good agreement with the actual diameter 10.3 cm. Its image is shown in Fig. 13.

VII. Conclusion

Transient scattered fields obtained from convex targets illuminated by a ramp incident waveform have been shown to provide a reasonable representation of the transverse physical cross section (or profile function) of the penetrable target as a function of range to the shadow boundary. However, a much less accurate representation is obtained beyond the shadow boundary. But even then, it is still a useful representation.

Ramp responses for several targets have been given, and it is shown that they do give good agreement with the target profile function. These cases include some land mines. Images have been generated for these targets, of different sizes and shapes, in free-space using the ramp response data from three orthogonal incidence angles.

A means of generating these orthogonal incidence angles for a shallow buried target is suggested that would make use of a modified version of the current dielectric rod antenna. However, this requires a rather cumbersome measurement. Therefore, a means of generating ramp images of a buried target is suggested using the broadside look angle ramp profile function, provided the shape of the target transverse to the direction of propagation is known. The approximate transverse shape of the target can be determined by observing the target in rotated planes about the center of the unknown response.

Preliminary approximate images have been generated for rotationally symmetric targets, such as a PMA-3 AP mine buried in dry clay at a JUXOCO site. The maximum value of the ramp profile function, obtained from the measured data, yields the dimension of the target in the transverse direction. Again, even in this cluttered medium, a realistic image of the target has been generated.

References

[1] E. M. Kennaugh and D. L. Moffatt, "Transient and impulse response approximations," *Proc. IEEE*, pp. 893–901, Aug. 1965.
[2] J. D. Young, "Target imaging from multiple frequency radar returns," Ph.D. dissertation, The Ohio State University, Columbus, June 1971.
[3] J. D. Young, K. A. Shubert, and D. L. Moffatt, "Synthetic radar imagery," *IEEE Trans. Antennas Propagat.*, vol. AP-24, May 1976.
[4] S. Nag and L. Peters Jr., "Ramp response signatures for dielectric targets," in *Proc. Soc. Photo-Opt. Instrument. Eng. Conf.*, Orlando, FL, Apr. 1998.
[5] C.-C. Chen, K. R. Rao, and R. Lee, "A new ultra-wide bandwidth dielectric rod antenna for ground-penetrating-radar applications," *IEEE Trans. Antennas Propagat.*, to be published.
[6] D. L. Moffatt, "A physical optics approximation of the scattering for axial incidence from rotationally symmetric targets," Contract AF 30(602)-2042, Antenna Lab., The Ohio State Univ. Res. Foundation, Columbus, Rep. 925-6, Feb. 1961.
[7] S. Nag, "Ramp response signatures of dielectric scatterers," Ph.D. dissertation, The Ohio State University, 1999.
[8] S. Nag, L. Peters Jr., I. J. Gupta, and C.-C. Chen, "Ramp response signatures for the detection of anti-personnel mines," in *Proc. Conf. Soc. Photo-Opt. Instrument. Eng.*, Orlando, FL, Apr. 1999.
[9] C.-C. Chen, "Design and development of enhanced ground penetrating radar systems for the detection and classification of unexploded ordnances and land mines," Ph. D. dissertation, The Ohio State Univ., Columbus, 1997.
[10] E. H. Newman, "A User's Manual for the Electromagnetic Surface Patch Code: Preliminary Version ESP5.0," unpublished, June 1998.

[11] C.-C. Chen, S. Nag, W. D. Burnside, J. Halman, K. Shubert, and L. Peters Jr., "A stand-off, focussed-beam land mine radar," *IEEE Trans. GeoSci. Remote Sensing*, to be published.

[12] A. van der Merwe, I. J. Gupta, and L. Peters Jr., "A clutter reduction technique for GPR data from mine like targets," in *Proc. Conf. Soc. Photo-Opt. Instrument. Eng.*, Orlando, FL, Apr. 1999.



Soumya Nag (M'00) [redacted] He received the B.E.E. degree from Jadavpur University, Calcutta, in 1991, the M.Eng. degree from Memorial University of Newfoundland, St. John's, Canada, in 1994, and the Ph.D. degree from The Ohio State University (OSU), Columbus, in 1999, all in electrical engineering.

From 1994 through 1999, he was a Graduate Research Associate with the ElectroScience Laboratory, OSU. He is currently a Signal Processing Engineer with Time Domain Corporation, Huntsville, AL. His research interests are in the fields of ultrawide-band radar, signal processing, electromagnetic scattering, and target identification.



Leon Peters, Jr. (S'50–A'51–SM'60–F'81–LF'89) was born in Columbus, OH. He received the B.E.E., M.Sc., and Ph.D. degrees from The Ohio State University (OSU), Columbus, in 1950, 1954, and 1959, respectively.

In 1950, he became a Research Associate in the OSU ElectroScience Laboratory (ESL) (formerly Antenna Laboratory), Department of Electrical Engineering. In 1960, he joined the teaching staff at OSU and has served as Technical Area Director for electromagnetic and remote sensing and as the Associate Department Chairman for research. From 1983 to 1994, he was the Director of the ESL. He is now a Professor Emeritus. His primary interest is now focused on ground-penetrating radar.

Dr. Peters is a member of Sigma Xi and Commission B of the International Union of Radio Science. He received the Distinguished Achievement Award from the Antenna Measurement Techniques Association in 1998.

Electromagnetic Resonances of Immersed Dielectric Spheres

Chi-Chih Chen, *Member, IEEE*

***Abstract*—The complex natural resonances (CNR) for lossless dielectric spheres in a lossless dielectric medium are investigated. Significant differences between the external and internal resonances are presented. The external resonances are related to the external creeping waves and the internal resonances to the internally reflected waves. The internal resonances are more important in practice because of their smaller damping factors. Simple physical interpretation for predicting the resonance behavior of a general dielectric sphere is obtained.**

***Index Terms*—Electromagnetic scattering, resonance.**

I. Introduction

THE need for a better understanding of dielectric target resonances was found in a study of plastic land-mine classification using the complex natural resonance (CNR) signatures. The CNR's of an immersed dielectric sphere were investigated because of the available exact solutions. Note that the scattering mechanisms of even a simple sphere can be very complicated. Rainbow and glory ray scattering, for example, have been studied for decades [1], [2]. The current paper will focus on the resonant properties of dielectric spheres. Both the sphere and its ambient medium are assumed to be homogeneous, isotropic, and lossless. The CNR's are related to the singularities of expansion coefficients. These CNR's used as a signature are far less damped when the dielectric constant of the sphere exceeds that of the ambient medium. This case will be referred to as a dielectric sphere, whereas the reverse case is designated as a dielectric bubble. The resonances associated with either a dielectric sphere or bubble can be separated into internal and external modes. The internal resonances are caused by the internal waves that experience multiple internal reflections, whereas the external modes are caused by the surface creeping waves. The physical mechanisms associated with the internal and external resonances are explored by examining the resonances of impenetrable spheres and bubbles such as perfect electrical conducting (PEC) and perfect magnetic conducting (PMC) spheres and spherical cavities. The resonance mechanisms of general dielectric spheres or bubbles are highly related to those of impenetrable spheres and bubbles.

Manuscript received September 30, 1997; revised December 12, 1997. This work was supported by the Office of Special Technology (OST) and the Army Research Office's Multi-Disciplinary University Research Initiative (MURI) for humanitarian demining. The opinions or conclusions of this paper do not necessarily represent those of these organizations.

The author is with the ElectroScience Laboratory, The Ohio State University, Columbus, OH 43212 USA.

Publisher Item Identifier S 0018-926X(98)05787-1.

The CNR's of impenetrable cylinders have been examined by numerous authors. These poles are related to surface propagating waves (creeping waves) [3]. The same wave mechanisms are also called "Franz waves" for acoustic scattering from a "rigid sphere" in fluids [4], [5]. This type of wave propagates in the external medium and is mainly determined by the surface geometry instead of the interior material property. Ashkin [6] has shown experimentally that these creeping waves are launched by grazing incident rays. Fahlen and Bryant [7] also provided an astonishing visualization of surface waves on water droplets. Once these creeping waves are launched, they propagate along the surface with attenuation due to the continuous radiation in the tangent direction. The diffraction coefficients of a smooth transparent object has been derived by Chen [8]. Franz and Beckmann [9] also studied the creeping waves for objects of finite conductivity. Chan *et al.* [10] and Barber *et al.* [11] used resonance signatures to characterize dielectric objects. An excellent discussion about the high-frequency scattering by a dielectric sphere is given by Nussenzveig [2], [12]. Although Conwell *et al.* [13] also discussed the resonances of dielectric spheres, they did not separate the internal and external modes. They dealt with only a sphere with refractive index of 1.4. This paper investigates much more general situations and also provides some physical understanding of the resonances associated with dielectric spheres and bubbles. Throughout this chapter, an $(e^{j\omega t})$ time dependence is used.

II. CNR's for Impenetrable Spheres and Bubbles

An impenetrable sphere is defined as a solid PEC or PMC sphere, whereas a impenetrable bubble is a spherical cavity immersed in a infinite PEC or PMC space. The characteristic equations of CNR's for a PEC sphere are written as [14]

$$H^{(2)}_{n+1/2}(k_x a) = 0 \quad (\text{TE}) \qquad \text{and}$$
$$H^{(2)'}_{n+1/2}(k_x a) = 0 \quad (\text{TM}) \tag{1}$$

where "a" is the sphere radius, $H^{(2)}_{n+1/2}$ is the $(n+1/2)_{\text{th}}$ order Hankel function of the second kind, and $k_x = \omega\sqrt{\epsilon_x \mu_x}$, ϵ_x and μ_x are the wavenumber, permittivity, and permeability of the ambient medium. The subscript "x" designates the external medium. The wave whose electrical fields are perpendicular to the radial directions are designated as TE modes. The field components with magnetic fields perpendicular to the radial directions are designated as TM modes. Fig. 1 shows the poles in the complex $k_x a$ space obtained from the above

0018–926X/98$10.00 © 1998 IEEE

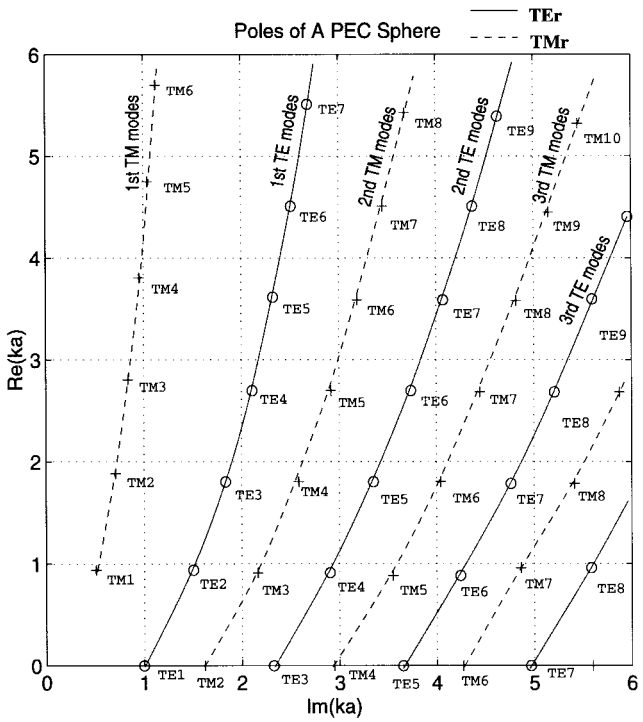


Fig. 1. Normalized complex poles for a PEC sphere. TE and TM poles are marked by "∘" and "+," respectively; the numbers indicate the order.

characteristic equations. The real and imaginary parts of each pole are related to a CNR's resonant frequency and damping factor, respectively. Note that only poles satisfying $\mathrm{Im}(k_x a) \geq 0$ are selected to ensure the damping behavior. A PMC sphere has the same set of characteristic equations as a PEC sphere except that the roll of TM and TE are interchanged. It is well known that the CNR of such impenetrable spheres are caused by the creeping waves as shown by Dragonette and Flax [15]. Since these are impenetrable spheres, the CNR's depend only on the external dielectric constant.

The characteristic equations of CNR's for a PEC bubble are also well known [16] and can be written in terms of Bessel functions and their derivatives such that

$$\begin{aligned} J_{n+1/2}(k_d a) &= 0 \quad (\mathrm{TE}) \qquad \text{and} \\ J'_{n+1/2}(k_d a) &= 0 \quad (\mathrm{TM}) \end{aligned} \tag{2}$$

where $k_d = \omega\sqrt{\epsilon_d \mu_d}$, ϵ_d, and μ_d are the wavenumber, permittivity, and permeability of the internal medium. The subscript "d" designates the internal medium. The roots of the above equations are pure real numbers and the CNR's are pure real, i.e., the resonances are nondamped. This is expected since the cavity is made of PEC. Tables of some the the roots can also be found in [16]. The characteristic equations and the CNR's for PMC bubble can be obtained by simply interchanging the roll of TE and TM. The physical interpretation of resonances inside an impenetrable bubble can be understood by realizing that the Bessel functions represent standing waves formed by incoming and outgoing waves in the radial direction such that

$$J_\lambda(z) = [H_\lambda^{(1)}(z) + H_\lambda^{(2)}(z)]/2 \tag{3}$$

where $\lambda = n + 1/2, z = k_d r$. Using the asymptotic expansion

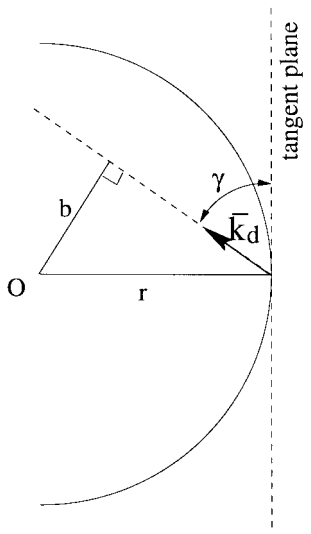


Fig. 2. Internal bounced waves.

for $H_\lambda^{(1)}(z)$ as $z \gg \lambda$ given by [17], one finds

$$\begin{aligned} H_\lambda^{(1)}(z) \sim \sqrt{\frac{2}{\pi^2 z \sin\gamma}} e^{-j(\pi/4)} e^{jz(\sin\gamma - \gamma\cos\gamma)} \\ \cdot \sum_{n=0}^{\infty} A_n \frac{\Gamma\left(n+\frac{1}{2}\right) e^{j(n\pi/2)}}{\left(\frac{z\sin\gamma}{2}\right)^n} \end{aligned} \tag{4}$$

where A_n are coefficients, Γ is the Gamma function, and $\gamma = \cos^{-1}(\lambda/z)$. The factor $e^{jn\pi/2}$ is a result of waves propagating through n caustics. The similar expansion can be readily obtained for $H_\lambda^{(2)}(z)$ by changing the sign of each "j" term. The phase variation in the radial direction is given by the second exponential term. As shown in [18], this phase variation represents an internal wave reflected away from the tangent plane at an angle of γ, as illustrated in Fig. 2.

Similarly, $H_\lambda^{(2)}(k_d r)$ are related to the outgoing portion waves which propagate toward the surface. It is interesting to find that the shortest distance to the sphere center is $b = \lambda/k_d$, for λ mode waves is found to be λ/k_d, which also agrees with the predictions of Nussenzveig [12] using the semiclassical scattering mechanics. In this study, the distance "b" refers to an "impact parameter" associated with the semiclassical collision problem. Also notice that if z increases with λ fixed, then γ increases and b decreases. The suggests that the large argument approximation corresponds to waves propagating closer the radial directions. Under such condition, the roots $j_{\lambda,s}$ and $j'_{\lambda,s}$ associated with (2) can be obtained using the McMahon's expansion [17] for $s \gg \lambda$, such that

$$\begin{aligned} j_{n+(1/2),s} &\sim \pi\left(\frac{n}{2} + s\right) \quad \text{and} \\ j'_{n+(1/2),s} &\sim \pi\left(\frac{n-1}{2} + s\right). \end{aligned} \tag{5}$$

Thus, the corresponding resonant frequencies are

$$\begin{aligned} f_{n,s}^{\mathrm{TE}} &\sim \frac{v_d}{2a}\left(\frac{n}{2} + s\right) \quad \text{and} \\ f_{n,s}^{\mathrm{TM}} &\sim \frac{v_d}{2a}\left(\frac{n-1}{2} + s\right) \end{aligned} \tag{6}$$

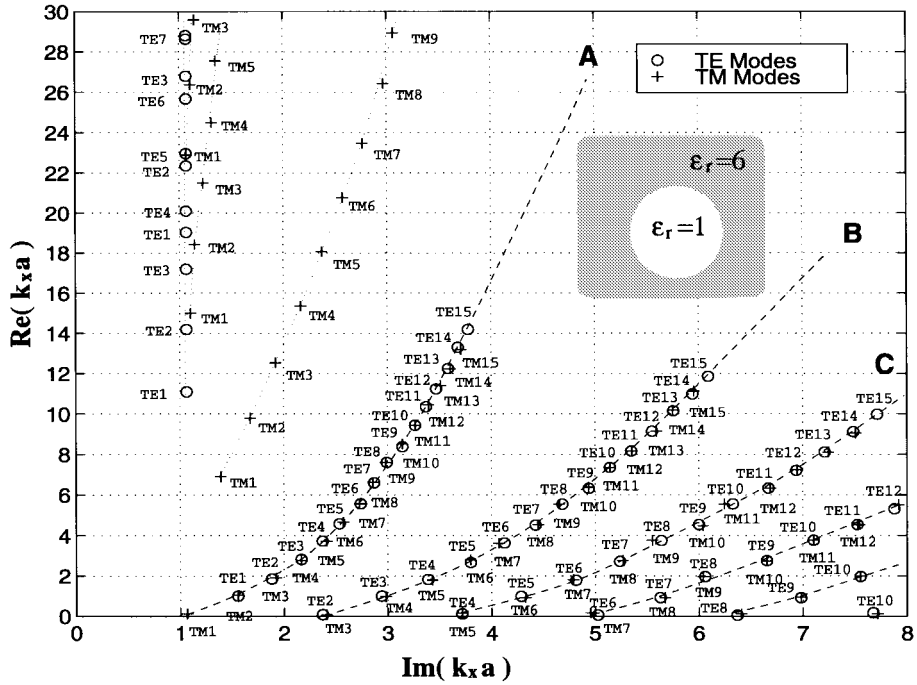


Fig. 3. Normalized poles for a dielectric bubble.

where v_d is the internal plane wave phase velocity. These resonant frequencies are the same as those found in a confocal case Fabry–Perot resonator [19].

III. CNR's For General Dielectric Bubbles

The characteristic equations of the CNR's for general dielectric spheres can be obtained from the denominators of the expansion coefficients in the normal mode-field expansions given by Harrington [20]. They are

$$\sqrt{\epsilon_d\mu_x}\hat{H}_n^{(2)}(k_x a)\hat{J}_n'(k_d a) - \sqrt{\epsilon_x\mu_d}\hat{H}_n^{(2)'}(k_x a)\hat{J}_n(k_d a) = 0, \quad \text{(TE)} \quad \text{and} \tag{7}$$

$$\sqrt{\epsilon_x\mu_d}\hat{H}_n^{(2)}(k_x a)\hat{J}_n'(k_d a) - \sqrt{\epsilon_d\mu_x}\hat{H}_n^{(2)'}(k_x a)\hat{J}_n(k_d a) = 0, \quad \text{(TM)} \tag{8}$$

where $\hat{J}_n$ and $\hat{H}_n$ are nth order modified spherical Bessel and Hankel functions. In this paper, it is assumed that $\mu_x = \mu_d = \mu_0$ (free-space permeability).

The CNR behavior for a dielectric bubble whose internal permittivity is less than that of the ambient medium is now studied. The poles in ka-space can be found from (7) nad (8). As an example, we consider $\epsilon_d = \epsilon_0, \epsilon_x = 6\epsilon_0$ and $\mu_d = \mu_x = \mu_0$. The corresponding poles are plotted in Fig. 3, where TE- and TM-mode pole are marked by "$\circ$" and "$+$," respectively. This figure is currently plotted in the $k_x a$ space, i.e., it is normalized by the external medium. Later, the poles will be revisited in the $k_d a$ space to study the internal resonant modes. Two different groups of poles can be clearly distinguished by their different vertical separation between adjacent modes. The first group of poles which have smaller vertical separations are linked by dashed lines, A, B, C, $\cdots$, etc. Each line corresponds to the multiple poles associated with an order number n. The second group of poles which have larger vertical separations are linked by dotted lines. Each group will be examined separately.

A. External Resonances

Let us focus on the first group of poles (on lines A, B, $\cdots$, etc.) and compare them with the poles for a PMC sphere immersed in the same medium as shown in Fig. 4. The TE and TM poles for the PMC sphere are marked by "$\times$" and "$*$," respectively. It is observed that the TM_n poles for the current dielectric bubble (marked by "$+$") are located very close to the TM_n mode poles (marked by "$\times$") for the PMC sphere and that the TE_n mode poles are located close to TM_{n+1} mode poles for the PMC sphere [18]. If one allows $\epsilon_x \to \infty$ but with $\sqrt{\epsilon_x\mu_x}$ fixed, the ambient becomes PMC and, thus, the TE_n mode poles of the dielectric bubble will approach to the TE_n mode poles of a PMC sphere.

From the above comparison, it is obvious that the CNR's of this group are external resonances and similar to those found for the an impenetrable sphere. Since these CNR's are caused by the surface creeping waves and are mainly determined by the geometry and the property of the ambient medium they are fairly insensitive to the internal property. This is clearly demonstrated in Fig. 5 where the CNR's for bubbles with internal dielectric constants of 1 and 2 are compared. As one can see, both cases have similar external CNR's locations.

B. Internal Resonances

Now the behavior of the second group of CNR's found in Fig. 3 will be examined by replotting the poles in the k_d space as shown in Fig. 6. Notice that most external poles are not included here. Three major features are observed from this group of CNR's are: 1) all the poles have imaginary parts greater than 4.3 indicating a low bound of damping factors;

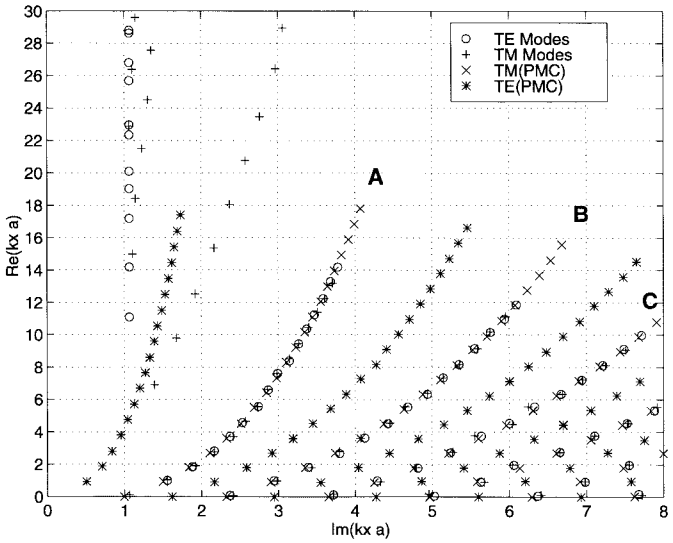


Fig. 4. Pole comparison for dielectric bubbles and for PMC spheres.

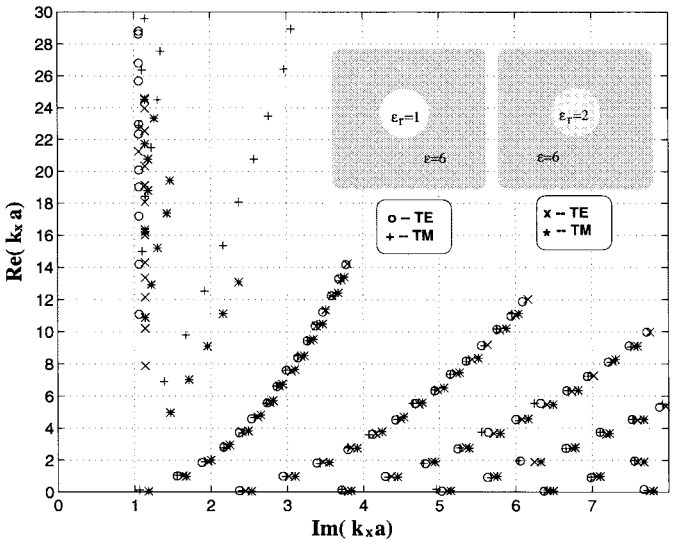


Fig. 5. Comparison of surface wave poles for dielectric bubbles with different internal permittivity.

2) smaller imaginary parts compared to most external poles studied earlier suggest lower damped resonances and more practical use; 3) TE-mode poles have smaller imaginary parts than those of TM-mode poles. In fact, all TE-mode poles tends to have the same imaginary parts. Further discussion of this behavior is explored below.

Physically, a dielectric bubble and a PEC spherical cavity of the same size and internal material are similar as far as the internal resonances are concerned. However, the former experiences damping due to energy leakage through the penetrable surface and the latter is undamped. The resonant frequency is determined by the coherent condition including the propagation phase, the phase of surface reflection at each bounce and the phase jumps due to caustics. Since all these properties are the same between a dielectric bubble and a PEC spherical cavity, one would expect them to have a very similar resonant spectrum. This prediction is verified by comparing their resonant frequencies of the first two modes $(s = 1, 2)$ for the lowest five orders $(n = 1, \cdots, 5)$, as shown in Table I. Indeed, very similar resonant frequencies are observed for both cases. This indicates that this group of CNR's shares the same resonant mechanism as were previously found for a spherical PEC cavity. This important discovery suggests that these CNR's are internal type and are insensitive to the external medium.

The damping factors for a spherical PEC cavity and a dielectric bubble are completely different. The CNR's for the PEC cavity are nondamped, however, the CNR's for a dielectric bubble are damped due to the energy leakage through the partial transmission process occurring at each internal reflection. In order to understand their damping behavior, the spherical interface is locally approximated by a plane interface near the reflection points. At each internal reflection, partial energy leaks out the sphere through the transmission process and the magnitude of the remained wave decreases according to the reflection coefficient, which is determined by the incident angle, the polarization, and the internal/external dielectric contrast. Assume that a particular internal resonant mode is formed by a plane wave experiencing M internal reflections as illustrated in Fig. 7. It is easy to show that the time period T is $M(2a\cos\theta_i)/v_d$, where a is the radius, θ_i is the incident angle, and $v_d = 1/\sqrt{\epsilon_d \mu_d}$ is the internal phase velocity. Define an *effective damping factor* (EDF) α_e and the *normalized effective damping factor* (NEDF) χ as

$$e^{-\alpha_e T} \equiv |R|^M \quad \text{i.e.}$$
$$\alpha_e \equiv -\frac{v_d}{2a\cos\theta_i}\ln|R| \quad \text{and} \tag{9}$$
$$\chi \equiv \frac{a}{v_d}\alpha_e = -\frac{1}{2\cos\theta_i}\ln|R|. \tag{10}$$

The reflection coefficient R is defined as

$$R^{\mathrm{TE}} = \frac{Z_x/\cos\theta_t - Z_d/\cos\theta_i}{Z_x/\cos\theta_t + Z_d/\cos\theta_i}, \quad \text{or}$$
$$R^{\mathrm{TM}} = \frac{Z_x\cos\theta_t - Z_d\cos\theta_i}{Z_x\cos\theta_t + Z_d\cos\theta_i} \tag{11}$$

where $Z_x = \sqrt{\mu_x/\epsilon_x}$, $Z_d = \sqrt{\mu_d/\epsilon_d}$, and θ_t is the refraction angle.

The NEDF for the current bubble case whose internal and external dielectric constants are one and six, respectively, is shown in Fig. 8. Note that χ for both TE and TM mode are always larger than 0.43 which corresponds to waves bounced back and forth along the radial direction, similar to a "Fabry–Perot resonator." Note that TE modes have much smaller damping than TM modes. The behavior of the NEDF predicts correctly the imaginary parts of the actual CNR's observed in Fig. 6. The damping behavior of the internal CNR's of a dielectric bubble can be predicted by a simple model involving multiple reflections. As the dielectric contrast increases, the amplitude of the reflection coefficient also increases. The resonance becomes less damped.

IV. CNR's For General Dielectric Spheres

The CNR behavior for general dielectric spheres is now investigated by considering the internal and external resonances

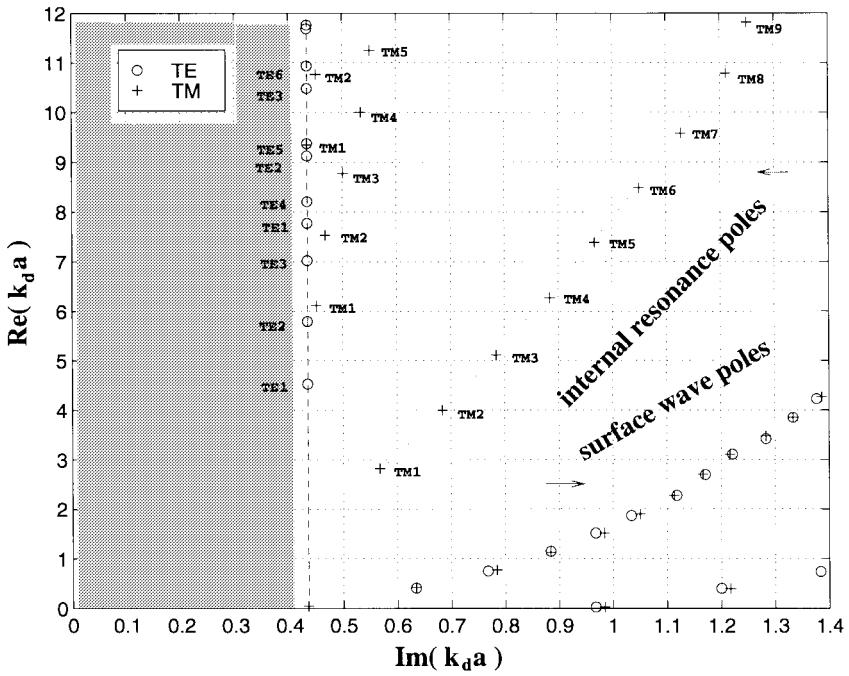


Fig. 6. Normalized poles ($k_d a$ space) for a dielectric bubble.

TABLE I
COMPARING THE NORMALIZED RESONANT FREQUENCIES $\mathrm{Re}(k_d a)$ OF THE FIRST AND SECOND MODES ($s = 1, 2$) POLES ($n = 1, \cdots, 5$) FOR A DIELECTRIC BUBBLE WITH THOSE FOR A PEC BUBBLE

$k_d a$	Dielectric Bubble				PEC Spherical Cavity			
n	$TE_{n,1}$	$TE_{n,2}$	$TM_{n,1}$	$TM_{n,2}$	$TE_{n,1}$	$TE_{n,2}$	$TM_{n,1}$	$TM_{n,2}$
1	4.52	7.77	2.81	6.12	4.49	7.73	2.74	6.12
2	5.80	9.13	4.00	7.51	5.76	9.10	3.87	7.44
3	7.02	10.48	5.12	8.77	6.99	10.42	4.97	8.72
4	8.20	11.68	6.28	9.98	8.18	11.71	6.06	9.97
5	9.38	12.99	7.38	11.24	9.36	12.97	7.14	11.19

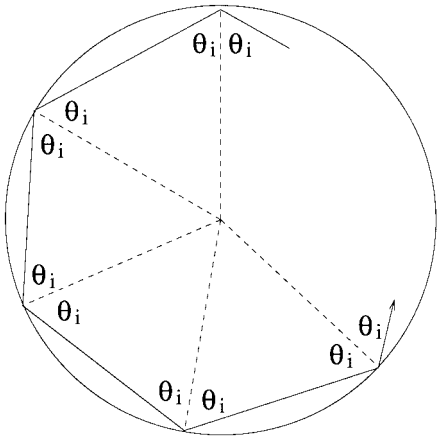


Fig. 7. An internal wave going through multiple internal reflections.

separately. Let us first examine the CNR's for a dielectric sphere ($\epsilon_r = 9$) in free-space, as shown in Fig. 9. The poles are plotted in $k_d a$ space to focus on the internal CNR's.

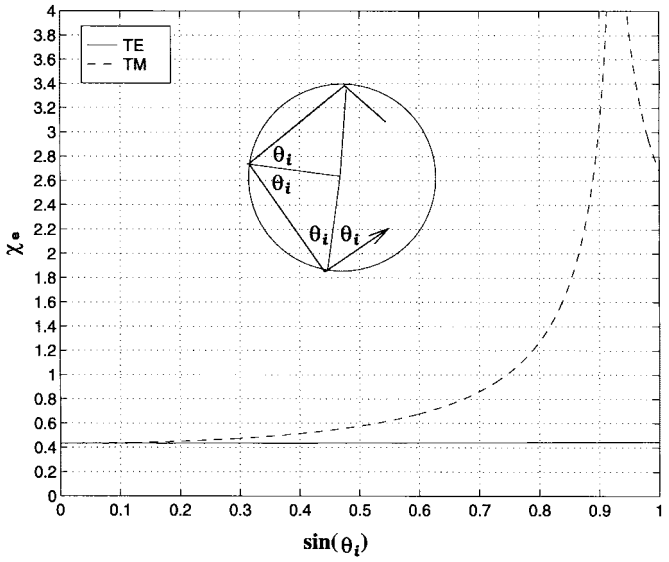


Fig. 8. Normalized effective damping factors (NEDF's) χ for dielectric bubble with internal and external dielectric being one and six, respectively.

Only a few of the external CNR's are visible as indicated by "surface wave poles." These CNR's show quite different behavior compared to those for the bubbles.

A. Internal Resonances

Several noticeable features are observed from Fig. 7. Both TE and TM modes can have very small imaginary parts and, thus, small damping. This is very important for practical applications. Second, all internal TE-mode CNR's are located to the left of a certain vertical line $\mathrm{Im}(k_d a) \approx 0.38$. Third, the imaginary parts of all internal TM-mode CNR's increase first as their order numbers increase and then decrease after

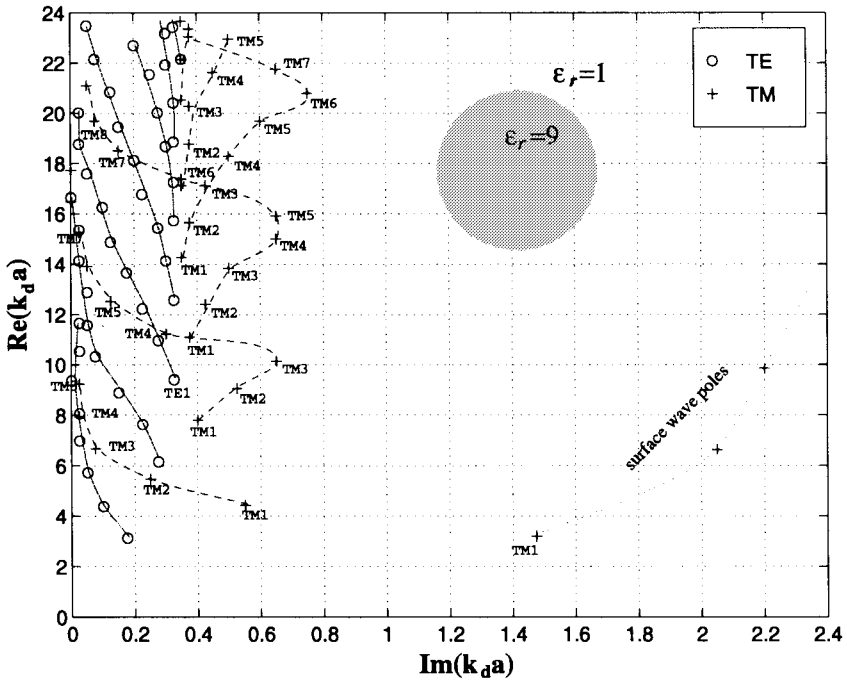


Fig. 9. Normalized poles for a dielectric sphere ($\epsilon_r = 9$) in free-space.

TABLE II
COMPARING THE NORMALIZED RESONANT FREQUENCIES $Re(d_d a)$ OF THE FIRST MODE ($s = 1$) POLES OF ORDER $n = 1, \cdots, 5$, FOR A DIELECTRIC SPHERE ($\epsilon_r = 9$) WITH THOSE FOR A PMC SPHERICAL CAVITY

$k_d a$	Dielectric Sphere				PMC bubble			
n	$TM_{n,1}$	$TM_{n,2}$	$TE_{n,1}$	$TE_{n,2}$	$TM_{n,1}$	$TM_{n,2}$	$TE_{n,1}$	$TE_{n,2}$
1	4.37	7.75	3.05	6.18	4.49	7.73	2.74	6.12
2	5.32	9.08	4.37	7.60	5.76	9.10	3.87	7.44
3	6.71	10.15	5.69	8.90	6.99	10.42	4.97	8.72
4	8.03	11.23	6.95	10.31	8.18	11.71	6.06	9.97
5	9.23	12.52	8.09	11.57	9.36	12.97	7.14	11.19

a certain order of numbers. The lowest branch TM modes become less damped after the very first pole. All these features will be examined below.

It was shown that the resonant frequencies inside a dielectric bubble are very close to those inside a PEC spherical cavity (see Tables I) since both have similar geometrical properties except that the dielectric bubble has a penetrable boundary, which results in damped resonances. Similarly, one would expect the internal resonant frequencies for a dielectric sphere to be close to those for a PMC spherical cavity. This prediction is immediately verified by the close agreement when comparing the resonant frequencies of the first two modes ($s = 1, 2$) for the five lowest orders ($n = 1, \cdots, 5$), as shown in Table II. The internal dielectric constant for both cases is equal to nine. Therefore, the internal CNR's for a dielectric sphere are caused by the same internal reflection mechanism as a in dielectric-filled PMC bubble.

It was shown that the internal CNR damping behavior associated with a dielectric bubble can be predicted from normalized effective damping factor (NEDF). Its damping was due to the partially reflected waves. Here, the NEDF is used again to predict the internal CNR damping behavior associated with a dielectric sphere. The new NEDF plot is shown in Fig. 10.

Total reflections occurs when the incident angle θ_i is greater than the critical angle θ_c, where $\sin\theta_c = 1/\sqrt{9} \approx 0.33$. It is also noticed that for TM mode, a zero reflection occurs at the Brewster angle θ_B where $\sin\theta_B = 0.3163$. It is also noticed that χ for the TE modes are always smaller than 0.35 and decrease as $\sin\theta_i$ increases until θ_i reaches the critical angle θ_c. Beyond the critical angle χ becomes zero and the CNR becomes undamped. For the TM mode, χ first increases from 0.35 as θ_i increases in the $\theta_i < \theta_B$ region. In $\theta_B < \theta_i < \theta_c$ region, χ decreases to zero again and remains to be zero for $\theta_i > \theta_c$. It is intriguing to notice that this unique TM mode "turning back" behavior agrees with what observed in Fig. 6. Further investigation about the turning point using the previous relation $\sin\theta_i \approx (n + 1/2/\mathrm{Re}(k_d a))$ with $\mathrm{Im}(k_d a)$ much less than $\mathrm{Re}(k_d a))$ shows that the turning points occur when $\theta_i \approx \theta_B$.

B. Internal and External CNR Coupling

In Fig. 9, the internal and external CNR's for dielectric sphere ($\epsilon_r = 9$) in free-space are well separated due to the very different imaginary part (or damping factor). As the dielectric difference between the sphere and the ambient medium decreases, the internal reflection coefficient decrease and, thus, the damping increases. This will move the internal CNR's closer to the external CNR's. Figs. 11 and 12 plot the CNR's for dielectric spheres whose dielectric constants are six and three, respectively. Comparing these with Fig. 9, one can clearly see that the internal CNR's become more damped.

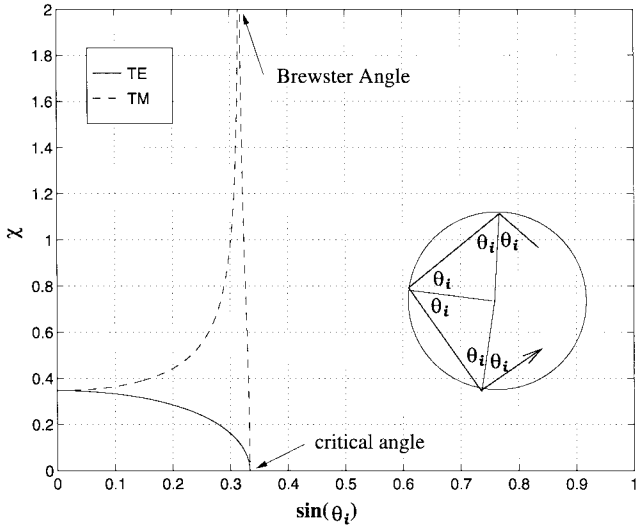


Fig. 10. Normalized effective damping factor (NEDF) χ for dielectric sphere with internal and external dielectric constant being nine and one, respectively.

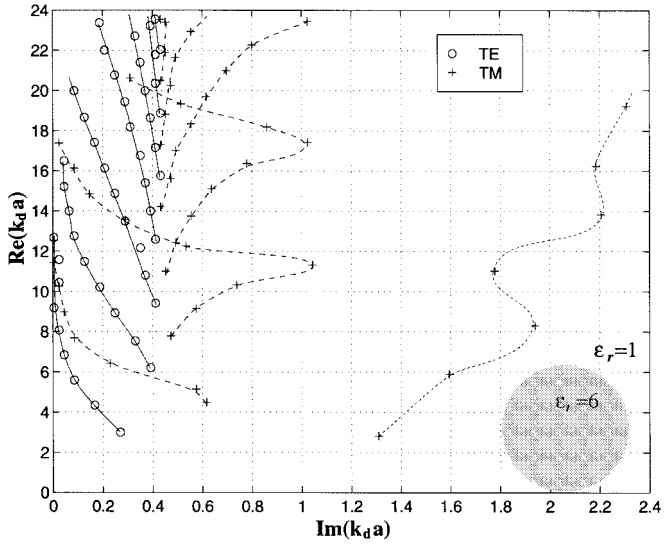


Fig. 11. Normalized poles for a dielectric sphere $(\epsilon_r = 6)$ in free-space.

Fig. 11 shows that some external pole locations belonging to the left-most string have been affected. This results in an unusual properties of this string. Much stronger coupling is also seen in Fig. 12 where the behavior of the first string of external CNR's have been completely altered as shown by the dotted line. The vertical spacing of this string is no longer the same as would be expected from a purely external resonances. The spacing of the high-order poles becomes similar to that of internal resonances. Strictly speaking, one can no longer classify them as either internal or external resonances.

C. External Resonances

It has been shown that the resonant frequencies and damping factors of external modes for a dielectric bubble are very similar to those for an impenetrable PMC sphere due to the similar surface wave mechanism caused by the continuous curvature diffraction. It was also shown that those external modes are insensitive to the internal medium. Similar properties will be shown to exist for the external resonances associated with a dielectric sphere except that the PMC sphere is replaced by a PEC sphere and the damping factors of the TE modes are more sensitive to the internal medium. Fig. 13 plots the poles for a dielectric sphere $(\epsilon_r = 9)$ in free-space in the $k_x a$ space. The TE- and TM-mode poles are marked by "$\circ$" and "$+$," respectively. For comparison, the normalized TE and TM poles associated with a PEC sphere are also plotted and marked by "$\times$" and "$*$," respectively. The following discussion focuses on the external resonant poles indicated by the strings marked by $A, B, C, \cdots$, etc. As one can see, the TM poles are located very close to those for a PEC sphere except for string A whose locations are affected by the internal resonance poles, as discussed earlier. It is also observed that all the external TE-mode poles deviate from the PEC poles toward higher damping factors (i.e., right of the figure). Nevertheless, all TE-mode resonant frequencies remain close to those associated with a PEC sphere whose resonant frequencies are determined by the phase velocities of the surface creeping waves and independent of internal material. Therefore, it is concluded that the external resonances of a dielectric sphere are caused by creeping waves similar to those on a PEC sphere except that the former have larger damping factors. For impenetrable spheres, the damping is due to the continuous power lost by radiating waves away from the spherical surface along the tangent direction. For a dielectric sphere, in addition to the radiation tangent to the surface, part of energy is continuously shedding into the sphere at the critical angle direction. The existence of such a wave mechanism was first suggested by van der Hulst [21] while studying the resonant peaks observed in the scattered fields from rain droplets. This unique process does not occur in the dielectric bubble case since no internal critical angle exists.

As the dielectric contrast decreases, the A-string pole locations become severely distorted as the internal CNR poles move further away from the real axis as discussed earlier. This is demonstrated by Fig. 14. All other surface wave poles also move farther away from the PEC poles to the right, as shown in Fig. 14.

V. Summary and Conclusion

The main contribution of this study is to provide a simple physical interpretation for predicting the resonance behavior of a general dielectric sphere. This is achieved by: 1) interpreting the internal standing waves as internal bouncing waves whose incident angles upon the sphere boundary are determined by $k_d a$ and n; 2) relating the internal and external resonances to those for impenetrable cavities and spheres, respectively; and 3) using a simple flat interface reflection/transmission model and the newly defined effective damping factor (EDF) to correctly predict the damping behavior of the internal resonances.

The resonances of a dielectric sphere (or bubble) were separated into external modes and internal modes. The former are caused by the surface creeping waves and have been shown to be very similar to PEC (or PMC) spheres. These external

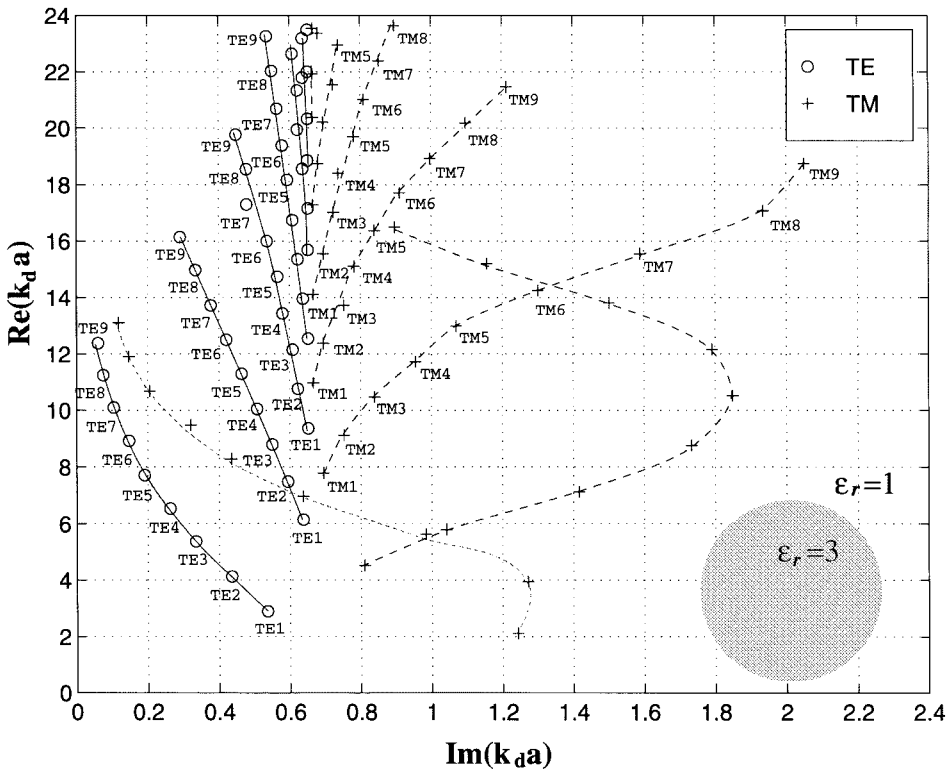


Fig. 12. Normalized poles for a dielectric sphere ($\epsilon_r = 3$) in free-space.

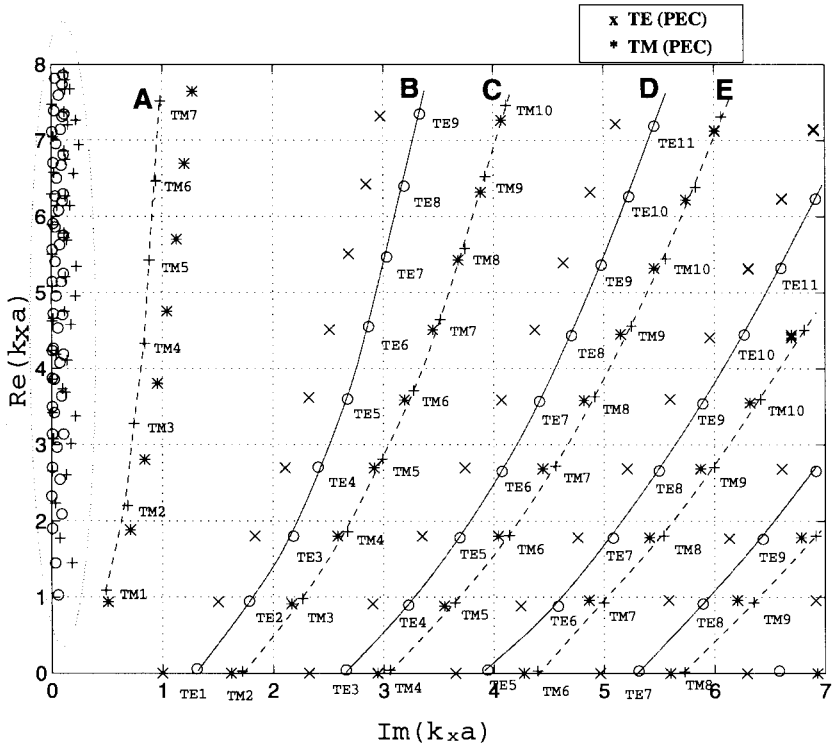


Fig. 13. Normalized poles in the $k_x a$ space for a dielectric sphere ($\epsilon_d = 9\epsilon_0$) in free-space. ($\circ$) TE modes ($+$) TM modes (x) TE modes for PEC sphere ($*$) TM modes for PEC sphere.

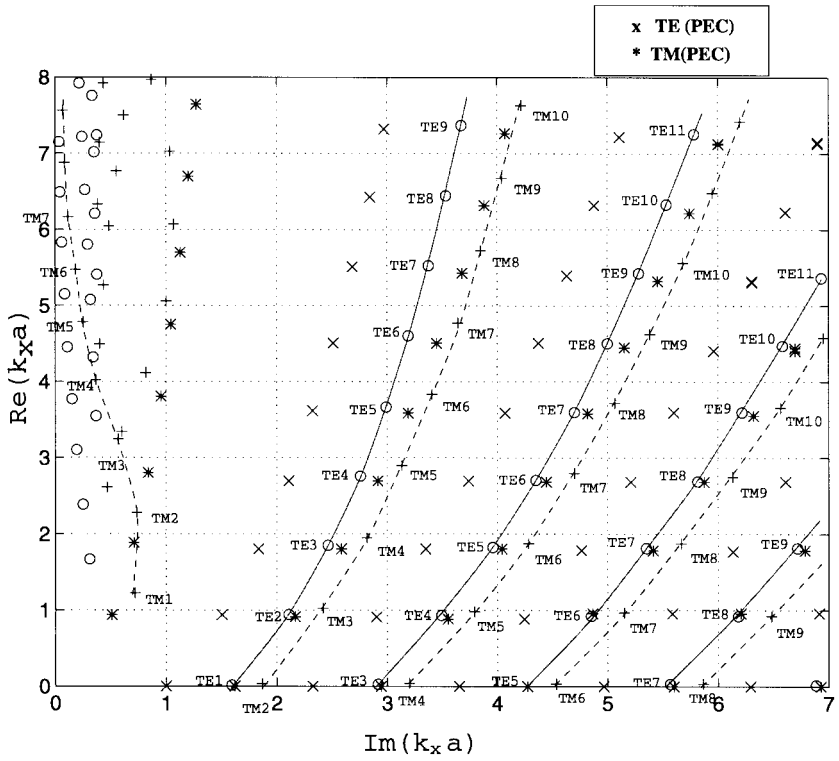


Fig. 14. Normalized poles for a dielectric sphere ($\epsilon_d = 3\epsilon_0$) in free-space. ($\circ$) TE modes ($+$) TM modes (x) TE modes for PEC sphere ($*$) TM modes for PEC sphere.

resonances are fairly independent of internal material. They are also found to have much larger damping factors compared to the internal resonances and, thus, are less important in practice. The internal resonances associated with a dielectric sphere (or bubble) were shown to be related to internal bouncing waves which experience multiple reflections. Their resonant frequencies were shown to be close to those of an impenetrable PMC (or PEC) bubble filled with the same dielectric material. It was also found that the internal resonances associated with a dielectric bubble are mainly in the radial direction. The damping of internal resonances is determined by the amount of energy transmitted out at each internal reflection. For both dielectric sphere and bubble, it was found that the internal TE modes are usually less damped than the TM modes. For dielectric spheres, it is interesting to find that higher TM modes could have very small damping factors when the internal incident angle is larger than the Brewster angle. A final note here is that almost all external resonance poles and off-axis internal resonance poles do not satisfy the large argument approximation, i.e., $ka \gg n$. Therefore, they are often not included in the literature that studies the resonances using large argument approximations.

From this study, it was found that the internal CNR's are less damped than the external ones. It was also found that the internal CNR's for a dielectric sphere are much less damped than those for a dielectric bubble. Therefore, it would be expected that the CNR's for buried plastic land mines would correspond to the bubble case and would be highly damped.

ACKNOWLEDGMENT

The author would like to thank Dr. L. Peters, Jr. for motivating this study and for his helpful discussions.

REFERENCES

[1] L. Peters and W. G. Swarner, "Approximations for dielectric or plasma scatters," *Proc. IEEE*, vol. 53, pp. 882–892, Aug. 1965.
[2] H. M. Nussenzveig, "High-frequency scattering by a transparent sphere—Part I: Direct reflection and transmission; Part II: Theory of the rainbow and the glory," *J. Math. Phys.*, vol. 10, pp. 82–177, Jan. 1969.
[3] W. Franz, "Uber die Greenschen funktionen des zylinders und der kugel," *Z. Naturforsch.*, vol. A-9, pp. 705–716, 1954.
[4] G. V. Frisk, J. W. Dickey, and H. Überall, "Surface wave modes on elastic cylinders," *J. Acoust. Soc. Amer.*, vol. 58, pp. 996–1008, Nov. 1975.
[5] G. V. Frisk and H. Überall, "Creeping waves and lateral waves in acoustic scattering by large elastic cylinders," *J. Acoust. Soc. Amer.*, vol. 59, pp. 46–53, Jan. 1976.
[6] A. Ashkin and J. M. Dziedzic, "Observation of optical resonances of dielectric spheres by light scattering," *J. Appl. Optics*, vol. 20, pp. 1803–1814, May 1981.
[7] T. S. Fahlen and H. C. Bryant, "Direct observation of surface waves on water droplets," *J. Optical Soc. Amer.*, vol. 56, pp. 1635–1636, Nov. 1966.
[8] Y. M. Chen, "Diffraction by a smooth transparent object," *J. Math. Phys.*, vol. 5, pp. 820–832, June 1964.
[9] W. Franz and P. Beckmann, "Creeping waves for objects of finite conductivity," *IRE Trans. Antennas Propagat.*, vol. AP-4, pp. 203–208, 1956.
[10] L. C. Chan, D. L. Moffatt, and L. Peters, "A characterization of subsurface radar targets," *Proc. IEEE*, vol. 67, pp. 991–1000, July 1979.
[11] P. Barber, J. F. Owen, and R. K. Chang, "Resonant scattering for characterization of axisymmetric dielectric objects," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 168–172, Mar. 1982.
[12] H. M. Nussenzveig, *Diffraction Effects in Semiclassical Scattering*. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[13] P. R. Conwell, P. W. Barber, and C. K. Rushforth, "Resonant spectra of dielectric spheres," *J. Acoust. Soc. Amer.*, vol. 1, pp. 62–67, Jan. 1984.
[14] R. F. Harrington, *Time-harmonic Electromagnetic Fields*. New York: McGraw-Hill, 1961, p. 294.
[15] L. R. Dragonette and L. Flax, "Relation between creeping waves and normal modes of vibration of a curved body," *J. Acoust. Soc. Amer.*, vol. 61, pp. 711–715, Mar. 1977.
[16] R. F. Harrington, *Time-harmonic Electromagnetic Fields*. New York: McGraw-Hill, 1961, pp. 269–270.
[17] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1972.
[18] C.-C. Chen, "Design and development of ground penetrating radar systems for the detection and classification of unexploded ordnances and land mines," Ph.D. dissertation, Ohio State Univ., Columbus, Aug. 1997.
[19] H. A. Haus, *Waves and Fields in Optoelectronics*. Englewood Cliffs, NJ: Prentice-Hall, 1984, p. 127.
[20] R. F. Harrington, *Time-harmonic Electromagnetic Fields*. New York: McGraw-Hill, 1961, pp. 292–298.
[21] H. C. van de Hulst, *Light Scattering by Small Particles*. New York: Wiley, 1957.



Chi-Chih Chen (M'92) was born in Taiwan, R.O.C. He received the B.E.E. degree from the National Taiwan University, in 1988, and the M.Sc. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, in 1993 and 1997, respectively.

In 1992 he became a Graduate Research Associate in the ElectroScience Laboratory, Department of Electrical Engineering, The Ohio State University, where he is now a Postdoctoral Researcher. His primary interests are in the fields of antenna design, signal processing, and radar measurements. He is currently involved in ground-penetrating radar applications for detecting and classifying unexploded ordnance and antipersonnel mines.

Dr. Chen is a member of Sigma Xi and Phi Kappa Phi honor societies.

A Novel Signal Processing Technique for Clutter Reduction in GPR Measurements of Small, Shallow Land Mines

Andria van der Merwe and Inder J. Gupta

***Abstract*—In this paper, a signal processing technique is developed to reduce clutter due to ground bounce in ground penetrating radar (GPR) measurements. This technique is especially useful when a GPR is used to detect subsurface antipersonnel mines. The GPR clutter is modeled using a simple parametric model. Buried mine and clutter contributions are separated through a pair of coupled iterative procedures. The algorithm outperforms existing clutter reduction approaches and also yields target features that are useful for detection and identification of these mines. The proposed technique effectively reduces clutter resulting in a significant decrease in false alarm rates.**

***Index Terms*—Clutter, ground penetrating radar (GPR), land mines.**

I. Introduction

THE global land mine crisis is creating immense social and economic problems worldwide [1], [2]. Safe and cost effective methods for clearing these mines are therefore needed. A promising sensor for detecting anti personnel mines is a ground penetrating radar [3]. However, GPR performs inadequately due to clutter, which dominates the data and obscures the mine information (the dominant clutter contributors being the strong returns received from the rough ground surface). These clutter contributions vary with soil conditions and/or surface roughness and lead to uncertainty in the measurements.

Antipersonnel mines are buried at shallow depths. Thus, returns from the mines and that from the rough ground interface overlap in time. It is very difficult to discriminate these coincident responses. Mines are made out of plastic materials and buried in lossy soil. The relative returns from the mines are therefore very low in energy. Commonly used complex natural resonance frequencies are highly damped due to the lack of dielectric contrast between the soil and the mine and are thus not suitable for detection [4]. The GPR also receives returns from other subsurface inhomogeneities, for example rocks, tree roots, or small pieces of metal in the ground, which leads to high levels of false alarms. Due to these physical limitations, clutter reduction and detection of small shallow land mines are very challenging.

In this paper, a new signal processing technique, which can be used to reduce clutter through data preprocessing, is presented. As opposed to most current clutter reduction techniques, which model clutter statistically [5]–[7], the new clutter reduction technique models clutter using parametric modeling. Frequency domain basis functions are developed to represent the clutter and the mine contributions in GPR data. An iterative signal processing algorithm is developed to estimate the unknown parameters in the basis functions and reduce the clutter. The algorithm developed in this paper improves existing signal processing techniques by incorporating an adaptive basis function for clutter representation. The new algorithm is robust to the variability of clutter between measurements and accounts for the uncertainty in GPR clutter characteristics. The fact that land mines are buried at shallow depths and that the returns from land mines are small relative to that of the clutter are compensated for in the new processing technique. Note that the returns from subsurface inhomogeneities, for example, rocks and tree roots are also treated as part of the clutter here. To access the performance of the clutter reduction technique, the cross correlation between a reference signature of the target of interest and the clutter reduced data is used. To quantify the performance of the clutter reduction processing, improvement in the detection is accessed through detector receiver-operating-characteristic (ROC) curves. Note that the clutter reduction processing precedes the detector. A simple matched filter, formulated to account for the uncertainty in the placement of the mine is used as a detector.

The proposed technique is applied to some simulated as well as some measured GPR data. A two-dimensional (2-D) finite-difference-time-domain (FDTD) algorithm is implemented to simulate land mines buried in a lossy clay medium under a rough ground surface. The FDTD technique is chosen here because it is ideally suited for modeling scatterers surrounded by inhomogeneous media, and does not require complicated Green's functions, which is typically used in the method of moments. Another advantage of using the FDTD technique is that the wide band nature of the GPR sensor can easily be incorporated into the formulation.

The rest of this paper is organized as follows. Some existing clutter reduction techniques are discussed in Section II. The proposed iterative clutter reduction technique is developed in Section III. A brief description of the FDTD simulation is given in Section IV. In Section V clutter reduction is shown using both

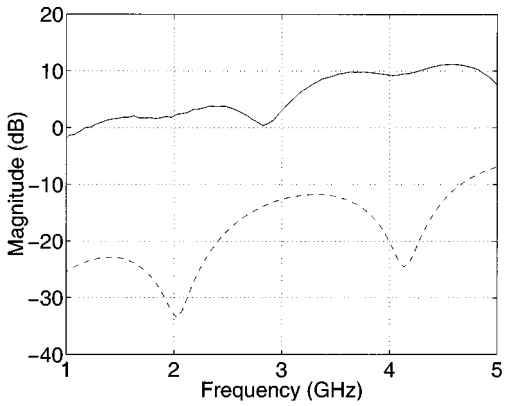
Manuscript received April 26, 1999; revised March 16, 2000. This work was supported by Duke University under an award from the ARO (OSD MURI Program). The findings, opinions, and recommendations expressed therein are those of the authors and are not necessarily those of Duke University or the ARO.

A. van der Merwe was with The Ohio State University, Columbus, OH 43210 USA, and is now with Andrew Corporation, Dallas, TX 75201 USA.

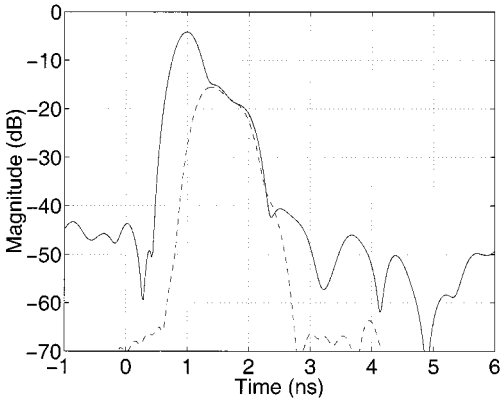
I. J. Gupta is with the ElectroScience Laboratory, Department of Electrical Engineering, The Ohio State University, Columbus, OH 43210 USA.

Publisher Item Identifier S 0196-2892(00)07155-2.

0196–2892/00$10.00 © 2000 IEEE



(a)



(b)

Fig. 1. Illustrations of the total simulated data and the reference target signature (a) in the frequency domain and (b) in the time domain. The reference signature is shown in the broken line. The total data in the frequency domain has been displaced by 10 dB.

simulated as well as measured GPR data. A quantitative performance evaluation of the new clutter reduction technique is presented in Section VI. Section VII contains conclusions.

II. Existing Data Processing Techniques

Some state-of-the-art clutter reduction techniques are discussed in this section. These techniques proceed by estimating the coherent component of the clutter. The coherent component will generally be the dominant clutter contributor and it will be confined to early observation times.

The coherent clutter component originates at the ground surface. Early time gating can therefore be used to reduce this component of the clutter. However, for shallow targets the responses from the targets and that from the clutter overlap and choosing an appropriate time gate is very challenging. This can clearly be seen from the time domain curves shown in Fig. 1(b). The reference target is buried at a depth of 2.5 cm. Due to the overlap between the return from the shallow target and that from the ground surface response, the target contributions will be removed using time gating. Another drawback of this approach is that the side lobes of the ground return dominate the late time that will obscure contributions from deeper targets.

Another method of clutter reduction is complex average subtraction. An estimate of the average is obtained by calculating the mean of a number of measurements without a buried object [8]. However, in reality, the soil is an inhomogeneous medium and its statistical properties vary with position along the surface. Thus, subtracting the average may not lead to sufficient clutter reduction. Subtracting a moving average estimate of the background is another possible technique [9]. The background estimate will however be affected by the presence of an object, and the estimate will not be accurate. Other techniques for estimating the background also performs marginally [10] due to variations and inhomogeneities of the ground.

Peak subtraction has also been proposed for reducing the ground surface return. For example, the early time peak can be synthesized in the time domain using a sinc-approximation and subsequently subtracted [11]. However, the presence of shallow targets leads to a bias in the amplitude and location of the early time estimate. Therefore, clutter reduction may not be accurate.

An improved peak subtraction approach has been suggested where the early time contributions are approximated using a superposition of damped exponentials in the frequency domain [12]. The TLS-Prony technique [13] is used to estimate the necessary exponential parameters. The estimated early time clutter is then subtracted from the data to reduce its contribution. This method is referred to as early time subtraction (ETS) and is equivalent to a single iteration of the proposed algorithm.

The clutter reduction techniques discussed above may reduce the coherent component of the ground clutter, but the incoherent component of the clutter, which is typically spread out in time, is still present in the data. Assuming that the incoherent clutter can be represented as a random process with known statistics, a whitening filter can be used to account for this component of the clutter. The whitening filter is applied to the GPR problem in [5]. The clutter statistics are determined from many realizations of the rough surface. The whitening filter is implemented as a linear prediction error filter [14]. A limitation of this technique is the fact that in reality, the soil is an inhomogeneous medium and therefore, its statistical properties vary with position along the surface. Any mismatch in the clutter statistics used to design the filter and that of the GPR data degrades the performance of the whitening filter. Clutter is therefore present in the data after processing.

Accurate estimation of the clutter is crucial in all the above techniques. Finding an unbiased estimate is indeed a very challenging task due to the uncertainty and variation in the ground scattering. Approaches that incorporate estimates that are robust to changes in the scattering from environment and take the variability of the medium into account show more promise in this application. Such a technique is developed in this paper and is discussed next.

III. Proposed Clutter Reduction Technique

In the case of high resolution, ultra-wideband (UWB) GPR where only a small spot on the ground surface is illuminated, it is reasonable to assume that the scattered response is composed of a superposition of responses from individual scattering mechanisms. Based on this observation the total scattered field received by the GPR can be represented as

$$S(\omega) = H_c(\omega) + T(\omega) + n(\omega) \tag{1}$$

where $H_c(\omega)$ represents the clutter contributions, $T(\omega)$ represents the contribution from the desired target, $n(\omega)$ represents additive Gaussian noise, and $\omega = 2\pi f$. It is important to note that $H_c(\omega)$ represents all the unwanted contributions, for example, scattering from the rough surface as well as interaction terms and contributions from subsurface soil inhomogeneities. These unwanted contributions are collectively referred to as clutter. Furthermore, it is assumed that scattering from only a single target contributes to the response and that the effects of neighboring subsurface scatterers are negligible. These assumptions are reasonable in the case of the small illuminated area of interest here, but need to be addressed in future research.

The individual contributions in (1) are represented using parametric basis functions. First consider the development of the functional form of the clutter basis function. An exact electromagnetic model of the scattering should include parameters such as the physical optics reflection coefficient and radius of curvature to succinctly describe the local scattering features. However, the objective of the present study is clutter reduction and a parametric function is therefore chosen that can be adjusted to match the frequency dependent behavior of the clutter in a given GPR measurement. A possible functional form of the basis function is found from the frequency dependence of canonical scatterers as predicted by the geometrical theory of diffraction [15]. This has been used to extract features from SAR measurements of airborne targets by Moses and Potter [16]. Another possibility is to use a damped exponential model [17]. The latter model is suggested here because it is simple and computational attractive algorithms are available for estimating its parameters. Using the damped exponential model

$$H_c(\omega) = \sum_{k=1}^{K} A_k e^{-\omega\alpha_k} e^{-j\omega t_k} \tag{2}$$

where K is the total number of damped exponential terms needed to model the clutter, α_k is the damping factor, and t_k is the time delay of the kth term for $k = 1, \cdots, K$. Although the unknown parameters cannot be associated with any physical scattering mechanisms, these effectively model the clutter contributions. This will be shown using multiple examples in Section V.

The targets of interest have a regular, fixed shape with a known geometry. It is assumed that a reference signature $T_r(\omega)$ for the target of interest is known or can easily be obtained. The reference signature can be obtained from a simulation of the target in a homogeneous medium. Another method is to simulate the target buried deep underneath a flat ground surface. The flat surface response is subtracted and time gating is used to separate the return from the target. In the case of measured data, the reference signature is extracted from a calibration measurement of the desired target buried deep below the ground surface. The buried target return in (1) differs from this reference signature in two ways. First, the depth of the subsurface target is unknown and second, the level of the scattering may be different from that of the reference response due to the attenuation of the wave propagating though the soil. Plane wave propagation through a lossy, medium [18] is utilized to develop parametric functions to be used in conjunction with the reference signature to form a basis function for the target response. The wave propagating toward the target will be delayed with respect to the time origin defined at the ground surface. A linear phase factor $e^{-j\omega t_r}$ is used to account for this time delay t_r, which is related to the depth of the target. To account for the frequency independent component of the received signal, a complex amplitude factor A_r is included in the model. To account for the frequency dependence in the loss, a damping factor $e^{-\omega\gamma_r}$ is used where γ_r represents the damping factor. These factors are combined to yield the basis function for the target

$$T(\omega) = A_r e^{-\omega\gamma_r} e^{-j\omega t_r} T_r(\omega). \tag{3}$$

Note that (3) is not an exact electromagnetic model for the response from a subsurface target buried in a realistic soil medium, but it can be used to approximate the frequency dependence of the scattering from the target.

Using the basis functions defined in (2) and (3), the GPR data $S(\omega)$ can be represented by

$$S(\omega) = \sum_{k=1}^{K} A_k e^{-\omega\alpha_k} e^{-j\omega t_k} + A_r e^{-\omega\gamma_r} e^{-j\omega t_r} T_r(\omega) + n(\omega) \tag{4}$$

with the parameters, as defined above. The unknown parameters can be estimated and substituted into (4) to reconstruct the target and clutter contributions. Conceptually, the estimated clutter can be subtracted from the GPR data to minimize its contribution. However, direct minimization using (4) is computationally intensive and involves nonlinear minimization over a high-dimensional parameter space. To avoid the computational complexity, an iterative processing technique is developed. Apart from the computational advantages of using this procedure, the use of an adaptive estimation of the unknown parameters, accomplished by the iterative process, is also essential in determining an unbiased clutter estimate because the clutter estimate may be biased due the presence of a shallow target.

A conceptual block diagram of the iterative technique is shown in Fig. 2. Note that the technique involves two coupled iterative loops. Subscript m indicates the iteration number of the outer loop, and subscript n indicates the iteration number of the inner loop. The inputs to the algorithm are the total data in the frequency domain $S(\omega)$, and the reference signature for the target of interest $T_r(\omega)$. The inner loop is discussed first.

Define the following data set at the m^{th}, n^{th} iteration

$$Q^{m,n}(\omega) = S(\omega) - \hat{T}^{m,n-1}(\omega) \tag{5}$$

where $S(\omega)$ represents the scattered field data in (1), and $\hat{T}^{m,n-1}(\omega)$ represents the target estimated in the previous iteration. Note also that an estimate of the target is required and it is assumed that $\hat{T}^{1,0}(\omega) = 0$.

Represent the data set $Q^{m,n}(\omega)$ using a superposition of complex damped exponentials as follows:

$$Q^{m,n}(\omega) = \sum_{k=1}^{I_t} A_k^{m,n} e^{-\omega\alpha_k^{m,n}} e^{-j\omega t_k^{m,n}} \tag{6}$$

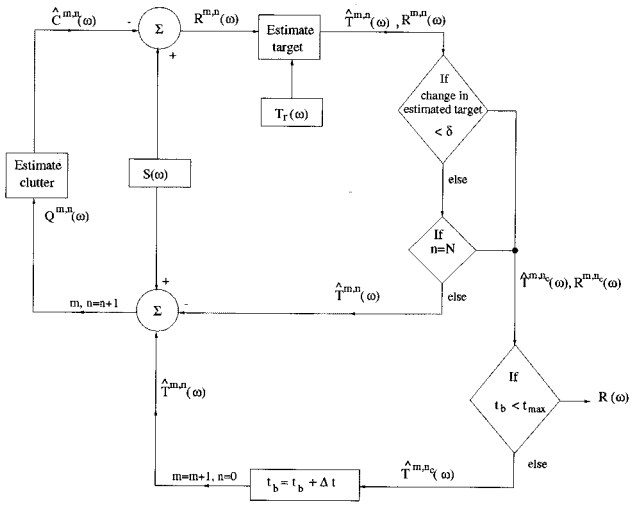


Fig. 2. Block diagram of the iterative algorithm.

where I_t is the total number of damped exponential terms needed to model the data set, $\alpha_k^{m,n}$ is the damping factor, and $t_k^{m,n}$ is the time delay of the kth term for $k = 1, \cdots, I_t$ at the nth, and mth iteration. The total least square (TLS) Prony Technique [13] is used to estimate these parameters. Some of the estimated complex damped exponentials represent the ground return and interaction terms, while others represent the target, subsurface inhomogeneities, and noise. At this point, it is desired to estimate only the clutter contributions. Thus, the poles that represent the clutter should carefully be identified. The fact that the dominant clutter resides in the earlier time contributions is used in choosing an appropriate time window T_g. Only the poles with estimated time locations within this time window are selected. At first, when m is small, a very narrow window is selected to ensure that the contribution from the shallow target is not removed. However, the first window has to be wide enough to include at least one high energy clutter pole. The time window can be adjusted to remove later time clutter as well. This is achieved in the outer loop and it will be discussed shortly. The selected complex exponential parameters are used to reconstruct the estimated clutter

$$\hat{C}^{m,n}(\omega) = \sum_{k=1}^{I_{m,n}} A_k^{m,n} e^{-\omega \alpha_k^{m,n}} e^{-j\omega t_k^{m,n}} \tag{7}$$

where $I_{m,n} \leq I_t$ and $A_k^{m,n}$, $\alpha_k^{m,n}$ and $t_k^{m,n}$ represent the complex exponential parameters of the kth pole within the selected time window at the nth, mth iteration. Note that the number of complex poles as well as the estimated parameters may change from one iteration to the next. The number of exponential terms $I_{m,n}$ needed to model the clutter in a particular measurement typically depends on the surface roughness. A higher number of terms is required to model the scattering from a rougher surface.

Subtracting the estimated clutter from the GPR data yields the residual data set defined as

$$R^{m,n}(\omega) = S(\omega) - \hat{C}^{m,n}(\omega) \tag{8}$$

which contains the leftover clutter contributions and/or the target contribution. The residual data set is used to obtain an estimate of the target contributions. The unknown parameters in the target basis function are determined by minimizing the least squares error (LSE) between the target model and the residual data set

$$\text{Error}^{m,n}(\overline{\theta}) = \left[\sum_{\omega_1}^{\omega_N} |R^{m,n}(\omega) - A_r^{m,n} e^{-\omega\gamma_r^{m,n}} e^{-j\omega t_r^{m,n}} T_r(\omega)|^2\right] \tag{9}$$

where ω_1 and ω_N are the end band frequencies and $\overline{\theta}$ is the parameter vector defined as $\overline{\theta} = [A_r^{m,n}, \gamma_r^{m,n}, t_r^{m,n}]$. A simplex downhill method [19] is used to minimize (9). The simplex method is desirable because it is numerically stable and does not require a gradient of Hessian of the cost function. At convergence, the simplex downhill optimization yields estimates of the unknown parameters which are used in (3) to reconstruct the target response at the mth, nth iteration

$$\hat{T}^{m,n}(\omega) = \hat{A}_r^{m,n} e^{-\omega\hat{\gamma}_r^{m,n}} e^{-j\omega\hat{t}_r^{m,n}} T_r(\omega). \tag{10}$$

The inner loop iterations are terminated either when the change in the estimated target[1] is less than a predefined threshold δ or when the maximum number of iterations are exceeded. Typical values for the threshold is $\delta \approx 10^{-3}$ while N is selected to be ten. The following output data are obtained at convergence of the inner loop:

$R^{m,n_c}(\omega)$	residual data set;
$\hat{T}^{m,n_c}(\omega)$	estimate of the target response;
$\hat{t}_r^{m,n_c}$	estimate of the time delay;
$\hat{\gamma}_r^{m,n_c}$	estimate of the damping factor.

Here, n_c indicates the number of the iteration at which the algorithm converged for a particular time window T_g and thus, a particular iteration m. The output available at convergence of the inner loop is the input to the outer loop. The outer loop is used to increase the width of the time window T_g. As alluded to above, the first time window is selected very narrow to ensure estimation of the early time clutter only. Thus, $T_g = [t_a, t_b]$ is selected such that $t_a < 0$ and t_b are small.[2] However, it may be desirable to remove the clutter at later observation times as well. This can be achieved by increasing the time window iteratively with increments Δt until the desired size of the time window $T_g = [t_a, t_{\max}]$ is reached. The increment Δt is small at small values of m to prevent the removal of shallow target contributions. At larger values of m, when most of the target contributions have been removed, the increment can be increased. Note that for $m \neq 1$, $\hat{T}^{m,0}(\omega) = \hat{T}^{m-1,n_c}$, while for $m = 1$, $\hat{T}^{1,0}(\omega) = 0$.

At convergence of the iterative procedure, two decomposed data sets are obtained, namely, the estimated clutter contribution and a residual data set. The residual data set is dominated by the buried target signature which can be further processed to yield target identification or detection. Note that the estimated

[1] $(1/E^{m,n}) \sum_{\omega_1}^{\omega_N} |\hat{T}^{m,n}(\omega) - \hat{T}^{m,n-1}(\omega)|^2 < \delta$, where E^n represents the energy in the estimated target at the current iteration.

[2] The reader is reminded that the time origin is defined at the ground surface.

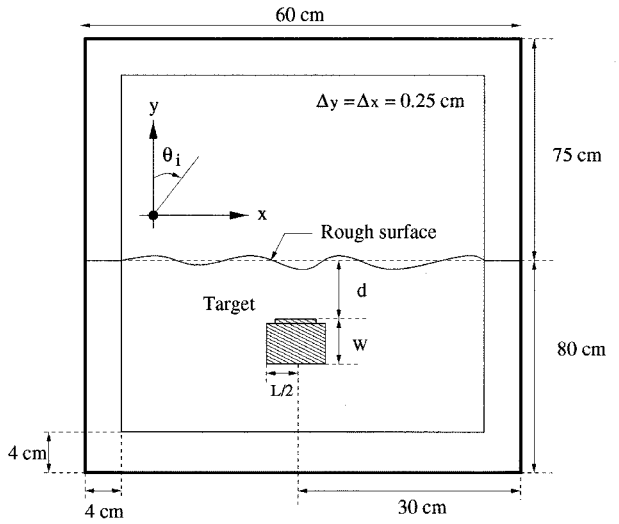


Fig. 3. Schematic of the FDTD computational domain.

target is not used in further processing since it may lead to unwanted detections for measurements without the target of interest. In some cases the residual data may also contain the contributions from subsurface soil inhomogeneities. However, further processing can be done to classify these as false alarms. The algorithm also provides an estimate of the depth of the buried target. This estimate provides valuable information regarding the placement of the mines in the ground. An estimate of the amplitude of the scattering from the target as well as an estimate of the loss in the soil is also obtained. These additional estimates may be used as features in target identification.

The performance of the proposed technique has been evaluated using some simulated as well as some measured GPR data. A FDTD algorithm was implemented to simulate mines buried in a lossy clay medium underneath a rough ground surface. A brief description of the FDTD algorithm is given next.

IV. FDTD Simulation

The two-dimensional (2-D) FDTD algorithm incorporates a lossy half space, rough surface and a buried dielectric target or a buried randomly shaped inhomogeneity. The algorithm is implemented in the time domain, starting from Maxwell's equations and employing the classical Yee algorithm [20]. The geometry and dimensions of the FDTD grid used here are shown in Fig. 3.

To ensure stability of the FDTD computations, the time step needs to satisfy the Courant Stability condition [20], which relates the time step Δt to the spatial increments as follows:

$$\Delta t \leq \frac{1}{v_g(1/\Delta^2 x + 1/\Delta^2 y)} \tag{11}$$

where Δx and Δy are of the spatial increments on the 2-D grid, and v_g is the speed of wave propagation in the soil. Numerical dispersion causes the wave to propagate in the FDTD mesh at a speed less than the physical speed of the wave, giving rise to a non physical distortion. To combat this effect, the spatial discretization is chosen to be more than ten samples per wavelength at the highest frequency of interest [21].

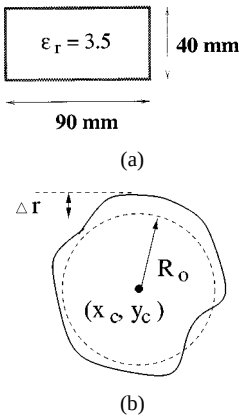


Fig. 4. Illustrations of (a) the simulated target and (b) the subsurface inhomogeneity.

The FDTD grid is surrounded by an absorbing boundary condition to absorb outwardly propagating waves at the edges of the finite computational domain [22]–[25]. The perfectly matched layer (PML) is implemented here. Note that the original PML formulation is adjusted when applied to the lossy soil medium. The parameters for each PML layer is chosen to yield an impedance match with the adjacent medium.

Plane wave incidence is studied here. Due to the finite size of the FDTD computational domain, the amount of energy in the reflected waves is finite and will be determined by the size of the finite illuminated spot on the ground surface which is in turn a function of the size of the grid along the ground surface. The plane wave has a Gaussian pulse shape. The parameters of the Gaussian pulse are chosen such that the incident field has most of its energy in the frequency band between 1 GHz and 5 GHz. Results are presented for normal incidence and TE_z polarization. All scattered fields are observed in backscatter.

In our simulation, the interface between the soil and the free space is modeled as a rough surface. The surface roughness is assumed to have a Gaussian spectrum [26].[3] The spatial correlation function for the rough surface is defined as

$$\rho(x) = h^2 e^{-x^2/l^2} \tag{12}$$

where h is the root mean square (RMS) surface height, and l is the correlation length. The interface (surface profile) varies randomly between surface generations.

A deterministic target, which is composed of a homogeneous, loss less dielectric medium, is also considered in these simulations. The relative dielectric constant of the target is fixed at $\epsilon_r = 3.5$, which is a typical value for plastic land mines [27]. The cross section is chosen to be representative of commercial plastic land mines. A schematic of the dimensions and geometry of the target is shown in Fig. 4. The depth of the target is defined as the distance d between the top of the target and the mean position of the rough surface as shown in Fig. 3. The depth is varied between 1.25 cm and 10 cm here.

Scattering from randomly shaped subsurface inhomogeneities is also studied here. Soil inhomogeneities, for

[3]The actual distribution for the surface roughness is an area of current research. Note that the proposed clutter reduction technique is independent of this distribution.

example, refilled trench holes, rocks, tree roots, and in particular, disturbed clay introduces a distortion of the received waveform, which depends on the location and characteristics of the inhomogeneity. The intensity of the scattering depends on the soil type and depth of the anomaly [28].

A random shape is generated using the following perturbed circle formulation:

$$\sqrt{((x-x_c)^2+(y-y_c)^2)} \leq R_o + \Delta r(x,y) \qquad (13)$$

where (x_c, y_c) are the coordinates of the center of the circle with a radius R_o. The perturbation to the circle at each (x, y) location is defined as

$$\Delta r(x,y) = \sum_{i=1}^{5} a_i \cos\left(\frac{2\pi}{T_o} i\theta\right) \qquad (14)$$

where $\theta = \arctan(y/x)$ and $T_o = 2\pi R_o$. The coefficients $a_i, i = 1, 2, \cdots, 5$ are random variables chosen from a uniform distribution between -0.5 and 0.5. In the current study, R_o is selected to be 2.5 cm. A schematic representation of the geometry of the inhomogeneity is also shown in Fig. 4.

V. Clutter Reduction Examples

In the first example, the iterative technique is applied to a simulated data set generated using the time domain FDTD algorithm. The data set is transformed into the frequency domain to simulate a wide band GPR measurement between 1 GHz and 5 GHz. The frequency sampling is $\Delta f = 50$ MHz. The reference signature for the target is obtained as discussed in Section III.

The target used in this example has a rectangular cross section which resembles the target shown in Fig. 4. It is composed of a homogeneous, lossless dielectric with a relative dielectric constant of $\epsilon_r = 3.5$. The surrounding soil resembles a dry clay medium with a relative dielectric constant of $\epsilon_r = 9$ and a conductivity of $\sigma_g = 0.05$ S/m. The target is buried at a depth of 2.5 cm beneath the mean position of a rough surface in the lossy soil. The Gaussian surface roughness parameters are $h = 2$ cm and l = 35 cm. White Gaussian noise is also added to the simulated data.[4]

The total data is shown in Fig. 1(a) and (b) in the frequency and time domains, respectively. Note that the curve that represents the total field in the frequency domain is displaced by 10 dB in all these results. The reference signature is also shown in the same figures. The reference signature and clutter reduced data sets at SNR's of 40 dB, 30 dB and 20 dB respectively are next compared. The residual data set after average subtraction, a single iteration (ETS), and multiple iterations of the algorithm at 40 dB are shown in Fig. 5. The simulated data set is also shown. When data is preprocessed using average subtraction and ETS, not all the clutter is removed. This leads to a mismatch between the residual data and the reference signature. The iterative technique sufficiently reduces the clutter and a very good agreement between the latter two signatures is obtained. Similar results are shown in Figs. 6 and 7 for SNR's of 30 dB and 20 dB, respectively. The performance of the average subtraction and ETS degrades at lower SNR values as expected. However, even at higher levels of noise there is still a good agreement between the reference signature and the residual data after application of the iterative algorithm.

[4]The SNR is defined in the frequency domain as follows, SNR = 10 $\log_{10}(E_s/E_n)$ dB where E_s and E_n denotes the energy in the simulated data and Gaussian noise respectively.

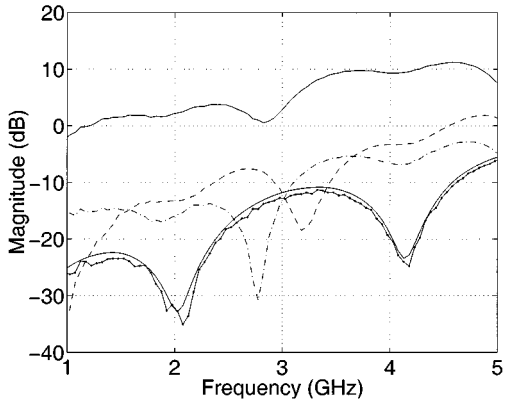


Fig. 5. Comparison between the reference signature and the residual data sets obtained after clutter reduction for SNR = 40 dB. The residual data after average subtraction is shown in the dashed line, after ETS in the dashed-dotted line and after the iterative technique in the solid-dotted line.

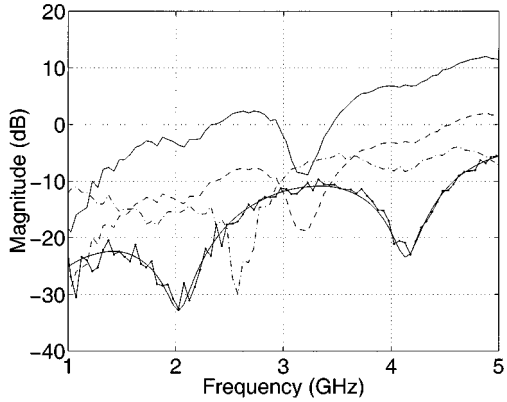


Fig. 6. Comparison between the reference signature and the residual data sets obtained after clutter reduction for SNR = 30 dB. The residual data after average subtraction is shown in the dashed line after ETS in the dashed-dotted line and after the iterative technique in the solid-dotted line.

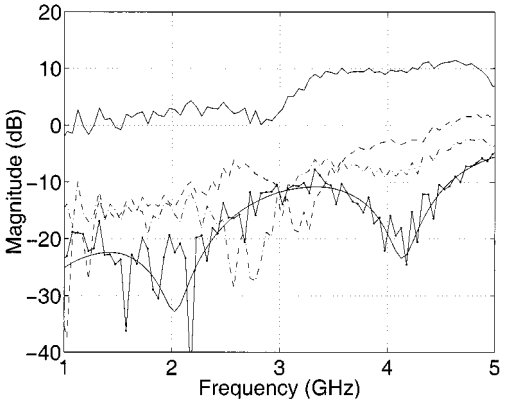


Fig. 7. Comparison between the reference signature and the residual data sets obtained after clutter reduction for SNR = 20 dB. The residual data after average subtraction is shown in the dashed line, after ETS in the dashed-dotted line and after the iterative technique in the solid-dotted line.

In the next example, the clutter reduction techniques are applied to 20 Monte Carlo realizations, which simulate GPR data collected along the ground surface over a surrogate minefield. In the first example, each Monte Carlo realization contains either a rough ground surface without any buried objects or a rough ground surface with the dielectric target buried at various depths. The depth varies between 1.25 cm and 5 cm. The statistics of the rough surface are the same as before. The surrounding soil resembles a dry clay medium with the same parameters as defined above. White Gaussian noise is added to the simulated data to yield a SNR $=$ 30 dB. The rough surface profile and additive noise is changed from one realization to the next. The reference signature is obtained as before. The time domain data for the 20 realizations are shown in Fig. 8. Note that some of the time domain responses contain subsurface targets. From these results, it can be seen that there is still some clutter left after applying the average subtraction and the ETS, as shown in Fig. 8(b) and (c). The clutter is, however, significantly reduced when using the iterative technique, as shown in Fig. 8(d). Also note that the target signature is not distorted in the latter case, and that the clutter can be reduced even in the cases where the target is buried at a very shallow depth.

Next, consider a data set where some of the Monte Carlo realizations contain subsurface inhomogeneities. Note that the latter contributions may be treated as clutter or false alarms. To simulate the uncertainty in these scatterers, the shape and dielectric composition of the subsurface inhomogeneities are randomly varied between realizations. The relative dielectric constant is uniformly varied between 4.5 and 7.5. The conductivity is set equal to that of the surrounding medium which is $\sigma_g = 0.05$ S/m. The shape is varied using the formulation given in (13). The value of the radius R_o used here is 25 mm. The depth of these inhomogeneities, defined pictorially in Fig. 4(b), is 2.5 cm. The realizations in the time domain are shown in Fig. 9. As before, it can be seen that after preprocessing using average subtraction or time gating, most of the clutter as well as the all the inhomogeneities are still present. After processing using the iterative approach, the clutter contributions from the ground as well some of the inhomogeneities are removed. We have extensively studied the performance of the proposed clutter reduction technique for targets buried in different media under different surface roughness statistics. In some simulations, the target was tilted with respect to the ground surface. The proposed technique performed very well in all cases [29].

Next, the iterative technique is applied to some experimental GPR data collected at Fort A.P. Hill, VA. Measurements were performed using a near field probe antenna. Data is acquired between 1 GHz and 5 GHz in increments of 100 MHz. Several scans of the test grid are shown in the time domain in Fig. 10. The mine of interest in these measurements is the TS-50 plastic land mine, shown in Fig. 11. The reference signature is obtained from a measurement of this mine in a calibration grid. The iterative technique is used to reduce clutter from the frequency domain data as discussed before. The processed data is transformed to the time domain using a Fourier transform with uniform weights and the result is shown in Fig. 12. The TS-50 mine can clearly be seen after clutter reduction.

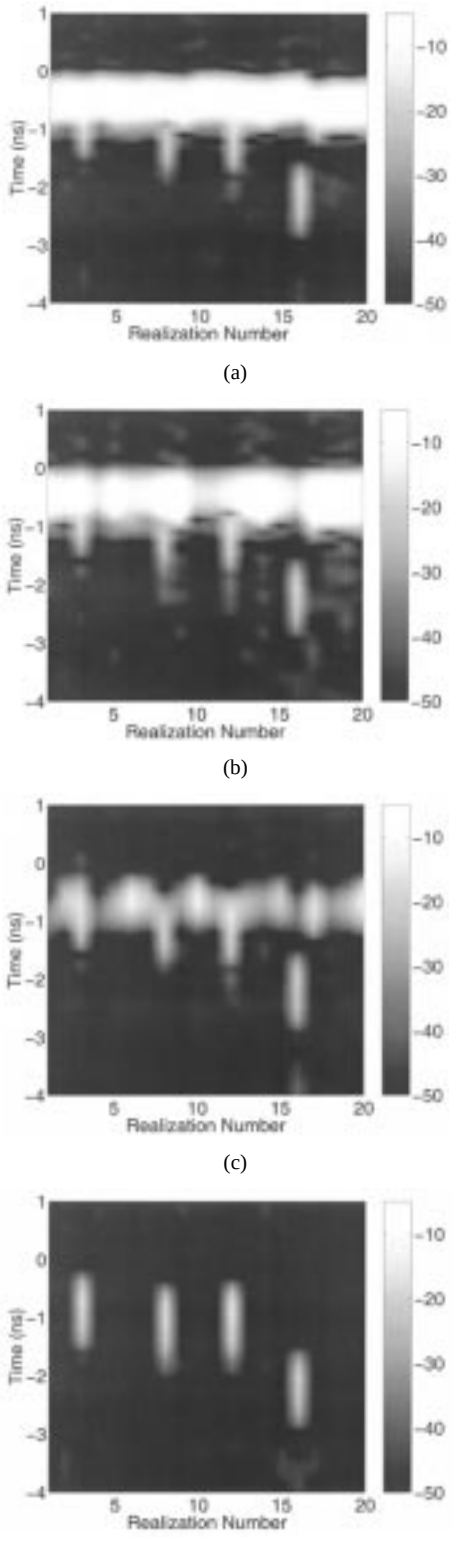


Fig. 8. Time domain response for multiple Monte Carlo realizations where some realizations contain subsurface targets. (a) Raw data, (b) clutter reduction using averaging, (c) clutter reduction using ETS, and (d) clutter reduction using the proposed technique.

A second set of the measured GPR data is shown in Fig. 13. This data set does not contain any mine. Some of the higher energy contributions close to the surface can easily be mistaken for shallow targets. The result after application of the iterative

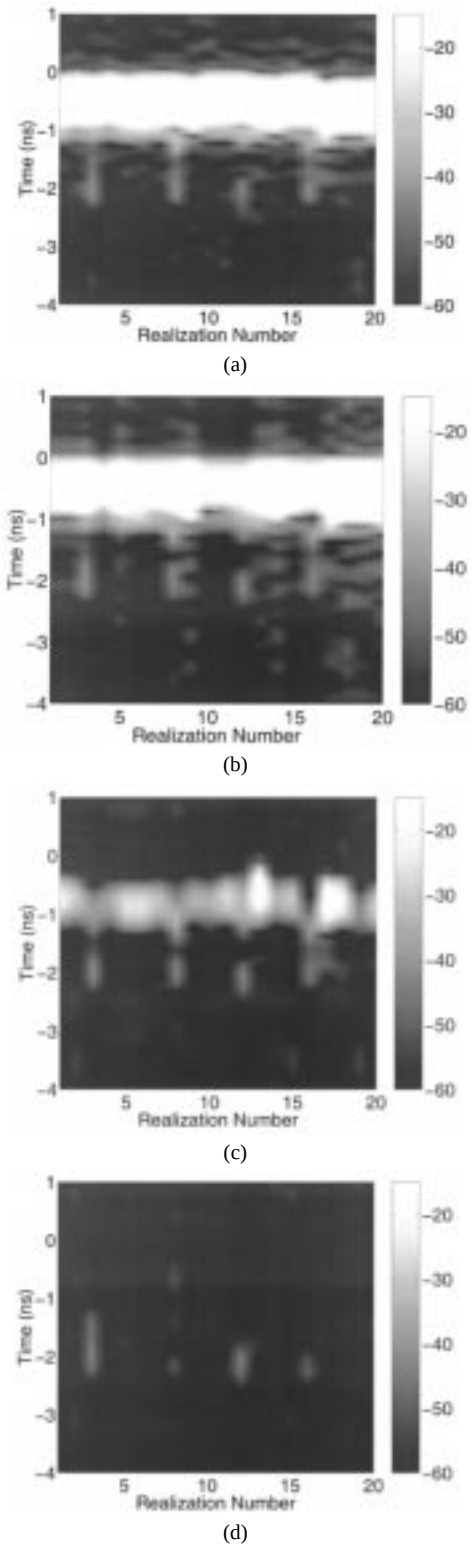


Fig. 9. Time domain response for multiple Monte Carlo realizations where some realizations contain subsurface inhomogeneities. (a) Raw data, (b) clutter reduction using averaging, (c) clutter reduction using ETS, and (d) clutter reduction using the proposed technique.

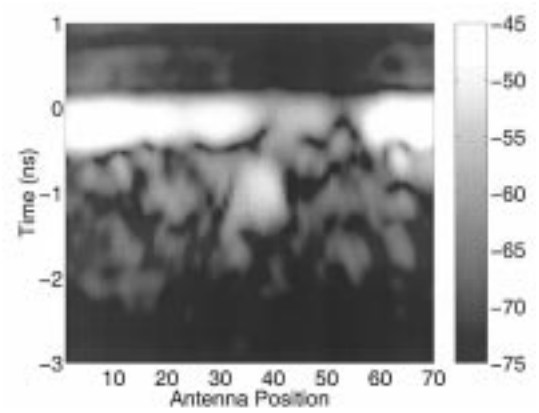


Fig. 10. Measured data versus position of the GPR antenna. This scan contains a target.

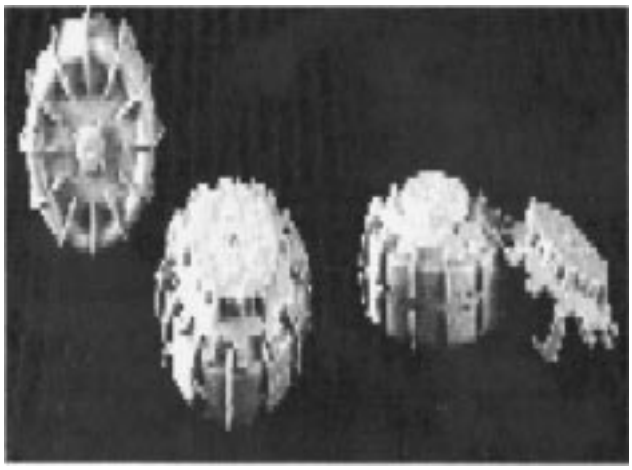


Fig. 11. TS-50 land mine.

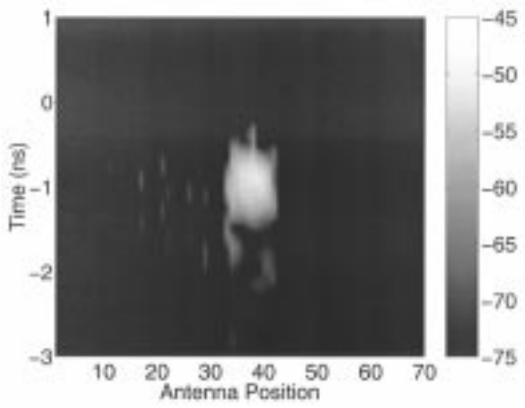


Fig. 12. Measured data versus position of the GPR antenna after clutter reduction using the proposed technique. The target is a TS-50 antipersonnel land mine.

approach is shown in Fig. 14. This result demonstrates that the new technique reduces the surface clutter and also has potential for lowering the false alarm rates. In [29], more examples with the measured data are included.

VI. Detection of Subsurface Targets

The clutter reduction techniques are evaluated next using some quantitative performance measures. To motivate the choice of the performance measure, consider the following.

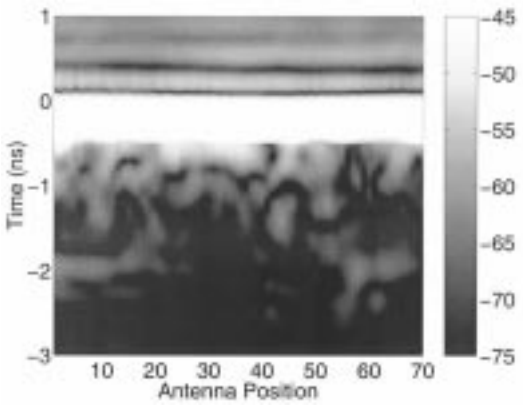


Fig. 13. Measured data versus position of the GPR antenna. This scan does not contain any target.

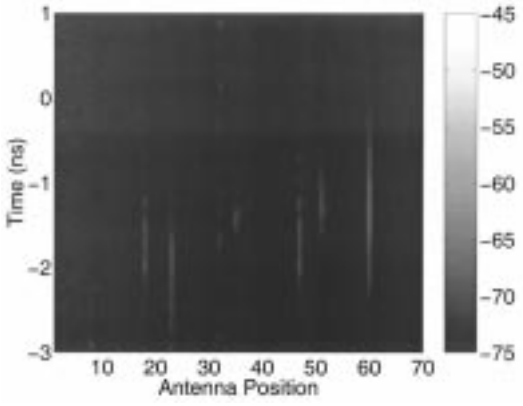


Fig. 14. Measured data versus position of the GPR antenna after clutter reduction using the proposed technique. No targets are present in these measurements.

The information of interest in GPR measurements is the target response. This response however, is obscured by the clutter. If the clutter can be reduced sufficiently, the mine response will dominate the returned signal, and it is then easier to detect the mine. The correlation between the residual data set after clutter reduction and a reference signature for the target of interest can be used as an indication of the clutter reduction performance where higher correlation values indicating better clutter reduction. This leads to the concept of a matched filter, which is formulated in such a way that the output is essentially the correlation between the residual data and a target reference signature. This correlation will be high if the clutter is sufficiently reduced. To completely quantify the performance, matched filter performance is assessed through receiver-operating-characteristic curves.

The procedure used for detection is formulated using a simple binary hypothesis test [30]. The hypothesis H_o occurs in the absence of the desired target when there is only clutter and noise while the hypothesis H_1 occurs when the desired target is present.

The matched filter used here is formulated as

$$\rho(\tau) = \frac{1}{2\pi\Delta}\left|\int_{-\infty}^{\infty} S(\omega)H^{\star}(\omega)e^{-j\omega\tau}\,d\omega\right| \tag{15}$$

where $H(\omega)$ is the reference signature, and $S(\omega)$ is the total GPR data in the frequency domain. The normalization factor is defined as $\Delta = \sqrt{E_S E_H}$, where E_S and E_H represent the energy in the respective signatures. Note however, that the dependence of the buried target return on depth is not taken into account in this definition. A more accurate formulation includes the factor $H(\omega, t_d)$, where t_d, related to the depth d of the buried target, is given by $t_d = (2d/v_g)$ and where v_g is the speed of propagation in the soil. However, the approximation used in (15) is justified because at the frequency band and soil losses considered, the soil is not dispersive, and the speed of wave propagation can be approximated as $v_g \approx (1/\sqrt{(\mu\epsilon)})$.

Since it is a controlled experiment, the target depth d is known,[5] and it can be used to calculate the time delay τ_d where $\tau_d = (2d/v_g)$. The output of the matched filter, defined in (15) at the delay $\tau = \tau_d$, is compared to the threshold as follows:

$$\rho(\tau = \tau_d) > \beta, H_1 \; < \beta, H_o \tag{16}$$

where β denotes the decision threshold. Note that the value of the threshold is different for the various preprocessing techniques.

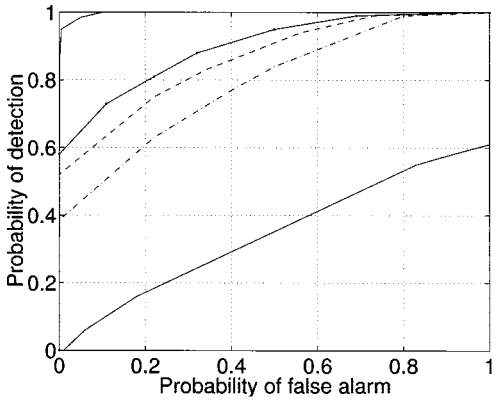
The matched filter is designed to correspond to the target of interest. The performance measure used here are the detector ROC curves. Monte Carlo simulations are performed with the same soil and surface roughness parameters as before. Each realization contains one of the following:

1) a rough ground surface without any buried objects;
2) a rough ground surface with a buried dielectric target at various depths;
3) a rough ground surface with a randomly shaped, dielectric inhomogeneity at various depths with varying shape and relative dielectric constant

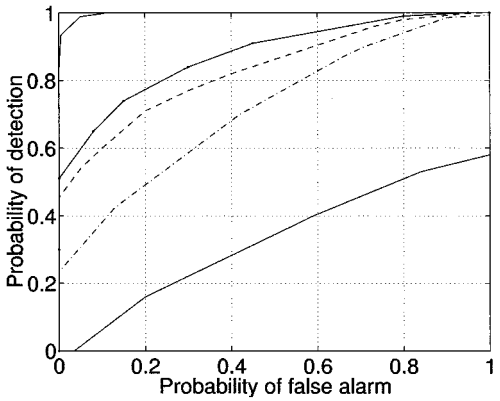
The depth of the subsurface target as well as that of the inhomogeneities are varied between 2.5 cm and 10 cm. The geometry of the inhomogeneity is given in Section V. As before, the rough surface profile, additive white noise and shape of the subsurface inhomogeneities are varied between realizations.

The clutter reduction techniques compared in these examples are: average subtraction, ETS, average subtraction followed by the whitening filter, ETS followed by the whitening filter, and the iterative reduction algorithm. The ROC curves for the various techniques are shown in Fig. 15(a)–(c) for SNR values of 40 dB, 30 db, and 20 dB, respectively. Performance of the techniques degrades with a decrease in SNR. The matched filter with average subtraction does not perform satisfactorily. This can be expected due to the high levels of clutter that typically occur in subsurface target measurements. For average subtraction followed by whitening, the performance of the detector improves to some extent. However, more improvement is obtained using ETS and even more if the whitening filter is combined with ETS. Note that in neither of the latter techniques is any knowledge of the buried target used. However, the performance of the whitening filter depends on the similarity between the measured clutter statistics and that used to design the whitening filter. Any discrepancies between these degrade the performance

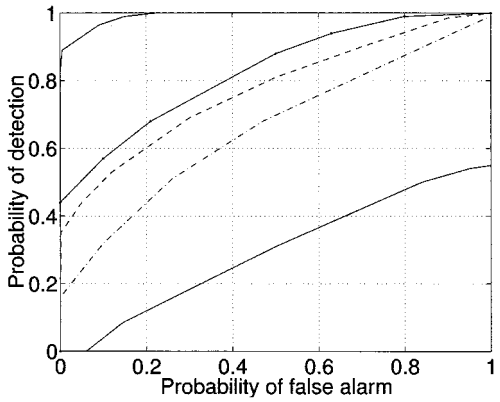
[5]In general, the target depth can be estimated using the proposed iterative clutter reduction technique.



(a)



(b)



(c)

Fig. 15. ROC curves for the matched filter at (a) SNR = 40 dB, (b) SNR = 30 dB, and (c) SNR = 20 dB. The curves for the various preprocessors are indicated by the following line types (starting from the bottom of the curve): the lower solid line indicates average subtraction, the dash-dotted line indicates average subtraction followed by the whitening filter, the dashed line indicates ETS, the solid-dotted line indicates ETS followed by whitening, and the top solid line indicates the proposed technique.

of the whitening filter. From these curves, it is also clear that the highest P_D for any P_F is obtained when the data is preprocessed using the iterative algorithm. This may be attributed to the fact that this technique removes clutter from the rough surface as well as due to ground surface-target interaction terms and also some of the false alarms due to subsurface inhomogeneities. The clutter-reduced data set therefore contains a very good signature of the subsurface target. For this reason, it is very easy to distinguish between the desired target and false alarms using a simple matched filter.

TABLE I
NUMBER OF FALSE ALARMS FOR 100% DETECTION

Clutter Reduction Technique	SNR=40dB	SNR=30dB	SNR=20dB
Early time subtraction (ETS)	92	100	100
ETS + Whitening filter	85	93	95
Iterative algorithm	11	12	18

The reduction in the number of false alarms due to efficient clutter reduction can also be seen from the data given in Table I, where the number of false alarms for 100% detection are shown. The maximum possible number of false alarms is 100. (40 realizations do not contain any buried objects, whereas 60 realizations contain subsurface inhomogeneities). The number of false alarms for the three best clutter reduction techniques are shown in the table. Note that when the clutter is reduced using the proposed algorithm, the number of false alarms is reduced approximately five to eight times at 100% detection. These values clearly illustrate the advantages of using an effective clutter reduction algorithm. Another point that should be mentioned here is that for the iterative algorithm, most of the false alarms are due to the subsurface inhomogeneities.

VII. SUMMARY AND CONCLUSIONS

A new signal processing technique to reduce clutter in GPR data is proposed here. It is demonstrated that the proposed algorithm reduces the clutter due to rough surface scattering, surface-target interaction terms as well as some subsurface inhomogeneities. This is true even for targets buried at very shallow depths under a rough, ground surface where the responses from the clutter and that from the target overlap in time. The iterative algorithm decomposes the data into its clutter and target contributions. The target is the dominant contributor in the clutter reduced data set and therefore, it can easily be distinguished from false alarms using a matched filter. The new clutter reduction technique also yields a set of parameters for the target, namely, an estimate of the depth of the subsurface target and an estimate of the propagation loss.

In the simulated results presented in this paper, the reference signature and the return from the subsurface target are both viewed from the same angle. However, a deviation between the reference signature and the target return may occur due to tilt in the target or variations in the view angles. The latter will typically occur in measured GPR data. From the measured results shown here, it is clear that the clutter reduction algorithm still performs well. Thus, the proposed technique has potential for real world applications.

ACKNOWLEDGMENT

The authors would like to thank Dr. Baertlein, Dr. C. Chen, Prof. J. T. Johnson, Prof. R. Lee, and Prof. L. Peters, Jr. for many

insightful discussions. Prof. Lee assisted in the implementation of the FDTD algorithm and Dr. Chen provided the measured GPR data.

REFERENCES

[1] K. Tsipis, "Report on the landmine brainstorming workshop," Tech. Rep., Sci. Technol. Int. Security Prog., Mass. Inst. Technol., Cambridge, 1996.

[2] J. Hurst, "Ten million tragedies, one step at a time," *Bull. At. Scientists*, pp. 362–367, July/Aug. 1993.

[3] L. Peters, J. J. Daniels, and J. D. Young, "Ground penetrating radar as a subsurface environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1820, Dec. 1994.

[4] S. Vitebskiy and L. Carin, "Resonances of perfectly conducting wires and bodies of revolution buried in a lossy dispersive half-space," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, Dec. 1996.

[5] T. Dogaru and L. Carin, "Time-domain sensing of targets buried under a rough air-ground interface," *IEEE Trans. Antennas Propagation*, vol. 46, pp. 360–372, Mar. 1998.

[6] G. Zhang, L. Tsang, and Y. Kuga, "Studies of the angular correlation function of scattering by random rough surfaces with and without a buried object," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 444–453, Mar. 1997.

[7] L. Tsang, G. Zhang, and K. Pak, "Numerical study of a buried object under a single random rough surface with angular correlation function," in *Proc. IGARSS'96*, vol. 4, Lincoln, NE, 1996, pp. 2143–2145.

[8] R. Caldecott, J. D. Young, J. P. Hall, and A. J. Terzuoli, "An underground obstacle detection and mapping system," Tech. Rep. EL-3984, ElectroSci. Lab., The Ohio State Univ. and the Elect. Power Res. Inst., Palo Alto, CA, May 1985.

[9] H. Brunzell, "Detection of shallowly buried objects using impulse radar," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 875–886, Mar. 1999.

[10] D. Carevic, M. Craig, and I. Chant, "Modeling GPR echos from land mines using linear combinations of exponentially damped sinusoids," *Proc. SPIE*, vol. 3079, pp. 568–580, 1997.

[11] C. C. Chen, F. Paynter, J. D. Young, and J. Peters, "Use of ground penetrating radar (GPR) for detection of unexploded ordnance," Tech. Rep. 727 388-2, ElectroSci. Lab., The Ohio State Univ., Columbus, OH, Feb. 1994.

[12] I. J. Gupta, A. van der Merwe, and C. C. Chen, "Extraction of complex resonances associated with buried targets," *Proc. SPIE*, vol. 3392, pp. 1022–1032, 1998.

[13] M. Rahman and K.-B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1440–1545, Oct. 1987.

[14] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[15] R. G. Kouyoumjian and P. H. Pathak, "A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface," *Proc. IEEE*, pp. 1448–1496, Nov. 1971.

[16] L. C. Potter and R. L. Moses, "Attributed scattering centers for SAR ATR," *IEEE Trans. Image Processing*, vol. 6, pp. 79–91, Jan. 1997.

[17] M. P. Hurst and R. Mitra, "Scattering center analysis via Prony's method," *IEEE Trans. Antennas Propagat.*, vol. AP-35, pp. 986–988, Aug. 1987.

[18] C. A. Balanis, *Advanced Engineering of Electromagnetics*: Wiley, 1989.

[19] P. E. Gill, W. Murray, and M. H. Wright, *Numerical Linear Algebra and Optimization*. Reading, MA: Addison-Wesley, 1991.

[20] A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Norwell, MA: Artech, 1995.

[21] P. G. Petropoulos, "Stability and phase error analysis of FDTD in dispersive dielectrics," *IEEE Trans. Antennas Propagat.*, vol. 42, pp. 62–69, Jan. 1994.

[22] B. Engquist and A. Majda, "Absorbing boundary conditions for the numerical simulation of waves," *Math. Comput.*, vol. 31, pp. 629–651, July 1977.

[23] G. Mur, "Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations," *IEEE Trans. Electromagn. Compat.*, pp. 377–382, Nov. 1981.

[24] R. Higdon, "Numerical absorbing conditions for the wave equation," *Math. Comput.*, vol. 49, pp. 65–90, July 1987.

[25] J. P. Berenger, "Perfectly matched layer for the FDTD solution of wave-structure interaction problems," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 110–117, Jan. 1996.

[26] E. I. Thorsos, "The validity of the Kirchoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Amer.*, vol. 83, pp. 78–92, 1988.

[27] National Ground Intelligence Center, "Mine Facts version 1.2: A CD-ROM of interactive database programs," . Charlottesville, VA: U.S. Dept. Defense, 1997.

[28] Y. Miyazaki, Y. Jyonori, and J. Sonoda, "Statistical characteristics of reflections and scatterings of electromagnetic radar pulses by buried objects in random underground media," in *6th Int. Conf. Ground Penetrating Radar*, vol. 1, Sendai, Japan, Oct. 1996, pp. 151–154.

[29] A. van der Merwe, "A Novel Signal Processing Technique for Clutter Reduction in GPR Measurements of Small, Shallow Land Mines," Ph.D. dissertation, The Ohio State Univ., Columbus, OH, 1999.

[30] H. L. van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1969.



Andria van der Merwe received the B.Ing. and M.Ing. degrees, both in electrical engineering, from the University of Pretoria, Pretoria, South Africa, in 1992 and 1995, respectively. She received the Ph.D. degree in electrical engineering from the Ohio State University, Columbus, OH, in 1999.

Currently, she is with Andrew Corporation, Dallas, TX. Her research interests include digital signal processing, parametric modeling, and stochastic estimation.

Dr. van der Merwe received the Ohio State University Edward F. Hayes Graduate Research Award in 1999 and is a member of Phi Kappa Phi and Sigma Xi.



Inder J. Gupta received the B.Sc. degree in electronics engineering from Panjab University, Chandigarh, India, in 1975, the M.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1977, and the Ph.D. degree in electrical engineering from The Ohio State University, Columbus, OH, in 1982.

Since 1979, he has been with the Electro Science Laboratory, Ohio State University, where, he is currently a Senior Research Scientist. He has been a Research Engineer with IIT Kanpur, Kanpur, India, and a Graduate Research Associate with the University of Tennessee Space Institute, Tullahoma, TN. He has done extensive research on adaptive antenna systems and compact range measurement systems and has written several research articles in these areas. His current research interests include adaptive antenna arrays, antijam techniques, compact range technology, radar imaging, and target recognition.

Dr. Gupta was the President of Antenna Measurement Techniques Association in 1997. He received the IEEE Antennas and Propagation Society's *H. A. Wheeler Applications Prize Paper Award* in 1991. He is also the recipient of The Ohio State University College of Engineering *Lumley Research Award* for 1991 and 1998.

Radiometric Model for the Infrared Signature of a Land Mine Buried Under a Rough Surface

İbrahim Kürşat Şendur and Brian A. Baertlein

The authors are with the ElectroScience Laboratory, Department of Electrical Engineering, The Ohio State University, Columbus, OH 43212-1191 USA. E-mail: baertlein.1@osu.edu .

Abstract

A radiometric model is presented for the infrared (IR) signature of a buried land mine. The soil surface temperature is predicted with a finite element method (FEM) based thermal model previously described by the authors. Enhancements to that thermal model are presented including a rough surface description of the soil-air boundary. The radiometric model addresses both the spatial and spectral characteristics of the soil. The atmospheric propagation modeling code MODTRAN is used to predict the intensity and spectral dependence of the incident radiometric components. Simulated IR imagery is presented for land mines buried under both smooth and rough surfaces. The effects of random surface emissivity variations are also demonstrated. Validation of the combined thermal-radiometric model is presented using experimental data. It is found that the signature shape and temperature contrast are fairly well predicted.

Keywords

Thermal infrared imagery, land mines, radiometric modeling, surface roughness, heat transfer, thermal modeling, numerical simulation, finite element method.

I. Introduction

In a recent work Sendur et al. [1] described a FEM based three-dimensional thermal model to predict the soil surface temperature over a buried land mine. That model is capable of simulating realistically shaped land mines and inhomogeneous soil. They presented the temporal evolution of the surface temperature for various mine types.

In addition to the mine's thermal signature, an IR image of the surface will also contain undesirable reflected light from natural sources. The latter comprises a form of clutter, which has a detrimental effect on mine detection performance. Natural sources are also wavelength dependent, which requires attention to the spectral content of those reflected components. Therefore, to better understand the IR signatures of buried land mines and clutter, a spectral radiometric model is required. There has been little work in radiometric modeling of mine signatures. Flynn et al. [2] developed a surface mine simulation tool for the UV-VNIR regime, but no prior work on radiometric modeling for buried mines has appeared.

In this work we describe a radiometric model capable of predicting the IR signatures of buried land mines. The model addresses both the spatial and spectral characteristics of the environment. The work is organized as follows: In Sect. II we summarize some

enhancements to our FEM thermal model which permit more realistic modeling of the natural environment. The radiometric model is presented in Sect. III. The radiometric components of IR imagery are briefly discussed in Sect. III-A. Sunlight and skylight models are given in Sect. III-B, including their spectral dependence. The simulation of IR imagery is discussed in Sect. III-C. Numerical results for a cylindrical anti-tank mine are presented in Sect. IV. Examples of the temporal evolution of IR imagery for smooth and rough soil surfaces are given therein. Concluding remarks appear in Sect. V.

II. Thermal model improvements

A buried mine affects the flow of heat flux into and out of the soil, since its thermal characteristics are different than those of soil. This alteration of the heat flow results in a disturbance of the surface temperature, which is the primary source of the IR signatures of buried land mines[1]. The authors have previously described a technique for calculating the soil surface temperature over a mine, which may be consulted for details [1]. As improvements to that thermal model, in this study we suggest two major modifications: (1) a rough soil-air boundary and (2) a more sophisticated model of the solar insolation.

Prior experimental work suggests that a rough soil-air interface has a strong influence on the clutter signature of soil. This roughness complicates the thermal modeling as a result of spatial variations in the absorbed down-welling solar radiance (via the variable local surface normal). Surface roughness also affects the calculation of the reflected sunlight and skylight, which arises in the radiometric model discussed next. A rough surface is easily incorporated in our thermal model by varying the tilt in the triangular surface facets. A sample discretization of a volume is given in Fig. 1.

The second modification to our previous model is the incorporation of a more realistic solar insolation function. In previous studies we used a relatively simple analytic function that addresses the time of day, cloud extinction, atmospheric absorption, and soil albedo. That model employs atmospheric transmittance data given by Allen [3], approximated in a functional form by Watson [4]. In this study the more sophisticated capability provided

[1] Other phenomena that may contribute to buried mine signatures include the difference in soil moisture content (and the concommitant change in soil emissivity and thermal properties) due to the moisture barrier formed by the mine. This barrier also affects the vitality of any overlying vegetation.

by MODTRAN [5] has been incorporated. MODTRAN provides several benefits including more accurate calculation of solar position for a given latitude and longitude and atmospheric modeling that includes the effects of absorption bands and weather conditions.

III. Radiometric model

As noted above, the IR signature of a buried mine includes various soil reflected components in addition to the thermal components, and the inclusion of those components is the principal contribution of this work. In this section we review the additional components and our models for them.

A. Components of the IR image

The radiative components seen by the IR camera include (1) thermal emission from the soil surface (a portion of which comprises the desired signal), (2) soil-reflected sunlight, (3) soil-reflected skylight (thermal emission from the atmosphere and sunlight scattered by particles and air molecules), and (4) a negligible amount of thermal emission from the atmosphere directly to the camera. These components are illustrated in Fig. 2, and we write the total radiance received by the sensor as

$$L_{rec}(\lambda, \mathbf{r}) = L_{surf}(\lambda, \mathbf{r}) + \mathcal{R}(\lambda, \mathbf{r}) \left[L_{sun}(\lambda) + L_{sky}(\lambda) \right]. \tag{1}$$

In this work we model the soil surface as a diffuse reflector, which simplifies the analysis.

B. Sunlight and skylight models

Sunlight and skylight contribute to both the thermal and radiometric models. The short wavelength solar heating of the soil surface comprises the dominant source for the thermal model, and long wavelength reflected radiance is received by the IR sensor.

The radiance L_{SUN} can be modeled with reasonable accuracy by a blackbody radiator. Neckel and Labs [6] show that a blackbody source at $\approx$5800 [K] represents a good numerical fit to exoatmospheric measured radiance data in the visible and near IR bands. Thekaekara and Drummond [7] summarized data regarding components of the total radiation and their spectral distribution, and they proposed standard values for them. Measured solar radiation at different sites and models for meteorological and climatic factors are presented

in a summary by Dogniaux [8] prepared for the European Community Programme on Solar Energy. The measurement techniques and instruments are described in Coulson [9]. Figure 3 (a) illustrates the predicted solar radiance at the outer surface of the atmosphere using a blackbody emitter at 5800 [K].

The atmosphere attenuates both the down-welling solar radiance and skylight and the reflected and emitted radiance. In many demining applications the distance between the IR camera and the soil surface is small and, hence, the attenuation of the down-welling solar radiance is of greatest concern. For airborne mine detection, attenuation of the emitted and reflected signals must also be considered.

Sophisticated modeling tools have been developed to deal with the complexities of the solar radiance and atmospheric transmission, absorption, and scattering phenomena. These models predict transmittance and radiance for sensor systems under varying atmospheric conditions. Among these computer codes, LOWTRAN (Low spectral resolution transmission), MODTRAN (Moderate spectral resolution transmission), and FASCODE (Fast atmospheric signature code) are commonly used. LOWTRAN is a one-parameter, band-model code that predicts atmospheric absorption and scattering for systems with low spectral resolution. It is computationally efficient, and it has an atmospheric database that can model six reference atmospheres with various constituents. MODTRAN is a two-parameter, band-model code with a higher resolution than LOWTRAN (2 cm^{-1} for MODTRAN versus 20 cm^{-1} for LOWTRAN). FASCODE is a high-resolution code, which is usually reserved for studies involving very narrow bandwidth.

To model the atmosphere we have selected the 1976 US standard atmosphere. Temperature, pressure and water-vapor altitude profiles, and altitude profiles of relevant gases (ozone, methane, nitrous oxide, carbon monoxide, and others) were set to values of the 1976 US standard atmosphere, but we set the CO_2 mixing ratio to 365 ppmv, which is consistent with more recent data. The weather is assumed to be clear and sunny. To model boundary-layer aerosols, rural extinction was selected with a visibility to 23 km. (Visibility is related to aerosol extinction at 550 μm.) Figure 3 presents the solar spectral radiance at Columbus, OH on April 3. Figures 3 (b)-(f) show the solar radiance predictions with atmospheric effects at 10AM, 12PM, 2PM, 4PM, and 6PM, respectively.

The quantity L_{SKY} depends on the composition of the local atmosphere and is more difficult to measure. It comprises wide-angle Rayleigh scattering by molecular constituents (very weak at IR wavelengths), small-angle Mie scattering by aerosols, and thermal radiation from the warm atmosphere. Although there has been extensive work in developing models for L_{SKY}, local changes in atmospheric particulates and water vapor content can greatly affect model accuracy. In addition, a number of molecular species (H_2O, CO_2, CO, N_2O, O_3, and CH_4) have absorption bands in the infrared, which affect the incident solar radiation [10, §3.5.2]. The thermal contribution is most important at long wavelengths, and in this work we approximated L_{SKY} by a blackbody radiator at the local air temperature.

C. Surface emission

Thermal emission from the soil surface depends on the surface temperature and emissivity. The surface temperature is obtained using the aforementioned thermal model. The spectral radiance of the soil (a graybody emitter) is given by the product of the surface emissivity $\mathcal{E}$ and the blackbody spectral radiance $L_{BB}(\lambda, T)$ described by Planck's radiation law

$$L_{BB}(\lambda, T) = \frac{2c^2h}{\lambda^5} \frac{1}{e^{h\nu/kT} - 1} \qquad [\text{W m}^{-2}\text{sr}^{-1}\mu\text{m}^{-1}], \tag{2}$$

where $h = 6.63 \times 10^{-34}$ [J s] is Planck's constant, ν [Hz] is the frequency of the optical radiation, λ [m] is the wavelength of that radiation, $k = 1.38 \times 10^{-23}$ [J K^{-1}] is Boltzmann's constant, $c = 3 \times 10^8$ [m s^{-1}] is the speed of light and T [K] is the temperature.

The radiation diffusely emitted from a point $\mathbf{r}$ on the soil surface at temperature $T(\mathbf{r})$ can be written as the product

$$L_{surf}(\lambda, \mathbf{r}) = \mathcal{E}(\lambda, \mathbf{r}) L_{BB}(\lambda, T(\mathbf{r})), \tag{3}$$

where $\mathcal{E}(\lambda, \mathbf{r})$ is the spectral directional emissivity. In Eq. (3) polarization effects, which are often directional and may be significant, have been ignored.

From the above expressions it is clear that emissivity has a direct effect on the thermal signature. In addition, grass and other forms of ground cover have a strong effect on the heat flow process and on the reflection and radiation of thermal energy. The emissivity of natural materials has been studied by Salisbury and D'Aria [11, 12]. In those works the

emissivity of soils, rocks and vegetation was described for both the 3-5 μm and 8-14 μm bands. The spectral response of these materials is complicated, but in general over the 3-5 μm band rocks and soils exhibit reflectance values of 5% to 30%, while vegetation has reflectance values of 2% to 15%. For the 8-14 μm band they found rock reflectance values of 1% to 10%. Vegetation reflectance values were found to have approximately the same range. The emissivity of these materials can be obtained from the reflectivity values by using Kirchoff's law for an opaque body, $\mathcal{R} = 1 - \mathcal{E}$.

D. IR image formation

Equation (1) permits us to calculate the power incident on an IR detector surface D from the radiating soil surface S. The power incident on the detector can be found by

$$\Phi = \int_{\lambda_1}^{\lambda_2} d\lambda \int_D dD \int_S dS L_{rec}(\lambda, \mathbf{r}) \frac{\cos\theta_1 \cos\theta_2}{R^2}, \tag{4}$$

where R is the distance between a point $\mathbf{r}$ on surface S and a point on the detector D; $\cos\theta_1$ and $\cos\theta_2$ are projections of the normal vectors for surfaces D and S, respectively, in the direction of the radiation; and the spectral band of interest is λ_1 to λ_2. In Eq. (4) the detector area is typically small, and the flux incident on each detector pixel is essentially constant. As a result, integration over the detector area is reduced to the product of the detector area and the flux falling on it.

The integration over the spectral band of interest $[\lambda_1, \lambda_2]$ is one-dimensional and presents no challenges for a narrow-band sensor. This integral is evaluated using a Gaussian quadrature rule. For a wider spectral domain with molecular absorption bands, evaluation of this integral is more challenging due to possible rapid variations in the integrand. To avoid possible numerical inaccuracy for larger spectral domains, we used adaptive Gaussian quadrature.

In terms of the computational time and memory requirements, the most challenging aspect of Eq. (4) is evaluation of the integral over the surface S, which is the area of the rough surface seen by an individual image pixel. The projected surface area can be large for some combinations of camera height, location, and observation angle. In addition, the surface may be self shadowing. Two types of shadowing are taken into account in the radiometric model, and these are illustrated in Fig. 4. Some parts of the soil surface can

be invisible due to blocking as illustrated in Fig. 4 (a). To permit a general study of rough surfaces our radiometric model has been designed to deal with this form of shadowing, but practical sensor systems are often constrained to avoid these conditions, since they can lead to missed detections. Solar heating of the soil surface is also affected by shadowing as shown in Fig. 4 (b), and our model is also capable of handling this effect.

Integration over the projected surface areas is done using numerical quadrature techniques and is illustrated in Fig. 5 (a). As noted above, the soil surface is described by triangular facets with varying normal vectors. Every image pixel is divided into two triangles, and the integration domains comprise the projection of these triangles (or abscissas for numerical integration) onto the rough surfaces. The details of the selection of the integration abscissas and the integration rule are discussed below. The major difficulty at this point is the projection of the abscissas onto the surface facets. The challenge in this process is twofold: (1) deciding which surface facets contain the projections of the abscissas (shown Fig. 5 (a)), and (2) determining which of these facets will contribute to the integral at the specified integration point (shown in Fig. 5 (b)). This is a computationally challenging process due to the large number of facets and abscissas. For example, the relatively modest numerical simulations presented below in Sect. IV involve $48 \times 48 \times 2 = 4608$ surface triangles and $120 \times 160 \times 6 = 57600$ abscissa values. Visible surface determination is a computationally expensive process for this large number of points. Techniques offered by computer graphics are essential for computational efficiency. Specifically, a z-buffer algorithm [13] is utilized to determine the facet associated with an integration point. Once the projection of the integration point over a facet is determined, evaluation of the integrand at that integration point is required. The physical temperature, emissivity and surface normal direction are known at the integration points and, therefore, the integrand in Eq. (4) can be obtained with the aid of Eq. (1).

The accuracy of the integration over these triangular domains affects the fidelity of the simulation, and a number of tests were done to verify this accuracy. Different adaptive Gaussian quadrature rules were used over a simplex coordinate system. For our application the number of surface facets is smaller than the number of image pixels, which implies that a facet is in the field of view of several pixels. Furthermore, the functional variations

over these facets are smooth. Ultimately, we concluded that a non-adaptive three point quadrature rule produced acceptable results.

IV. Results

The thermal and radiometric models described above were used to simulate the temporal and spatial signatures of mines buried under smooth and rough surfaces. In all simulations a steady wind speed of $W(t) = 2$ [m s^{-1}] and an average air temperature of $\overline{T}_{air} = 289$ [K] were used. The thermal diffusivity (κ) and conductivity ($\mathcal{K}$) of soil were taken to be 5.0×10^{-7} [m^2 s^{-1}] and 2.6 [W m^{-1} K^{-1}], respectively. The mine was modeled as a homogeneous cylinder of TNT, for which we used $\kappa = 9.25 \times 10^{-8}$ [m^2 s^{-1}] and $\mathcal{K}$ =0.234 [W m^{-1} K^{-1}]. Mine dimensions vary significantly. As a simple but representative target, we selected a simulant anti-tank mine [14] developed by the US Army. The simulant mine has a diameter of 25 cm, a height of 8.33 cm, and was buried 6.64 cm

The radiometric model requires as input the triangular rough surface representation and the output of the thermal model. Using these data and the geometry given in Fig. 6, we constructed a virtual IR camera, the specification of which are given in Table I. The IR camera is aimed at point O (see Fig. 6), which is also the global coordinate origin for both the thermal and radiometric models. To cover the surface above the mine with sufficient resolution, an appropriate camera location and height must be identified. For this work we selected a camera height of 5.7 meters, and a horizontal standoff (camera center to point O) of 1 meter. The resulting ground-projected field of view encompasses most of the computational volume.

TABLE I

Specifications of the virtual camera used in the radiometric model.

Spectral range	MWIR (4.4-5 μm)
Array size	160 (h) by 120 (v) pixels
FOV	9.1° (h) by 6.8° (v)
IFOV	1 mrad

A. Mines under smooth surfaces

Figure 7 shows the simulation results for the surrogate mine buried under a smooth surface. The results are presented as a sequence of images evaluated at three hour time increments starting from sunrise. These results show phenomena identified previously, namely, the so-called thermal "cross-over" times at 11 AM and 9 PM. The IR signature of the land mine is clearly seen. Clutter-like variations, which are ubiquitous in experimental studies do not appear in the simulations.

B. Random rough surface modeling using experimental data

Rough surfaces were constructed by specifying random elevations for points in a rectilinear grid defined by the top surface of our FEM computational volume. The surface was then represented by triangles fitted to the grid points as shown in Fig. 1. The same representation was used by both the thermal and radiometric models.

To generate a random rough surface we require a description of the surface height statistics. In this work we used experimental data acquired by Salvati et al. [15] in support of a satellite remote sensing study. Those authors used a microtopographic laser scanner to measure surface heights on a grid of spacing 0.15 cm. The region sampled in their work was formerly cultivated farmland, which contained pronounced furrows. To simulate uncultivated soil we used data in the "down-furrow" direction and extended it isotropically. We approximated the resulting surface height spectrum using the following analytical expression

$$W(k_x, k_y) = \exp\left(- L_1(k_x^2 + k_y^2)^{1/2} - L_2(k_x^2 + k_y^2)^{1/4} \right) \tag{5}$$

where k_x and k_y are spectral frequencies in the $\hat{x}$ and $\hat{y}$ directions, respectively, and we estimated $L_1 = 0.025$ and $L_2 = 0.25$ from the data. The experimental and theoretical surface height spectra are plotted in Fig. 8. In addition, we compared the histogram of the data with a Gaussian distribution using the Kolmogorov-Simirnov (KS) test. The KS levels of significance deviate from unity by less than 4×10^{-3}, which strongly imply a Gaussian distribution.

The problem of generating a rough surface with a zero-mean, Gaussian distributed, height profile $z = f(x, y)$ can be addressed by using the surface height spectrum $W(k_x, k_y)$

given by Eq. (5). Each spectral component is multiplied by a complex random number of unit magnitude and uniformly distributed phase $\psi_{m,n}$. We form the quantity

$$P(m,n) = e^{+j\psi_{m,n}} \frac{2\pi N^2}{L} \sqrt{W\left(\frac{2m\pi}{LN}, \frac{2n\pi}{LN}\right)}, \tag{6}$$

where N^2 is the number of points in the discretization of the surface profile and L is the edge-length of the surface. An example random surface generated in this manner is shown in Fig. 9.

Figure 10 shows the temporal evolution of a mine signature for the rough surface plotted in Fig. 9. Again, the results are presented as a sequence of images evaluated at three hour time increments starting from sunrise. In these figures it is harder to identify the mine signature, because the circular shape is lost. In addition, the imagery shows significant clutter, which could easily be mis-detected.

C. Random surface emissivity modeling

As noted in Sect. III-C, the radiation emitted from a surface is a function of surface emissivity. In this section we investigate the effects of a random surface emissivity profile. Using studies by Salisbury and D'Aria [11,12] as a guide, we assumed soil emissivity values ranging from 0.8 to 0.9. To define the spatial distribution of the emissivity, we assumed a Gaussian distribution and a Gaussian spectrum given by

$$W(k_x, k_y) = \frac{L_x L_y e_s^2}{4\pi} \exp\left(- \frac{L_x^2 k_x^2}{4} - \frac{L_y^2 k_y^2}{4} \right), \tag{7}$$

where L_x and L_y are correlation lengths in the x and y directions respectively, and e_s is the surface rms emissivity. In what follows we assume an isotropic surface with $L_x = L_y = L$. Equation (7) is used with Eq. (6) to obtain a realization of the surface emissivity. The resulting values are scaled to the range $[0.8, 0.9]$ to obtain the desired distribution. In Fig. 11 we have plotted an emissivity realization with correlation length $L = 5$ cm. Figure 12 shows the temporal evolution of the IR mine signatures for this emissivity. The emissivity variations break up the mine signature and the imagery shows significant clutter.

D. Model Validation

Efforts have been made to validate the combined thermal-radiometric model using experimental data.[2] We used 8-12 μm calibrated images of AT mines acquired by TRW at times near 10:00 AM and 12:00 PM. The images were acquired at the US Army Fort A.P. Hill, Site 71A mine lanes #13 (dirt) and #14 (gravel). Collateral information was collected including soil temperature at multiple depths (0.5", 2", 4", 8"), wind speed and direction, soil moisture content (one depth), air temperature, barometric pressure, and down-welling and up-welling radiance in the 0.3-3 μm and 3-50 μm spectral bands.

To predict the measured data, it was necessary to make assumptions about several unknown environmental parameters. Specifically, we assumed soil thermal conductivity=1.8 [W/mK] and diffusivity=10^{-6} [m^2/s] for the given moisture content in a clay-loam soil. The mine (an "EM-12" mine surrogate) is known to have a 30 cm diameter and 5.2 cm height. The mine's internal contents are not known, and because it is part of an active test site it was not possible to excavate the mine and examine it (to avoid disturbing the target for subsequent tests). Collateral information suggested a filling of styrofoam pellets and carnuba wax with an overlying void, which we modeled as a good thermal insulator. The thermal model's FEM mesh was constructed of pentahedral elements with cell dimensions of 1.27 cm (height) by 2.56 cm (base). Other environmental parameters were derived from the available data and models. A surface emissivity of 0.98 was assumed. MODTRAN was used to estimate the incident radiance as a function of time, and it was found to replicate measured values to an accuracy of 10%.

The measured data appear in Fig. 13 (a). To reduce small-scale image clutter related to surface roughness and emissivity variations we applied a low-pass filter to the original data, which produced the imagery shown in Fig. 13 (b). The model results appear in Fig. 13 (c), the shape of which is seen to compare well with Fig. 13 (b). A one-dimensional cut through the data is shown in Fig. 13 (d), which helps to quantify the agreement. We see that the signature shape and temperature contrast are fairly well predicted, but the model temperature is approximately 2K below the measured data. There are several possible sources for this error including incorrect parameter estimates and our use of periodic

[2]Validation of the thermal model alone using independent models has been discussed in a separate work [16].

thermal boundary conditions in time. Warmer conditions the previous night can produce an offset that is not easily addressed without prior temperature data. The agreement in Fig. 13 is encouraging, but additional work in validation is clearly called for.

V. Conclusions

A radiometric model has been presented for the spatial and spectral IR signatures of buried land mines. Atmospheric effects and the spectral content of natural sources are taken into account via the MODTRAN model. Both the thermal and radiometric models used in this study incorporate a rough surface for the soil-air interface, which has implications for both thermal heating and for reflected radiometric components. The temperature distribution computed using the thermal model is combined with surface-reflected radiometric components to produce the image seen at an IR sensor. Simulations of mines under rough surfaces were performed, and it was found that surface roughness can add appreciable clutter. In addition, random variations in the surface emissivity profile were found to be another potential source of clutter. An effort was made to validate the model using experimental data. It was found that the signature shape and temperature contrast are fairly well predicted.

Acknowledgments

The authors would like to thank TRW for providing the experimental results, with special thanks to Ms. Sofia Monawer and Dr. Steve Campana for their help in interpreting the data. The assistance of Ms. Judy Salvati and Prof. Joel Johnson in dealing with the soil surface statistics is also gratefully acknowledged.

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

References

[1] I. K. Sendur, B. A. Baertlein, and K. Vafai, "Three-dimensional thermal modeling of land mine signatures over a diurnal cycle," *submitted for publication to the IEEE Trans. Geo. Rem. Sens.*

[2] B. T. Blume A. C. Dubey D. S. Flynn, D. A. Vechniski and N. H Witherspoon, "Minefield image synthesis

tool," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, Eds., 1995, vol. 2496, pp. 208–221.

[3] C. W. Allen, *Astrophysical Quantities*, First edition, The Athlone Press., London, England, 1963.

[4] K. Watson, "Geologic application of thermal infrared images," *Proc. IEEE*, vol. 63, no. 1, pp. 128–137, Jan. 1975.

[5] A. Berk, L. S. Bernstein, and D. C. Robertson, "MODTRAN: a moderate resolution model for LOWTRAN7," *Tech. Report GL-TR-89-0122, Air Force Geophysical Laboratory, Hanscorn AFB, MA, October 21 1989.*

[6] H. Neckel and D. Labs, "The solar radiation between 3300 and 12500 å," *Solar Phys.*, vol. 90, pp. 205–258, 1984.

[7] M. P. Thekaekara and A. J. Drummond, "Standard values for the solar constant and its spectral components," *Nature Physical Science*, vol. 229, no. 4, pp. 6–9, 1971.

[8] R. Dogniaux (ed.), *Prediction of Solar Radiation in Areas with a Specific Microclimate*, First edition, Kluwer Academic Publishers, Inc., The Netherlands, 1994.

[9] K. L. Coulson, *Solar and Terrestrial Radiation*, First edition, Academic Press, Inc., New York, NY, 1975.

[10] L. W. Wolfe and G. J. Zissis, *The Infrared Handbook*, revised edition, Environmental Research Institute of Michigan (ERIM), Ann Arbor, MI, 1989.

[11] J. W. Salisbury and D. M. D'Aria, "Emissivity of terrestrial materials in the 8-14 μm atmospheric window," *Remote Sensing Environment*, vol. 42, pp. 83–106, 1992.

[12] J. W. Salisbury and D. M. D'Aria, "Emissivity of terrestrial materials in the 3-5 μm atmospheric window," *Remote Sensing Environment*, vol. 47, pp. 345–361, 1994.

[13] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics*, Addison-Wesley Publishing Company, Inc., 1990.

[14] Anon., "Simulant mines (sims)," Tech. report, US Army Countermine and Demolition (Countermine Division), Fort Belvoir, VA, October 21 1998.

[15] J. L. Salvati, J. T. Johnson, C. Huang, and A. Klik, "Soil microtopography from the southern great plains hydrology experiment 1999," *Proc. Int. Geo. Rem. Sens. Symp.*, 2000.

[16] I. K. Sendur and B. A. Baertlein, "Integral equation solution for thermal signatures of buried land mines," *submitted for publication to the IEEE Trans. Geo. Rem. Sens.*

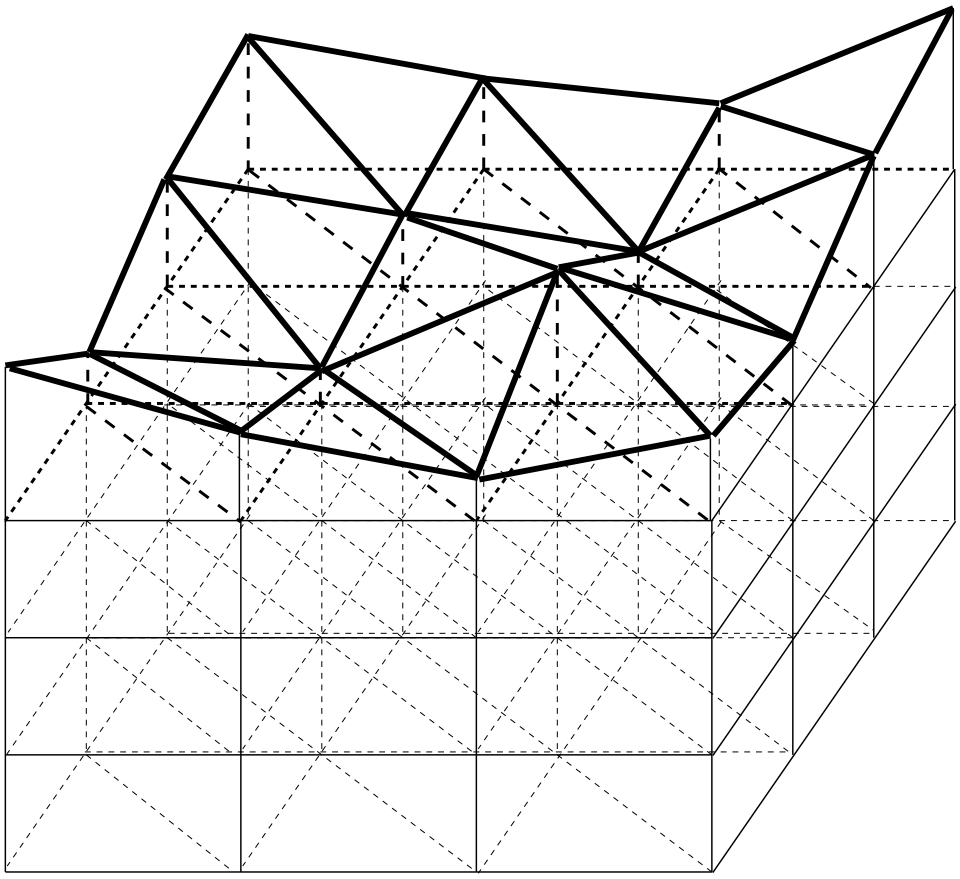


Fig. 1. Sample spatial discretization of the computational domain. The volume is subdivided by pentahedral elements resulting in a triangular mesh for the soil surface, which is emphasized with thicker lines.

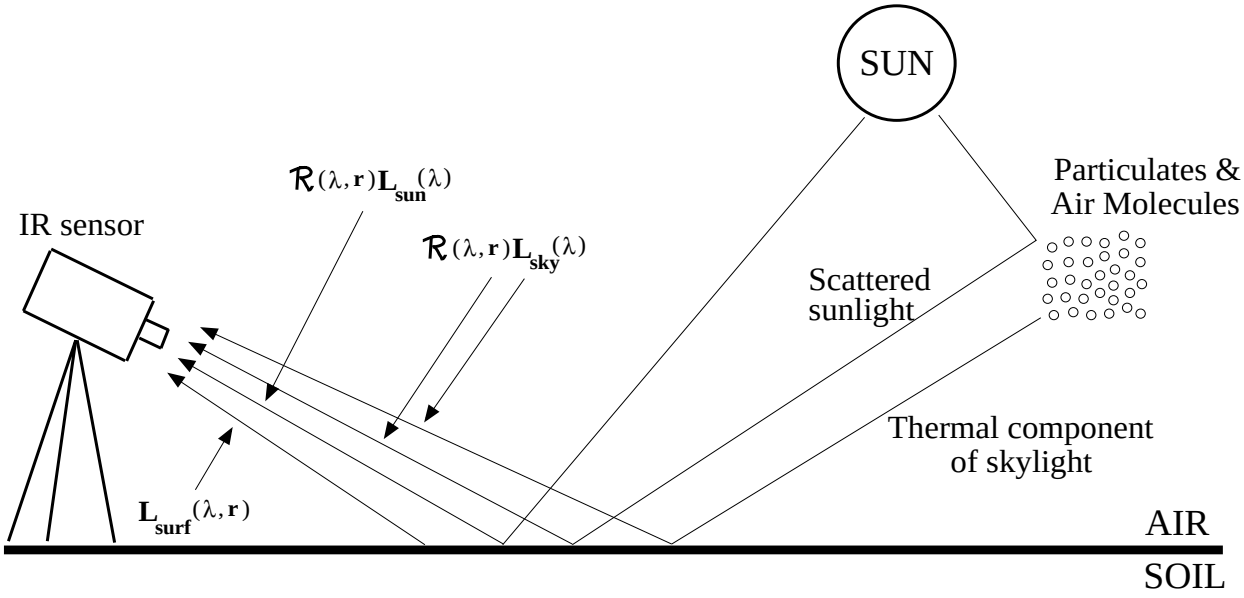


Fig. 2. Contributors to a thermal image of the ground include direct sunlight, aerosol-scattered sunlight, thermal emissions from the air reflected by the ground, and thermal emissions from the soil.

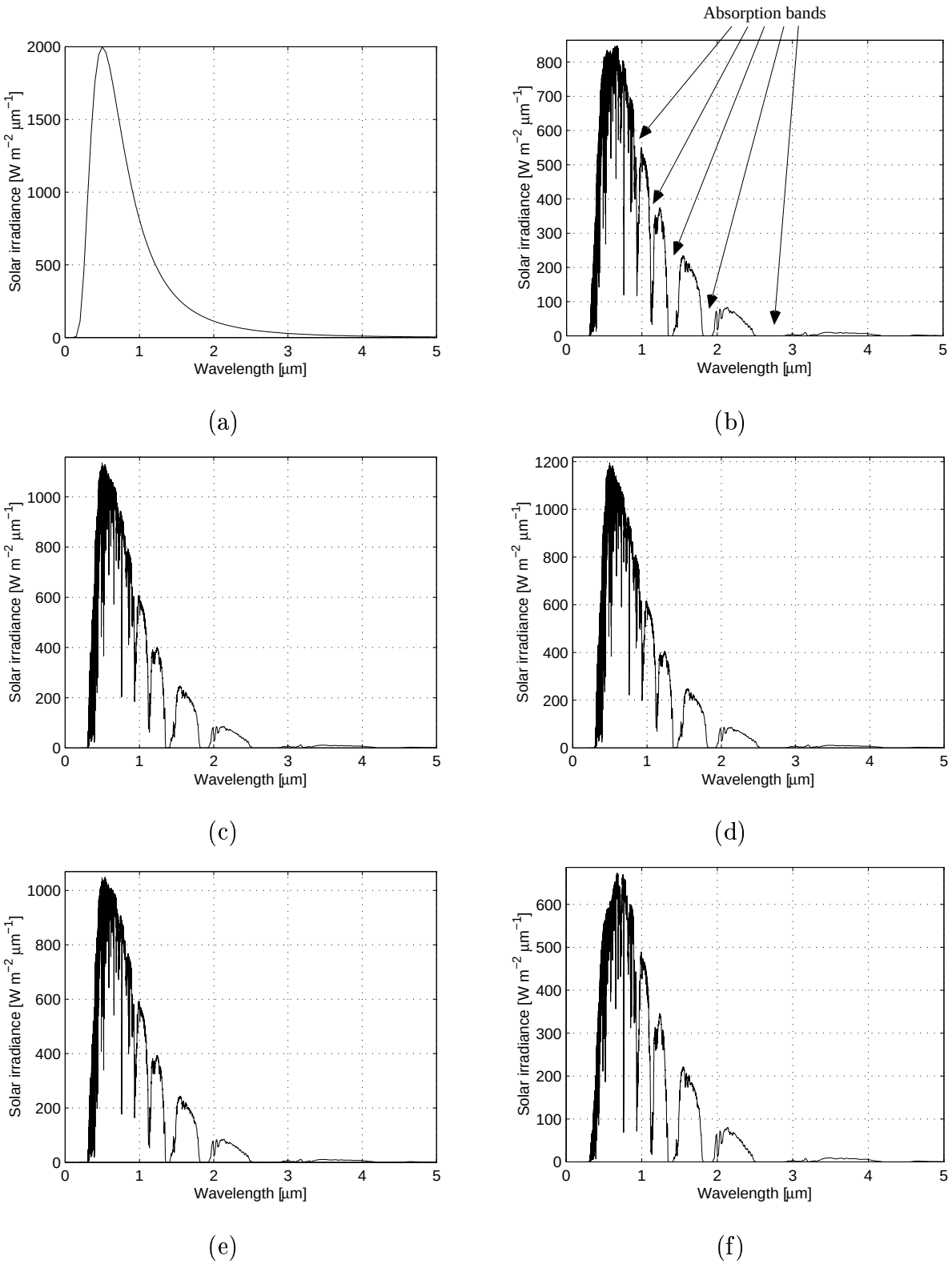


Fig. 3. Predicted spectral radiance at Columbus, OH on April 3. (a) Solar spectral radiance with no atmospheric effects. (b) Prediction by MODTRAN at 10 AM. (c) 12 PM (d) 2 PM (d) 4 PM (d) 6 PM

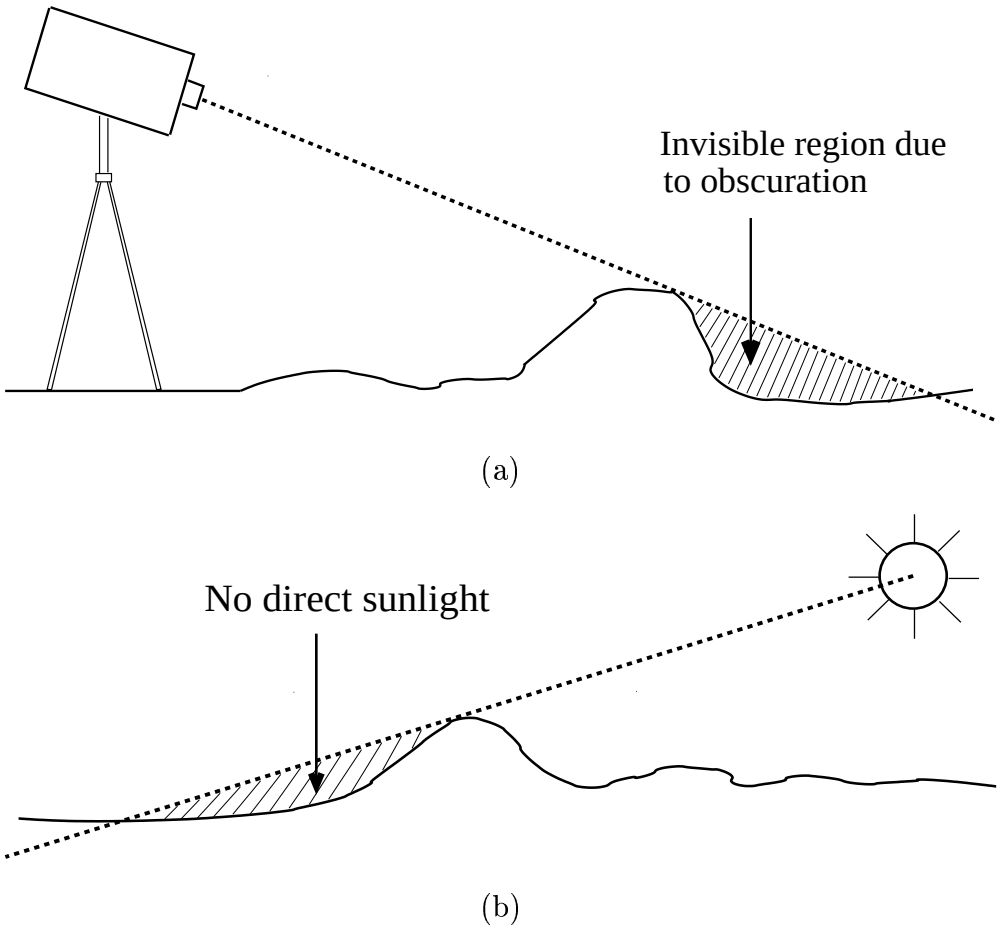


Fig. 4. (a) Surface roughness can produce obscuration. (b) Heating of the soil surface is affected by shadowing.

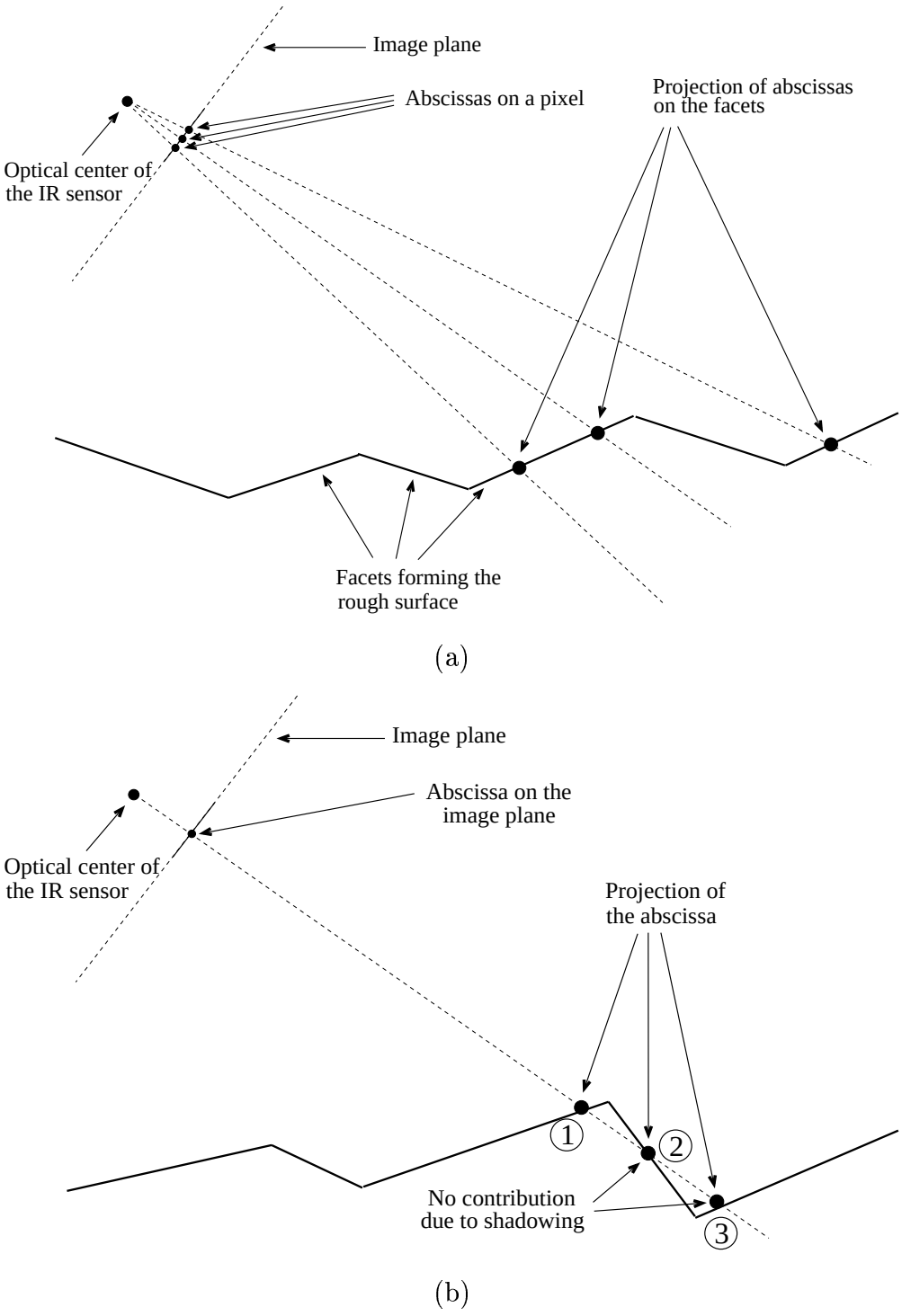


Fig. 5. (a) Projection of the integration abscissas onto the facets forming the rough surface. (b) Visible surface determination is required because of shadowing.

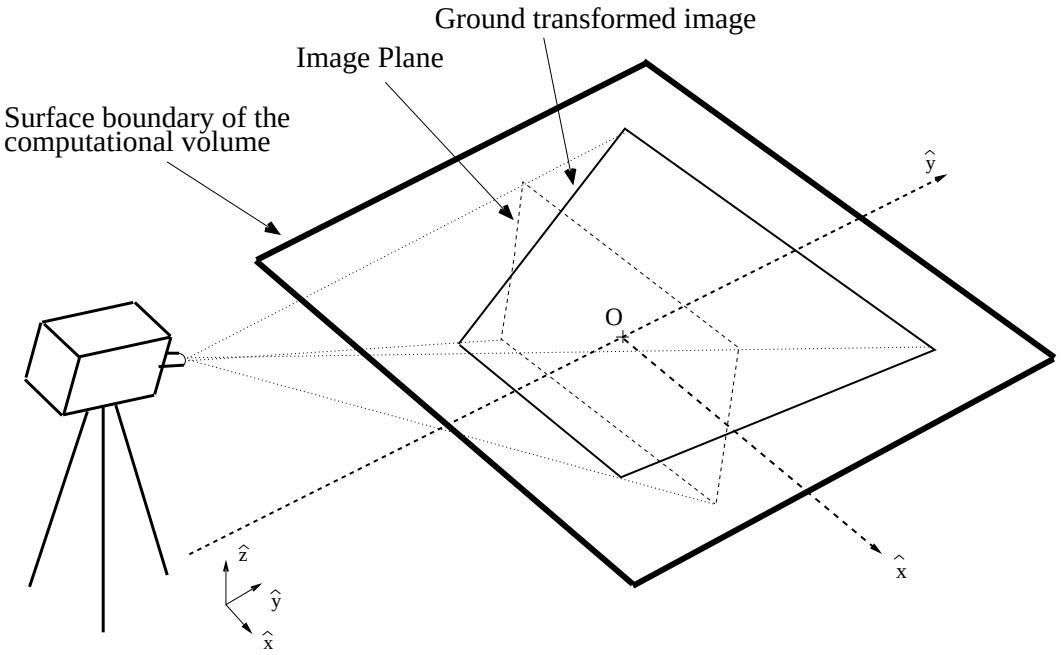


Fig. 6. Sketch of the region of interest including the IR camera. Point O is both the center of the field of view and the origin of coordinates. The image plane, the ground transformed image, the surface boundaries of the computational volume, and the global axis are illustrated.

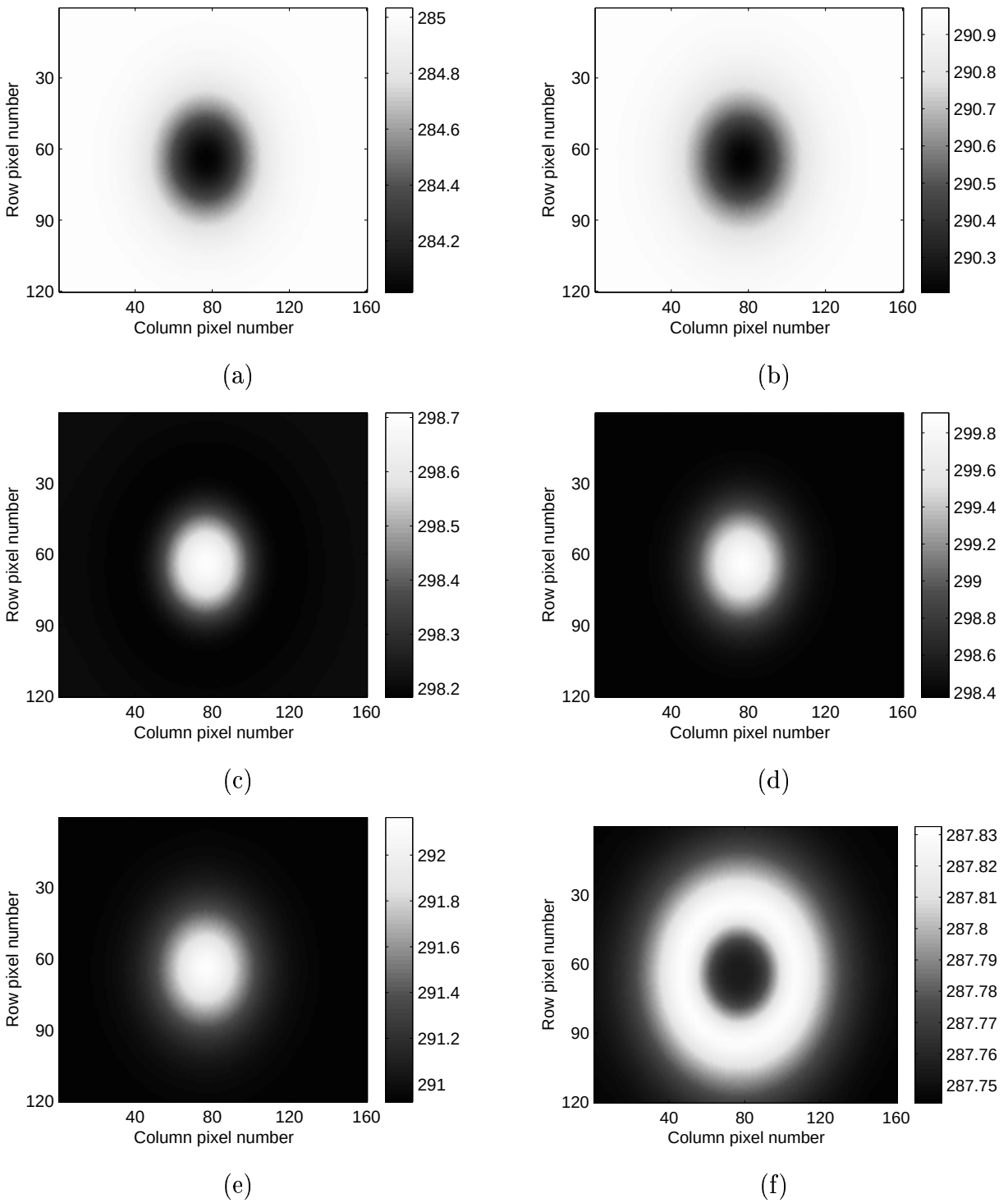


Fig. 7. Simulation of the IR signature of the SIM-25 mine buried 6.66 cm under a smooth soil surface for a sensor located 5.7 m above and 1 m to the right of the mine. The radiometric model is used to predict the response at different times of the day: (a) At dawn. (b) 3 hours after sunrise. (c) At noon. (d) 3 hours after noon. (e) At sunset. (f) 3 hours after sunset.

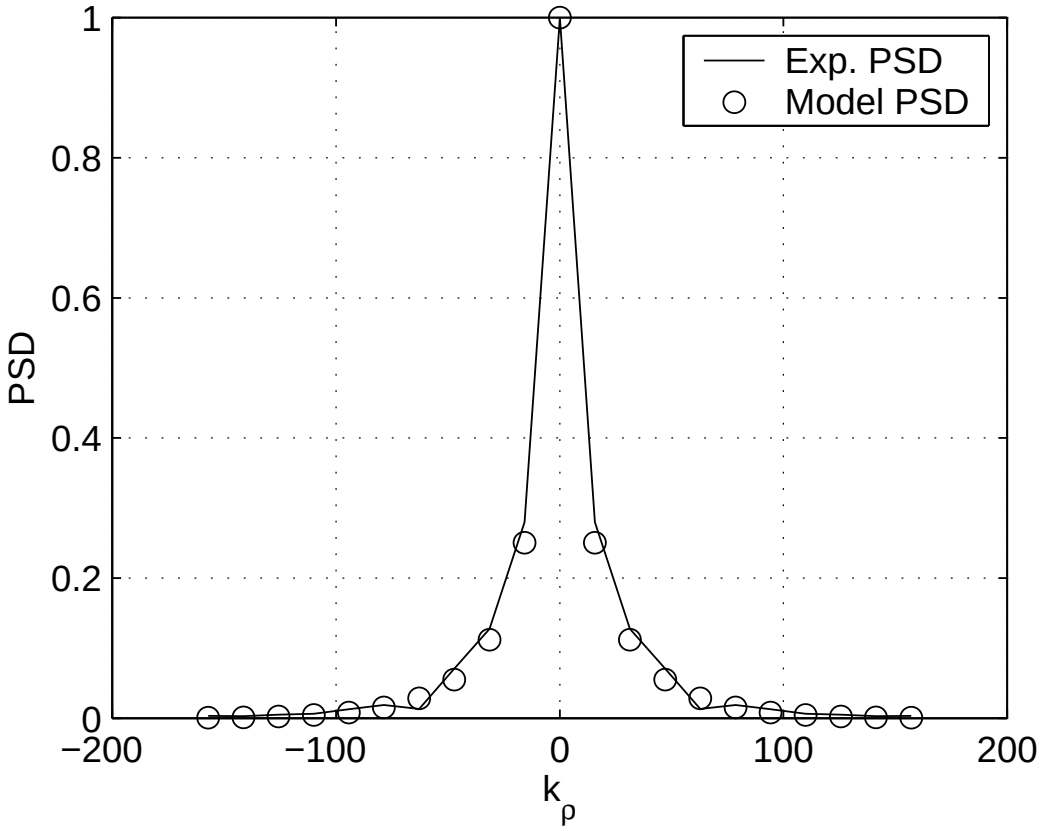


Fig. 8. Comparison of the experimental and model PSDs.

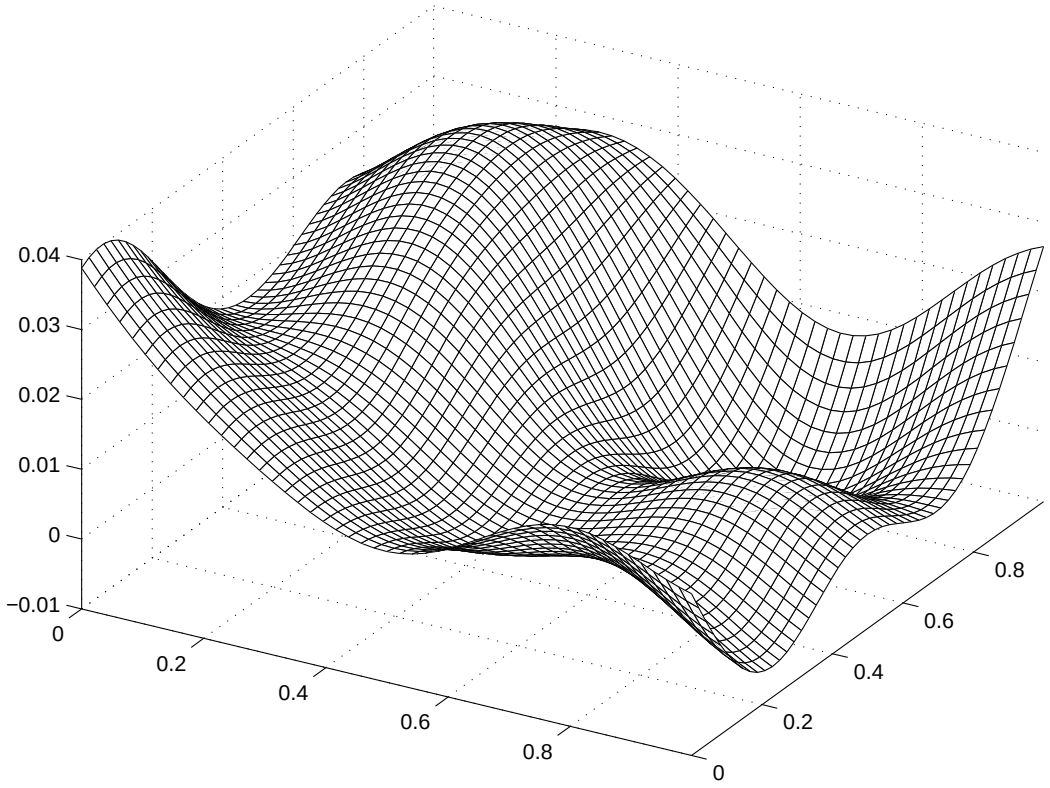


Fig. 9. An example of a rough surface realized using a Gaussian spectrum with correlation length $L = 20$ cm.

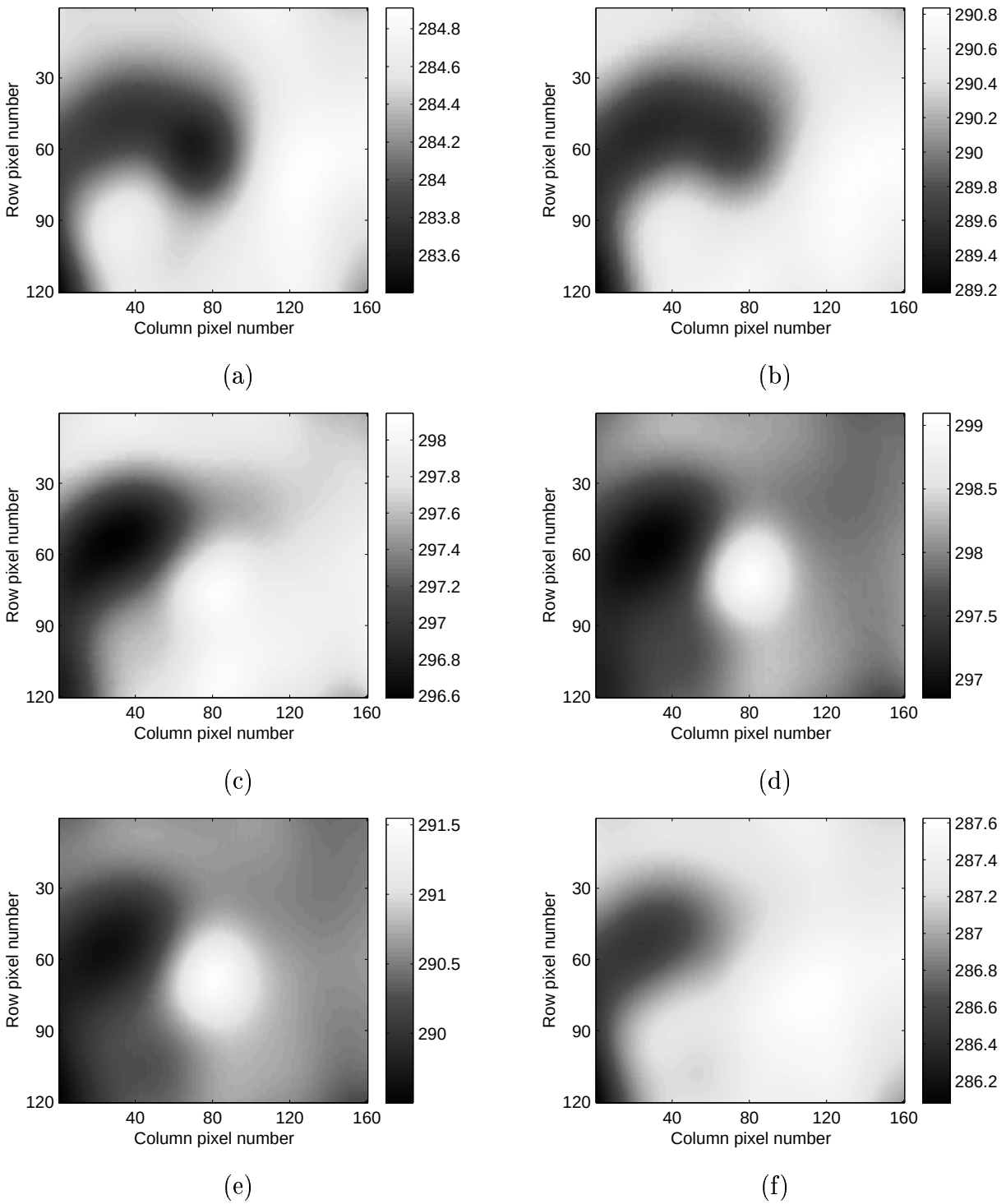


Fig. 10. Simulation of the IR signature of the SIM-25 mine buried 6.66 cm under a rough soil surface with peak to peak height variations of 5 cm. The radiometric model is used to predict the response at different times of the day: (a) At dawn. (b) 3 hours after sunrise. (c) At noon. (d) 3 hours after noon. (e) At sunset. (f) 3 hours after sunset.

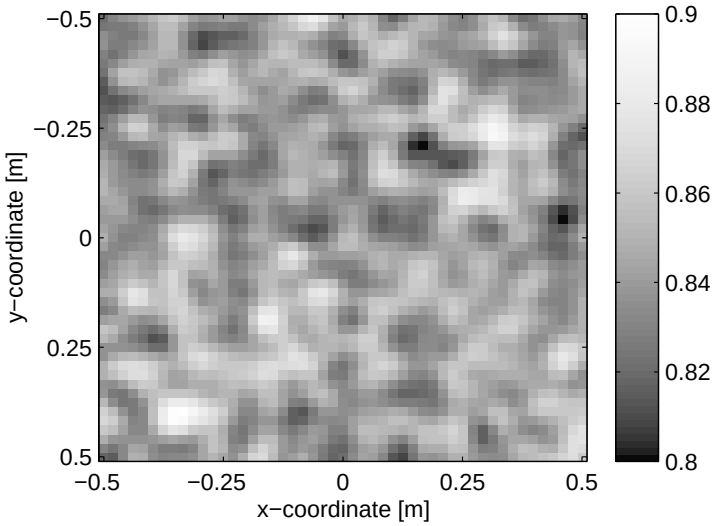


Fig. 11. An example surface emissivity realized using a Gaussian spectrum with correlation length $L = 5$ cm.

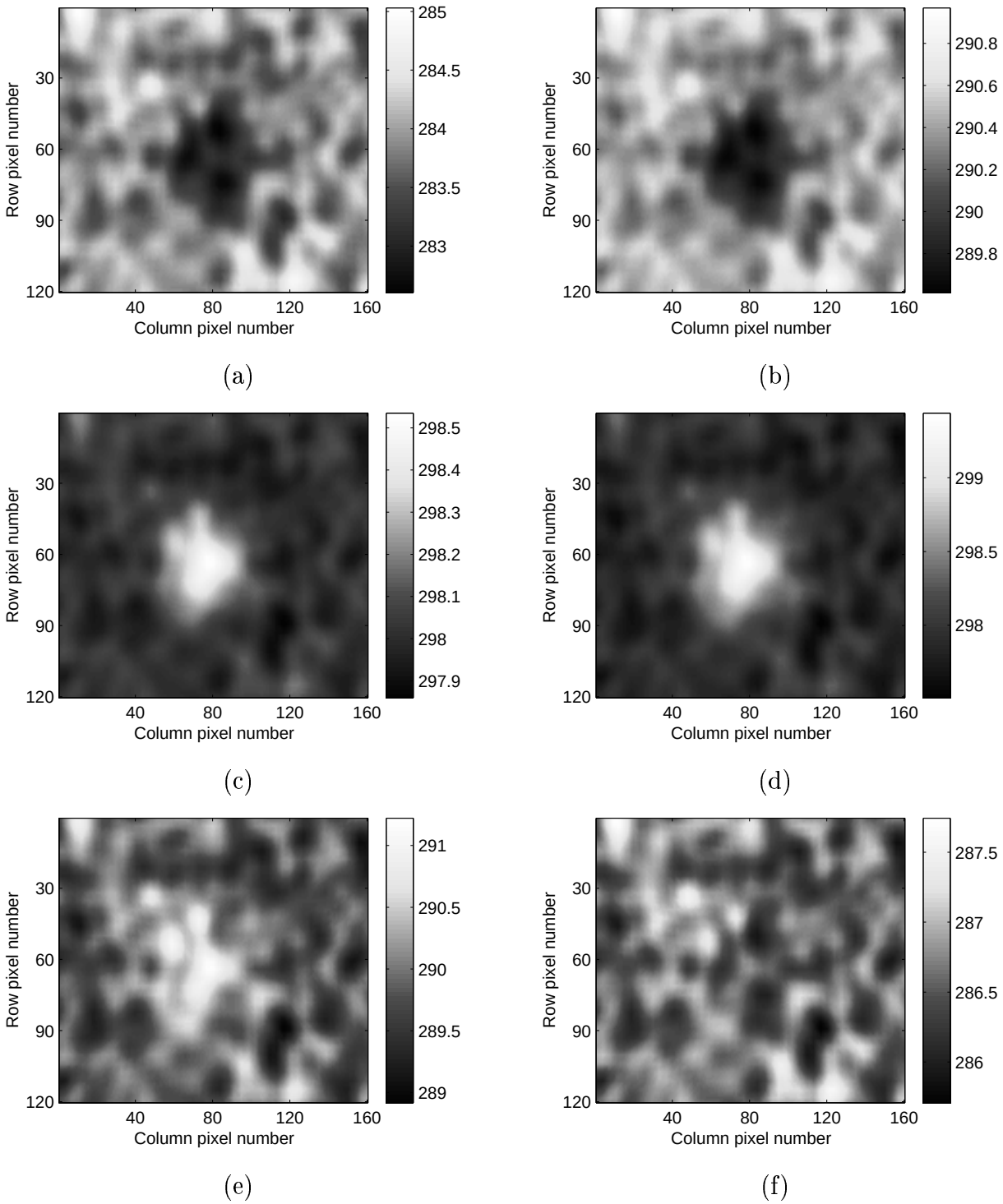


Fig. 12. Simulation of the IR camera response of the SIM-25 mine buried 6.66 cm under a smooth soil surface with a Gaussian surface emissivity profile. The radiometric model is used to predict the IR camera response at different times of the day (a) At dawn. (b) 3 hours after sunrise. (c) At noon. (d) 3 hours after noon. (e) At sunset. (f) 3 hours after sunset.

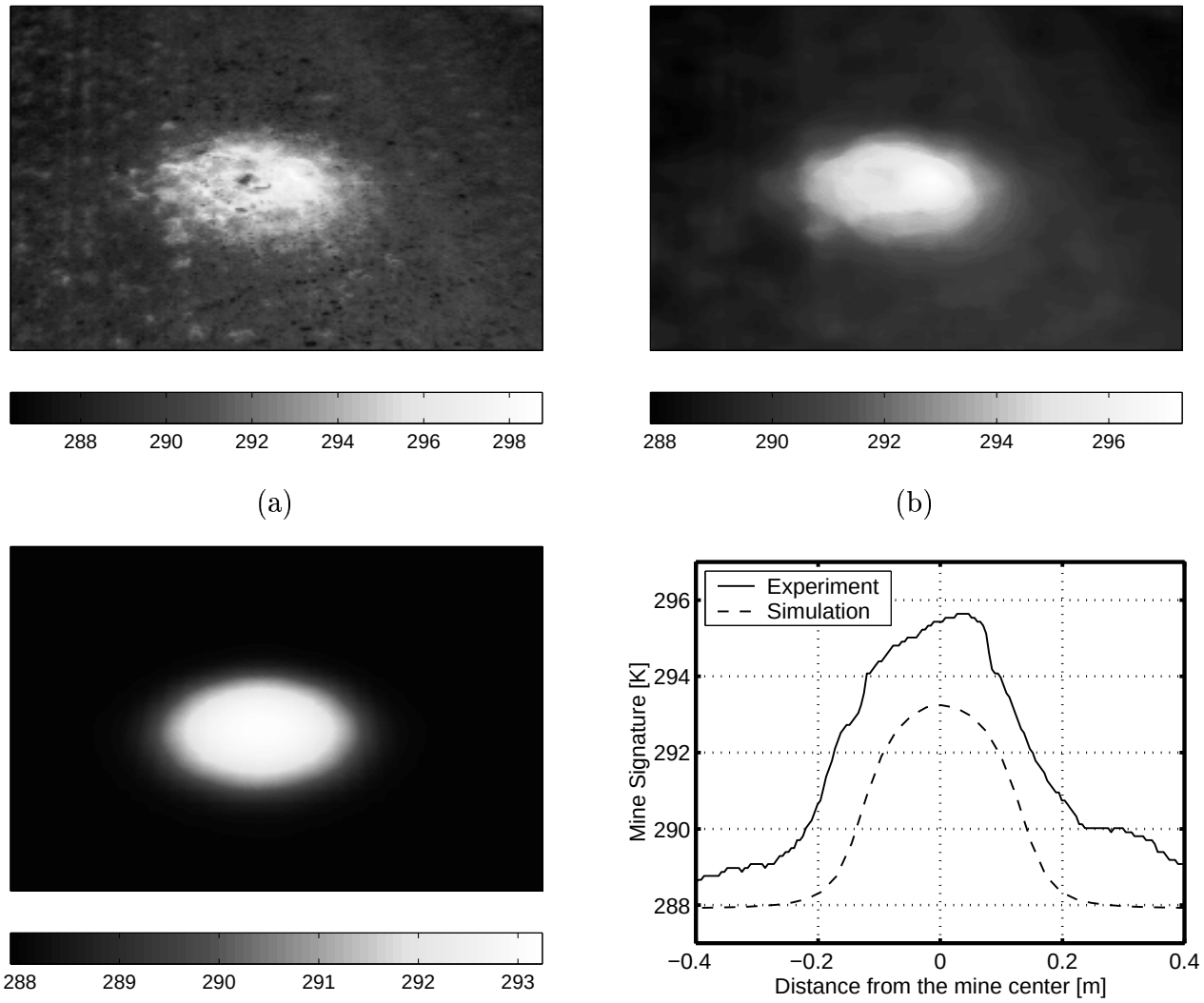


Fig. 13. (a) Raw experimental data. (b) Measured data after low-pass filtering to remove surface clutter. (c) Model prediction. (d) Comparison of model and experiment.

Integral Equation Solution for Thermal Signatures of Buried Land Mines

İbrahim Kürşat Şendur and Brian A. Baertlein

The authors are with the ElectroScience Laboratory, Department of Electrical Engineering, The Ohio State University, Columbus, OH 43212-1191 USA. E-mail: baertlein.1@osu.edu .

Abstract

Predicting the thermal signature of a buried land mine requires modeling the complicated inhomogeneous environment and the structurally complex mine. It is useful, both in checking such models and in making rough calculations of expected signatures, to have an accurate, easily computed reference solution for a relatively simple geometry. In this paper a reference solution is presented for the time varying surface temperature distribution over a homogeneous cylindrical body (the mine model) buried in an infinite homogeneous half space (the soil model) with a planar interface. In this work the convection coefficient and air temperature are assumed to be time invariant and the radiation condition is linearized about its mean value. Using a periodic boundary condition in time at the planar interface, the temperature distribution in the lower-half space is expanded in a Fourier series. A volume integral equation for the Fourier series coefficients is obtained via Green's second identity, and the Green's function for the Fourier coefficients is derived. The solution procedure uses the method of weighted residuals (MWR), in which the integral equation is reduced to a matrix equation and then solved for the unknown temperature distribution. The integral equation solution is compared with a finite element method based model.

Keywords

Thermal infrared imagery, land mines, heat transfer, integral equations, thermal model, numerical simulation, Green's function, singularity extraction.

I. Introduction

Research in infrared (IR) detection of land mines has been ongoing for several decades, and those sensors (especially multi-spectral instruments) are considered promising. Although many experimental and empirical studies of IR mine detection have been performed, the physics that define IR signatures are poorly understood. These signatures are dependent on several factors, including solar insolation, cloud cover, vegetation, surface irregularities, and past meteorological conditions. There is a critical need to understand the effects of these environmental factors so that detection performance and sensor utilization can be improved.

Models for IR signatures of buried mines are not well developed, but there has been progress in modeling related problems. Thermal IR remote sensing of soil provides useful information for terrestrial studies such as characterizing geological materials, monitoring effusive volcanism, detecting fractures, and hazard assessment. Thermal IR data can also

provide valuable information about temporal and spatial variations of soil temperature, and soil temperature prediction has been widely studied in the literature using both analytical and numerical techniques. Those models are one-dimensional, but they describe heat transfer at the soil-air interface and various environmental processes that are critical components of three-dimensional thermal signature models. Watson [1] used a Fourier series formulation [2] for periodic heating of a surface to obtain an analytical solution for the temperature distribution in homogeneous soil heated by diurnal solar insolation. That work included the effects of skylight, atmospheric absorption, thermal radiation, and thermal conduction in soil. England et al. [3] included the effects of convective heat transfer between soil and air. Kahle [4] obtained a solution of the one-dimensional heat flow equation, including sensible and latent heat transfer. Watson [5] determined the surface temperature over a homogeneous layer by using the Laplace transform to obtain a relation between surface flux and surface temperature. England [6] used a one-dimensional model to study the radiometric characteristics of diurnally heated freezing and thawing soils. Liou and England [7] developed diurnal and annual models for freezing and thawing moist soils subject to annual insolation, radiant heating and cooling, and sensible and latent heat exchanges with the atmosphere. Liou and England [8] used a one-dimensional coupled heat and moisture transport model for bare, unfrozen, moist soils. They later extended the model to freezing soils [9].

Recently, thermal models of IR mine signatures have been described that offer a more complete understanding of the underlying physics. Sendur and Baertlein [10] used a one-dimensional analytical thermal model to predict the surface temperature distribution over a mine. The land mine is assumed to be buried under a layered earth subjected to diurnal solar heating. The analytical formulation is based on a Fourier analysis of the periodic phenomenon. Closed form expressions were derived for the temperature distribution on the soil surface and at depth. That simple analysis confirmed a number of results that are seen in experimental studies, including the presence of two thermal "cross-over" times, a lag between the solar illumination and the soil temperature, the attenuation of the signature strength with depth, and a marked difference in the signatures of thermally conducting and insulating buried objects.

In a companion paper, Sendur and Baertlein [11] developed a three-dimensional thermal model to predict surface temperatures over buried mines. That work is based on the finite element method (FEM), which is capable of modeling realistically shaped mines and inhomogeneous soil. Although a FEM approach to the problem appears very attractive, the accuracy and efficiency of such model bears investigation because of artificial boundaries required at the surface of the computational volume and discretization of the mine shape. An accurate and easily computed solution for a relatively simple geometry is desirable for checking the more detailed FEM models and for making rough calculations of more complex environments. In this paper we present an integral equation solution for a homogeneous cylindrical body (the mine model) buried in an infinite homogeneous half-space with a planar interface (the soil model). The model presented in this work uses a time invariant convection coefficient and air temperature. The solution of the volume integral equation is obtained using the method of weighted residuals (MWR), a technique also known to researchers in electromagnetics as the "Method of Moments (MoM)". Although some derivations in this paper are inspired by work in electromagnetics, no knowledge of that subject is required here.

This work is organized as follows: Sect. II provides a description of the three-dimensional heat flow equation, including the convective and radiative boundary conditions at the soil-air interface. The Fourier series representation of the temperature distribution with appropriate boundary conditions is also presented in that section. In Sect. III integral representations for the temperature distribution in the soil and mine are obtained. The Green's function for this problem is derived in Sect. IV and subsequently simplified and converted into a computationally efficient form. The numerical solution procedure for the Fourier series coefficients of the temperature distribution is formulated in Sect. V. Numerical results are presented in Sect. VI, and the integral-equation and FEM solutions are compared. Concluding remarks appear in Sect. VII.

II. Fourier series coefficients of temperature distribution

The physical processes and the heat transfer mechanisms that produce IR signatures of buried land mines are discussed in a companion paper [11] and will not be repeated here. The key results are summarized below for the sake of completeness.

The temperature distribution in the soil and mine is described by the three-dimensional heat flow equation

$$C(\mathbf{r})\frac{\partial T(\mathbf{r},t)}{\partial t} = \nabla \cdot (\mathcal{K}(\mathbf{r})\nabla T(\mathbf{r},t)), \tag{1}$$

where $T(\mathbf{r},t)$ [K] is the temperature distribution, $\mathcal{K}$ [W m^{-1} K^{-1}] is the thermal conductivity of the material, and C [J m^{-3} K^{-1}] is the volumetric heat capacity of the material. Previous research on thermal modeling of homogeneous soil [1, 3–9] has produced the linearized boundary condition at the soil-air interface

$$\frac{\partial T(\mathbf{r},t)}{\partial z} \approx T(\mathbf{r},t)\frac{1}{\mathcal{K}_s}(h(t) + 4\mathcal{E}\sigma T_{sky}^3(t)) - \frac{1}{\mathcal{K}_s}(\mathcal{F}_{sun}(t) + h(t)T_{air}(t) + 4\mathcal{E}\sigma T_{sky}^4(t)). \tag{2}$$

where $\mathcal{F}_{sun}$ is the incident solar radiation reduced by cloud extinction, atmospheric absorption, soil albedo and the cosine of the zenith angle; T_{sky} and T_{air} are the sky and air temperatures, respectively; $\mathcal{K}_s$ [W m^{-1} K^{-1}] is the thermal conductivity of the soil; $\mathcal{E}$ [unitless] is the mean emissivity of the surface; $\sigma = 5.67 \times 10^{-8}$ [W m^{-2} K^{-4}] is the Stefan-Boltzmann constant; and $h(t)$ is a convection coefficient.

In this paper we present an integral equation based solution to Eq. (1). The boundary condition at the soil-air interface given by Eq. (2) is time-varying. If the convection coefficient $h(t)$ is approximated by its mean value $\overline{h}$, then this boundary condition becomes a periodic function at the diurnal rate. This approximation suggests that the temperature distribution $T(\mathbf{r},t)$ can be written as

$$T(\mathbf{r},t) = \sum_{n=-\infty}^{\infty} T_n(\mathbf{r})e^{i\omega nt} \tag{3}$$

where $\omega = 2\pi/(86400)$ [rad s^{-1}] is the radian frequency of a diurnal cycle and $T_n(\mathbf{r})$ is the nth Fourier coefficient of the temperature. Since $T(\mathbf{r},t)$ is a real quantity, the coefficients satisfy $T_n(\mathbf{r}) = T_{-n}^*(\mathbf{r})$ where * denotes the complex conjugate operator. Substituting Eq. (3) in Eq. (1) and assuming piecewise constant material properties, the Fourier coefficients of the three-dimensional heat-flux equation can be written as

$$\nabla^2 T_n(\mathbf{r}) - \frac{i\omega n}{\kappa_i(\mathbf{r})}T_n(\mathbf{r}) = 0, \qquad i = s, m \tag{4}$$

Solar heating is the dominant heat source in this problem, and this fact permits additional approximations. Replacing $T_{air}(t)$ with its mean value $\overline{T}_{air}$, and simplifying T_{sky} in

a similar manner yields

$$\frac{\partial T(\mathbf{r},t)}{\partial z} \approx T(\mathbf{r},t)\frac{1}{\mathcal{K}_s}\left(\overline{h}+4\mathcal{E}\sigma\overline{T}^3_{sky}\right)-\frac{1}{\mathcal{K}_s}\left(\mathcal{F}_{sun}(t)+\overline{hT}_{air}+4\mathcal{E}\sigma\overline{T}^4_{sky}\right) \quad (5)$$

By using a Fourier expansion of the solar insolation function

$$\mathcal{F}_{sun}(t)=\sum_{n=-\infty}^{\infty}\mathcal{F}_n e^{i\omega nt} \quad (6)$$

the boundary condition for the nth Fourier coefficient can be written as

$$\frac{\partial T_n(\mathbf{r})}{\partial z}=\alpha T_n(\mathbf{r})+\beta_n \quad (7)$$

where

$$\alpha=\left(\overline{h}+4\mathcal{E}\sigma\overline{T}^3_{sky}\right)\Big/\mathcal{K}_s \quad (8)$$

$$\beta_n=-\left(\mathcal{F}_n+\overline{hT}_{air}+4\mathcal{E}\sigma\overline{T}^4_{sky}\right)\Big/\mathcal{K}_s \quad (9)$$

III. Integral representation for the temperature distribution

In this section an integral representation for the Fourier coefficients of temperature in the lower half space will be obtained for a cylindrical mine with radius ρ_0 and thickness τ, buried at depth h under a smooth soil surface (see Fig. 1). Homogeneous soil with thermal diffusivity $\kappa_s(\mathbf{r})=\kappa_s$ and thermal conductivity $\mathcal{K}_s(\mathbf{r})=\mathcal{K}_s$ is assumed. The thermal properties of the mine need not be uniform in the following formulation, and the position ($\mathbf{r}$) dependence on the mine's thermal properties will be retained in this formulation.

A. Problem Formulation

For a mine that is a body of revolution, the problem has rotational symmetry. Therefore, Eq. (4) can be written as

$$\begin{aligned}\nabla^2_{\rho z}T^s_n(\mathbf{r})-k^2_n T^s_n(\mathbf{r}) &= 0 \quad ; \quad \mathbf{r}\in\text{soil}\\ \nabla^2_{\rho z}T^m_n(\mathbf{r})-\hat{k}^2_n(\mathbf{r})T^m_n(\mathbf{r}) &= 0 \quad ; \quad \mathbf{r}\in\text{mine}\end{aligned} \quad (10)$$

where

$$\nabla^2_{\rho z}=\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right)+\frac{\partial^2}{\partial z^2} \quad (11)$$

$$k_n^2 = \frac{i\omega n}{\kappa_s} \tag{12}$$

$$\hat{k}_n^2(\mathbf{r}) = \frac{i\omega n}{\kappa_m(\mathbf{r})} \tag{13}$$

In Eq. (10) $T_n^m(\mathbf{r})$ and $T_n^s(\mathbf{r})$ represent the Fourier coefficients in the mine and soil regions, respectively.

We employ a standard method [12] of solving the differential operator equation defined in Eq. (4). Forming the inner product of that equation and the Green's function G yields

$$\langle \nabla_{\rho z}^2 T_n(\rho', z'), G(\rho, z; \rho', z') \rangle = \langle \nabla_{\rho z}^2 G(\rho, z; \rho', z'), T_n(\rho', z') \rangle + \mathcal{R}(\rho, z) \tag{14}$$

where $\langle \cdot, \cdot \rangle$ is the inner product, which comprises integration over the entire lower half space. The Green's function satisfies the equation

$$(\nabla_{\rho z}^2 - k_n^2) G(\rho, z; \rho', z') = -\frac{\delta(\rho - \rho')}{\rho} \delta(z - z') \tag{15}$$

with boundary conditions defined below. The conjunct $\mathcal{R}$ is an integral over the problem boundaries [12, p. 198] and is also defined below. Using the above property of G and the differential equation satisfied by T_n we immediately obtain

$$T_n(\rho, z) = \int_{z_1}^{z_2} dz' \int_0^{\rho_0} d\rho' \rho' c(\rho', z') T_n(\rho', z') G(\rho, z; \rho', z') + \mathcal{R} \tag{16}$$

In this result $z_1 = h$, $z_2 = h + \tau$, we have exchanged primed and unprimed coordinates, and we have defined

$$c(\mathbf{r}) = \hat{k}_n^2(\mathbf{r}) - k_n^2 = i\omega n \frac{\kappa_s - \kappa_m(\mathbf{r})}{\kappa_s \kappa_m(\mathbf{r})} \tag{17}$$

Integration by parts is used to determine $\mathcal{R}$ from its definition in Eq. (14). In doing so, we exploit the thermal boundary conditions, namely that temperature is continuous

$$T_n^m(\mathbf{r} = \mathbf{r}_{ms}) = T_n^s(\mathbf{r} = \mathbf{r}_{ms}) \tag{18}$$

and thermal flux is continuous

$$\mathcal{K}_m \nabla T_n^m(\mathbf{r}) \Big|_{\mathbf{r}=\mathbf{r}_{ms}} = \mathcal{K}_s \nabla T_n^s(\mathbf{r}) \Big|_{\mathbf{r}=\mathbf{r}_{ms}} \tag{19}$$

where $\mathbf{r}_{ms}$ denotes any point on the mine-soil boundary. The following result is obtained

$$\begin{aligned}\mathcal{R}(\rho,z) &= -\int_0^\infty d\rho'\rho'\mathcal{Z}^s(\rho,z;\rho',0)+\rho_0\int_{z_1}^{z_2}dz'\mathcal{F}^m(\rho,z;\rho_0,z')\Big(1-\frac{\mathcal{K}_m(\rho_0,z')}{\mathcal{K}_s}\Big)\\ &+ \int_0^{\rho_0}d\rho'\rho'\Bigg(\mathcal{S}^m(\rho,z;\rho',z_2)\Big(1-\frac{\mathcal{K}_m(\rho',z_2)}{\mathcal{K}_s}\Big)\\ &\quad -\mathcal{S}^m(\rho,z;\rho',z_1)\Big(1-\frac{\mathcal{K}_m(\rho',z_1)}{\mathcal{K}_s}\Big)\Bigg)\end{aligned} \tag{20}$$

in which

$$\mathcal{Z}^i(\rho,z;\rho',z') = \mathcal{S}^i(\rho,z;\rho',z')-T_n^i(\rho',z')\frac{\partial G(\rho,z;\rho',z'')}{\partial z''}\bigg|_{z''=z'} \tag{21}$$

$$\mathcal{F}^i(\rho,z;\rho',z') = G(\rho,z;\rho',z')\frac{\partial T_n^i(\rho'',z')}{\partial\rho''}\bigg|_{\rho''=\rho'} \tag{22}$$

$$\mathcal{S}^i(\rho,z;\rho',z') = G(\rho,z;\rho',z')\frac{\partial T_n^i(\rho',z'')}{\partial z''}\bigg|_{z''=z'} \tag{23}$$

In Eqs. (21)-(23) the subscript i can be s or m, which refer to the soil and mine, respectively. By selecting the following boundary condition for the Green's function at the soil-air interface

$$\frac{\partial G(\rho,z;\rho',z')}{\partial z}=\alpha G(\rho,z;\rho',z') \tag{24}$$

the integral representation for the temperature distribution $T_n(\rho,z)$ in the lower half space can be obtained as

$$\begin{aligned}T_n(\rho,z) &= -\int_{z_1}^{z_2}dz'\int_0^{\rho_0}d\rho'\rho'c(\rho',z')T_n^m(\rho',z')G(\rho,z;\rho',z')\\ &- \beta_n\int_0^\infty d\rho'\rho'G(\rho,z;\rho',z'=0)+\rho_0\int_{z_1}^{z_2}dz'\mathcal{F}^m(\rho,z;\rho_0,z')\Big(1-\frac{\mathcal{K}_m(\rho_0,z')}{\mathcal{K}_s}\Big)\\ &+ \int_0^{\rho_0}d\rho'\rho'\Bigg(\mathcal{S}^m(\rho,z;\rho',z_2)\Big(1-\frac{\mathcal{K}_m(\rho',z_2)}{\mathcal{K}_s}\Big)\\ &\quad -\mathcal{S}^m(\rho,z;\rho',z_1)\Big(1-\frac{\mathcal{K}_m(\rho',z_1)}{\mathcal{K}_s}\Big)\Bigg)\end{aligned} \tag{25}$$

Equation (25) is an integral relation from which one can determine the temperature distribution anywhere in the lower half space by integrating the temperature distribution over the mine. If (ρ,z) is a point within the mine, this relation is a Fredholm's integral equation of the second kind for the unknown mine temperature distribution. In the absence of the

mine, i.e., for a homogeneous half space, the first, third, and last terms in Eq. (25) vanish and the temperature distribution can be found directly from

$$T_n(\rho, z) = -\beta_n \int_0^\infty d\rho' \rho' G(\rho, z; \rho', z' = 0) \tag{26}$$

The Green's function for Eq. (25) is derived in Sect. IV. The numerical solution procedure for Eq. (25) is presented in Sect. V.

B. Alternative Interpretation

An alternative view of the problem offers a physical interpretation of the expressions derived in Sect. III-A, and it permits analysis of mines containing media that are not piecewise constant. The key mathematical tool in this work is the volume equivalence theorem for the heat transfer equation. The volume equivalence theorem is widely used in electromagnetics [13–15] to determine the scattered electromagnetic fields in the presence of a material body in a homogeneous environment.

The original problem comprises a buried mine with thermal properties $\kappa_m(\mathbf{r})$ and $\mathcal{K}_m(\mathbf{r})$ in homogeneous soil with thermal properties $\kappa_s(\mathbf{r}) = \kappa_s$ and $\mathcal{K}_s(\mathbf{r}) = \mathcal{K}_s$. A new equivalent problem can be defined in which the mine is replaced by homogeneous soil and an equivalent heat source $\mathcal{Q}(\mathbf{r})$ as shown in Fig. 2. The temperature anomaly due to the inhomogeneity can be viewed as being generated by a so-called "induced" source $\mathcal{Q}(\mathbf{r})$, which is proportional to the temperature distribution in the inhomogeneous volume. The derivation is well known in electromagnetics, but somewhat involved. Using Eq. (3) in Eq. (1) and the properties of the divergence operator, we obtain

$$\mathcal{K}(\mathbf{r})\nabla^2 T_n(\mathbf{r}) + \nabla\mathcal{K}(\mathbf{r}) \cdot \nabla T_n(\mathbf{r}) - i\omega n C(\mathbf{r}) T_n(\mathbf{r}) = 0 \tag{27}$$

Dividing both sides of Eq. (27) by $\mathcal{K}(\mathbf{r})$ we get

$$\nabla^2 T_n(\mathbf{r}) + \frac{\nabla\mathcal{K}(\mathbf{r})}{\mathcal{K}(\mathbf{r})} \cdot \nabla T_n(\mathbf{r}) - k^2(\mathbf{r}) T_n(\mathbf{r}) = 0 \tag{28}$$

where

$$k^2(\mathbf{r}) = i\omega n \frac{C(\mathbf{r})}{\mathcal{K}(\mathbf{r})} = k_n^2 + (\hat{k}_n^2(\mathbf{r}) - k_n^2) U_{mine}(\mathbf{r}) \tag{29}$$

where k_n^2 and $\hat{k}_n^2(\mathbf{r})$ are defined in Sect. III-A and

$$U_{mine}(\mathbf{r}) = \begin{cases} 1 & ; \quad \mathbf{r} \in \text{mine} \\ 0 & ; \quad \mathbf{r} \in \text{soil} \end{cases} \tag{30}$$

Substituting Eq. (30) into Eq. (28) and rearranging the terms yields

$$\nabla^2 T_n(\mathbf{r}) - k_n^2 T_n(\mathbf{r}) = \mathcal{Q}(\mathbf{r}) \tag{31}$$

where $\mathcal{Q}(r)$ is the fictitious heat source. This is given by two terms

$$\mathcal{Q}(\mathbf{r}) \quad = \quad \mathcal{Q}_v(\mathbf{r}) + \mathcal{Q}_s(\mathbf{r}) \tag{32}$$

where

$$\mathcal{Q}_v(\mathbf{r}) \quad = \quad c(\mathbf{r}) T_n(\mathbf{r}) U_{mine}(\mathbf{r}) \tag{33}$$

$$\mathcal{Q}_s(\mathbf{r}) \quad = \quad -\frac{\nabla \mathcal{K}(\mathbf{r})}{\mathcal{K}(\mathbf{r})} \cdot \nabla T_n(\mathbf{r}) \tag{34}$$

The quantity $c(\mathbf{r})$ in Eq. (33) has been defined in Sect. III-A. Also note that

$$\begin{aligned} \frac{\nabla \mathcal{K}(\mathbf{r})}{\mathcal{K}(\mathbf{r})} \quad &= \quad \left(1 - \frac{\mathcal{K}_m(\rho, z_1)}{\mathcal{K}_s}\right) \delta(z - z_1) \Big(u(\rho) - u(\rho - \rho_0)\Big) \hat{z} \\ &- \quad \left(1 - \frac{\mathcal{K}_m(\rho, z_2)}{\mathcal{K}_s}\right) \delta(z - z_2) \Big(u(\rho) - u(\rho - \rho_0)\Big)\Big) \hat{z} \\ &- \quad \left(1 - \frac{\mathcal{K}_m(\rho_0, z)}{\mathcal{K}_s}\right) \delta(\rho - \rho_0) \Big(u(z - z_1) - u(z - z_2)\Big) \hat{\rho} \end{aligned} \tag{35}$$

where u is the unit step function. Eq. (31) can be manipulated to yield a volume integral equation, which can be solved for the unknown, induced heat source. Using this induced heat source, the temperature distribution in the lower half space can be found by integration. This volume formulation can be conveniently used to treat any inhomogeneous material, not just the piecewise constant media discussed here.

Applying Green's second identity to the functions $T_n(\rho, z)$ and $G(\rho, z; \rho', z')$ with a source distribution $\mathcal{Q}(\rho, z)$ in the mine gives

$$\begin{aligned} T_n(\rho, z) \quad &= \quad -\int_{z_1}^{z_2} dz' \int_0^{\rho_0} d\rho' \rho \mathcal{Q}(\rho', z') G(\rho, z; \rho', z') \\ &+ \quad \int_\Gamma \Big[T_n(\rho, z) \nabla G(\rho, z; \rho', z') - G(\rho, z; \rho', z') \nabla T_n(\rho, z) \Big] \cdot \hat{n} dS \end{aligned} \tag{36}$$

where Γ denotes the boundaries (at the soil-air interface and at the mine) and $\hat{n}$ is the unit normal on those boundaries. Using the definitions of the boundaries, Eq. (36) can be written as

$$T_n(\rho,z) \;=\; -\int_{z_1}^{z_2} dz' \int_0^{\rho_0} d\rho' \rho \mathcal{Q}(\rho',z') G(\rho,z;\rho',z') - \int_0^{\infty} d\rho' \rho' \mathcal{Z}^s(\rho,z;\rho',0) \quad (37)$$

Equation (37) can be transformed into Eq. (25) by selecting the boundary condition for G as given by Eq. (24).

IV. Green's Function

The Green's function G is the solution of the heat transfer equation for an internal point source of heat. When G is known, a solution for an arbitrary source can be obtained via Eq. (25). In this section a Green's function will be derived, which satisfies Eq. (15) with the boundary condition on the soil-air interface given by Eq. (24). In addition to this boundary condition, the Green's function must have a finite value for $G(\rho = 0, z; \rho', z')$ and must vanish as $\rho \to \infty$ and $z \to \infty$. In Sect. IV-A the derivation is presented. In Sect. IV-B the Green's function is simplified and transformed into a more appropriate form for numerical evaluation.

A. Derivation of the Green's Function

The Green's function is easily derived in the spectral domain. Taking the Hankel transform [12]

$$\hat{f}(k_\rho) = \int_0^{\infty} x J_0(k_\rho x) f(x) dx \quad (38)$$

of both sides of the Eq. (15) we obtain

$$\left(\frac{\partial^2}{\partial z^2} - (k_\rho^2 + k_n^2)\right) \widehat{G}(k_\rho, z; \rho', z') = -J_0(k_\rho \rho') \delta(z - z') \quad (39)$$

where $\widehat{G}(k_\rho, z; \rho', z')$ is the Hankel transform of the Green's function $G(\rho, z; \rho', z')$. In obtaining Eq. (39) the differential equation for the zeroth order Bessel function was used. The solution of the ordinary differential equation given by Eq. (39) with the boundary conditions at the interface can be obtained through well known methods. We can show

$$\frac{\widehat{G}(k_\rho, z; \rho', z')}{J_0(k_\rho \rho')} = \frac{1}{2\sqrt{k_n^2 + k_\rho^2}} \exp\left(-\sqrt{k_n^2 + k_\rho^2} z_>\right) \left[\exp\left(\sqrt{k_n^2 + k_\rho^2} z_<\right)\right.$$

$$+r_0(k_n,k_\rho,\alpha)\exp\Big(-\sqrt{k_n^2+k_\rho^2}z_<\Big)\Big] \quad (40)$$

where

$$r_0(k_n,k_\rho,\alpha)=\frac{\sqrt{k_n^2+k_\rho^2}-\alpha}{\sqrt{k_n^2+k_\rho^2}+\alpha} \quad (41)$$

The inverse Hankel transform, given by

$$f(x)=\int_0^\infty k_\rho J_0(k_\rho x)\hat{f}(k_\rho)dk_\rho \quad (42)$$

yields the desired result

$$\begin{aligned} G(\rho,z;\rho',z') &= \int_0^\infty dk_\rho J_0(k_\rho\rho')J_0(k_\rho\rho)\frac{k_\rho}{2\sqrt{k_n^2+k_\rho^2}}\exp\Big(-\sqrt{k_n^2+k_\rho^2}z_>\Big) \\ &\times \Big[\exp\Big(\sqrt{k_n^2+k_\rho^2}z_<\Big)+r_0(k_n,k_\rho,\alpha)\exp\Big(-\sqrt{k_n^2+k_\rho^2}z_<\Big)\Big] \end{aligned} \quad (43)$$

B. Simplification of the Green's Function

Numerical evaluation of the Green's function given by Eq. (43) is computationally challenging. The integral in Eq. (43) is similar to the Sommerfeld integrals [16], which have been studied extensively in the physics and electromagnetics literature. The major difficulties in the computation of these integrals can be summarized as follows:

- For large argument x, the function $J_0(x)$ oscillates rapidly, leading to slow convergence.
- For $|z-z'|\ll 1$ or $(z+z')\ll 1$ the exponential terms in the integrals decay slowly, again producing slow convergence.

Straightforward manipulations ameliorate some of these problems. The quantity $r_0(k_n,k_\rho,\alpha)$ can be expressed as

$$r_0(k_n,k_\rho,\alpha)=1-\frac{2\alpha}{\sqrt{k_n^2+k_\rho^2}+\alpha} \quad (44)$$

and, hence, the Green's function can be expressed as

$$G(\rho,z;\rho',z') = G_1(\rho,z;\rho',z')+G_2(\rho,z;\rho',z')+G_3(\rho,z;\rho',z') \quad (45)$$

where

$$G_1(\rho,z;\rho',z') = \int_0^\infty dk_\rho J_0(k_\rho\rho')J_0(k_\rho\rho)\frac{k_\rho}{2\sqrt{k_n^2+k_\rho^2}}\exp\Big(-\sqrt{k_n^2+k_\rho^2}(z_>-z_<)\Big) \quad (46)$$

$$G_2(\rho,z;\rho',z') = \int_0^\infty dk_\rho J_0(k_\rho\rho')J_0(k_\rho\rho)\frac{k_\rho}{2\sqrt{k_n^2+k_\rho^2}}\exp\Big(-\sqrt{k_n^2+k_\rho^2}(z+z')\Big) \quad (47)$$

$$G_3(\rho,z;\rho',z') = -\alpha\int_0^\infty dk_\rho J_0(k_\rho\rho')J_0(k_\rho\rho)\frac{k_\rho}{\sqrt{k_n^2+k_\rho^2}(\sqrt{k_n^2+k_\rho^2}+\alpha)}\cdot \exp\Big(-\sqrt{k_n^2+k_\rho^2}(z+z')\Big) \quad (48)$$

Employing the multiplication identity for the zeroth order Bessel functions used by Johnson [17]

$$J_0(k_\rho\rho')J_0(k_\rho\rho) = \frac{1}{2\pi}\int_0^{2\pi} J_0(k_\rho\rho_t)d\phi \quad (49)$$

where

$$\rho_t = \sqrt{\rho^2+\rho'^2-2\rho\rho'\cos\phi} \quad (50)$$

and changing the order of integration we can rewrite Eqs. (46) and (47) as

$$G_1(\rho,z;\rho',z') = \frac{1}{4\pi}\int_0^{2\pi}d\phi'\int_0^\infty dk_\rho J_0(k_\rho\rho_t)\frac{k_\rho}{\sqrt{k_n^2+k_\rho^2}}\exp\Big(-\sqrt{k_n^2+k_\rho^2}(z_>-z_<)\Big) \quad (51)$$

$$G_2(\rho,z;\rho',z') = \frac{1}{4\pi}\int_0^{2\pi}d\phi'\int_0^\infty dk_\rho J_0(k_\rho\rho_t)\frac{k_\rho}{\sqrt{k_n^2+k_\rho^2}}\exp\Big(-\sqrt{k_n^2+k_\rho^2}(z+z')\Big) \quad (52)$$

Using the following identity [12]

$$\frac{\exp(-k\sqrt{r^2+z^2})}{\sqrt{r^2+z^2}} = \int_0^\infty \gamma J_0(\gamma r)\frac{\exp(-|z|\sqrt{k^2+\gamma^2})}{\sqrt{k^2+\gamma^2}}d\gamma \quad (53)$$

permits Eqs. (51) and (52) to be simplified to

$$G_m(\rho,z;\rho',z') = \int_0^{2\pi}d\phi'\frac{\exp(ik_nR_m)}{4\pi R_m}; \qquad m=1,2 \quad (54)$$

where

$$R_1 = \sqrt{\rho'^2+\rho^2-2\rho'\rho\cos\phi'+(z-z')^2} \quad (55)$$

$$R_2 = \sqrt{\rho'^2+\rho^2-2\rho'\rho\cos\phi'+(z+z')^2} \quad (56)$$

Similarly, substituting Eq. (49) into Eqs. (57) and changing the order of integration we obtain

$$G_3(\rho,z;\rho',z') = -\frac{\alpha}{2\pi}\int_0^{2\pi}d\phi I_{G_3} \quad (57)$$

where

$$I_{G_3} = \int_0^\infty dk_\rho J_0(k_\rho \rho_t) \frac{k_\rho}{\sqrt{k_n^2 + k_\rho^2}(\sqrt{k_n^2 + k_\rho^2} + \alpha)} \exp\Big(- \sqrt{k_n^2 + k_\rho^2}(z + z')\Big) \tag{58}$$

We can express this in another form as suggested by Kuo and Mei [18]. Multiplying both sides of Eq. (58) by $\exp(-\alpha(z+z'))$ and differentiating with respect to $z + z'$ yields

$$\begin{aligned} \frac{\partial(I_{G_3} \exp(-\alpha(z+z')))}{\partial(z+z')} &= -\exp(-\alpha(z+z')) \int_0^\infty dk_\rho \\ & \qquad J_0(k_\rho \rho_t) \frac{k_\rho}{\sqrt{k_n^2 + k_\rho^2}} \exp\Big(- \sqrt{k_n^2 + k_\rho^2}(z + z')\Big) \end{aligned} \tag{59}$$

The integral in Eq. (59) can be evaluated using Eq. (53). After integrating both sides with respect to z over appropriate limits, I_{G_3} can be written as

$$I_{G_3} = \exp(-\alpha(z+z')) \int_{z+z'}^\infty \frac{\exp(ik_n R_3)}{R_3} \exp(-\alpha\zeta) d\zeta \tag{60}$$

where

$$R_3 = \sqrt{\rho'^2 + \rho^2 - 2\rho'\rho\cos\phi' + \zeta^2} \tag{61}$$

Using Eqs. (54), (57), and (60) in conjunction with Eq. (45) we obtain the final form of the Green's function as

$$\begin{aligned} G(\rho, z; \rho', z') = & \int_0^{2\pi} d\phi' \frac{\exp(ik_n R_1)}{4\pi R_1} + \int_0^{2\pi} d\phi' \frac{\exp(ik_n R_2)}{4\pi R_2} \\ & - \frac{\alpha}{2\pi} \int_0^{2\pi} d\phi \exp(-\alpha(z+z')) \int_{z+z'}^\infty d\zeta \frac{\exp(ik_n R_3)}{R_3} \exp(-\alpha\zeta) \end{aligned} \tag{62}$$

The last integral in Eq. (62) has infinite integration limits, but the integrand is a rapidly decaying function for typical values of α, and it is well approximated by an integral over a finite domain. When the source and observation points approach one another, we have $R_1 \to 0$, and the Green's function has a well-known integrable singularity. The treatment of this case is discussed in Sect. V-B and Appendix A.

V. NUMERICAL SOLUTION PROCEDURE

In this section we describe a numerical solution procedure for the integral equation in Eq. (25) to obtain the Fourier coefficients $T_n^m(\mathbf{r})$. Our approach is to solve for the temperature T_n^m within the mine, which can be used to evaluate Eq. (25) for the Fourier

coefficients at the surface. The time history of the temperature is then found by evaluating the Fourier series.

We employ the MWR in this work. As noted above, the electromagnetics community has developed an extensive body of knowledge on MWR solutions of integral equations in the temporal frequency domain under the guise of the so-called "method of moments" (MoM). The MoM discretizes the integral equation into a matrix equation, the solution of which is obtained using standard methods. Following the pioneering works of Richmond [19] and Harrington [20], an extensive literature has developed on this procedure, and good summary references are also available [21–23].

A. Matrix formulation

The solution of the integral equation using the MWR begins with a representation of the unknown temperature distribution over the buried mine using specified expansion functions and unknown coefficients. The problem at hand imposes constraints on these functions. The integral representation of the temperature distribution given in Eq. (25) requires both the temperature and its derivatives with respect to ρ and z. We employ expansion functions $\Lambda(\rho, z)$ that are linear in ρ and z, viz:

$$T_n^m(\rho, z) = \sum_{m=1}^{M} \sum_{n=1}^{N} A_{m,n} \Lambda_{nm}(\rho, z) \tag{63}$$

where M and N are the number of divisions in the z and ρ directions, respectively. This representation yields a continuous, piecewise-linear representation of T_n^m and a piecewise constant form for its derivatives.

A sample discretization of the buried mine with $N = 4$ subdivisions in the $\hat{\rho}$-direction and $M = 3$ subdivisions in the $\hat{z}$-direction is illustrated in Fig. 3. The thermal properties over each subsectional basis function are assumed constant and taken into account in the constant $c_{m,n}$, which can be defined as

$$c_{m,n} = c\left((\rho_n + \rho_{n-1})/2, (z_m + z_{m-1})/2\right) \tag{64}$$

where c was previously defined in Eq. (17) and the indexed values ρ_n and z_m are the subsection boundaries. By substituting the approximate temperature distribution of Eq. (63)

into the integral equation Eq. (25), we obtain

$$\begin{aligned}
\beta_n \int_0^\infty d\rho' \rho' G(\rho, z; \rho', z' = 0) &= -\sum_{m=1}^{M}\sum_{n=1}^{N} A_{mn} \left[\Lambda_{mn}(\rho, z) \right. \\
&\qquad \left. + c_{mn} \int\int dz' d\rho' \rho' G(\rho, z; \rho', z') \Lambda_{mn}(\rho', z') \right] \\
&+ \rho_0 \sum_{m=1}^{M} A_{mN} \int dz' G(\rho, z; \rho' = \rho_0, z') \cdot \\
&\qquad \frac{d}{d\rho'} \Lambda_{nm}(\rho', z') \left(1 - \frac{\mathcal{K}_m(\rho_0, z')}{\mathcal{K}_s}\right) \\
&+ \sum_{n=1}^{N} A_{Mn} \int d\rho' \rho' G(\rho, z; \rho', z' = z_2) \cdot \\
&\qquad \frac{d}{dz'} \Lambda_{nm}(\rho', z') \left(1 - \frac{\mathcal{K}_m(\rho', z_2)}{\mathcal{K}_s}\right) \\
&- \sum_{n=1}^{N} A_{1n} \int d\rho' \rho' G(\rho, z; \rho', z' = z_1) \cdot \\
&\qquad \frac{d}{dz'} \Lambda_{nm}(\rho', z') \left(1 - \frac{\mathcal{K}_m(\rho', z_1)}{\mathcal{K}_s}\right) \qquad (65)
\end{aligned}$$

To complete the formulation, we express the left-hand side of Eq. 65 using the summation in Eq. (63), multiply both sides of Eq. (65) by testing functions $w_{n'm'}(\rho, z)$ and integrate over the mine. Various testing functions have been used in the literature. Galerkin methods are known to be optimum in the least-square sense, but for integral equations they are also computationally expensive in terms of matrix-fill time. To reduce the computational cost, we have employed a point-matching technique. The result can be expressed as a linear system of equations

$$\overline{\mathbf{Z}}\,\mathbf{A} = \mathbf{V} \qquad (66)$$

for the constants A_{mn}, which we represent by the matrix $\mathbf{A}$. The kernel of the integral equation appears in $\overline{\mathbf{Z}}$ which we refer to as the "impedance" matrix. Solving Eq. (66) using an appropriate technique yields the unknown coefficient vector $\mathbf{A}$, which can be used to compute the temperature distribution over the buried mine via Eq. (63). If desired, the temperature distribution everywhere in the lower half-space can be computed using Eq. (25).

B. Impedance matrix, source vector, and singularity extraction

In this section evaluation of the matrix $\overline{\mathbf{Z}}$ is discussed. Rows of this matrix have indices derived from the (ρ, z) expansion functions indexed by m and n, with a similar relation for the columns involving ρ' and z'. To make the notation more concise, we use the index p to refer to a particular point (ρ, z), which we denote (ρ_{n_p}, z_{m_p}) with a similar convention for ρ' and z' using q. In general, Z_{pq} involves integrals of the form

$$\int_{z_{m_q-1}}^{z_{m_q}} dz' \int_{\rho_{n_q-1}}^{\rho_{n_q}} d\rho' G(\rho_{n_p}, z_{m_p}; \rho', z') f(\rho', z') \tag{67}$$

where $f(\rho, z)$ is a continuous (at most linear) function of the arguments. Consider first the off-diagonal matrix elements Z_{pq} with $p \neq q$. For this case, the testing point (ρ_{n_p}, z_{m_p}) is not in the integration domain and the Green's function is a smooth function. In this case Z_{pq} can be directly evaluated by numerical methods.

For the diagonal matrix elements Z_{pp}, however, the testing point (ρ_{n_p}, z_{m_p}) lies in the integration domain. This results in $R_1 \to 0$ which causes a singularity as R_1^{-1}. Consequently, an integrand regularization technique (e.g., singularity extraction) is necessary in numerical evaluation of the integrals. Singularity extraction techniques are widely used in MoM solutions in electromagnetics for both surface and volume formulations [24–27]. Techniques for bodies of revolution have also been studied [28,29], which are now described.

In brief, calculation of the diagonal elements requires that the singular parts of the integrand are evaluated separately by analytical methods. A careful examination suggests that the singular part of the Green's function $G(\rho, z; \rho', z')$ is due to function $G_1(\rho, z; \rho', z')$, which leads us to write

$$G_1(\rho, z; \rho', z') = G_1^R(\rho, z; \rho', z') + \int_0^{2\pi} d\phi' \frac{1}{4\pi R_1} f(\rho, z) \tag{68}$$

where $f(\rho, z)$ is a smooth function and

$$G_1^R(\rho, z; \rho', z') = \int_0^{2\pi} d\phi' \frac{\exp(ik_n R_1) - 1}{4\pi R_1} \tag{69}$$

The function $G_1^R(\rho, z; \rho', z')$ can be integrated numerically. For small elements a Taylor series expansion of the integrand yields

$$\lim_{R_1 \to 0} G_1^R(\rho, z; \rho', z') = \frac{k_n}{2} \tag{70}$$

The remaining singular integral in Eq. (68) has one of the following forms

$$\begin{bmatrix} I_c(\rho_l,\rho_u;\phi_l,\phi_u;z_l,z_u) \\ I_\rho(\rho_l,\rho_u;\phi_l,\phi_u;z_l,z_u) \\ I_z(\rho_l,\rho_u;\phi_l,\phi_u;z_l,z_u) \end{bmatrix} = \int_{\rho_l}^{\rho_u} d\rho'\rho' \int_{z_l}^{z_u} dz' \int_{\phi_l}^{\phi_u} d\phi' \cdot \frac{1}{\sqrt{\rho'^2+\rho^2-2\rho'\rho\cos\phi'+(z-z')^2}} \begin{bmatrix} 1 \\ \rho' \\ z' \end{bmatrix} \quad (71)$$

Evaluation of these functions is discussed in Appendix I.

Singularity extraction is not necessary in evaluating the source vector elements V_p, since $G(\rho,z;\rho',z')$ is evaluated at the testing point $z_{m_p}>0$ while $z'=0$.

C. Surface Temperature Distribution

The numerical procedure described in Sect. V-A yields the Fourier coefficients of the temperature distribution within the mine. The temperature distribution over the soil surface $z=0$ is found by using Eqs. (3) and (65). The procedure offers some challenges. When $z=z'=0$, the exponential factors in Eq. (62) are eliminated, leading to slow convergence. It is shown in Appendix II that an approximate but accurate and efficient form for the Green's function is

$$\begin{aligned} G(\rho,z=0;\rho',z'=0) &= \frac{1}{2\pi}\int_0^{2\pi} d\phi' \frac{\exp(ik_n\rho_t)}{\rho_t} - \alpha K_0(-ik_n\rho_>)I_0(-ik_n\rho_<) \\ &- \frac{\alpha^2}{2\pi}\int_0^{2\pi} d\phi' \frac{\exp(ik_n\rho_t)}{ik_n} - \frac{\alpha^3}{4\pi}\int_0^{2\pi} d\phi' \frac{\rho_t K_{-1}(-ik_n\rho_t)}{ik_n} \end{aligned} \quad (72)$$

The Green's function representation given by Eq. (72) is used to evaluate the third term in Eq. (65) when the observation point is on the soil surface.

VI. Results

In this section numerical simulation results are presented to compare the three-dimensional modeling capabilities of the integral equation solution with those of the FEM based solution. Additional results comparing these methods are given in a companion paper [11], where we considered several cases including a circular TNT-filled anti-tank mine simulant of diameter 20 cm and height 7.5 cm buried 4.5 cm under a perfectly smooth soil surface.

In this work, consider the simple case of circular cylinder of diameter 20 cm, height 7.5 cm, and depth 4.5 cm filled with a good thermal insulator (styrofoam).

An important parameter of this problem is the required number of harmonics in the Fourier expansion given by Eq. (3). The forcing function of this integral equation is the solar insolation function $\mathcal{F}_{sun}(t)$, the Fourier series of which is dominated by its first term. Therefore, we expect the temperature coefficients T_n to be rapidly convergent. Our numerical simulations support this expectation. An example Fourier harmonic distribution is given in Table I. Based on our numerical experiments, we kept the first five Fourier harmonics in the expansion given by Eq. (3).

TABLE I

FOURIER HARMONICS OF THE TEMPERATURE DISTRIBUTION AT A POINT OVER THE MINE

Harmonic #	T_n	$\lvert T_n\rvert/\lvert T_1\rvert$
1	$2.5351 + j3.8669$	1
2	$0.3582 + j1.3392$	0.2998
3	$0.0042 + j0.0945$	0.0205
4	$0.0140 - j0.1289$	0.0280
5	$0.0059 - j0.0248$	0.0055
6	$-0.0058 + j0.0120$	0.0029
7	$-0.0043 + j0.0094$	0.0022

Figure 4 shows the surface temperature over the center of the mine as a function of time. The FEM results are compared with those obtained via the integral equation method. The results show good agreement between the codes. In the second set of results, we investigated spatial variations in the surface temperature, by plotting the temperature along a cut through the center of the mine. We present the temperature at times that correspond to maximum signature contrast. Figure 5 illustrates the spatial distribution results. The FEM and integral equation models show good agreement. The integral equation solution took about 2 hours on a 300 MHz Pentium II machine. Most of this time is used to fill and factor the $\overline{\mathbf{Z}}$ matrix. The FEM solution took about 1 hour 40

minutes in the same machine. No significant attempts have been made to optimize either code.

VII. Conclusions

An integral-equation has been formulated for the temperature over a buried land mine. Using the approximations of a time invariant convection coefficient and air temperature, and linearizing the radiation boundary condition about its average value, the case of periodic (diurnal) solar heating leads to a Fourier decomposition of the time variation. The integral equation for each Fourier component of the temperature distribution was solved using the method of weighted residuals. The resulting model provides a reference solution for a relatively simple geometry, which can be used to check more sophisticated FEM-based models.

Appendix

I. ANALYTICAL TREATMENT OF THE SINGULAR INTEGRAL

In this appendix we evaluate the functions I_c, I_ρ, and I_z which were introduced in Eq. (71). Singularity extraction is required in the evaluation of these functions if the testing point lies in the integration domain. The boundaries of this integral and the location of the singularity are illustrated in Fig. 6.

Consider a small volume around the singularity. Such a volume can be defined by the cylindrical coordinates $\rho \in (\rho - \epsilon_\rho, \rho + \epsilon_\rho)$, $\phi \in (\phi - \epsilon_\phi, \phi + \epsilon_\phi)$, and $z \in (z - \epsilon_z, z + \epsilon_z)$. Figure 7 demonstrates the top and side views of this volume. The functions can be written as

$$\begin{bmatrix} I_c(\rho_1, \rho_2; 0, 2\pi; z_1, z_2) \\ I_\rho(\rho_1, \rho_2; 0, 2\pi; z_1, z_2) \\ I_z(\rho_1, \rho_2; 0, 2\pi; z_1, z_2) \end{bmatrix} = \begin{bmatrix} I_c(\rho_1, \rho_2; 0, 2\pi; z_1, z - \epsilon_z) \\ I_\rho(\rho_1, \rho_2; 0, 2\pi; z_1, z - \epsilon_z) \\ I_z(\rho_1, \rho_2; 0, 2\pi; z_1, z - \epsilon_z) \end{bmatrix}$$

$$+ \begin{bmatrix} I_c(\rho_1, \rho_2; 0, 2\pi; z + \epsilon_z, z_2) \\ I_\rho(\rho_1, \rho_2; 0, 2\pi; z + \epsilon_z, z_2) \\ I_z(\rho_1, \rho_2; 0, 2\pi; z + \epsilon_z, z_2) \end{bmatrix}$$

$$
\begin{aligned}
&+ \begin{bmatrix} I_c(\rho_1, \rho-\epsilon_\rho; 0, 2\pi; z-\epsilon_z, z+\epsilon_z) \\ I_\rho(\rho_1, \rho-\epsilon_\rho; 0, 2\pi; z-\epsilon_z, z+\epsilon_z) \\ I_z(\rho_1, \rho-\epsilon_\rho; 0, 2\pi; z-\epsilon_z, z+\epsilon_z) \end{bmatrix} \\
&+ \begin{bmatrix} I_c(\rho+\epsilon_\rho, \rho_2; 0, 2\pi; z-\epsilon_z, z+\epsilon_z) \\ I_\rho(\rho+\epsilon_\rho, \rho_2; 0, 2\pi; z-\epsilon_z, z+\epsilon_z) \\ I_z(\rho+\epsilon_\rho, \rho_2; 0, 2\pi; z-\epsilon_z, z+\epsilon_z) \end{bmatrix} \\
&+ \begin{bmatrix} I_c(\rho-\epsilon_\rho, \rho+\epsilon_\rho; \epsilon_\phi, 2\pi-\epsilon_\phi; z-\epsilon_z, z+\epsilon_z) \\ I_\rho(\rho-\epsilon_\rho, \rho+\epsilon_\rho; \epsilon_\phi, 2\pi-\epsilon_\phi; z-\epsilon_z, z+\epsilon_z) \\ I_z(\rho-\epsilon_\rho, \rho+\epsilon_\rho; \epsilon_\phi, 2\pi-\epsilon_\phi; z-\epsilon_z, z+\epsilon_z) \end{bmatrix} \\
&+ \begin{bmatrix} I_c^{sing} \\ I_\rho^{sing} \\ I_z^{sing} \end{bmatrix}
\end{aligned} \tag{73}
$$

where I_c^{sing}, I_ρ^{sing}, and I_z^{sing} are integrals of the singular kernels in the small volume. The first five terms on the right-hand side involve no singularities and can be calculated numerically. The limits of the integrals in Eq. (73) can be better visualized with the aid of Fig. 7. The volume of the singular region is calculated as

$$V_{sing} = \int_{z-\epsilon_z}^{z+\epsilon_z} dz' \int_{\rho-\epsilon_\rho}^{\rho+\epsilon_\rho} d\rho' \rho' \int_{-\epsilon_\phi}^{\epsilon_\phi} d\phi' = 8\rho\epsilon_\rho\epsilon_\phi\epsilon_z \tag{74}$$

The singular region can be approximated by a small sphere centered at the singularity with a radius of

$$r_{sing} = \left(\frac{6}{\pi}\rho\epsilon_\rho\epsilon_\phi\epsilon_z\right)^{\frac{1}{3}} \tag{75}$$

which provides an equal volume. Assuming $\epsilon = \epsilon_\rho = \rho\epsilon_\phi = \epsilon_z$, the integral functions I_c^{sing}, I_ρ^{sing}, and I_z^{sing} can be calculated as

$$I_c^{sing} = \int_0^{r_{sing}} r'^2 dr' \int_0^{\pi} d\theta \sin\theta \int_0^{2\pi} d\phi' \frac{1}{r'} = 2\left(6\sqrt{\pi}\right)^{\frac{2}{3}} \epsilon^2 \tag{76}$$

$$I_\rho^{sing} = \int_0^{r_{sing}} r'^2 dr' \int_0^{\pi} d\theta \sin\theta \int_0^{2\pi} d\phi' \frac{\rho + r'\sin\theta}{r'} = 2\epsilon^2\left(\rho\left(6\sqrt{\pi}\right)^{\frac{2}{3}} + \pi\epsilon\right) \tag{77}$$

$$I_z^{sing} = \int_0^{r_{sing}} r'^2 dr' \int_0^{\pi} d\theta \sin\theta \int_0^{2\pi} d\phi' \frac{r'\cos\theta}{r'} = 2\pi\epsilon^3 \tag{78}$$

II. GREEN'S FUNCTION WHEN THE SOURCE AND OBSERVATION POINTS ARE ON THE SOIL SURFACE

When the source and observation points are on the soil surface, i.e., $z = 0$ and $z' = 0$, the Green's function given by Eq. (43) can be simplified as

$$G(\rho, z = 0; \rho', z' = 0) = \int_0^\infty dk_\rho J_0(k_\rho \rho') J_0(k_\rho \rho) \frac{k_\rho}{\sqrt{k_n^2 + k_\rho^2} + \alpha} \tag{79}$$

Without the exponential factors, this integral converges slowly. The following approximation is valid for small α

$$\begin{aligned} \frac{1}{\sqrt{k_n^2 + k_\rho^2} + \alpha} &= \frac{1}{\sqrt{k_n^2 + k_\rho^2}} \frac{1}{1 + \frac{\alpha}{\sqrt{k_n^2 + k_\rho^2}}} \\ &\approx \frac{1}{\sqrt{k_n^2 + k_\rho^2}} \left(1 - \frac{\alpha}{\sqrt{k_n^2 + k_\rho^2}} + \frac{\alpha^2}{k_n^2 + k_\rho^2} - \frac{\alpha^3}{\left(k_n^2 + k_\rho^2\right)^{\frac{3}{2}}} \right) \end{aligned} \tag{80}$$

Substituting Eq. (80) into Eq. (79), employing the multiplication identity for the zeroth order Bessel functions given by Eq. (49), and changing the orders of ϕ and k_ρ integrals we obtain

$$\begin{aligned} G(\rho, z = 0; \rho', z' = 0) &\approx \frac{1}{2\pi} \int_0^{2\pi} d\phi' \int_0^\infty dk_\rho J_0(k_\rho \rho_t) \frac{k_\rho}{\sqrt{k_n^2 + k_\rho^2}} \\ &- \alpha \int_0^\infty dk_\rho J_0(k_\rho \rho') J_0(k_\rho \rho) \frac{k_\rho}{k_n^2 + k_\rho^2} \\ &+ \frac{\alpha^2}{2\pi} \int_0^{2\pi} d\phi' \int_0^\infty dk_\rho J_0(k_\rho \rho_t) \frac{k_\rho}{\left(k_n^2 + k_\rho^2\right)^{\frac{3}{2}}} \\ &- \frac{\alpha^3}{2\pi} \int_0^{2\pi} d\phi' \int_0^\infty dk_\rho J_0(k_\rho \rho_t) \frac{k_\rho}{\left(k_n^2 + k_\rho^2\right)^2} \end{aligned} \tag{81}$$

The k_ρ integrals in Eq. (81) can be evaluated in closed form [30–32], leading to

$$\begin{aligned} G(\rho, z = 0; \rho', z' = 0) &\approx \frac{1}{2\pi} \int_0^{2\pi} d\phi' \frac{\exp(ik_n \rho_t)}{\rho_t} - \alpha K_0(-ik_n \rho_>) I_0(-ik_n \rho_<) \\ &- \frac{\alpha^2}{2\pi} \int_0^{2\pi} d\phi' \frac{\exp(ik_n \rho_t)}{ik_n} - \frac{\alpha^3}{4\pi} \int_0^{2\pi} d\phi' \frac{\rho_t K_{-1}(-ik_n \rho_t)}{ik_n} \end{aligned} \tag{82}$$

where K_0, K_{-1}, and I_0 are modified Bessel functions.

Acknowledgments

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed herein are those of the author and are not necessarily those of Duke University or the ARO.

References

[1] K. Watson, "Geologic application of thermal infrared images," *Proc. IEEE*, vol. 63, no. 1, pp. 128–137, Jan. 1975.

[2] H. S. Carslaw and J. C. Yaeger, *Conduction of Heat in Soils*, Oxford Univ. Press, New York, NY, 1953.

[3] A. W. England, J. F. Galantowicz, and M. S. Schretter, "The radiobrightness thermal inertia measure of soil moisture," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, no. 1, pp. 132–139, Jan. 1992.

[4] A. B. Kahle, "A simple thermal model of the earth's surface for geologic mapping by remote sensing," *Journal of Geophysical Research*, vol. 82, pp. 1673–1680, 1977.

[5] K. Watson, "Periodic heating of a layer over a semi-infinite solid," *Journal of Geophysical Research*, vol. 78, no. 26, pp. 5904–5910, Sept. 1973.

[6] Y.-A. Liou and A. W. England, "Annual temperature and radiobrightness signatures for bare soils," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, no. 4, pp. 981–990, July 1996.

[7] A. W. England, "Radiobrightness of diurnally heated, freezing soil," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, no. 4, pp. 464–476, July 1990.

[8] Y.-A. Liou and A. W. England, "A land-surface process/radiobrightness model with coupled heat and moisture transport in soil," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, no. 1, pp. 273–286, Jan. 1998.

[9] Y.-A. Liou and A. W. England, "A land-surface process/radiobrightness model with coupled heat and moisture transport for freezing soils," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, no. 2, pp. 669–677, March 1998.

[10] I. K. Sendur and B. A. Baertlein, "Thermal analysis of IR signatures for buried land mines," Progress Report to US Army Research Office on Contract 97-SC-ARO-1015, The Ohio State University ElectroScience Laboratory, Columbus, OH, March 1999.

[11] I. K. Sendur and B. A. Baertlein, "Three-dimensional thermal modeling of land mine signatures over a diurnal cycle," *submitted to IEEE Trans. Geo. Rem. Sens.*

[12] I. Stakgold, *Green's Functions and Boundary Value Problems*, Academic Press, 1980.

[13] C. A. Balanis, *Advanced Engineering Electromagnetics*, John Wiley and Sons, 1989.

[14] W. C. Chew, *Waves and Fields in Inhomogeneous Media*, IEEE Press, 1995.

[15] J. A. Kong, *Electromagnetic Wave Theory*, John Wiley and Sons, 1986.

[16] A. Sommerfeld, *Partial Differential Equation in Physics*, Academic Press, 1949.

[17] W. A. Johnson, "Analysis of a vertical, tubular cylinder which penetrates an air-dielectric interface and which is excited by an azimuthally symmetric source," *Radio Science*, vol. 18, no. 6, pp. p. 1273–1281, 1983.

[18] W. C. Kuo and K. K. Mei, "Numerical approximations of the Sommerfeld integral for fast convergence," *Radio Science*, vol. 13, no. 3, pp. p. 407–415, 1978.

[19] J. H. Richmond, "Digital computer solutions of the rigorous equations for scattering problems," *Proc. IEEE*, vol. 53, pp. p. 796–804, 1965.

[20] R. F. Harrington, "Matrix methods for field problems," *Proc. IEEE*, vol. 55, no. 2, pp. p. 136–149, 1967.

[21] R. F. Harrington, *Field Computation by Moment Methods*, IEEE Press, 1993.

[22] E. K. Miller, L. Medgyesi-Mitschang, and E. H. Newman, Eds., *Computational Electromagnetics*, IEEE Press, 1992.

[23] R. C. Hansen, Ed., *Moment Methods in Antennas and Scattering*, Artech, 1990.

[24] D. R. Wilton, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak and C. M. Butler, "Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains," *IEEE Trans. Antennas Propagat.*, vol. 32, no. 2, pp. p. 276–281, 1984.

[25] D. E. Livesay, and K. M. Chen, "Electromagnetic fields induced inside arbitrarily shaped biological bodies," *IEEE Trans. Microwave Theory Tech.*, vol. 22, no. 12, pp. p. 1273–1280, 1974.

[26] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 30, pp. p. 409–418, 1982.

[27] D. H. Schaubert, D. R. Wilton, and A. W. Glisson, "A tetrahedral modeling method for electromagnetic scattering by arbitrarily shaped inhomogeneous dielectric bodies," *IEEE Trans. Antennas Propagat.*, vol. 32, no. 1, pp. p. 77–85, 1984.

[28] T. K. Sarkar, "A study of the various methods for computing electromagnetic field utilizing thin wire integral equations," *Radio Science*, vol. 18, pp. p. 29–38, 1983.

[29] M. G. Andreasen, "Scattering from bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 13, no. 2, pp. p. 303–310, 1965.

[30] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, 1980.

[31] A. Erdelyi, *Tables of Integral Transforms*, vol. 2, Mc-Graw Hill, 1954.

[32] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, 1945.

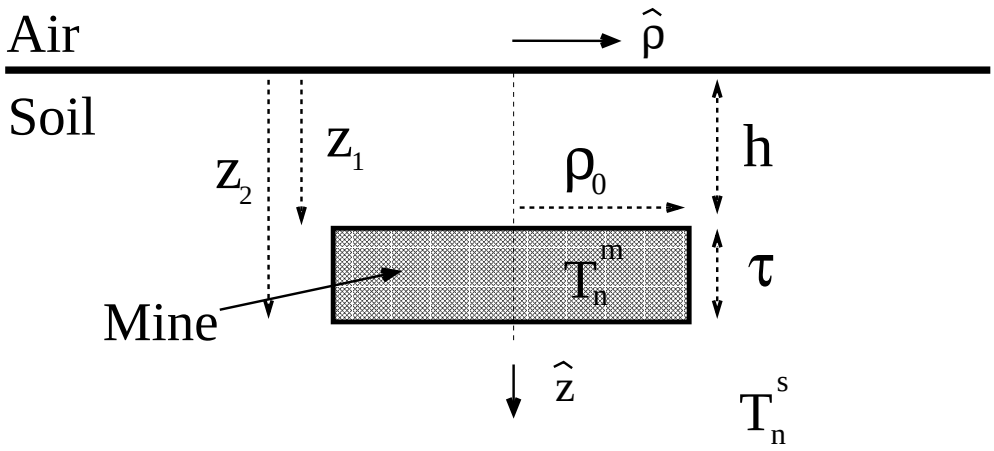


Fig. 1. A cylindrical mine with radius ρ_0 and thickness τ, buried at depth h under a smooth soil surface. The Fourier series coefficients in the mine and soil regions are represented with $T_n^m(\mathbf{r})$ and $T_n^s(\mathbf{r})$, respectively.

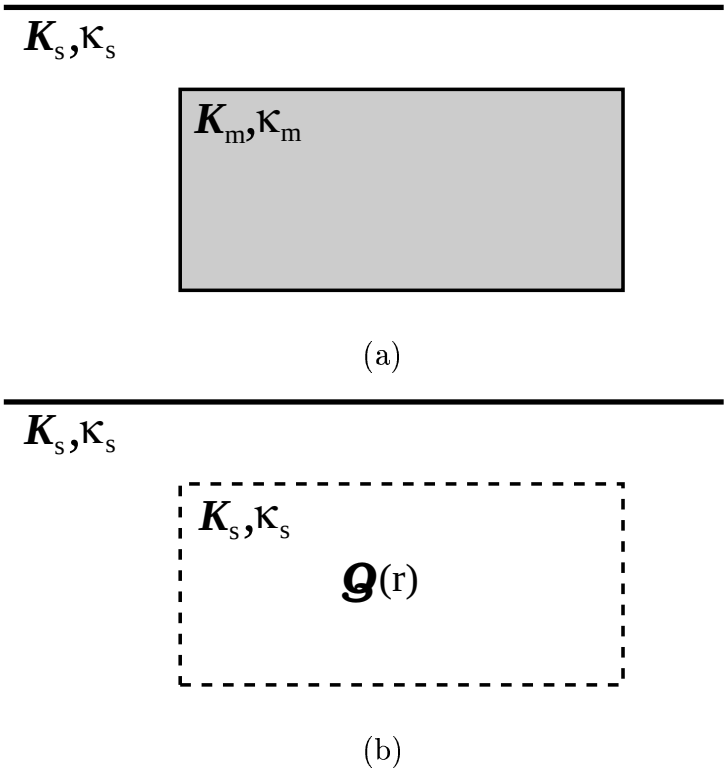


Fig. 2. Volume equivalence for heat transfer equation. (a) The original problem, in which a buried mine with thermal properties $\kappa_m(\mathbf{r})$ and $\mathcal{K}_m(\mathbf{r})$ is present in homogeneous soil with thermal properties $\kappa_s(\mathbf{r}) = \kappa_s$ and $\mathcal{K}_s(\mathbf{r}) = \mathcal{K}_s$. (b) The equivalent problem, in which the mine is replaced by homogeneous soil and an equivalent distributed heat source $\mathcal{Q}(r)$.

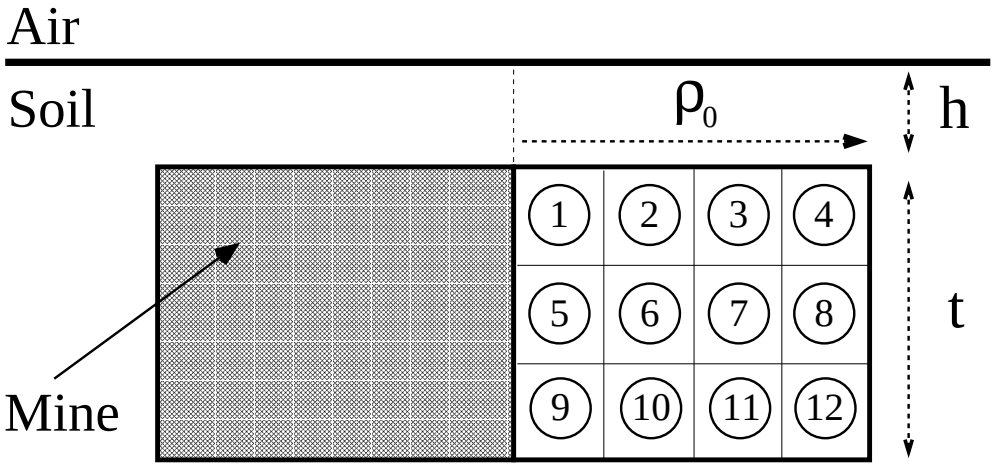


Fig. 3. An example discretization of the buried mine with 4 subdivisions in the $\hat{\rho}$-direction and 3 subdivisions in the $\hat{z}$-direction

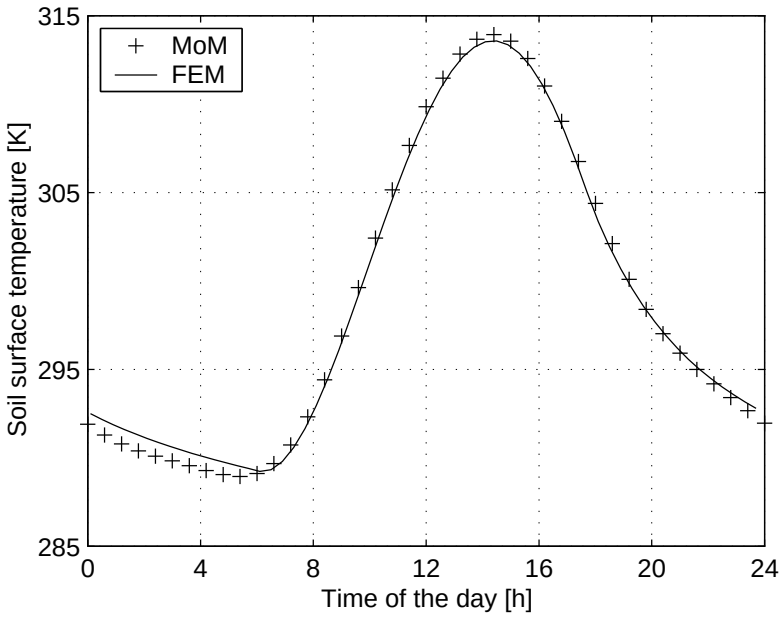
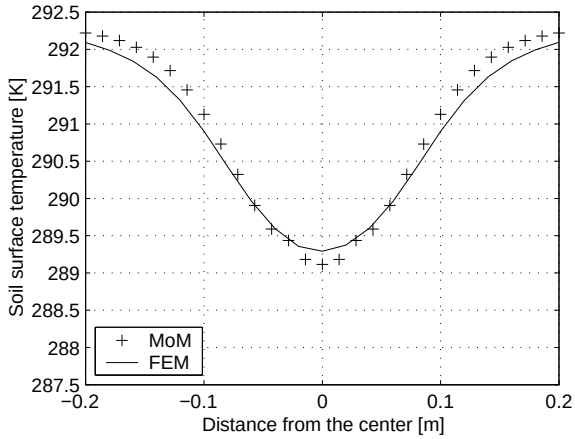
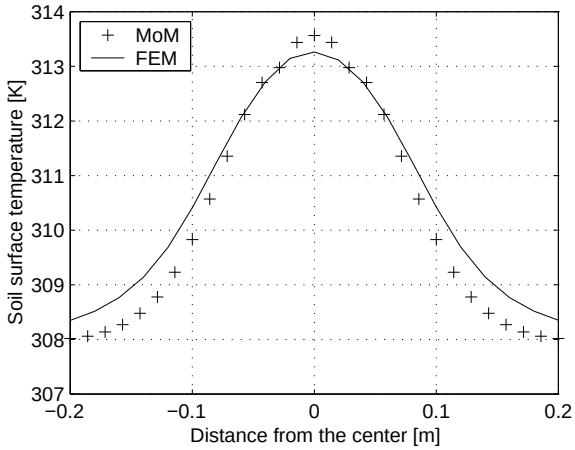


Fig. 4. The surface temperature distribution over the center of a cylindrical mine using the integral equation solution and a finite element solution.



(a)



(c)

Fig. 5. The spatial dependence of the surface temperature distribution: (a) When the mine signature has a maximum negative contrast, (b) When the mine signature has a maximum positive contrast.

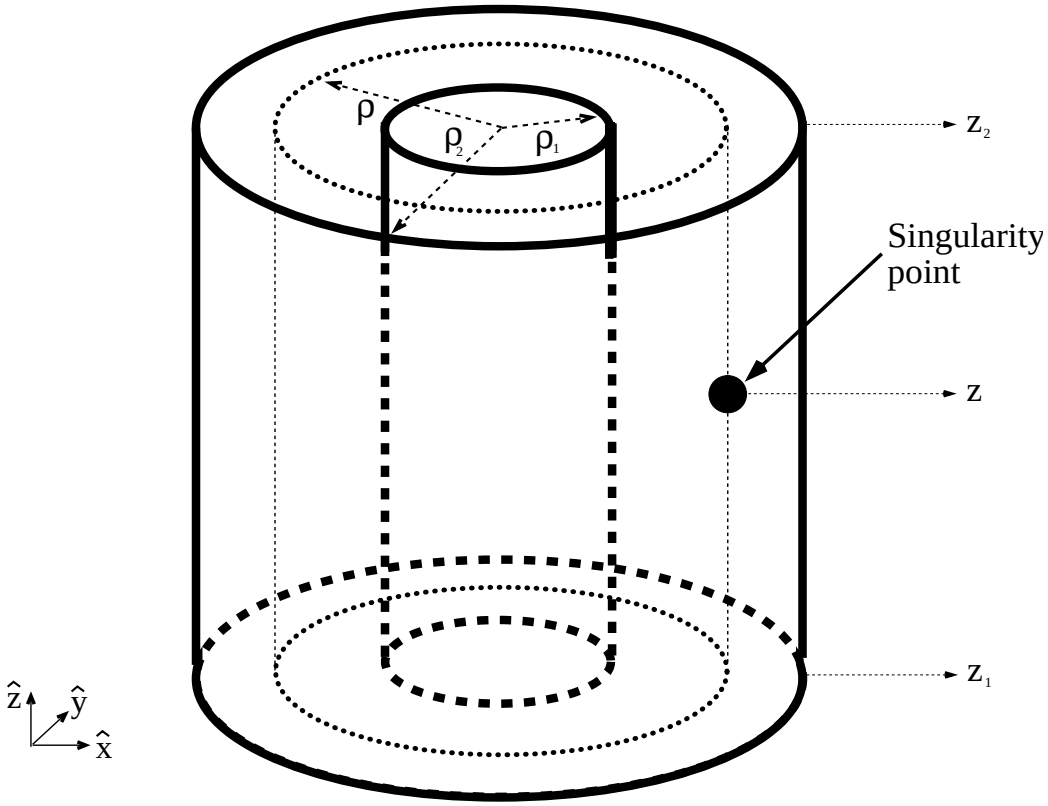


Fig. 6. The singularity point and the boundaries of the singular integral.

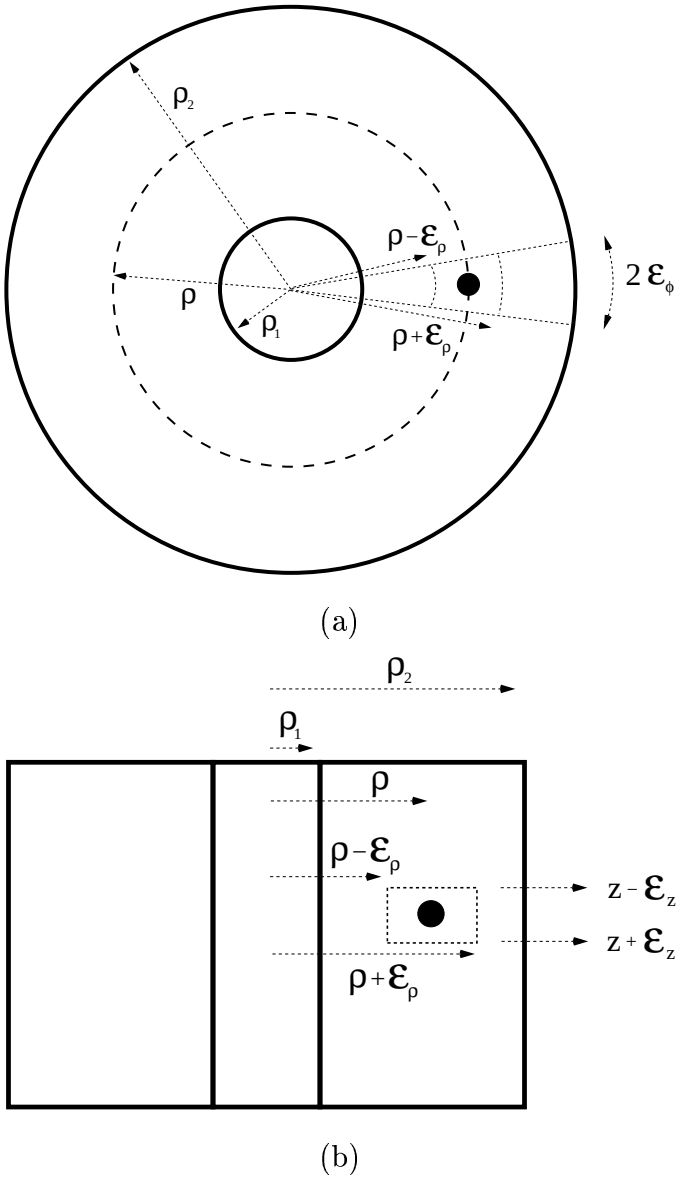


Fig. 7. The small volume around the singularity (a) Top view. (b) Side view.

Feature-Level and Decision-Level Fusion of Noncoincidently Sampled Sensors for Land Mine Detection

Ajith H. Gunatilaka, *Member*, *IEEE*, and Brian A. Baertlein, *Member*, *IEEE*

Abstract—We present and compare methods for feature-level (predetection) and decision-level (postdetection) fusion of multisensor data. This study emphasizes fusion techniques that are suitable for noncommensurate data sampled at noncoincident points. Decision-level fusion is most convenient for such data, but it is suboptimal in principle, since targets not detected by all sensors will not obtain the full benefits of fusion. A novel algorithm for feature-level fusion of noncommensurate, noncoincidently sampled data is described, in which a model is fitted to the sensor data and the model parameters are used as features. Formulations for both feature-level and decision-level fusion are described, along with some practical simplifications. A closed-form expression is available for feature-level fusion of normally distributed data and this expression is used with simulated data to study requirements for sample position accuracy in multisensor data. The performance of feature-level and decision-level fusion algorithms are compared for experimental data acquired by a metal detector, a ground-penetrating radar, and an infrared camera at a challenging test site containing surrogate mines. It is found that fusion of binary decisions does not perform significantly better than the best available sensor. The performance of feature-level fusion is significantly better than the individual sensors, as is decision-level fusion when detection confidence information is also available ("soft-decision" fusion).

Index Terms—Land mines, sensor fusion, infrared, ground penetrating radar, metal detectors.

1 Introduction

Buried land mines are an extremely hazardous fact of life for civilians in many countries. Various estimates put the number of buried landmines worldwide at over 80 million and growing. The extreme risk posed by mines implies that a mine sensor should have essentially no missed detections. This requirement and the clutter-rich environment in which mine sensors operate leads to high false alarm rates. Nonetheless, low false alarm rates are required because of the large land areas where mine clearance is necessary and the cost of clearing a putative mine. Several sensor technologies have been considered for the detection of mines, including electromagnetic induction (EMI), ground penetrating radar (GPR), infra-red (IR) imaging, quadrupole resonance (QR), various chemical detectors (e.g., electronic "sniffers"), and sensors of acoustically induced surface vibrations. At present, none of these technologies has proven to be acceptably effective and reliable for the wide range of mines and environmental conditions encountered in the field.

Because of the diverse physical properties measured by different sensor technologies, multisensor fusion is attractive as a means of improving mine detection. Unfortunately, demining sensor fusion is hindered by several factors. First, sensors used for mine detection typically produce noncommensurate data, i.e., data that are not of comparable form and, hence, are not easily combined. By way of example, three commmonly used mine sensors are IR cameras, which produce surface imagery; ground penetrating radars which produce depth profiles along a line scan; and metal detectors, which produce scalar measurements along a line scan. Even when the data are commensurate, the sensors may operate on different platforms, leading to noncoincident sampling and problems with data registration. Since fusion comprises combining information from multiple sources regarding a specified phenomenon (e.g., the presence of a mine) at a specific location, accomodating noncoincident samples is a necessity.

Some of these concerns can be alleviated by decision-level fusion, in which the operator performs detection at the sensor level and then combines (fuses) detections from individual sensors. This process reduces diverse sensor data to a common format (binary decisions or detection probabilities), which are readily combined using conventional statistical techniques. A variety of techniques have been developed for this approach [1], [2].

In principle, however, decision-level fusion is suboptimal since if a target is not detected by all sensors, it will not experience the full benefits of fusion. This situation and stressing performance requirements, led us to explore other methods of fusion. In this paper, we discuss a feature-level fusion approach to noncoincident sensor sampling. The data are represented by a model with unknown parameters (features) and random additive clutter. Optimization

- *A.H. Gunatilaka is with Lucent Technologies, 6200 E. Broad Street, Columbus, OH 43213.*
- *B.A. Baertlein is with The Ohio State University, ElectroScience Laboratory, 1320 Kinnear Road, Columbus, OH 43212. E-mail: baertlein.1@osu.edu.*

Manuscript received 8 Feb. 2000; revised 28 Nov. 2000; accepted 14 Dec. 2000.
Recommended for acceptance by R. Beveridge.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 111425.

0162-8828/01/$10.00 © 2001 IEEE

techniques are used to determine features from the available data, and classification is performed on the basis of the features. The effects of sample position uncertainty are also examined for the case of normally distributed features.

Fusion of demining sensor data has been investigated previously. Chauduri et al. [3] developed a demining system in which coregistered GPR and EMI sensors were fused using several decision-level fusion approaches. Brusmark et al. [4] demonstrated the decision-level fusion of coincidently sampled GPR and EMI sensor data collected over targets buried in a sand box. Clark et al. [5] fused two IR sensors of different wavelengths, which were registered using a set of fiducial markers. A neural network-based feature-level fusion approach was used in that work. Miao et al. considered the fusion of six coregistered IR images corresponding to different wavelengths [6]. Each image was separately classified with neural networks and the resulting decisions were fused using both majority voting and a consensus theory approach. Weisenseel et al. [7] studied decision-level fusion of GPR and EMI data using synthetic data. The problem was greatly simplified by assuming that the mines were detectable by both sensors, metallic clutter by EMI alone and nonmetallic clutter by GPR alone. Only an AND rule was considered for fusion. Breejen et al. [8] compared several methods for decision-level fusion of EMI, GPR, and IR data. They used the same data for both training and testing the classifiers, which often yields optimistic results. The performance of a vehicle mounted mine detector employing GPR, EMI, and video and IR cameras, developed by EG&G Management Systems, Incorporated, was reported in [9]. While surface emplaced antitank mines and mines with significant metal content were easily detected, buried plastic mines were found to be challenging. Development and testing of a vehicle-mounted multisensor demining platform that uses EMI, GPR, and IR for detection and a TNA sensor for target confirmation is underway at the Canadian Defence Research Establishment [10]. Gunatilaka and Baertlein [11] compared predetection (feature-level) and postdetection (hard and soft decision-level) fusion of EMI, GPR, and IR data. A benefit was observed for feature-level fusion, and fusion of soft decisions was found to be better than fusion of hard decisions. Gunatilaka and Baertlein [12] compared several methods for feature extraction and several fusion strategies. In all cases, feature-level fusion was found to be superior to decision-level fusion. Baertlein [13] recently investigated the problem of constructing fusion algorithms with limited training data. Using approaches based on Statistical Learning Theory for feature-level fusion and joint optimization of local decision thresholds for hard decision fusion, it was shown that algorithm performance could be improved. For the specific training data used in that work, superior performance was noted for hard decision fusion. Marble et al. [14] demonstrated the benefits of using hard decision-level fusion to combine EMI, GPR, and magnetometer data for detection of mines with large metal content. In that work, significant problems with misregistration of noncommensurate sensors were overcome by converting all sensor outputs to a scalar, image-like format and interpolating the result.

The work reported here complements and extends the prior studies in several respects. First, we consider both feature-level and decision-level fusion of experimental data from noncommensurate (EMI, GPR, and IR) sensors. Second, we investigate the benefits of fusion under more realistic conditions by using a test site containing large amounts of clutter and having an irregular surface. Finally, each sensor was used on a different platform, which led us to consider the effects of noncoincident sampling and positioning errors.

The work is organized in three major sections. In Section 2 we describe the theoretical basis for feature-level fusion of noncoincident samples. The basis for decision-level fusion is presented in Section 3. Experimental data and examples of fusion are given in Section 4. For a list of symbols referenced in this paper, see Table 1. Finally, concluding remarks appear in Section 5.

2 Feature-Level Fusion of Randomly Sampled Multisensor Data

2.1 Problem Definition

Suppose that a suite of N_S countermine sensors are used to acquire data in a region. Sensor i acquires J_i data samples $\mathbf{d}_i$ at locations $\mathbf{R}_{ij}$, which we denote $\mathbf{d}_{ij} \equiv \mathbf{d}_i(\mathbf{R}_{ij})$. Fusion is to be performed on the data set $\mathcal{D} = \{\mathbf{d}_i(\mathbf{R}_{ij}), i = 1, \ldots, N_s; j = 1, \ldots, J_i\}$. We assume that different sensors produce samples that differ in number J_i, sample positions $\mathbf{R}_{ij}$, format, and the dimensionality of their data.

Let $\mathbf{R}$ be a point where we wish to determine the presence of a mine. We refer to $\mathbf{R}$ as the "interrogation point." We form the K hypotheses $H_k(\mathbf{R})$, $k = 1, 2, \ldots, K$ for the fused suite regarding the presence or absence of various types of mines at $\mathbf{R}$. We may also include in the set $\{H_k(\mathbf{R})\}$ discrete clutter objects, e.g., buried rocks, metallic clutter, etc. The explicit dependence of these hypotheses on the interrogation point is omitted hereafter in the interest of brevity.

The decision criterion for the fusion processor is the Bayes risk. At each point $\mathbf{R}$, we determine the hypothesis H_k that minimizes the risk $\mathcal{R}$ conditioned on the data $\mathcal{D}$, viz:

$$k = \arg\min_j \mathcal{R}_j(\mathcal{D}) = \arg\min_j \sum_{\ell=1}^{K} C_{j\ell} \Pr(H_\ell|\mathcal{D}), \qquad (1)$$

where $\mathcal{R}_j(\mathcal{D})$ is the risk of selecting hypothesis H_j given data $\mathcal{D}$, and C_{ij} is the cost of choosing H_i when H_j is true. If all costs are equal or for the special case $K = 2$ (binary detection), it is sufficient to compute the likelihood ratio and to compare it to a threshold. In general, however, multiple hypotheses may be involved and equal costs are inappropriate for the mine-detection problem. It is difficult to assign costs in the mine-detection problem but, fortunately, fusion can be explored without knowledge of the costs. Results are reported here in a manner independent of cost, specifically, a posteriori probabilities $\Pr(H_k|\mathcal{D})$ and ROC curves are computed. When suitable costs are known, one can use the general formulation given above to

TABLE 1
List of Symbols

Symbol	Description
N_S	The number of demining sensors
$\mathbf{R}$	The location at which fusion is being used to determine the presence of a mine
H_k	The kth hypothesis regarding the object present at location $\mathbf{R}$
K	The number of possible hypotheses H_k
$\mathbf{R}_{ij}$	The location at which the jth sample of sensor i is acquired
$\mathbf{d}_{ij}$	The data sampled by sensor i at the jth sampling position
$\mathbf{d}_i$	The data acquired by sensor i
J_i	The number of data samples acquired by sensor i
M_i	The number of data samples from sensor i used in fusion at $\mathbf{R}$
N_{D_i}	The number of elements in each sample from sensor i
N_E	The number of elements from all sensors and all samples used in fusion $[= \sum_i M_i N_{D_i}]$
$\mathcal{D}$	The set of data acquired by all sensors at all sampling points
$\mathcal{R}_j(\mathcal{D})$	The risk of choosing hypotheis H_j given data $\mathcal{D}$
C_{jk}	The cost of choosing H_j when H_k is true
$\Pr(H_k\|\mathcal{D})$	The probability that H_k is true given data $\mathcal{D}$
θ_{ik}	The feature vector for sensor i given hypothesis H_k
$\mathbf{R}_{0,ik}$	The nominal center of the mine signature for sensor i under H_k
$\mathbf{\Theta}_{ik}$	$= [\theta_{ik}\, \mathbf{R}_{0,ik}]^T$
$\mathbf{\Theta}_k$	$= [\mathbf{\Theta}_{1k}\, \mathbf{\Theta}_{2k}\, \ldots\, \mathbf{\Theta}_{N_s k}]^T$
$\mathbf{g}_i(\mathbf{R}_{ij}; \mathbf{\Theta}_{ik})$	Noise-free (and clutter-free) signal acquired by sensor i at $\mathbf{R}_{ij}$ when H_k is true
$\mathbf{n}_i(\mathbf{R}_{ij})$	The contribution of noise and clutter to the signal measured by sensor i at $\mathbf{R}_{ij}$
$\mathbf{D}$	A vector comprising the subset of $\mathcal{D}$ relevant to making the target decision at $\mathbf{R}$
$\mathbf{N}$	The noise and clutter vector corresponding to $\mathbf{D}$
$\mathbf{G}_k(\mathbf{\Theta}_k)$	The signal vector corresponding to $\mathbf{D}$
h_{iq}	The q-th hypothesis of sensor i
Q_i	The number of distinct hypotheses that can be reported by sensor i
$\mathbf{H}$	$= [h_1\, h_2\, \ldots\, h_{N_S}]^T$
t_i	The decision threshold used by sensor i
$\mathbf{T}$	$= [t_1\, t_2\, \ldots\, t_{N_S}]^T$
u_i	The decision reached by sensor i in response to the data
u_0	The decision resulting from decision-level fusion of $\mathbf{U}$
$\mathbf{U}$	$= [u_1\, u_2\, \ldots\, u_{N_S}]^T$
$P_{F_i}(t_i), P_{M_i}(t_i)$	The probabilities of false and missed detection respectively for sensor i
P_D, P_F, P_M	The respective probabilities of detection, false detection and missed detection after fusion
$\Lambda(\mathbf{D})$	The likelihood ratio
$\Lambda_G(\mathbf{D})$	The likelihood ratio for the generalized likelihood ratio test

derive the minimum risk decision from the a posteriori probabilities. Similarly, knowledge of the costs permits one to readily determine minimum risk operating points on a ROC curve.

2.2 Fusion of Noncoincidently Sampled Features

We assume that sensor i acquires at a point $\mathbf{R}_{ij}$ data $\mathbf{d}_i(\mathbf{R}_{ij})$, which lies in the vicinity of a putative target of type k at $\mathbf{R}$. This data is modeled as a signal $\mathbf{g}_i(\mathbf{R}_{ij}; \mathbf{\Theta}_{ik})$ corrupted by additive clutter (and noise) $\mathbf{n}_i(\mathbf{R}_{ij})$, viz:

$$\mathbf{d}_i(\mathbf{R}_{ij})|H_k = \mathbf{g}_i(\mathbf{R}_{ij}; \mathbf{\Theta}_{ik}) + \mathbf{n}_i(\mathbf{R}_{ij}), \tag{2}$$

where $\mathbf{\Theta}_{ik}$ is a feature vector (i.e., a set of model parameters) that describes the signature of targets of type k for sensor i (see Fig. 1). We assume that the clutter $\mathbf{n}_i(\mathbf{R})$ is independent of the parameter $\mathbf{\Theta}_{ik}$. Note that, if the clutter discretes are also assigned specific hypotheses H_j, the phenomena modeled by $\mathbf{n}_i(\mathbf{R})$ can be relatively smooth.

For each sensor, the presence of a mine produces a measurable response over a finite spatial region and, hence, only a subset of $\mathcal{D}$ may be relevant to interrogation of the region around $\mathbf{R}$. Let the relevant data for sensor i be given by subsequences $j_1, j_2, \ldots, j_{M_i}$ with length M_i. It is convenient to assemble this reduced data set into the vectors

$$\mathbf{D} = \Big[\mathbf{d}_1(\mathbf{R}_{1j_1}) \cdots \mathbf{d}_1(\mathbf{R}_{1j_{M_1}}) \; \mathbf{d}_2(\mathbf{R}_{2j_1}) \cdots \mathbf{d}_{N_s}(\mathbf{R}_{N_s j_{M_{N_s}}})\Big]^T \tag{3}$$

$$\mathbf{N} = \Big[\mathbf{n}_1(\mathbf{R}_{1j_1}) \cdots \mathbf{n}_1(\mathbf{R}_{1j_{M_1}}) \; \mathbf{n}_2(\mathbf{R}_{2j_1}) \cdots \mathbf{n}_{N_s}(\mathbf{R}_{N_s j_{M_{N_s}}})\Big]^T \tag{4}$$

$$\mathbf{G}_k(\mathbf{\Theta}_k) = \Big[\mathbf{g}_{1k}(\mathbf{R}_{1j_1}; \mathbf{\Theta}_{1k}) \cdots \mathbf{g}_{N_s k}(\mathbf{R}_{N_s j_{M_{N_s}}}; \mathbf{\Theta}_{N_s k})\Big]^T. \tag{5}$$

If a single datum $\mathbf{d}_i(\mathbf{R})$ from sensor i has length N_{D_i}, then the length of these column vectors is $N_E = \sum_i N_{D_i} M_i$. Using this notation, we have for the combined signal model $\mathbf{D}|H_k = \mathbf{G}_k(\mathbf{\Theta}_k) + \mathbf{N}$, where the combined (concatenated) feature vector for all sensors is

$$\mathbf{\Theta}_k = [\Theta_{1k}\, \Theta_{2k} \cdots \Theta_{N_s k}]^T. \tag{6}$$

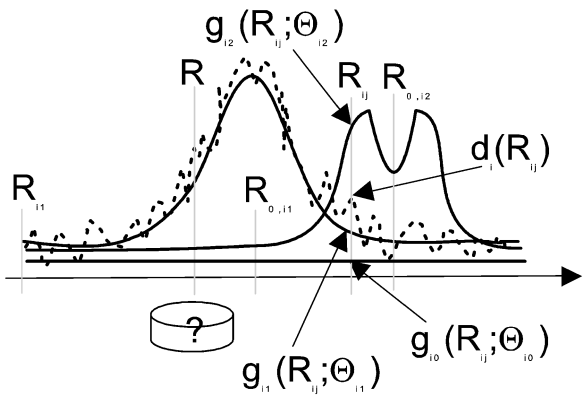


Fig. 1. The data model for three hypotheses $k = 0, 1, 2$. For each sensor i data $\mathbf{d}_i(\mathbf{R}_{ij})$ are available. For hypothesis k, the sensor data is represented by a known function $\mathbf{g}_{ik}(\mathbf{R}; \Theta_{ik}))$ of position $\mathbf{R}$ and random parameters Θ_{ik}. The reference position of signature $\mathbf{g}_{ik}$ is given by $\mathbf{R}_{0,ik}$. The signature parameters Θ_{ik}, which include the reference position $\mathbf{R}_{0,ik}$ are estimated from the data. In this example, $k = 0$ corresponds to the no-target case, $k = 1$ is a target with a unimodal signature, and $k = 2$ is a target with a bimodal signature. The objective of fusion is to determine if a target exists at any point $\mathbf{R}$ on the basis of data from multiple sensors.

To support both decision-level fusion and feature-level fusion with uncertain sample locations, we include in $\mathbf{\Theta}_{ik}$ a position offset $\mathbf{R}_{0,ik}$ that describes the nominal "center" of the mine signature with respect to $\mathbf{R}$ and we write $\mathbf{\Theta}_{ik} = \left[\theta_{ik}\ \mathbf{R}_{0,ik}\right]^T$. In general, we take θ_{ik} and $\mathbf{R}_{0,ik}$ to be independent in what follows.

2.3 Approximate Forms

Feature-level fusion of the foregoing data can be formulated along two lines. Consider first the direct approach using the a posteriori probabilities, which avoids the problem of estimating the a priori probabilities. Minimizing the risk in (1) leads us to consider the a posteriori probabilities as an integral over the classifier $\Pr(H_k|\mathbf{\Theta}_k)$

$$\Pr(H_k|\mathbf{D}) = \int d\mathbf{\Theta}_k \Pr(H_k|\mathbf{\Theta}_k) f_{\mathbf{\Theta}|\mathbf{D}}(\mathbf{\Theta}_k), \tag{7}$$

in which we have assumed that H_k is determined by Θ_k alone and, hence, $\Pr(H_k|\Theta_k, \mathbf{D}) = \Pr(H_k|\Theta_k)$. This expression is exact, but it requires knowledge of the density $f_{\mathbf{\Theta}|\mathbf{D}}$ and that we perform an integral over a feature space of possibly high dimensions. To avoid these issues, it is attractive to employ the maximum a posteriori (MAP) approximation. If the data $\mathbf{D}$ strongly imply the value $\hat{\mathbf{\Theta}}$ and the features $\mathbf{\Theta}_k$ are not too specific (i.e., the maximum of $\Pr(H_k|\Theta_k)$ is relatively broad), then $f_{\mathbf{\Theta}|\mathbf{D}}$ will be strongly peaked about $\hat{\mathbf{\Theta}}$ and we have

$$\Pr(H_k|\mathbf{D}) \approx \Pr(H_k|\hat{\mathbf{\Theta}}) \int d\mathbf{\Theta}_k f_{\mathbf{\Theta}_k|\mathbf{D}}(\mathbf{\Theta}_k) = \Pr(H_k|\hat{\mathbf{\Theta}}). \tag{8}$$

Conversely, if $\mathbf{D}$ provides no information about $\mathbf{\Theta}$, then $f_{\mathbf{\Theta}|\mathbf{D}}(\mathbf{\Theta}) = f_{\mathbf{\Theta}}(\mathbf{\Theta})$ and $\Pr(H_k|\mathbf{D})$ reduces to the a priori probability.

Feature-level fusion can also be formulated using class-conditioned densities, and this approach will be used in our study of sensor position errors. Bayes' rule leads to

$$\mathcal{R}_k(\mathbf{D}) = \frac{1}{f_{\mathbf{D}}} \sum_{j=1}^{K} C_{jk} \Pr(H_j) f_{\mathbf{D}|H_j}(\mathbf{D}). \tag{9}$$

The a priori probabilities $\Pr(H_j)$ are required but, since those probabilities are seldom known, this approach is somewhat less attractive in practice. The factor $[f_{\mathbf{D}}]^{-1}$ is common to all $\mathcal{R}_k(\mathbf{D})$ and, hence, our primary task is to evaluate the class-conditioned probabilities $f_{\mathbf{D}|H_j}(\mathbf{D})$. Since $\mathbf{G}_k$ is presumed known, the data $\mathbf{D}$ conditioned on the parameter $\mathbf{\Theta}_k$ has the same density as the clutter $\mathbf{N}$, which leads to $f_{\mathbf{D}|\mathbf{\Theta}_k,H_k}(\mathbf{D}) = f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta}_k))$ and

$$f_{\mathbf{D}|H_k}(\mathbf{D}) = \int d\mathbf{\Theta}_k f_{\mathbf{\Theta}_k|H_k}(\mathbf{\Theta}_k) f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta}_k)). \tag{10}$$

We approximate this exact result as done in (8). Assuming that $f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}(\mathbf{\Theta}_k))$ has a well-defined maximum for $\mathbf{\Theta}_k = \hat{\mathbf{\Theta}}_k$ and that $f_{\mathbf{\Theta}_k|H_k}$ is not strongly peaked leads to

$$\begin{aligned} f_{\mathbf{D}|H_k}(\mathbf{D}) &\approx f_{\mathbf{\Theta}_k|H_k}(\hat{\mathbf{\Theta}}_k) \int d\mathbf{\Theta}_k f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta}_k)) \\ &\approx E_k f_{\theta_k|H_k}(\hat{\theta}_k) f_{\mathbf{R}o|H_k}(\hat{\mathbf{R}}_{0,k}), \end{aligned} \tag{11}$$

where we have approximated θ_{ik} and $\mathbf{R}_{0,ik}$ as conditionally independent, and the constant factor

$$E_k = \int d\mathbf{\Theta}_k f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta}_k)) \tag{12}$$

is related to the residual error.

For the two class problem ($K = 2$), a likelihood ratio formulation is possible. The likelihood ratio for a known signal with unknown parameters is given by

$$\Lambda(\mathbf{D}) = \frac{f_{\mathbf{D}|H_1}(\mathbf{D})}{f_{\mathbf{D}|H_0}(\mathbf{D})} = \frac{\int d\mathbf{\Theta} f_{\mathbf{\Theta}|H_1}(\mathbf{\Theta}) f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_1(\mathbf{\Theta}))}{\int d\mathbf{\Theta} f_{\mathbf{\Theta}|H_0}(\mathbf{\Theta}) f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_0(\mathbf{\Theta}))}. \tag{13}$$

If the same model is used for both hypotheses, then approximating the results as proposed above yields

$$\Lambda(\mathbf{D}) \approx \frac{f_{\mathbf{\Theta}|H_1}(\hat{\mathbf{\Theta}}_1) \int d\mathbf{\Theta} f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_1(\mathbf{\Theta}))}{f_{\mathbf{\Theta}|H_0}(\hat{\mathbf{\Theta}}_0) \int d\mathbf{\Theta} f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_0(\mathbf{\Theta}))} = \frac{f_{\mathbf{\Theta}|H_1}(\hat{\mathbf{\Theta}}_1)}{f_{\mathbf{\Theta}|H_0}(\hat{\mathbf{\Theta}}_1)} \frac{E_1}{E_0}, \tag{14}$$

where $\hat{\mathbf{\Theta}}_k$ is the estimate of $\mathbf{\Theta}$ that maximizes $f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_k(\mathbf{\Theta}_k))$ under H_k. Detection is performed on the basis of the class-conditioned densities of $\mathbf{\Theta}$.

It is interesting to compare this formulation to the more conventional generalized likelihood ratio test (GLRT) [15, p. 38]. In the GLRT formulation, the model parameters $\mathbf{\Theta}$ are estimated using the maximum likelihood technique. This leads to the expression

$$\Lambda_G(\mathbf{D}) \approx \frac{\max_{\mathbf{\Theta}_1} f_{\mathbf{D}|\mathbf{\Theta}_1,H_1}}{\max_{\mathbf{\Theta}_0} f_{\mathbf{D}|\mathbf{\Theta}_0,H_0}} = \frac{f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_1(\hat{\mathbf{\Theta}}_1))}{f_{\mathbf{N}}(\mathbf{D} - \mathbf{G}_0(\hat{\mathbf{\Theta}}_0))}. \tag{15}$$

Thus, GLRT detection is based primarily on the distribution of the model error. The model parameters are considered indirectly, via their contribution to the error. The estimator-correlator method [15, p. 79] uses a similar approach, in which the parameters are estimated and the resulting model signature is correlated with the data. For Gaussian statistics,

detection is based on the correlation. The feature-level fusion algorithm described above could be referred to as an "estimator-classifier" technique in that detection (or classification) is based on the parameters rather than the residual error in the model. We will also find it useful to include in the parameter vector Θ_k, a factor related to the integrated residual error density E_k, typically, in the form of the error energy.

2.4 Gaussian Approximation

The case of Gaussian statistics is of interest because it provides insight into the problems of feature-level fusion and because it can be used to efficiently simulate the performance of certain fusion strategies. With certain assumptions, the expression for $f_{\mathbf{D}|H_k}$ in (10) can be evaluated in closed form.

Consider first, the case in which only the clutter $\mathbf{N}$ has a Gaussian density. That is, the N_s random variables $\mathbf{N}_i(\mathbf{R}_{ij})$ (each of which are evaluated at M_i locations) have a density that is jointly normal with mean $\mu_{\mathbf{N}}$ and covariance matrix $\mathbf{C}_N$.

Now, make the additional assumption that the mine signature shape matrix $\mathbf{G}_k$ is known (including the underlying mine positions $\mathbf{R}_{0,ik}$) up to an unknown signature amplitude. Taking θ_k to be this amplitude, we make the linear approximation $\mathbf{G}_k(\mathbf{\Theta}_k) = \tilde{\mathbf{G}}_k\theta_k$ in which $\tilde{\mathbf{G}}_k$ is a known constant vector independent of θ_k. If the features θ_k are also Gaussian distributed with density $\theta_k|H_k \sim \mathcal{N}(\mu_{\theta_k}, \mathbf{C}_{\theta_k})$, we obtain the Bayesian linear model [16, p. 326], which leads to

$$f_{\mathbf{D}}(\mathbf{D}|H_k) = \frac{(2\pi)^{-N_{\mathrm{E}}/2} e^{-Q_k/2}}{\det(\mathbf{C}_N)^{1/2}\det(\mathbf{C}_{\theta_k})^{1/2}\det(\mathbf{C}_{\hat{\theta}_k}^{-1} + \mathbf{C}_{\theta_k}^{-1})^{1/2}}, \quad (16)$$

where N_E is the length of these vectors (defined previously) and

$$\begin{aligned} Q_k = &\left(\mathbf{D} - \tilde{\mathbf{G}}_k\hat{\theta}_k - \mu_N\right)^T \mathbf{C}_N^{-1}\left(\mathbf{D} - \tilde{\mathbf{G}}_k\hat{\theta}_k - \mu_N\right) + \\ &\left(\hat{\theta}_k - \mu_{\theta_k}\right)^T \left(\mathbf{C}_{\hat{\theta}} + \mathbf{C}_{\theta_k}\right)^{-1}\left(\hat{\theta}_k - \mu_{\theta_k}\right). \end{aligned} \quad (17)$$

In this result, the quantity $\hat{\theta}_k$ is the maximum-likelihood estimate (MLE) for the feature vector θ_k, given by

$$\hat{\theta}_k = \mathbf{C}_{\hat{\theta}_k}\tilde{\mathbf{G}}_k^T\mathbf{C}_N^{-1}(\mathbf{D} - \mu_N) \quad (18)$$

and $\mathbf{C}_{\hat{\theta}_k}^{-1} = \tilde{\mathbf{G}}_k^T\mathbf{C}_N^{-1}\tilde{\mathbf{G}}_k$ is its covariance. The terms of (17), we interpret as errors due to random clutter and to feature variability, respectively. Since the first term in (17) is identical to that arising in $f_{\mathbf{D}|\hat{\mathbf{\Theta}},H_k}$, it is clear that the GLRT approximation of (15) ignores the second term in (17), which describes the feature distribution.

2.5 Sample Position Errors

We noted earlier that to fuse multisensor data, it is necessary to know what locations in the environment are represented by the data, so that one can make a statement about the presence of a target the "interrogation" point. In practice, errors in sample position are largely unavoidable, especially when the sensors are located on different platforms. Some common types of sample errors include uniform offsets produced by a baseline error, linearly increasing offsets produced by a drift in position baseline, isotropically distributed errors due to sample-to-sample positioning inaccuracy, and combinations of the above.

When position errors are present, the sample position $\mathbf{R}_{ij}$ has the form $\hat{\mathbf{R}}_{ij} + \mathbf{r}_{ij}$, where $\hat{\mathbf{R}}_{ij}$ is the intended position and $\mathbf{r}_{ij}$ is an unknown error. Thus,

$$\mathbf{d}_i(\mathbf{R}_{ij})|H_k = \mathbf{g}_k(\hat{\mathbf{R}}_{ij} + \mathbf{r}_{ij}; \mathbf{\Theta}) + \mathbf{n}(\hat{\mathbf{R}}_{ij} + \mathbf{r}_{ij}).$$

Using a Bayesian approach and making the additional assumption that the sample positions are independent of the mine signature $\mathbf{\Theta}_k$ and of the hypothesis H_k, these random variables are integrated out as follows:

$$\begin{aligned} &f_{\mathbf{D}|H_k}(\mathbf{D}) = \\ &\int d\mathbf{r}_{1j_1}\int d\mathbf{r}_{1j_2}\cdots\int d\mathbf{r}_{N_s j_{M_{N_s}}} f_{\mathbf{r}_{1j_1},\mathbf{r}_{1j_2},\ldots,\mathbf{r}_{N_s j_{M_{N_s}}}}(\mathbf{r}_{1j_1}, \mathbf{r}_{1j_2}, \ldots, \mathbf{r}_{N_s j_{M_{N_s}}}) \\ &\int d\mathbf{\Theta}_k f_{\mathbf{\Theta}_k|H_k}(\mathbf{\Theta}_k) f_{\mathbf{N}_{1j_1},\ldots,\mathbf{N}_{N_s j_{M_{N_s}}}}\left[\mathbf{d}_{1j_1} - \mathbf{g}(\hat{\mathbf{R}}_{1j_1} + \mathbf{r}_{1j_1}; \mathbf{\Theta}_{1k}), \ldots,\right. \\ &\qquad \left.\mathbf{d}_{N_s j_{M_{N_s}}} - \mathbf{g}(\hat{\mathbf{R}}_{N_s j_{M_{N_s}}} + \mathbf{r}_{N_s j_{M_{N_s}}}; \mathbf{\Theta}_{N_s k})\right]. \end{aligned} \quad (19)$$

The effect of position errors on system performance for feature-level fusion was studied using this result and the Bayesian linear model described in (16). If modest numbers of sensor data points are being fused at each interrogation point, the resulting integrals can be numerically evaluated using Monte Carlo techniques. Gaussian-distributed features were simulated for three sensors with isotropically distributed errors around an intended sample point. The position errors were normally distributed with variance σ^2. The sensor signature models $\mathbf{g}_{ik}$ were taken to be zero when the mine was absent and Gaussian functions with variance Δ_i^2 when the mine was present. A single feature was simulated for each sensor and a common value $\Delta_i \equiv \Delta$ was used for all sensors. The "feature-to-noise" ratio $\sqrt{\mathrm{var}(\theta_{ik})/\mathrm{var}(N_{ik})}$ (an analog of the signal-to-noise ratio for this problem) was set to two.

It was found that system performance in the presence of isotropic position errors could be roughly characterized by the ratio σ/Δ. Figs. 2, 3, and 4 show the results for $\sigma/\Delta = 0.5$, 1, and 1.5, respectively. (The case of a random offset was also examined. The results were similar and are not shown.) We see that for $\sigma/\Delta < 1$ the performance of the classifier is not significantly affected by position offsets, but for $\sigma/\Delta > 1$ there is a significant degradation in performance. These results suggest the intuitively reasonable finding that the performance of feature-level fusion systems will not be adversely effected if the positioning errors are smaller than the extent of the measurable target signature. When this condition is violated, fusion becomes less effective since all sensors may not see the target.

It is worth noting that position-related errors can sometimes be overcome through ingenuity on the part of the analyst and, hence, these results should be viewed as suggestive only. For example, when the position errors comprise constant offsets and the individual sensor performance is good (i.e., high probability of detection

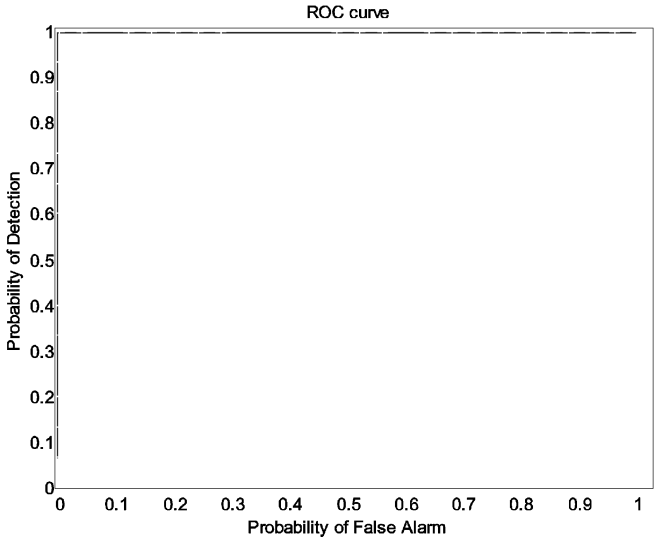


Fig. 2. The ROC curve for simulated data with normalized position errors $\sigma/\Delta = 0.5$. Performance is essentially perfect.

and low false alarm rates), the offset might be estimated by registering the detections with those of other sensors. Such conditions, however, are uncommon in mine detection.

3 Decision-Level Fusion

The formulation of an optimal decision-level fusion algorithm is similar to that defined in Section 2 for feature-level fusion. As before, each sensor i collects data $\{\mathbf{d}_i(\mathbf{R}_{i1}), \ldots, \mathbf{d}_i(\mathbf{R}_{iM_i})\}$. On the basis of these data, it makes an intermediate declaration u_i regarding the truth of intermediate sensor-specific hypotheses $h_{iq}(\mathbf{R})$ for $q = 1, 2, \ldots, Q_i$. For example, a sensor suite that comprises EMI, GPR, and IR sensors might use $Q_i = 2$ for all i with $h_{11/12} =$ "metallic object present/not present," $h_{21/22} =$ "dielectric discontinuity present/not present," and $h_{31/32} =$ "thermal anomaly present/not present" rather

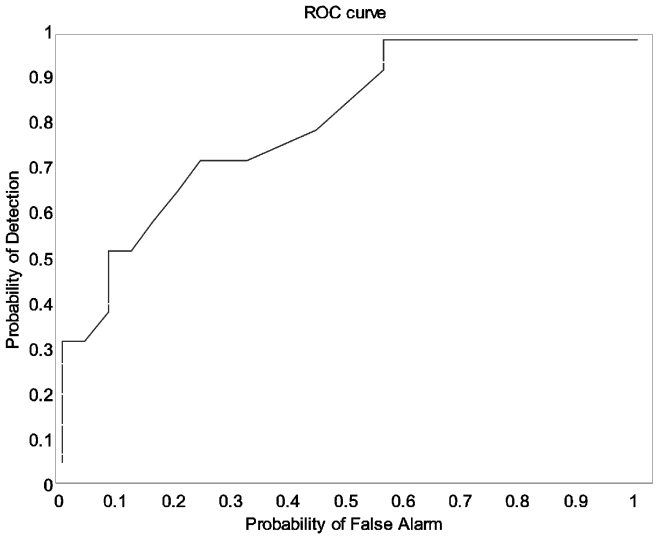


Fig. 3. The ROC curve for simulated data with normalized position errors $\sigma/\Delta = 1$. Performance degrades as a result of sensor position errors.

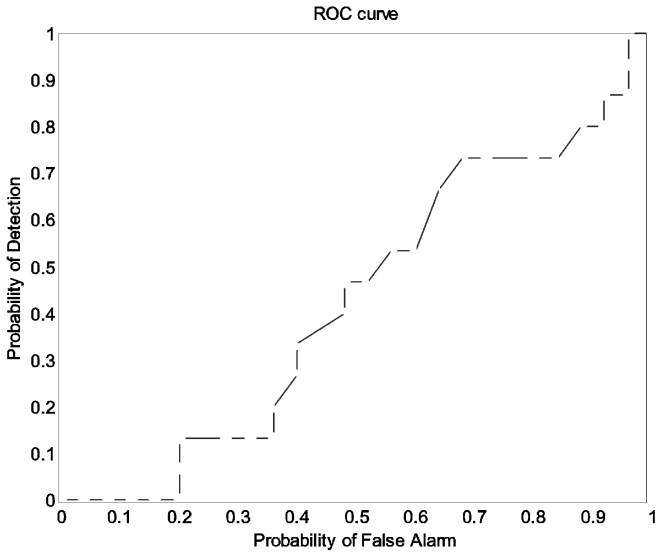


Fig. 4. The ROC curve for simulated data with normalized position errors $\sigma/\Delta = 1.5$. Sensor position errors are so large that fusion is ineffective.

than simply $H_{1/2} =$ "mine present/not present." In what follows, we restrict our attention to the case of equal numbers of intermediate hypotheses $Q_i \equiv Q$.

3.1 General Formulation

The goal of fusion is to combine these declarations into a fused declaration u_0 that addresses hypotheses H_k, $k = 1, 2, \ldots, K$. Each detection u_i also involves a threshold t_i, which must be specified by the operator. To simplify the notation, form the vector of individual declarations given by $\mathbf{U} \equiv [u_1\ u_2\ \ldots\ u_{N_s}]^T$ and the threshold vector $\mathbf{T} \equiv [t_1\ t_2\ \ldots\ t_{N_s}]^T$. The declarations $\mathbf{U}$ and thresholds $\mathbf{T}$ are assumed to be the only data on which the decision u_0 is based. In particular, it is assumed that the fusion processor has no knowledge of the sensor data $\mathbf{D}$. It is further assumed that, for each sensor i, the detection confidences $\Pr(u_i|h_{iq})$ and the a priori probabilities $\Pr(H_k)$ are known for all k.

For optimal decision-level fusion, we seek a rule for fusing $\mathbf{U}$ into u_0 such that the Bayes risk is minimized. We can write this risk in the form

$$\mathcal{R} = \sum_{\mathbf{U}} \left[\sum_{j=1}^{K} \sum_{k=1}^{K} C_{jk} \Pr(H_k) \Pr(u_0 = j|\mathbf{U}) \Pr(\mathbf{U}|H_k; \mathbf{T}) \right], \tag{20}$$

which makes explicit the fusion rule $\Pr(u_0 = j|\mathbf{U})$. In this result, the summation on $\mathbf{U}$ is over all possible combinations of sensor outcomes and the decision u_0 is based only on the local decisions u_i. The individual decisions u_i address the hypotheses h_{iq}, which are related to H_k in a manner that implies the following:

$$\Pr(\mathbf{U}|H_k; \mathbf{T}) = \sum_{q=1}^{Q} \Pr(\mathbf{U}|\mathbf{H}_q; \mathbf{T}) \Pr(\mathbf{H}_q|H_k), \tag{21}$$

where $\mathbf{H}_q = \left[h_{1q}\ h_{2q}\ \ldots\ h_{N_s q}\right]$ and we have used the fact that decision u_i depends on h_{iq} and not on H_k. It remains to

specify the fusion rule $\Pr(u_0 = j|\mathbf{U})$ that will minimize the bracketed quantity in (20). Some approaches are discussed in the following sections.

3.2 An Optimal Hard-Decision Fusion Rule

For sensors that produce "hard" decisions (i.e., all-or-nothing declarations), the fusion rule is a mapping from the N_s-fold product of the space $\{1, 2, \ldots, Q\}$ to one of K output hypotheses. There exist $K^{Q^{N_s}}$ such mappings and, from this (possibly large) set, one must determine the most effective rule. In general, the most effective mapping will depend on the effectiveness of the individual sensors. Ad hoc mappings (AND, OR, majority voting, etc.) have been used, some with considerable success.

In some situations, there exist optimal mappings for decision-level fusion [1, §§ 3.3, 3.4]. Consider the simplest case in which the local decision rules (i.e., the mappings from $\mathbf{D}$ to the local decisions u_i) are fixed and the sensors output hard decisions. For $K = Q = 2$, the minimum in (20) reduces to a well-known likelihood ratio test, given by

$$\frac{\Pr(\mathbf{U}|H_1;\mathbf{T})}{\Pr(\mathbf{U}|H_0;\mathbf{T})} \underset{u_0=0}{\overset{u_0=1}{\gtrless}} \eta, \tag{22}$$

where η is the threshold for the fused decision and is independent of $\mathbf{T}$. If the declarations u_i are conditionally independent for each sensor,[1] then we can factor the joint probabilites into a product. Defining miss and false alarm probabilities for each sensor as follows:

$$P_{Mi}(t_i) \equiv \Pr(u_i = 0|H_1;t_i) \tag{23}$$

$$P_{Fi}(t_i) \equiv \Pr(u_i = 1|H_0;t_i) \tag{24}$$

and taking the logarithm of the likelihood-ratio test, we can write this optimum test as

$$\sum_{i=1}^{N_s}\left[u_i \log\left(\frac{1-P_{Mi}(t_i)}{P_{Fi}(t_i)}\right) + (1-u_i)\log\left(\frac{P_{Mi}(t_i)}{1-P_{Fi}(t_i)}\right)\right] \underset{u_0=0}{\overset{u_0=1}{\gtrless}} \eta. \tag{25}$$

It is easy to see that the logarithms in this expression have a large magnitude for reliable sensors and, hence, the optimal decision fusion rule makes fused decisions that are dominated by the reliable decisions.

Since the Bayes risk can also be written in terms of the a posteriori probabilities $\Pr(H_k|\mathbf{U})$, we can use another approach to derive the fusion rule. A simple (but not necessarily optimal) approach is to select u_0 such that the a posteriori probability $\Pr(H_k|\mathbf{U})$ is maximized [17, § 6.3]. We write the a posteriori probability as follows:

$$\Pr(H_k|\mathbf{U};\mathbf{T}) = \left[\prod_{i=1}^{N_s}\Pr(u_i|H_k;t_i)\right]\Pr(H_k)/\Pr(\mathbf{U};\mathbf{T}), \tag{26}$$

where we have invoked the conditional independence of the sensors. This result can be evaluated using known values for $\Pr(u_i|H_l;t_i)$, which are specified by $P_{F_i}(t_i)$ and $P_{M_i}(t_i)$.

1. Some physical phenomena (e.g., a small depression in the soil) can have an effect on several sensors and, hence, this assumption can be difficult to justify. Nonetheless, it is commonly used in fusion studies.

Note that in both formulations $\Pr(H_k|\mathbf{U})$ depends implicitly on the thresholds $\mathbf{T}$. Optimal fusion requires that we select $\mathbf{T}$ to produce the best performance. For modest numbers of sensors, a search over all combinations of $\mathbf{T}$ is possible, but, in general, the search may be impractical. In such cases, it is more attractive to select a value of t_i that produces "good" sensor performance and to use that threshold for the calculations described above.

3.3 Soft Decision-Level Fusion

Consider the situation in which sensor i reaches decision u_i on the basis of features $\mathbf{\Theta}_i$ computed from its data using a deterministic transformation, e.g., $\mathbf{F}_i(\mathbf{D}_i) = \mathbf{\Theta}_i$. The hard decision-level fusion rules in (25) and (26) use global sensor reliability information in the form of $\Pr(u_i|H_k;t_i)$ and the prior probabilities $\Pr(H_k)$, but they do not utilize information about the confidence each sensor places in its individual decisions. Such information is quantified by $\Pr(u_i|\mathbf{\Theta}_i;t_i)$. When the sensors output this local confidence information, the outputs are equivalent to "soft" decisions and we can perform fusion using an approach similar to that described above for feature-level fusion. Treating the $\mathbf{U}$ as features returned by each sensor, a discrete analog of (7) yields

$$\Pr(H_k|\mathbf{\Theta}) = \sum_{u_1=1}^{Q}\cdots\sum_{u_{N_s}=1}^{Q}\Pr(H_k|\mathbf{U})\Pr(u_1|\mathbf{\Theta}_1;t_1)\Pr(u_2|\mathbf{\Theta}_2;t_2)\ldots\Pr(u_{N_s}|\mathbf{\Theta}_{N_s};t_{N_s}), \tag{27}$$

where we have assumed that the fused hypothesis H_k depends only on the individual sensor decisions u_i. The quantity $\Pr(H_k|\mathbf{U})$ is computed as described in the hard decision case.

4 Example Results: Fusion of Experimental Data

To test the foregoing algorithms, a sensor suite was used to acquire experimental data over a surrogate mine field, and the resulting data were fused using feature-level and decision-level techniques. In this section, we describe the sensors, the mine field, and the fused system performance.

4.1 Sensors

The sensor suite used in this work comprised a commercial electromagnetic induction (EMI) sensor, a ground penetrating radar (GPR), and a commercial infrared (IR) camera. The sensors and the signal processing used in feature extraction are documented below.

4.1.1 Electromagnetic Induction (EMI) Sensor

Virtually all land mines in use today contain some metal, which may be detected by a metal detector. The quantity of metal used varies widely from several kilograms to a few hundred milligrams. Excessive false alarm rates arise when detecting mines with small metal content because of the large amount of metallic clutter of anthropic origin found in most minefields.

The EMI sensor used in this work was a commercially available pulsed-induction metal detector, the Schiebel AN-19/2 (Schiebel Instruments, Incorporated, Washington,

D.C.) Internal signals within the device were sampled with a digital oscilloscope. The sensor was fitted to a linear scanner, which was used to acquire 56 samples over a 55 inch scan. Scans were performed at regularly spaced intervals (nominally every six inches) over the region of interest to obtain a grid of sample points.

Signal processing for the EMI sensor is a multistep process. The sensor response acquired over soil known to be metal-free, is subtracted from the waveforms. The integral of the resulting difference signal is an effective detector of buried metal. The integrated values acquired along a scan form spatial signatures. For targets not too near the sensor head and not too large compared to the loop radius, the sensor's spatial response has a monopolar shape that is reasonably well approximated by the function

$$G_1(\mathbf{R};\boldsymbol{\Theta}_1) = B + \frac{S}{1+(|\mathbf{R}-\mathbf{R}_0|/a)^b}, \tag{28}$$

where $\mathbf{R}_0$ is the location of the target's centroid, B is the response of the background, S is the amplitude of the target response, and a and b are shape parameters. A more sophisticated physics-based model of the response of a buried sphere has been presented by Das et al. [18] and could be used as an alternative model, but the simple empirical form above was adequate to approximate the limited set of targets considered here.

The second step in EMI signal processing involves the estimation of the feature vector $\boldsymbol{\Theta}_1$ given by $\boldsymbol{\Theta}_1 = [a\ b\ S\ E\ \mathbf{R}_0]^T$. Note that the background level B is not used in $\boldsymbol{\Theta}_1$ and we have added the quantity $E = |\mathbf{D}-\mathbf{G}|^2/|\mathbf{D}|^2$, which is the residual error in the estimate normalized to the signal energy. To compute $\boldsymbol{\Theta}_1$ at each sample position $\mathbf{R}$, a spatial window was defined and the data $\mathbf{D}_1$ within this window were fitted to the model using a nonlinear optimization technique. Since stronger signals tend to persist over larger distances, the size of the window was adjusted on the basis of the signal amplitude.

4.1.2 Ground Penetrating Radar

A ground penetrating radar detects the presence of reflecting boundaries in the soil. Such boundaries occur at the surface of the ground and at most buried objects. A GPR will detect both metallic and metal-free objects and, hence, it is a useful complement to the metal detector. Unfortunately, voids, stones, changes in soil density and moisture content, and many other subsurface natural objects are also detected, leading to a high false alarm rate.

The GPR used in this effort was developed at The Ohio State University (OSU), ElectroScience Laboratory (ESL) [19]. It employs a novel dielectric rod antenna that was scanned horizontally over the earth at a fixed height (nominally two inches). A network analyzer was used to measure the complex reflection coefficient at 51 frequencies between 1 and 6 GHz. Approximately 100 samples were acquired on each scan, which had a length of 55 inches. The antenna and its supporting platform were advanced in two inch intervals, resulting in a grid of samples.

Calibration of the radar was performed via a two step process. A "background" signature (the response of the radar when the antenna points into free space) was acquired and subtracted from all data. The temporal impulse response of the radar was then determined by measuring the (background-free) field scattered from a reference target (a short cylinder) in free space. The ideal scattered field for this target is known from numerical calculations and the ratio of the measured and ideal frequency-domain responses yields the impulse response of the radar. This response is then deconvolved from the measured data using frequency-domain division. By design, the spectral response of the reference target contains no zeroes, and regularization was not found to be necessary in this work.

Processing of the GPR data begins with calibration, in which the response of the system is deconvolved from the measured data. A Fourier transform of the calibrated data is then used to produce a map of the reflection coefficient as a function of along-scan position and time (or depth). The data is contaminated by a surface-reflected waveform, which we remove using an OSU-developed technique [11] based on a local, spatially varying maximum-likelihood estimate of the ground position and reflection coefficient. The resulting data are modeled as a point scatterer plus clutter, the signature of which at a typical depth in homogeneous soil is known and defines $\mathbf{G}_2(\mathbf{R})$. Target depth has an effect on both the signature shape and time delay, but, because the range of mine burial depths is restricted, the effect of depth errors on target shape was found to be insignificant. The GPR feature vector is given by $\boldsymbol{\Theta}_2 = [\rho\ d\ \mathbf{R}_0]^T$, where ρ is the peak correlation coefficient for the model signature and data, and d and $\mathbf{R}_0$ are, respectively, the depth and horizontal position indicated by the correlation peak. Although false targets are often seen in GPR scans, they tend not to persist across adjacent scans and, hence, the correlation was computed over a three-dimensional (depth, along-scan, and cross-scan) data window. GPR data tend to be highly correlated along a scan, because of returns from plane-stratified media and antenna ringing not eliminated by calibration. A modest amount of whitening was performed prior to correlation.

4.1.3 Infrared Camera

The natural heating and cooling of soil that occurs over a diurnal cycle causes thermal energy to flow into and out of the earth. The presence of a buried object changes this thermal flow, leading to temperature anomalies above the buried object at certain times in the diurnal cycle. These anomalies can be detected in soil temperature maps created by a suitable infrared (IR) camera. Since a mine's thermal properties are only weakly related to its metal content and electromagnetic properties, passive IR imaging is presumed to be a good adjunct to EMI and GPR sensors. In addition to buried mines, naturally occuring buried objects (e.g., stones), variations in soil mineral and water content and variations in solar illumination (due to shadowing by foliage, for example) also produce temperature anomalies. Our sensor was a commercially available MWIR camera, the IRRIS 160ST, (Cincinnati Electronics, Cincinnati, OH).

The IR data was processed by first remapping the imagery to ground coordinates using a standard perspective transformation [20], followed by bilinear interpolation of the distorted pixels and resampling to a uniform pixel size. An image chip was extracted at each interrogation point $\mathbf{R}$.

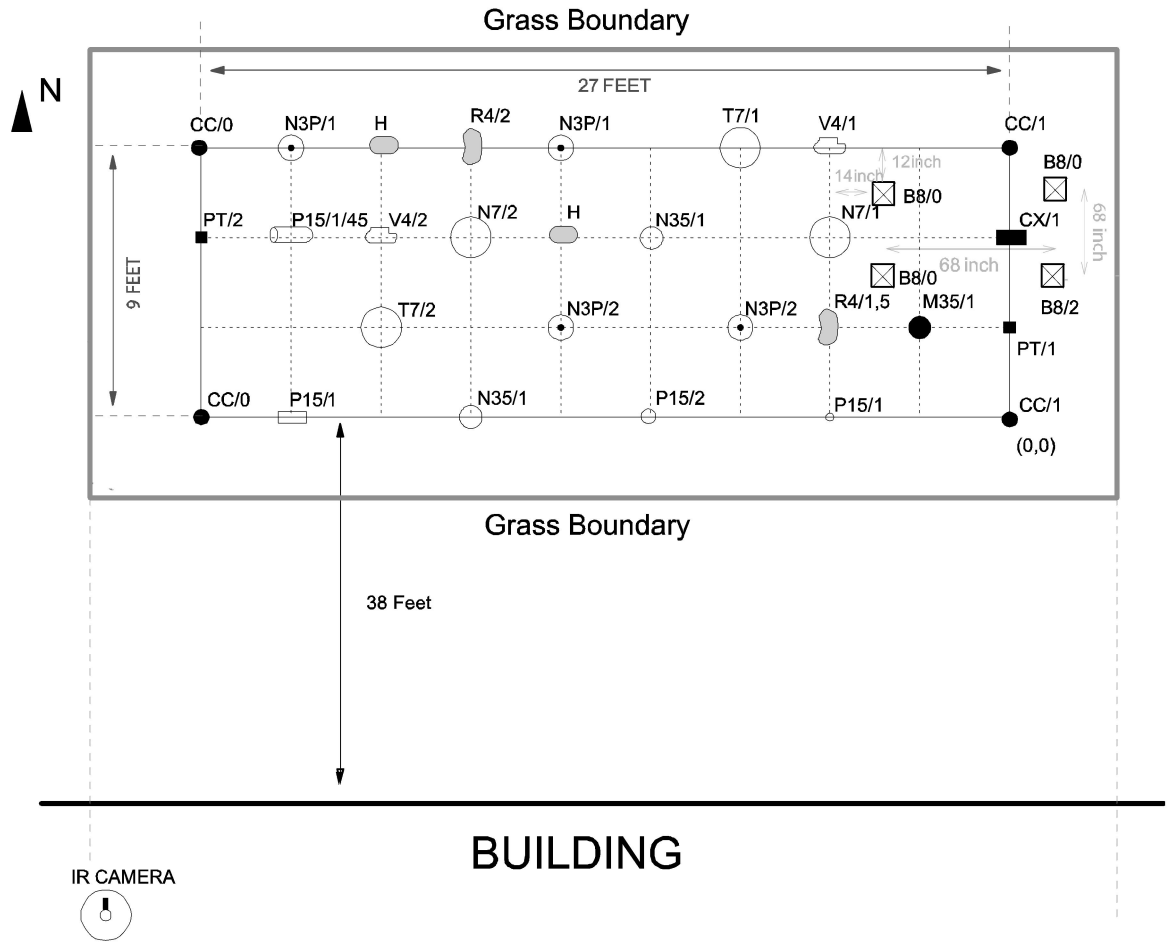


Fig. 5. Plan view of the surrogate mine and clutter test grid. The surface of the mine grid is bare soil. The surrounding area is covered with grass.

Based on empirical observations, we employed a signature model identical to that used for the EMI sensor, namely,

$$\mathbf{G}_3(\mathbf{R};\mathbf{\Theta}) = B + \frac{S}{1+(|\mathbf{R}-\mathbf{R}_0|/a)^b}. \tag{29}$$

The feature vector is $\mathbf{\Theta}_3 = [a\ b\ S\ E\ \mathbf{R}_0]^T$, where E is the normalized residual error energy.

4.2 Test Site

A surrogate mine field was created to test our sensors under conditions that approximate those found in field situations. The site location, adjacent to our laboratory, is former farm land. A total of 40 mine-like and clutter-like targets were emplaced in a 4×10 grid, as shown in Fig. 5. An abbreviation identifying the buried object and the depth in inches is shown near each object in the map. The identities of the buried objects are given in Table 2.

This site presented a surprising number of challenges to mine detection and data from it are valuable in that they illustrate problems that could be encountered in real demining operations. The target set was deliberately selected to be stressing to our sensors. In contrast to real mines, essentially all of which contain some metal, roughly half the mine surrogates were metal-free. In addition, some of the surrogate mines are quite small (only 1.5 inches in diameter) and undetectable to all of the sensors. These objects were redesignated as clutter. The site has been intermittently used for various engineering experiments for more than 40 years, which produced a significant amount of small metallic debris. Consistent with our goal of achieving realistic environmental conditions, the test area was not cleared prior to emplacing the targets. As a result, during two years of testing, roughly 100 metallic fragments and a number of stones were located in (and, subsequently, removed from) the region. When the targets were initially emplaced, a lush grass cover was present in the area, which was subsequently removed to improve the performance of the IR camera. The soil contains a large amount of clay, which has a negative effect on GPR performance. The surface of the mine field contains a number of topological irregularities but it is planar to within an estimated variation of ± 2 inches. The objects in this field had been in situ for more that 18 months at the time these data were acquired.

Data were acquired along the paths shown in Fig. 6. The small circles are the target grid points shown previously. The GPR samples were acquired at ~ 0.5 inch intervals along north-south (N-S) lines spaced every two inches east-west (E-W). The EMI sensor acquired samples at ~ 1 inch intervals along N-S lines spaced six inches E-W. The IR camera was positioned on the roof of a two-story building adjacent to the test site. The camera's field of view did not permit us to sample the entire mine field with an acceptable pixel density and a sequence of eight images was necessary as indicated by the trapezoids. Styrofoam fiducial markers

TABLE 2
Description of the Objects Buried in the Mine Test Grid

Abbr.	Object	Dia. (in.)	Ht. (in.)	Metal Content?	Identity	Quantity
CC	Soda can	2.5	5	Y	clutter	4
CX	Crushed soda can*	3	5	Y	clutter	1
PT	Pop top*	0.5	1	Y	clutter	2
M35	Aluminum disk	3.5	5/8	Y	surrogate	1
P15	Plexiglas disk	1.5	2	N	clutter	4
R4	Rock	4	2	N	clutter	2
V4	Void (Styrofoam)	4	2	N	clutter	2
H4	Refilled hole	6	2	N	clutter	2
N35	Nylon disk	3.5	5/8	N	surrogate	2
N3P	Nylon disk with steel pin	3	1	Y	surrogate	4
T7	Teflon disk	7	1	N	surrogate	2
N7	Nylon disk	7	1	N	surrogate	2
B8	Concrete pylon*	8	8	N	clutter	4
-	Blank	-	-	?	clutter	12

For the planar objects indicated by an asterisk, dimensions are length and width.

were placed on the region to permit later image warping and registration, which are indicated by squares in Fig. 6.

4.3 Performance of Individual Sensors

For each sensor i, feature vectors $\mathbf{\Theta}_i$ were extracted using the signal processing algorithms discussed above and supplied to sensor-specific classifiers $\Pr(u_i|\mathbf{\Theta}_i, t_i)$. Results for the binary detection problem involving hypotheses $H_1 = \text{"nonmine"}$ and $H_2 = \text{"mine"}$ are reported here. The prior probabilities for this site are $\Pr(H_0) = 28/40$ and $\Pr(H_1) = 12/40$. It is important to note that the number of mine surrogates being sampled is small. The results obtained must be interpretted in light of this fact and, hence, it is difficult to draw general conclusions.

To compute P_D and P_F, the test area was divided into a large number of square cells (280). A cell was considered to contain a detection if at least four adjacent detections were present. We computed the false alarm probability as the number of false detections divided by the total number of cells not occupied by mines.

To permit comparison between sensors, we have taken all intermediate sensor-specific hypotheses h_{ki} to be H_k, even though some surrogates are undetectable to some sensors. It is known that with sufficient training data backpropagation neural networks will approximate an a posteriori density [21], and the outputs of such a network (normalized to a unit sum) were used to approximate the required a posteriori probabilities. The number of input and output nodes for these networks were defined by the sizes of $\mathbf{\Theta}_i$ and K, respectively. The number of hidden nodes were 3, 7, and 6 for the GPR, EMI, and IR sensors, respectively. Target detections u_i were produced by thresholding the network outputs at level t_i. Leave-one-out training and validation was used in all of the results. Because networks trained with random weight initializations may converge to spurious local error minima, several trials were conducted and the best results were retained.

Figs. 7, 8, and 9 present binary detection maps for each sensor, which were generated by choosing appropriate detection thresholds t_i and discarding regions with small number of adjacent detections. The dark masses on these maps indicate detections, while circles represent the locations of the declared mine surrogates. Detection cells

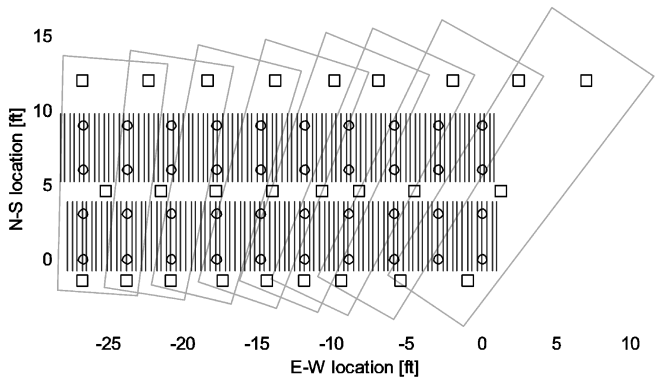


Fig. 6. Locations of sample points for the sensor suite.

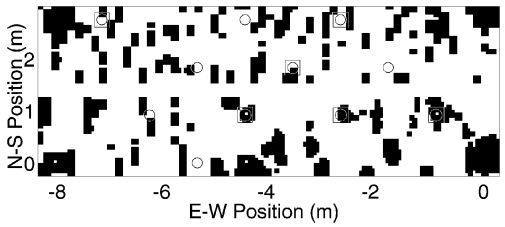


Fig. 7. Detection map obtained using EMI sensor.

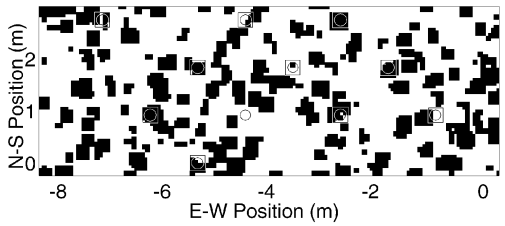


Fig. 8. Detection map obtained using IR sensor.

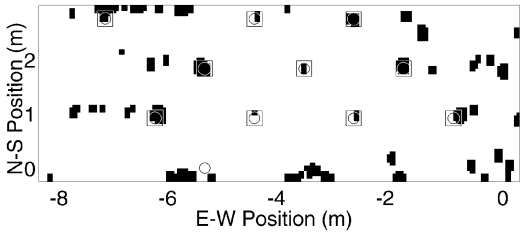


Fig. 9. Detection map obtained using the GPR sensor.

for the targets are indicated by squares. It is clear from these results that the EMI and IR sensors have a high false alarm rate when all targets are detected. The poor performance of the EMI sensor was anticipated because most of the mine surrogates in our test grid are nonmetallic and the site contains a large quantity of metallic clutter. The poor performance of the IR sensor was not anticipated, but may be due to the relatively large number of surface irregularities. For the threshold chosen here, the GPR sensor shows reasonably good performance, with only a single missed detection and a moderate false alarm rate. Detection of the missed mine, however, is problematic. The performance of these sensors can be compared quantitatively using receiver operating characteristic (ROC) curves, as shown in Fig. 10. These results confirm the findings shown in the detection maps. In particular, the inability of the GPR to detect all mines at a reasonable false alarm rate is clear.

4.4 Performance of Fusion Algorithms

For decision-level fusion, we used the "good threshold" approximation described in Section 3.2 to avoid the search over all possible sensor thresholds t_i implied in (25). The EMI and IR sensor thresholds were selected such that all expected targets (i.e., targets with metal for the EMI sensor and targets that present a sizable thermal discontinuity for the IR sensor) were detected while minimizing clutter detections. The criterion used for selecting the GPR threshold was the "knee" in the ROC curve in Fig. 10. The resulting individual sensor P_{M_i} and P_{F_i} values are shown in Table 3. The detections u_i for these thresholds comprise the inputs u_i for the decision-level fusion algorithms and are shown in the foregoing individual sensor detection maps. The local confidences $\Pr(u_i|\mathbf{\Theta}_i, t_i)$ required for the soft-decision fusion scheme of (27) were set to the a posteriori probability estimates produced by the corresponding classifier output. The ROC curve for decision-level fusion appears in Fig. 11.

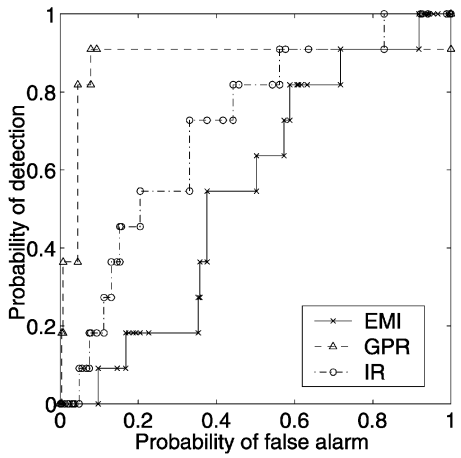


Fig. 10. Receiver operating characteristic (ROC) curves of individual sensors.

TABLE 3
Individual Sensor P_{F} and P_{M} Using "Good" Thresholds

Sensor	P_{F_i}	P_{M_i}
GPR	0.0929	0.0909
EMI	0.3755	0.4545
IR	0.5613	0.0909

Comparing Figs. 10 and 11, we see that hard-decision fusion based on (25) offers little benefit over the GPR sensor alone. Based on the form of (25), which most strongly weights the most reliable sensors, this result was expected. The performance of soft-decision fusion, however, is somewhat better. It is clear that with soft-decision fusion, $P_D = 1$ can be achieved with a smaller P_F, which is a significant improvement over the best individual sensor performance. The detection map obtained for soft-decision fusion is shown in Fig. 12. The detection map for hard-fusion is not shown, because it is nearly identical to the GPR detection map.

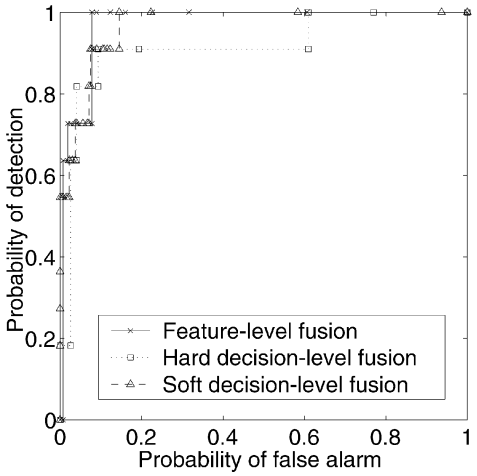


Fig. 11. Receiver operating characteristic (ROC) curves for three types of fusion.

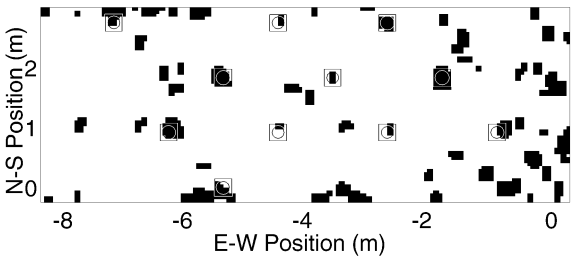


Fig. 12. Detection map obtained using soft-decision fusion.

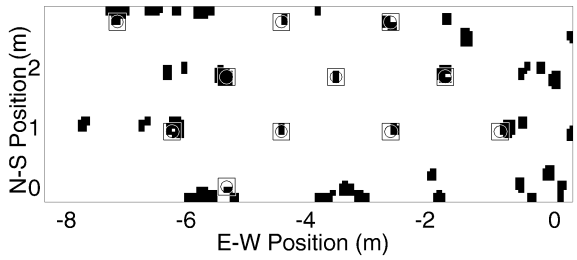


Fig. 13. Detection map obtained using feature-level fusion.

For feature-level fusion, the sensor feature vectors $\boldsymbol{\Theta}_i$ were concatentated and used in a single classifier (cf. (8)), resulting in the ROC curve shown in Fig. 11. As done for the individual sensors, the classifier $\Pr(H_k|\hat{\boldsymbol{\Theta}})$ was approximated by a backpropagation neural network. A total of 14 hidden nodes were used in this network.

Using as a figure of merit, the smallest P_F at which $P_D = 1$, we see that the feature-fused system is significantly better than the individual sensors and marginally better than the soft-decision fusion case. The detection map produced by the feature-fused system, corresponding to the largest threshold that can achieve $P_D = 1$, is given in Fig. 13. Comparing the soft-decision case in Fig. 12 and the feature-level case in Fig. 13, the reduction in false alarm rate indicated by the ROC curve (roughly 40 percent) is evident.

5 Summary and Conclusions

We have developed a feature-level fusion technique for noncommensurate sensor data sampled at noncoincident locations. The features used for fusion were extracted from spatial signatures of targets by applying nonlinear optimization techniques to data collected in a neighborhood of the location of interest. The maximum a posteriori approximation was used to simplify the exact fusion formulation to a form that was practical to implement. For the special case of normal clutter and normal linear features, fusion can be expressed as the Bayesian linear model for which a compact closed form expression exists. This result was used to demonstrate the importance of accurate sensor position information. We also discussed several decision-level fusion schemes that can be used for these data.

We tested the fusion algorithms on noncommensurate noncoincidently sampled data acquired by EMI, GPR, and IR sensors over a test field containing mine surrogates and both deliberate and unknown clutter. The targets and the clutter present in the test site were both very challenging. Although the number of mine surrogates sampled was small, the results presented here indicate that the performance of feature-level fusion and soft-decision fusion can be significantly better than that of the individual sensors. This performance enhancement is encouraging, especially when we recognize that the IR and EMI sensors were only marginally effective due to both the heavy clutter at the test site and our selection of surrogate targets. As expected, hard-decision fusion showed little improvement over the performance of the best individual sensor since the algorithm most strongly weights the most efffective sensors.

Overall, we have shown that feature-level and soft-decision level fusion of noncoincidently sampled EMI, IR, and GPR data can result in significant improvements in mine detection performance. We found a small additional benefit in using feature-level fusion, which provided a 40 percent reduction in the false alarm rate beyond the soft-decision fusion case. The feature-level fusion algorithm also provides an explicit method for dealing with position errors, which is not available through the decision-level fusion algorithms. The computational costs of decision-level and feature-level fusion algorithms are approximately equal since the computational expense involved in fusion is dominated by features estimation, and all methods of detection must perform this processing. There is, however, a small penalty for feature-level fusion since the classifier must be somewhat more complex to account for the greater number of inputs.

These results are encouraging but the small sample size of our data set suggests that a larger study be done to confirm these results. A data collection has been done at a larger minefield operated by the US Army at Fort A.P. Hill, Virginia. That study is now complete, and the results will be reported in the future.

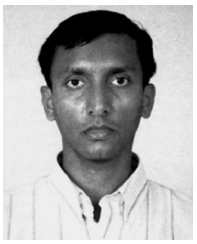
Acknowledgments

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions, and recommendations expressed therein are those of the authors and are not necessarily those of Duke University or the ARO. A preliminary version of this paper was presented at the 1999 SPIE AeroSense Conference, Orlando, Florida.

References

[1] P.K. Varshney, *Distributed Detection and Data Fusion.* New York: Springer-Verlag, 1997.
[2] B.V. Dasarathy, *Decision Fusion.* Los Alamitos, Calif.: IEEE CS, 1994.
[3] S. Chauduri, A. Crandall, and D. Reidy, "Multisensor Data Fusion for Mine Detection," *Proc. SPIE: Sensor Fusion III,* R.C. Harney, ed., vol. 1306, pp. 187-203, 1990.
[4] B. Brusmark, A. Christiansen, P. Jägerbro, and A. Lauberts, "Combination of Ground Penetrating Radar and Metal Detector Data for Mine Detection," *Proc. Seventh Int'l Conf. Ground Penetrating Radar,* 1998.
[5] G.A. Clark, S.K. Sengupta, R.J. Sherwood, J.E. Hernandez, M.R. Buhl, P.C. Schaich, R.J. Kane, M.J. Barth, and N.K. Del Grande, "Sensor Feature Fusion for Detecting Buried Objects," *Proc. SPIE: Underground and Obscured-Object Imaging and Detection,* N.K. Del Grande, I. Cindrich, and P.B. Johnson, eds., vol. 1942, pp. 178-188, 1993.
[6] X. Miao, M.R. Azimi-Sadjadi, B. Tian, A.C. Dubey, and N.H. Witherspoon, "Detection of Mines and Minelike Targets Using Principal Component and Neural Methods," *IEEE Trans. Neural Networks,* vol. 9, no. 3, pp. 454-463, May 1998.
[7] R.A. Weisenseel, W.C. Karl, D.A. Castanon, E. Miller, C. Rappaport, and C.A. Dimarzio, "Statistical Fusion of GPR and EMI Data," *Proc. SPIE: Detection and Remediation Technologies for Mines and Minelike Targets IV,* A.C. Dubey, J.F. Harvey, J.T. Broach, and R.E. Dugan, eds., vol. 3710, pp. 1179-1187, 1999.
[8] E. den Breejen, K. Schutte, and F. Cremer, "Sensor Fusion for Anti-Personnel Landmine Detection, a Case Study," *Proc. SPIE: Detection and Remediation Technologies for Mines and Minelike Targets IV,* A.C. Dubey, J.F. Harvey, J.T. Broach, and R.E. Dugan, Eds., vol. 3710, pp. 1235-1245, 1999.

[9] P.G. Johnson and P. Howard, "Performance Results of the EG&G Vehicle Mounted Mine Detector," *Proc. SPIE: Detection and Remediation Technologies for Mines and Minelike Targets IV,* A.C. Dubey, J.F. Harvey, J.T. Broach, and R.E. Dugan, eds., vol. 3710, pp. 1149-1159, 1999.
[10] J.E. McFee, V.C. Aitken, R. Chesney, Y. Das, and K.L. Russell, "Multisensor Vehicle-Mounted Teleoperated Mine Detector with Data Fusion," *Proc. SPIE: Detection and Remediation Technologies for Mines and Mine-Like Targets,* A.C. Dubey, J.F. Harvey, and J. Broach, eds., vol. 3392, pp. 1082-1093, 1998.
[11] A. Gunatilaka and B.A. Baertlein, "Comparison of Pre-Detection and Post-Detection Fusion for Mine Detection," *Proc. SPIE: Detection and Remediation Technologies for Mines and Mine-Like Targets,* A.C. Dubey, J.F. Harvey, J.T. Broach, and R.E. Dugan, eds., vol. 3710, pp. 1212-1223, 1999,
[12] A. Gunatilaka and B.A. Baertlein, "A Subspace Decomposition Technique to Improve GPR Imaging of Anti-Personnel Mines," *Proc. SPIE: Detection and Remediation Technologies for Mines and Minelike Targets V,* A.C. Dubey, J.F. Harvey, J.T. Broach, and R.E. Dugan, eds., vol. 4038, 2000.
[13] B.A. Baertlein, "Neural-Network Based Multi-Sensor Fusion for Mine Detection," Technical Report N61331-93-C-0050, Ballena Systems Corporation, 1994.
[14] J.A. Marble, J.G. Ackenhusen, J.W. Wegrzyn, J. Mancuso, and C. Dwan, "Sensor Fusion Performance Gain for Buried Mine/uxo Detection Using GPR, EMI, and MAG Sensors," *Proc. SPIE: Detection and Remediation Technologies for Mines and Mine-Like Targets V,* A.C. Dubey, J.F. Harvey, J.T. Broach, and R.E. Dugan, eds., vol. 4038, pp. 1473-1484, Apr. 2000.
[15] H.V. Poor, *An Introduction to Signal Detection and Estimation,* second ed. New York: Springer-Verlag, 1994.
[16] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1993.
[17] D.L. Hall, *Mathematical Techniques in Multisensor Data Fusion.* Boston: Artech House, 1992.
[18] Y. Das, J.E. McFee, J. Toews, and G.C. Stuart, "Analysis of an Electromagnetic Induction Detector for Real-Time Location of Buried Objects," *IEEE Trans. Geoscience Remote Sensing,* vol. 28, no. 3, pp. 278-287, May 1990.
[19] S. Nag, L. Peters Jr., I.J. Gupta, and C.C. Chen, "Ramp Response for the Detection of Anti-Personel Mines," *Proc. SPIE: Detection and Remediation Technologies for Mines and Minelike Targets IV,* A.C. Dubey, J.F. Harvey, J.T. Broach, and R.E. Dugan, eds., vol. 3710, pp. 1313-1322, 1999.
[20] I.K. Sendur and B.A. Baertlein, "Techniques for Improving Buried Mine Detection in Thermal IR Imagery," *Proc. SPIE: Detection and Remediation Technologies for Mines and Minelike Targets IV,* A.C. Dubey, J.F. Harvey, J.T. Broach, and R.E. Dugan, eds., vol. 3710, pp. 1272-1283, 1999.
[21] M.D. Richard and R.P. Lippmann, "Neural Network Classifiers Estimate Bayesian a Posteriori Probabilities," *Neural Computation,* vol. 3, pp. 461-483, 1991.



Ajith H. Gunatilaka received the BSc degree in electronic and telecommunication engineering from the University of Moratuwa, Sri Lanka, in 1990. In 1994, he obtained the MS degree in optical sciences and engineering from the University of Central Florida. He received the PhD degree in electrical engineering from the Ohio State University in March 1990. He was a graduate research associate at the Electro-Science Laboratory of The Ohio State University from 1994 to 2000. He is currently a member of technical staff at Lucent Technologies in Columbus, Ohio. He is a member of the IEEE.



Brian A. Baertlein received the PhD in electrical engineering from the University of Arizona in 1988. His professional career has included analyses of sensor systems of various types, development of detection and estimation algorithms, and work in sensor fusion, scattering and propagation phenomena, antennas, and electromagnetic compatibility. He is currently a research scientist and adjunct associate professor of electrical engineering at The Ohio State University (OSU) Electro-Science Laboratory (ESL). Before joining OSU, he was a senior scientist with several small businesses doing work for the US Department of Defense and the Department of Energy. He is am member of the IEEE.

▷ **For further information on this or any computing topic, please visit our Digital Library at** http://computer.org/publications/dlib.

Three-Dimensional Thermal Modeling of Land Mine Signatures over a Diurnal Cycle

İbrahim Kürşat Şendur[1], Brian A. Baertlein[1], and Kambiz Vafai[2]

[1]ElectroScience Laboratory, Department of Electrical Engineering, The Ohio State University, Columbus, OH 43212-1191 USA. E-mail: baertlein.1@osu.edu .

[2]Department of Mechanical Engineering, University of California, Riverside, CA 92521-0925 USA. E-mail: vafai@engr.ucr.edu .

Abstract

A three-dimensional thermal model has been developed to study the passive infrared (IR) signature of a land mine buried under a smooth soil surface. The physical processes and heat transfer mechanisms that produce thermal IR signatures are briefly discussed, and a mathematical description of the processes is presented. The finite element method (FEM) is used to study the mine signature and environmental effects. Natural solar insolation (both direct and atmospheric-emitted components) and the effects of convective and radiative heat transfer are represented by linearized boundary conditions. The Crank-Nicholson time-stepping algorithm is used for the temporal evolution of the simulation. A periodic diurnal cycle is simulated by running the model to steady state. Thermal modeling results for a homogeneous soil half-space and a layered earth are compared to those of a previously developed one-dimensional analytical model. Three-dimensional modeling results are compared with results from an integral equation solution presented in a companion paper, and good agreement is found. Numerical simulations are presented for an anti-personal mine, a square anti-tank mine, and a circular anti-tank mine buried under a smooth soil surface. The temporal evolution of the temperature distribution is presented both at the surface and as a function of depth.

Keywords

Thermal infrared imagery, land mines, heat transfer, thermal model, numerical simulation, finite element method.

I. Introduction

With tens of millions of land mines buried in over 70 countries, humanitarian demining requires robust, effective, and expeditious mine detection sensors. It has long been known that a host of sensor technologies are capable of detecting mines [1]. Among these sensors are ground penetrating radars [2–4] (GPRs), metal detectors [5], IR sensors, chemical sensors [6], condensed phase techniques [7, 8], and acoustic sensors [9–11].

IR sensors have several unique and highly desirable properties for demining. Chief among these are their abilities to detect mines at long range and to rapidly scan large areas. In addition, since the IR sensor is independent of the electrical properties of the target, it complements a metal detector or a GPR in a sensor-fused system.

Detection of land mines via thermal IR imagery is affected by several factors, including time of day, cloud cover, vegetation, surface irregularities, and past meteorological conditions. There is a critical need to understand the effects of these environmental factors

to improve detection performance and to better utilize IR sensors. The IR signature of a buried mine involves both thermal and radiometric phenomena. In this work we address the thermal issues. Radiometric issues are treated in a separate work [12].

For many years, experimental studies have been conducted to assess the detection performance of thermal IR imagery and to better understand the underlying physics. Among these studies we note the work of LeSchack and Del Grande [13] who describe the effects of emissivity on target signatures in the context of a two-color IR system. Janssen et al. [14] reported a study of basic phenomenology in which several sensors were used to examine the time evolution of signatures for both buried and surface mines. A related study is that of Maksymomko et al. [15] who measured the temperatures of both live and surrogate mines through multiple diurnal cycles. Simard [16] found a linear relation between the apparent temperature of buried mines and the soil temperature gradient on gravel roads. Russel et al. [17] showed that this gradient could be predicted from remote soil temperature measurements, making it possible to remotely assess the potential effectiveness of an IR sensor. Pregowski and Swiderski [18] presented a qualitative discussion of the effects of soil cover and water content. In a more recent study, Pregowski et al. [19] developed a two-dimensional model to predict the thermal IR signatures of buried land mines. Those numerical simulations were based on a finite-difference time-domain (FDTD) formulation, and they included water transport in the soil medium. Pregowski et al. concluded that a single IR image taken at a single spectral band may not be sufficient for IR detection, because of clutter. Another work, obliquely related to the effort proposed here, is that of Del Grande et al. [20], who discussed the use of computed tomography and a finite difference thermal analysis code (TOPAZ3D) for imaging structural flaws in materials.

In this work a three-dimensional thermal model based on a FEM formulation is presented to predict IR signatures of land mines. This work starts with a description and mathematical formulation of the processes generating IR mine signatures in Sect. II. Heat transfer mechanisms are described and the convective and radiative boundary conditions at the soil-air interface are discussed in that section. The FEM formulation of the problem is given in Sect. III. The accuracy and efficiency of the model developed herein are demonstrated in Sect. IV by comparing FEM results with results of other techniques. Simulation

results for several land mines are also presented in Sect. IV. A summary and concluding remarks appear in Sect. V.

II. Physical description of the process

An IR sensor measures the sum of emission and reflection from a surface, and the contributions of the individual components are not easily separated. Depending on the sensor application, either the emitted or reflected component may constitute clutter. In detection of buried objects, the emitted component carries the target-related information and, therefore, the temperature of the soil surface is of primary interest. Thermal emission from the soil depends on the surface temperature and emissivity. A buried mine disturbs the surface temperature when the solar-driven diurnal flow of thermal energy into and out of the soil is affected by the mine. The surface temperature distribution can be found through the solution of the three-dimensional heat transfer equation in the soil and the mine. In this section, we discuss the heat transfer mechanisms involved and the boundary conditions applied to the problem.

Three heat transfer mechanisms affect the surface temperature distribution as shown in Fig. 1. Conduction occurs in the soil and at the soil-mine interface. At the soil-air interface convective and radiative heat transfer mechanisms are present. Convection involves the transfer of heat between the soil and air, and is greatly affected by wind. Radiative heat transfer involves direct solar radiation, atmosphere-scattered solar radiation, and thermal radiation between the earth and atmosphere. All of these processes are time varying (and herein, assumed periodic) over the diurnal cycle. Mathematical descriptions of the phenomena are as follows:

The temperature distribution in the soil and mine is described by the three-dimensional heat flow equation

$$C(\mathbf{r})\frac{\partial T(\mathbf{r},t)}{\partial t} = \nabla \cdot (\mathcal{K}(\mathbf{r})\nabla T(\mathbf{r},t)), \tag{1}$$

where $T(\mathbf{r},t)$ [K] is the temperature, $\mathcal{K}$ [W m^{-1} K^{-1}] is the thermal conductivity of the material, and C [J m^{-3} K^{-1}] is the volumetric heat capacity of the material. Thermal diffusivity κ [m^2 s^{-1}] is defined as the ratio of the thermal conductivity to the volumetric

heat capacity, which can be written as

$$\kappa_i(\mathbf{r}) = \frac{\mathcal{K}_i(\mathbf{r})}{C_i(\mathbf{r})}, \qquad i = s, m \tag{2}$$

where the subscripts s and m refer to the soil and mine, respectively. For piecewise constant properties in the soil and mine regions, Eq. (1) can be rewritten as

$$\nabla^2 T(\mathbf{r}, t) - \frac{1}{\kappa_i(\mathbf{r})} \frac{\partial T(\mathbf{r}, t)}{\partial t} = 0, \qquad i = s, m \tag{3}$$

To complete the thermal problem description, boundary condition at the soil-air interface must be specified. At this boundary, convective and radiative heat transfer mechanisms are present. Numerous models have been presented in the literature to represent one-dimensional heat transfer at a soil-air interface over a diurnal cycle [21–26], and they are used here. The net heat flux into the ground $\mathcal{F}_{net}$ can be written as (for z positive downward)

$$\mathcal{F}_{net}(t) = -\mathcal{K} \frac{\partial T(\mathbf{r}, t)}{\partial z}\bigg|_{z=0} = \mathcal{F}_{sun}(t) + \mathcal{F}_{sky}(t) - \mathcal{F}_{sh}(t) - \mathcal{F}_{gr}(t), \tag{4}$$

where $\mathcal{F}_{sun}$ is the incident solar radiation reduced by cloud extinction, atmospheric absorption, soil albedo and the cosine of the zenith angle; $\mathcal{F}_{sky}$ is the sky brightness with a correction for cloud cover; $\mathcal{F}_{sh}$ is the sensible heat transfer from land to atmosphere due to convection; and $\mathcal{F}_{gr}$ is the gray body emission from the soil surface.[1] The incident solar radiation can be measured experimentally or obtained from sophisticated models (e.g., MODTRAN4 [27]). A simple closed form approximation is often sufficient for exploratory studies. Watson [21] used

$$\mathcal{F}_{sun}(t) = S_0(1 - A)(1 - Cl)\mathcal{H}(t) \tag{5}$$

to approximate the short-wavelength solar flux. In Eq. (5) $S_0 = 1353$ [W/m^2] is the solar constant, A is the ground albedo, Cl is a factor that accounts for the reduction in solar flux due to cloud cover, and $\mathcal{H}(t)$ is the local insolation. A convenient approximation for $\mathcal{H}(t)$ is [21, 28]

$$H(t) = \begin{cases} M(Z(t)) \cos Z'(t) & -t_r < t < t_s \\ 0 & t_s < t < t_r \end{cases} \tag{6}$$

[1]In formulating this expression we have ignored evaporation and other movement of soil moisture, as well as changes in the state in the medium.

where t_s and t_r [h] are the local sunset and sunrise times with respect to midnight, $M(Z)$ is the atmospheric transmission as a function of zenith angle Z, and Z' is the local zenith angle for the inclined surface. The function M is modeled as

$$M(Z) = 1 - 0.2\sqrt{\sec Z} \tag{7}$$

and Z is determined from

$$\cos Z = \cos\lambda \ \cos\delta \cos\omega t + \sin\lambda \ \sin\delta. \tag{8}$$

in which λ is the local latitude and δ is the declination angle of the sun

$$\delta = -23.433^\circ \cos[2\pi(\text{month})/12] \tag{9}$$

The local zenith angle Z' can be determined from

$$\cos Z' = \cos d \cos Z - \sin d(\sin\psi \cos\delta \sin\omega t - \cos\psi \sin\delta \cos\lambda - \sin\delta \sin\lambda \cos\omega t) \tag{10}$$

where d is the surface slope angle measured downward from the horizontal, and ψ is the azimuth of the slope angle measured counterclockwise from north. For flat ground, assumed here, $d = 0$.

The long-wavelength radiation from the atmosphere, $\mathcal{F}_{sky}$, is given as

$$\mathcal{F}_{sky}(t) = \mathcal{E}\sigma T_{sky}^4(t), \tag{11}$$

where $\sigma = 5.67 \times 10^{-8}$ [W m^{-2} K^{-4}] is the Stefan-Boltzmann, constant and T_{sky} [K] is an effective sky radiance temperature. The heat loss due to ground radiation is given by Stefan's Law

$$\mathcal{F}_{gr}(t) = \mathcal{E}\sigma T^4(t, z = 0), \tag{12}$$

where $\mathcal{E}$ [unitless] is the mean emissivity of the surface and $T(t, z = 0)$ is the soil temperature at the soil-air interface. The sensible heat transfer between the surface and atmosphere is approximated by

$$\mathcal{F}_{sh}(t) = h(t)(T_{air}(t) - T(t, z = 0)). \tag{13}$$

where h [W m^{-2} K^{-1}] is a convection coefficient. Kahle [23] offered the following model for this parameter

$$h(t) = \rho_a c_{pa} C_d (W(t) + 2) \tag{14}$$

where $\rho_a = 1.16$ [kg/m^3] is the density of air, $c_{pa} = 1007$ [J kg^{-1} K^{-1}] is the specific heat of air, C_d [unitless] is the wind drag coefficient chosen to be 0.002, and $W(t)$ [m/s] is the wind speed. In this work we used a time-averaged (i.e., $W(t) = const.$) convective heat transfer coefficient. Kahle [23], drawing on measurements by Kondratyev [29], formulated an empirical model for the air temperature, which was later modified by England [24] to yield the expression

$$T_{air}(t) = T_{0,air} - T_{del}\cos(2\pi(t-2)/24), \tag{15}$$

where t is the local time as above, and $T_{0,air}$ and T_{del} are estimated from local meteorological data. Kahle also proposed a model for the sky temperature, which was modified by England [24] using Brundt's formula [23]

$$T_{sky}(t) = T_{air}(t)(0.61 + 0.05\sqrt{w})^{0.25}, \tag{16}$$

where w [mmHg] is the water vapor pressure.

Using Eqs. (5) and (11)-(13) in Eq. (4) completes the specification of the boundary condition at the soil-air interface. This condition involves the nonlinear function of the surface temperature T^4, but over the limited temperature range that arises in mine detection (on the order of 10 K), the result can be linearized using a technique described by Watson [21], which yields

$$\frac{\partial T(\mathbf{r},t)}{\partial z} \approx T(\mathbf{r},t)\frac{1}{\mathcal{K}_s}(h(t) + 4\mathcal{E}\sigma T_{sky}^3(t)) - \frac{1}{\mathcal{K}_s}(\mathcal{F}_{sun}(t) + h(t)T_{air}(t) + 4\mathcal{E}\sigma T_{sky}^4(t)). \tag{17}$$

III. FEM Formulation

The FEM is a well-known, efficient computational technique widely used for the solution of heat transfer problems [30,31]. The FEM permits modeling of arbitrarily shapes, which allows one to incorporate a rough soil surface and realistically shaped mines and anomalies. Furthermore, different thermal parameter values may be used over each spatial element, which permits modeling of inhomogeneous soil and mines. A spatial FEM formulation can be combined with a time-stepping scheme to obtain a solution for time-dependent heat transfer problems. The Galerkin formulation used here yields the best approximation in the variational sense. In this section, the FEM formulation of the three-dimensional heat transfer equation is given, including the spatial and temporal discretizations of the problem.

An important part of any FEM implementation is the spatial discretization of the computational domain. In what follows we denote the computational volume by Ω and its boundary by Γ. We take this volume to have a rectangular parallelepiped (box) shape. The boundary planes in x and y are denoted Γ_x and Γ_y respectively. The lower boundary plane is denoted Γ_z. The volume and its boundary of element e are represented by Ω_e and Γ_e, respectively. The element boundaries on the soil-air interface are represented by $\Gamma_{e,air}$.

We require the solution of Eq. (1) with the boundary conditions in Eq. (17) when driven by a periodic source. As noted previously, the temperature ranges of interest are relatively small, and over this limited range the thermal properties of soil and mine are assumed to be independent of temperature and of time. Hence, phenomena such as drying and soil water movement are neglected in the current model. We also assume that the temporal and spatial dependencies of $T(\mathbf{r},t)$ are separable within an element.

Over each element Ω^e we approximate the temperature distribution as

$$T^e(\mathbf{r},t) = \sum_{i=1}^{N_e} T_i^e(t)\phi_i^e(\mathbf{r}), \tag{18}$$

where $\phi_i^e(\mathbf{r})$ are specified linear interpolation functions, N_e denotes the number of nodes over the subdivided finite element, and $T_i^e(t)$ are the nodal temperatures, which comprise the unknown coefficients in our representation. Using this expression in Eq. (1) and using a Galerkin formulation to enforce weak equality we obtain

$$\int_{\Omega_e} \phi_j^e(\mathbf{r}) \left[C(\mathbf{r}) \frac{\partial T^e(\mathbf{r},t)}{\partial t} - \nabla \cdot (\mathcal{K}(\mathbf{r}) \nabla T^e(\mathbf{r},t)) \right] d\mathbf{r} = 0 \qquad \forall \Omega^e \in \Omega. \tag{19}$$

Continuity in temperature and flux are enforced on the surface Γ^e shared by adjacent elements.

The boundary conditions imposed on external boundaries Γ are determined from physical considerations. The boundary condition at the soil-air interface has been defined in Eq. (17). For the transverse coordinates x and y, we assume the presence of an infinite rectilinear array of mines (e.g., a mine field), one of which is centered in our computational domain. As a result of symmetry arguments, we have

$$\left.\frac{T(\mathbf{r},t)}{\partial x}\right|_{\Gamma_x} = 0 \quad \text{and} \quad \left.\frac{T(\mathbf{r},t)}{\partial y}\right|_{\Gamma_y} = 0 \tag{20}$$

in x and y, respectively. One can show numerically that, although an infinite array of mines is being analyzed, for a sufficiently large computational volume, there is negligible interaction between mines, and the results are a good approximation to the response of an isolated mine. For the boundary condition on Γ_z (at the bottom face of the computational volume) we observe that at sufficiently large depths (a few tens of centimeters) the temperature is independent of time over a diurnal cycle. The required dept h is on the order of the "diurnal depth", which is given by $D = \sqrt{\kappa L_p/\pi}$, where $L_p = 24$ hours, can be predicted from one-dimensional models of soil [32] using thermal characteristics found in the literature [33, 34]. Furthermore, at sufficiently large depths the spatial derivative of the temperature is zero. On the basis of these findings the boundary condition at Γ_z can be taken to be either

$$T(\mathbf{r}, t)\Big|_{\Gamma_z} = const. \tag{21}$$

or

$$\frac{\partial T(\mathbf{r}, t)}{\partial z}\Big|_{\Gamma_z} = 0. \tag{22}$$

The constant in Eq. (21) can be derived from a one-dimensional model, noting that at sufficiently large depths the steady state solution exists. In this work we use the boundary condition given by Eq. (22), since the zero flux constraint is more natural for FEM formulations. Over the course of a year ($L_p = 365$ days), variations in the temperature distribution may extend to tens of meters [23], and if necessary, this lower boundary condition can be further modified by a small constant flux representing the seasonal trend or geothermal flux [21, 23].

After substituting Eq. (18) in Eq. (19) and incorporating the boundary conditions given by Eqs. (20), (21), and (17) the unknown nodal coefficients $T_i(t)$ can be expressed as a matrix equation

$$\overline{\mathbf{M}}\,\dot{\mathbf{T}} + \overline{\mathbf{K}}\,\mathbf{T} = \mathbf{F}, \tag{23}$$

in which boldface letters represent vectors, over-lined boldface letters represent matrices and the superposed dot on $\mathbf{T}$ denotes the derivative with respect to time. The elements of these matrices and vectors are given by

$$M_{ij} = \int_{\Omega_e} C(\mathbf{r})\phi_i^e(\mathbf{r})\phi_j^e(\mathbf{r})d\mathbf{r}, \tag{24}$$

$$K_{ij} = \int_{\Omega_e} \mathcal{K}(\mathbf{r}) \left[\frac{\partial \phi_i^e(\mathbf{r})}{\partial x} \frac{\partial \phi_j^e(\mathbf{r})}{\partial x} + \frac{\partial \phi_i^e(\mathbf{r})}{\partial y} \frac{\partial \phi_j^e(\mathbf{r})}{\partial y} + \frac{\partial \phi_i^e(\mathbf{r})}{\partial z} \frac{\partial \phi_j^e(\mathbf{r})}{\partial z} \right] d\mathbf{r} + \int_{\Gamma_{e,air}} \left[4\mathcal{E}\sigma T_{sky}^3(t) + h(t) \right] \phi_i^e(\mathbf{r}) \phi_j^e(\mathbf{r}) dS, \quad (25)$$

and

$$F_i = \int_{\Gamma_{e,air}} \left[4\mathcal{E}\sigma T_{sky}^4(t) + \sigma h(t) T_{air}(t) + \mathcal{F}_{sun}(t) \right] \phi_i^e(\mathbf{r}) dS. \quad (26)$$

The spatial discretization was accomplished using pentahedral ("prism"-shaped) elements. Linear interpolation functions $\phi_i^e(\mathbf{r})$ were chosen.

Temporal discretization of Eq. (23) leads to a finite difference formulation, which is solved using the Crank-Nicholson scheme. We have

$$(2\,\overline{\mathbf{M}} + \Delta t\,\overline{\mathbf{K}}_{t+\Delta t})\mathbf{T}_{t+\Delta t} = \Delta t\,(\mathbf{F_t} + \mathbf{F}_{t+\Delta t}) + (2\,\overline{\mathbf{M}} - \Delta t\,\overline{\mathbf{K}}_t)\mathbf{T}_t. \quad (27)$$

Eq. (27) is solved at each time step to evaluate the nodal temperatures $\mathbf{T}_{t+\Delta t}$. Employing a dense LU decomposition technique for this solution is inefficient, because the matrices $\overline{\mathbf{M}}$ and $\overline{\mathbf{K}}$ are sparse. Banded matrix storage provides a partial remedy for this problem, but limitations in node numbering usually results in inefficient memory use. An efficient sparse solver is employed in our work. The sparse matrices are stored as vectors with non-zero elements, so that inefficient memory usage is avoided. For very large problems, however, an iterative technique will be more efficient. For large numbers of elements and small time increments the matrix equation can become ill-conditioned, but this effect did not arise for the element sizes and time increments chosen in this study.

An initial condition at $t = 0$ is also required. Although the solution procedure described here will converge to a steady state value using $T(\mathbf{r}, t = 0) = const.$ as the initial condition, the computation time can be reduced if the average depth-dependent temperature of homogeneous soil is used as a starting value. A suitable estimate is derived from a one-dimensional steady state model for homogeneous soil

$$T(\mathbf{r}, t = 0) = \left(\frac{1}{L_p} \int_{L_p} \mathcal{F}_{sun}(t) dt + 4\mathcal{E}\sigma \overline{T}_{sky}^4 + \overline{hT}_{air} \right) \Big/ \left(4\mathcal{E}\sigma \overline{T}_{sky}^3 + \overline{h} \right) \quad (28)$$

in which overbar indicates a time averaged value and L_p denotes the 24 hour period. Equation (28) is derived by using a Fourier series expansion of the temperature with the boundary condition given by Eq. (17) and solving only for the dc term of the series. A similar expression neglecting the convection term was derived by Watson [21].

IV. Results

In this section numerical simulation results are presented to demonstrate the temporal and spatial dependence of IR mine signatures. The results are organized in three groups. In the first group the temperature of homogeneous soil is computed using three different techniques, namely, an analytical model, an integral-equation based body-of-revolution model, and the above-described FEM model. This 1-D solution provides a check on the temporal behavior. The integral equation solution is described in a companion paper [12]. In the second group of results, the three-dimensional modeling capabilities of the FEM and integral-equation solutions are compared for a simple homogeneous target. In the third group the temporal and spatial temperature distributions of several land mines are studied and suggestions are presented for mine detection.

A. Thermal modeling of homogeneous soil

We used an analytical model developed by Watson [21] as a benchmark solution to evaluate the integral-equation and FEM solutions for a homogeneous half-space comprising damp soil. The predicted soil surface temperature is given in Fig. 2, which shows good agreement among the three solutions. The thermal properties of the soil and of other materials used in the mine models are given in Table I. Here and throughout this work we use the model parameters $S_0 = 1353$ [W/m^2], $\lambda = 35°$, $\delta = 0°$, $A = 0.3$, $Cl = 0.2$, $d = 0°$, $\psi = 0°$, $\epsilon = 1$, $t_R = 6{:}00$ [h], $t_S = 18{:}00$ [h], $w = 0.8$ [mmHg], $T_{sky} = 260$ K, and $T_{air} = 289$ K. These values are appropriate for an experiment performed on the vernal equinox, on which the celestial equator intersects the ecliptic, resulting in zero declination of the sun. The duration of day and night are equal, which is used to select the sunrise and the sunset times. An average air temperature of 289 K is assumed, which is a typical value for Columbus, OH in spring.

B. IR signature of a cylindrically shaped homogeneous mine

In this section a circular anti-tank mine simulant buried 4.5 cm under a perfectly smooth soil surface is studied. The simulant mine has a diameter of 20 cm and a height of 7.5 cm, and it is filled with trinitrotoluene (TNT), an explosive used in most land mines. The thermal properties of the materials used in the mine models are given in Table I. It is

TABLE I

THERMAL DIFFUSIVITY (κ), CONDUCTIVITY ($\mathcal{K}$) AND DIURNAL DEPTH (D) FOR MATERIALS USED IN MODELING MINES

Material	κ [$\mathrm{m^2 s^{-1}}$]	$\mathcal{K}$ [$\mathrm{W\ m^{-1}\ K^{-1}}$]	D [m]
Soil	5.0×10^{-7}	2.6	0.117
TNT	9.25×10^{-8}	0.234	0.0504
Air	0.0225	26.3	24.87
Plastic	1.4×10^{-7}	0.24	0.0621

noteworthy that TNT is a better thermal insulator than soil, and as we describe below, this fact has a significant effect on the signature. Figure 3 shows the surface temperature distribution at the center of the mine as a function of time. The results show good agreement between the codes. In the second set of results (Fig. 4) we investigated spatial variations in the surface temperature distribution. For this purpose we selected a cut through the center of the mine, and we present the temperature distribution as a function of radial distance from the center of the mine at different times. Figure 4 (a) and (b) illustrate the results at times when the mine signatures have maximum negative and positive contrast, respectively. Again, the FEM and integral equation models show good agreement.

C. IR signatures of anti-tank and anti-personal land mines

In this section we present the temperature distributions of several land mines using the FEM model. Numerical results are presented for a rectangular anti-personal (AP) mine, a rectangular anti-tank (AT) mine, and a circular AT mine, all of which are buried under a smooth soil surface. The temporal evolution of the temperature distribution is presented both at the surface and as a function of depth.

First, a square AT mine buried 3.74 cm under a smooth soil surface was considered. The computational volume, which had dimensions 1.4798 m $\times$ 1.4798 m $\times$ 0.3572 m, was discretized by pentahedral elements formed on a rectangular grid of spacing 3.02 cm,

3.02 cm, and 1.88 cm in the x, y, and z directions, respectively. The mine dimensions are similar to those of an M-19 mine, which was modeled as a rectangular parallelepiped with edge sizes 33.2 cm $\times$ 33.2 cm $\times$ 9.4 cm. The mine was assumed to be composed of homogeneous TNT as shown in Fig. 5 (a). The overall computational domain was discretized using 91238 finite elements. This discretization results in 50000 spatial nodes, which also determines the size of the resulting matrix equation. The memory requirement of the FEM code for this discretization is 249 MB. A time increment of 360 seconds was chosen and the FEM solution was observed to reach steady after only two simulated diurnal cycles. The Crank-Nicholson scheme is unconditionally stable, which permits us substantial freedom in choosing the time increment. The time step size was selected to give good resolution over a 24 hour period. The simulation takes about 4.5 hours on a 300 MHz Pentium II machine.

Figure 6 shows the surface temperature distribution over the rectangular AT mine. The results are presented as a sequence of images evaluated at three hour time increments starting from sunrise. As is commonly observed in experimental imagery, contrast changes occur twice daily at the thermal "cross-over" times, and these events appear in our simulations. The results suggest that the physical temperature differences during the day peak at roughly 4.6 K. The temperature difference has a maximum at the center of the mine and diminishes as we move away from the center. The temperature distribution as a function of depth is presented in Fig. 7. The location of the mine is indicated with a rectangle. The temperature distribution is again given every three hours starting at sunrise. The figures present the evolution of heat flow into the ground and the influence of the mine on this heat flow. The surface temperature over the mine is cooler at dawn, and it warms as time proceeds. As a result of its insulating properties (relative to soil) the mine tends to block the flow of heat into the soil, causing the overlying soil to become hotter during the day. During the night, the mine blocks the upward flow of energy in the soil, permitting the layer of soil above the mine to cool more rapidly.

In the second set of simulations, a rectangular AP mine buried at 1 cm depth is investigated. AP mines are typically placed closed to the surface to permit triggering by the small weight of a human. The size of the AP mine was chosen to be comparable to

that of an M-14 mine (a circular mine). The number of spatial elements used was the same as the foregoing AT mine simulation, but the element dimensions were reduced to $0.5 \times 0.5 \times 0.5$ cm^3. The AP mine is a smaller mine and we used a more detailed model for it, as shown in Fig. 5 (b). The outer case of the AP mine is modeled by a 0.5 cm thick plastic. The interior is modeled by TNT and air. This air is intended to represent the gap over the explosive occupied (in part) by the triggering mechanism. Figure 8 shows the surface temperature distribution over this mine, which suggests that temperature differences up to roughly 2 K will be observed. The temperature distribution as a function of depth is presented in Fig. 9 at three hour time increments. The AP mine signature demonstrates characteristics similar to those of the AT mine. Our previous observations regarding thermal cross-over times and obstruction of heat flow by the mine are also applicable here. Two additional observations can be made regarding AP mines. First, although the amplitude of the AP mine signature is smaller than that of an AT mine, the AP mine still gives a detectable signature[2]. Second, the air gap over the explosive has a significant effect on the heat flow as shown in Fig. 9.

We next consider the signature of a simulant circular AT mine [35] buried 6.64 cm. The simulant mine has a diameter of 25 cm and a height of 8.33 cm. The mine is assumed to be composed of TNT as shown in Fig. 5 (c). The computational volume has the dimensions $0.9984 \times 0.9984 \times 0.3154$ m^3 and was discretized with elements of size $2.08 \times 2.08 \times 1.66$ cm^3 in the x, y, and z directions, respectively. This discretization results in 87552 finite elements and 48020 nodes. The surface temperature distribution and the temperature distribution as a function of depth are given in Figs. 10 and 11, respectively. Figure 10 suggests a temperature difference of at most 1.73 K during the day.

In the previous three simulations, the mine has a rectangular shape, whereas in this simulation the mine is circular. We observe, however, that surface signatures of rectangular mines are not rectangular, but rather have somewhat rounded shapes. In Fig. 12, we study this blurring of the mine shape by showing isothermal contours for the circular and rectangular AT mines at several times. The results in Fig. 12 are for 16:00, which corresponds to the median between two thermal cross-over times. Figures 12 (a) and (b)

[2]The noise-equivalent temperature difference in modern IR cameras is less than 0.1 K.

present isothermal contours on the top surface of the buried mines, where the isotherms have the shape as the mine. The small distortion in Fig. 12 (a) from a perfect cylinder is due to approximation of a cylinder by pentahedral elements. Figures 12 (c) and (d) present the isothermal contours at a depth of half the mine burial depth. As we can see, isotherms for the cylindrical mine preserve their circular shape, but isotherms for the rectangular mine are rounded. In Figs. 12 (e) and (f) the isotherms are plotted on the soil surface, where additional rounding of the rectangular mine signature is observed.

Finally, we investigate the dependence of the signatures on burial depth and time. For this purpose, we utilized results from the cylindrical AT mine shown in Fig. 5 (c). In the simulations, the mine is assumed to be buried at depths of 1.66 cm, 3.32 cm, 4.98 cm, and 6.64 cm. The signatures are plotted as a function of distance from the center. Figure 13 (a) shows signatures for mines buried at different depths at 15:00. The signatures have a Gaussian-like spatial dependence, which is little affected by different burial depths, although the amplitudes of the mine signatures are smaller for deeper mines. Figure 13 (b) shows the signatures at 21:00, and the results are in accordance with our foregoing observations. It can be seen that the thermal cross-over time depends on the burial depth. For the mine buried at 6.66 cm the mine signature is almost completely lost, but for other burial depths the signatures are still detectable. Figures 13 (c) and (d) illustrates that (except for a time dependent amplitude) the spatial dependence of the signature is largely independent of time.

V. Summary and Conclusions

A FEM-based three-dimensional thermal model has been developed to investigate the IR signatures of buried land mines. The behavior of this FEM model was compared with other formulations, including an analytical model of 1-D geometries and an integral equation solution (presented in a companion paper). Good agreement is observed in these comparisons. Numerical simulations using different land mines and a variety of depths were presented.

Our work with this model supports the following observations for IR mine detection:

- TNT, the explosive material used in most land mines, is a better insulator than soil. It obstructs the heat flow into and out of the soil. As a result, soil above the mine is

hotter than the background soil during the day, but cooler during the night.

- The simulated peak contrast is consistent with experimental observations. For all observation times the mine-generated contrast has a peak at the center of the mine and it diminishes as we move away from the mine center.
- The peak temperature difference is reduced as the burial depth is increased.
- Two thermal "cross-over" times occur during a day, at which times contrast changes occur in the imagery. The mine signatures are completely lost at thermal cross-over times, and data acquisition should be avoided near those times, but, unfortunately, the time of these events depends on mine composition and burial depth.
- The spatial dependence of a mine signature is (with the exception of an amplitude scaling) largely independent of image acquisition time and burial depth.
- The soil overlying the land mine has the effect of a spatial low-pass filter on the IR image. Signatures of square mines buried several inches will appear similar to those of circular land mines.

An experimental validation of the model described here is of great interest, but it requires a radiometric model to include the effects of natural sources. A radiometric model has been developed, and validation of the combined thermal-radiometric model is presented in a separate work [36]. That work also describes the effects of surface roughness on mine signatures.

Acknowledgments

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

References

[1] C. Stewart, "Summary of mine detection research (U)," Technical Report 1636-TR, US Army Engineer Research and Development Laboratories, Corps of Engineers, Fort Belvoir, VA, Volume 1, DTIC AD320124, May 1960.

[2] D. J. Daniels, D. J. Gunton, and H. E. Scott, "Special issue on subsurface radar," *IEE Proceedings, Part F*, vol. 135, pp. 277–392, 1988.

[3] L. Peters Jr., J. J. Daniels, and J. D. Young, "Ground penetrating radar as a subsurface environmental sensing tool," *Proc. IEEE*, vol. 82, pp. 1802–1820, 1994.

[4] D. Giglio and J. Ralston, "Perspective on underground and obscured target detection and imaging," in *Underground and Obscured-Object Imaging and Detection*, N. K. del Grande, I. Cindrich, and P. B. Johnson, Eds., 1993, vol. 1942, pp. 2–11.

[5] G. Sower and S. Cave, "Perspective on underground and obscured target detection and imaging," in *Detection Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. A. Rigano, Eds., 1993, vol. 1942, pp. 2–11.

[6] N. S. Lewis, M. C. Longergan, E. J. Severin, B. J. Doleman, and R. H. Grubbs, "Array-based vapor sensing using chemically sensitive carbon black-polymer resistors," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, A. C. Dubey and R. L. Barnard, Eds., 1997, vol. 3079, pp. 660–670.

[7] A. D. Hibbs, G. A. Barrall, S. Beevor, L. J. Burnett, K. Derby, A. J. Drew, D. Gregory, C. S. Hawkins, S. Huo, A. Karunaratne, D. K. Lathrop, Y. K. Lee, R. Matthews, S. Milberger, B. Oehmen, T. Petrov, D. C. Skvoretz, S. A. Vierkotter, D. O. Walsh, and C. Wu, "Field test results of a nuclear quadrupole resonance land mine detection system," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, Eds., 2000, vol. 4038, pp. 564–571.

[8] F. Liu, S. L. Tantum, L. M. Collins, and L. Carin, "Statistical signal processing for detection of buried land mines using quadrupole resonance," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, Eds., 2000, vol. 4038, pp. 572–577.

[9] C. G. Don, "Using acoustic impulses to identify a buried nonmetallic object," *127th Meeting of the Acoustical Society of America, M. I. T.*, 1994.

[10] N. Xiang and J. M. Sabatier, "Land mine detection measurements using acoustic-to-seismic coupling," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, Eds., 2000, vol. 4038, pp. 645–655.

[11] Jr. W. R. Scott, G. D. Larson, and J. S. Martin, "Simultaneous use of elastic and electromagnetic waves for the detection of buried land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, Eds., 2000, vol. 4038, pp. 667–678.

[12] I. K. Sendur and B. A. Baertlein, "Integral equation solution for thermal signatures of buried land mines," *submitted to IEEE Trans. Geo. Rem. Sens.*

[13] L. A. LeSchack and N. K. Del Grande, "A dual-wavelength thermal infrared scanner as a potential airborne geophysical exploration tool," *Geophys.*, vol. 41, no. 6, pp. p. 1318–1336, 1976.

[14] Y. H. L. Janssen, A. N. de Jong, H. Winkel, and F. J. M. van Putten, "Detection of surface laid and buried mines with IR and CCD cameras, an evaluation based on measurements," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, R. L. Barnard, C. J. Lowe, and J. E. McFee, Eds., 1996, vol. 2765, pp. 448–459.

[15] G. Maksymomko, B. Ware, and D. Poole, "A characterization of diurnal and environmental effects on mines and the factors influencing the performance of mine detecting ATR algorithms," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, Eds., 1995, vol. 2496, pp. 140–151.

[16] J. R. Simard, "Improved lanmine detection capability(ILDC): Systematic approach to the detection of buried mines using passive IR imaging," in *Detection and Remediation Technologies for Mines and Minelike Targets*, A. C. Dubey, R. L. Barnard, C. J. Lowe, and J. E. McFee, Eds., 1996, vol. 2765, pp. 489–500.

[17] K. Russel, J. McFee, and W. Sirovyak, "Remote performance prediction for infrared imaging of buried mines," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, A. C. Dubey and R. L. Barnard, Eds., 1997, vol. 3079, pp. 762–769.

[18] P. Pregowski and W. Swiderski, "Using of comparison method in IR thermal detection of buried mines," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, A. C. Dubey, J. F. Harvey, and J. T. Broach, Eds., 1998, vol. 3392, pp. 1241–1248.

[19] P. Pregowski, W. Swiderski, R. T. Walczak, and K. Lamorski, "Buried mine and soil temperature prediction by numerical model," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, Eds., 2000, vol. 4038, pp. 1392–1403.

[20] N. K. Del Grande, K. W. Dolan, P. F. Durbin, M. R. Gorvad, B. T. Kornblum, D. E. Perkins, D. J. Schneberk, and A. B. Shapiro, "Three-dimensional thermal imaging of structural flaws by dual-band infrared computed tomography," in *Underground and Obscured Object Imaging and Detection*, N. K. Del Grande, I. Cindrich, and P. B. Johnson, Eds., 1993, vol. 1942, pp. 207–215.

[21] K. Watson, "Geologic application of thermal infrared images," *Proc. IEEE*, vol. 63, no. 1, pp. 128–137, Jan. 1975.

[22] A. W. England, J. F. Galantowicz, and M. S. Schretter, "The radiobrightness thermal inertia measure of soil moisture," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, no. 1, pp. 132–139, Jan. 1992.

[23] A. B. Kahle , "A simple thermal model of the earth's surface for geologic mapping by remote sensing," *Journal of Geophysical Research*, vol. 82, pp. 1673–1680, 1977.

[24] A. W. England, "Radiobrightness of diurnally heated, freezing soil," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, no. 4, pp. 464–476, July 1990.

[25] Y.-A. Liou and A. W. England, "A land-surface process/radiobrightness model with coupled heat and moisture transport in soil," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, no. 1, pp. 273–286, Jan. 1998.

[26] Y.-A. Liou and A. W. England, "A land-surface process/radiobrightness model with coupled heat and moisture transport for freezing soils," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, no. 2, pp. 669–677, March 1998.

[27] A. Berk, L. S. Bernstein, and D. C. Robertson, "MODTRAN a moderate resolution model for lowtran7," *Tech. Report GL-TR-89-0122, Air Force Geophysical Laboratory, Hanscorn AFB, MA, October 21 1989.*

[28] B. A. Baertlein and I. K. Sendur, "Thermal infrared detection of buried land mines," Progress Report to US Army Research Office on Contract 97-SC-ARO-1015, The Ohio State University ElectroScience Laboratory, Columbus, OH, December 1997.

[29] K. Ya. Kondratyev, *Radiation in the Atmosphere*, Academic Press, New York, NY, 1969.

[30] J. E. Akin, *Finite Elements for Analysis and Design*, Academic Press, San Diego, CA, 1994.

[31] J. N. Reddy and D. K. Gartling, *The Finite Element Method in Heat Transfer and Fluid Dynamics*, CRC Press, Inc., Boca Raton, FL, 1994.

[32] L. W. Wolfe and G. J. Zissis, *The Infrared Handbook*, revised edition, Environmental Research Institute of Michigan (ERIM), Ann Arbor, MI, 1989.

[33] W. M. Rohsensow, J. P. Hartnett, and E. N. Ganic (eds.), *Handbook of Heat Transfer Fundementals*, 2nd ed., McGraw Hill, New York, NY, 1985.

[34] R. G. Reeves, A. Anson, and D. Landen (eds.), *Manual of Remote Sensing*, 1st ed., American Society of Photogrammetry, Falls Church, VA, 1973.

[35] Anon., "Simulant mines (sims)," Tech. report, US Army Countermine and Demolition (Countermine Division), Fort Belvoir, VA, October 21 1998.

[36] I. K. Sendur and B. A. Baertlein, "Radiometric model for the infrared signature of a land mine buried under a rough surface," *submitted to IEEE Trans. Geo. Rem. Sens.*

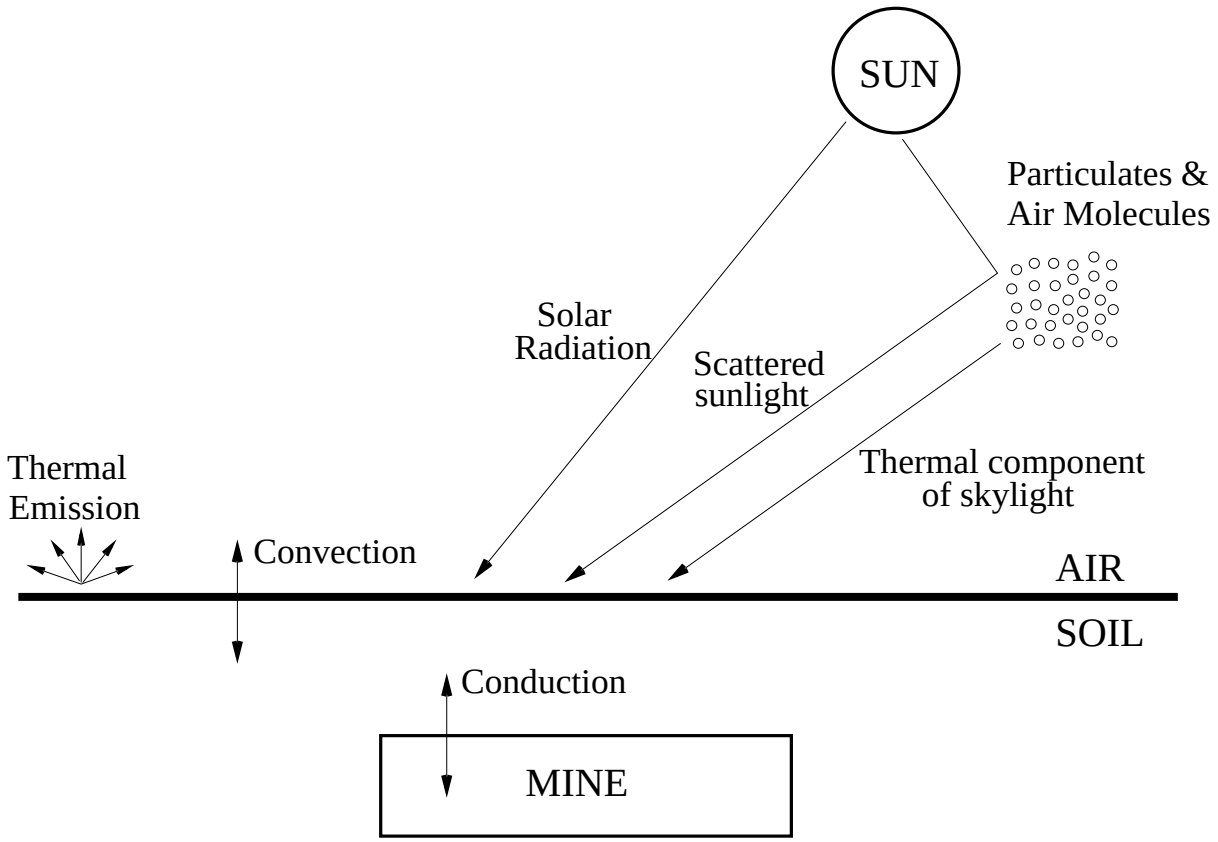


Fig. 1. The heat transfer mechanisms in the soil, at the soil-mine interface, and at the soil-air interface.

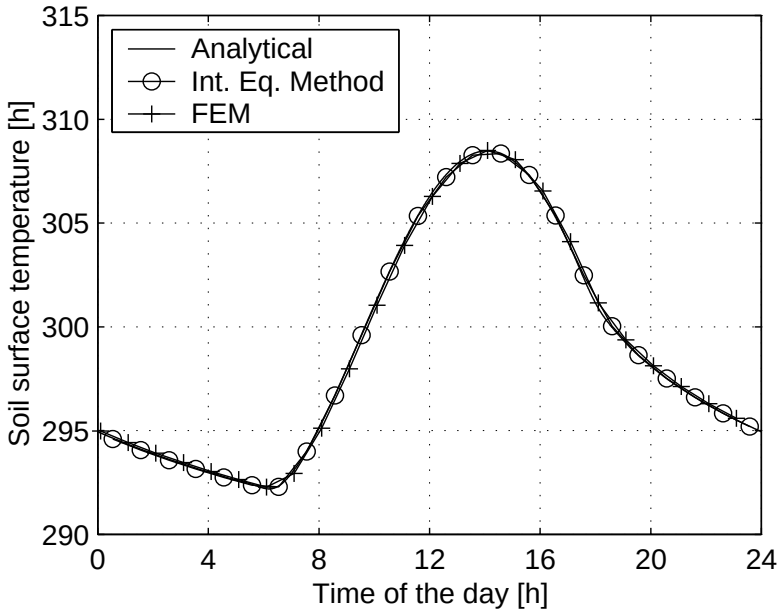


Fig. 2. The surface temperature over homogeneous soil as a function of time using an analytical solution, the integral equation solution described herein, and a FEM solution.

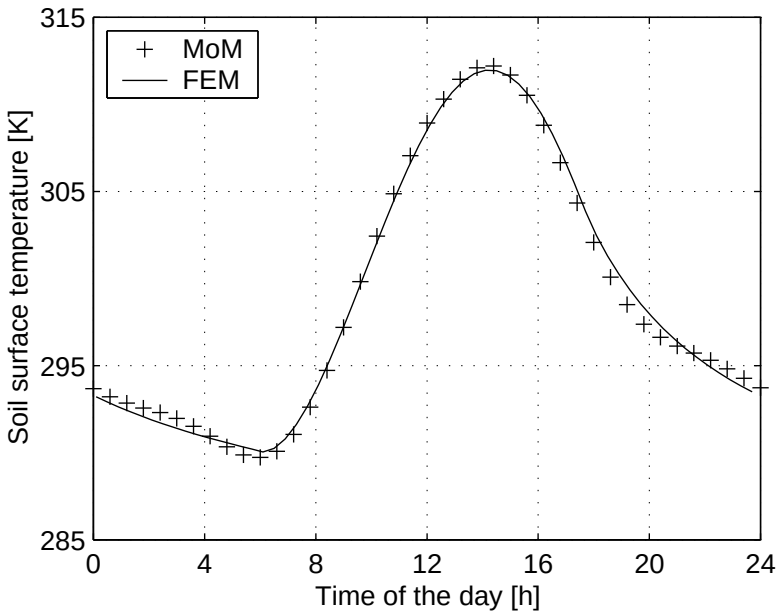
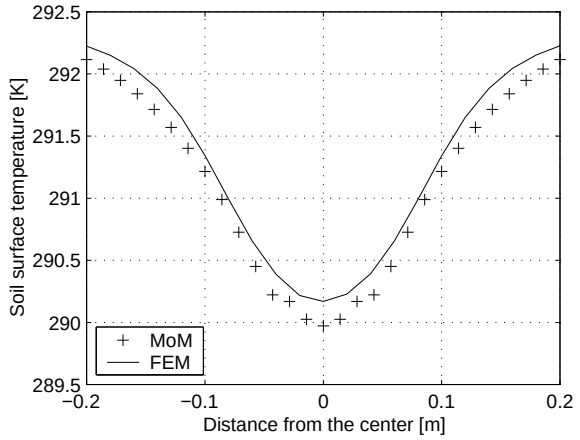
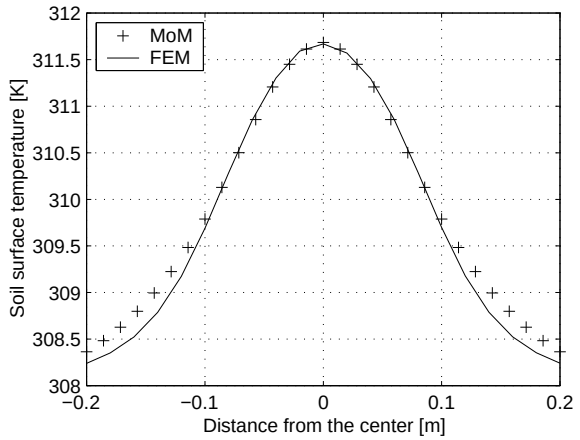


Fig. 3. The surface temperature distribution over the center of a cylindrical mine using the integral equation solution and a finite element solution.



(a)



(b)

Fig. 4. The spatial dependence of the surface temperature distribution: (a) When the mine signature has a maximum negative contrast, (b) When the mine signature has a maximum positive contrast.

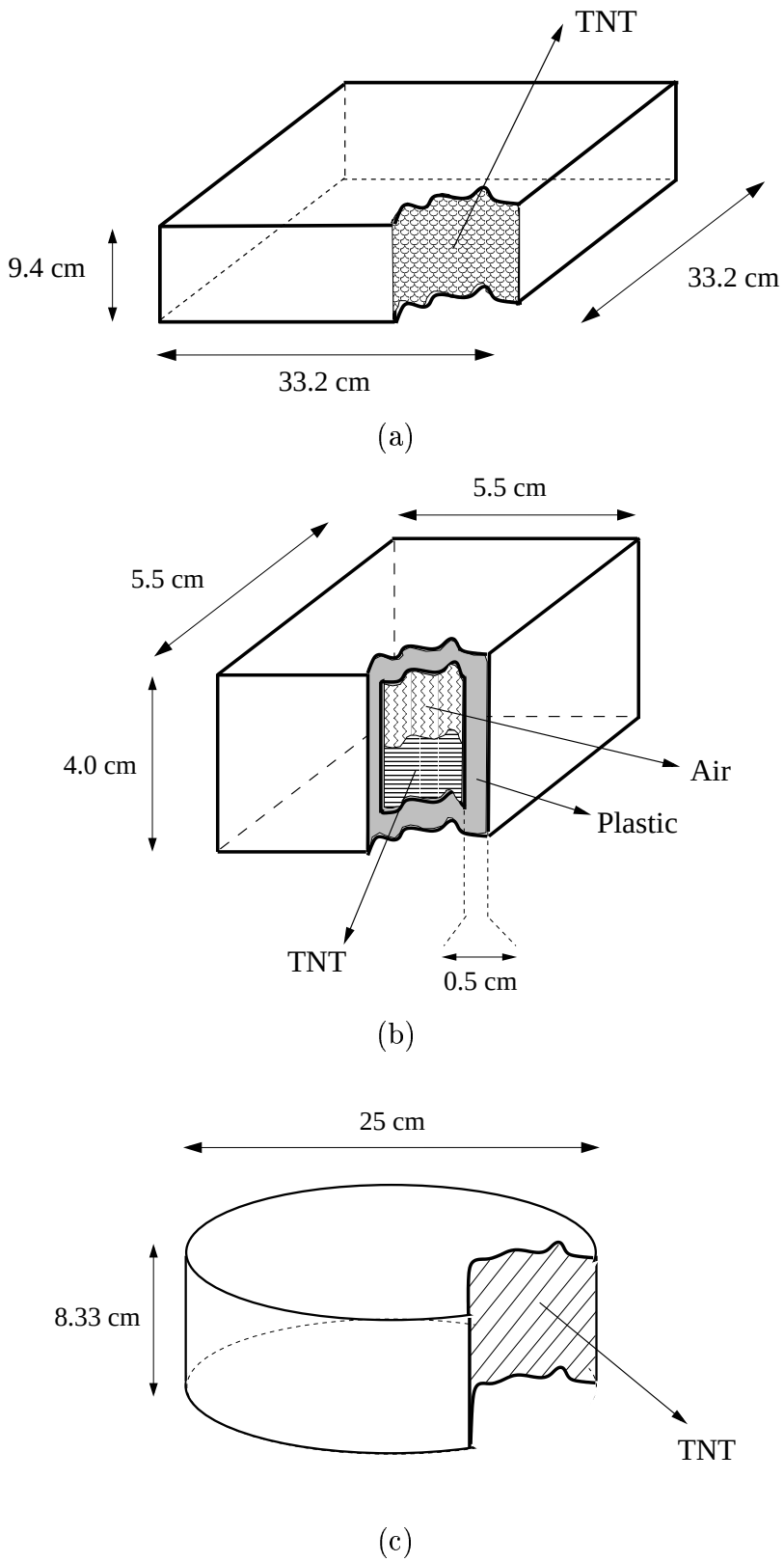


Fig. 5. Models for the surrogate mines used in the simulations (a) Rectangular anti-tank mine. (b) Rectangular anti-personal mine. (c) Cylindrical anti-tank mine.

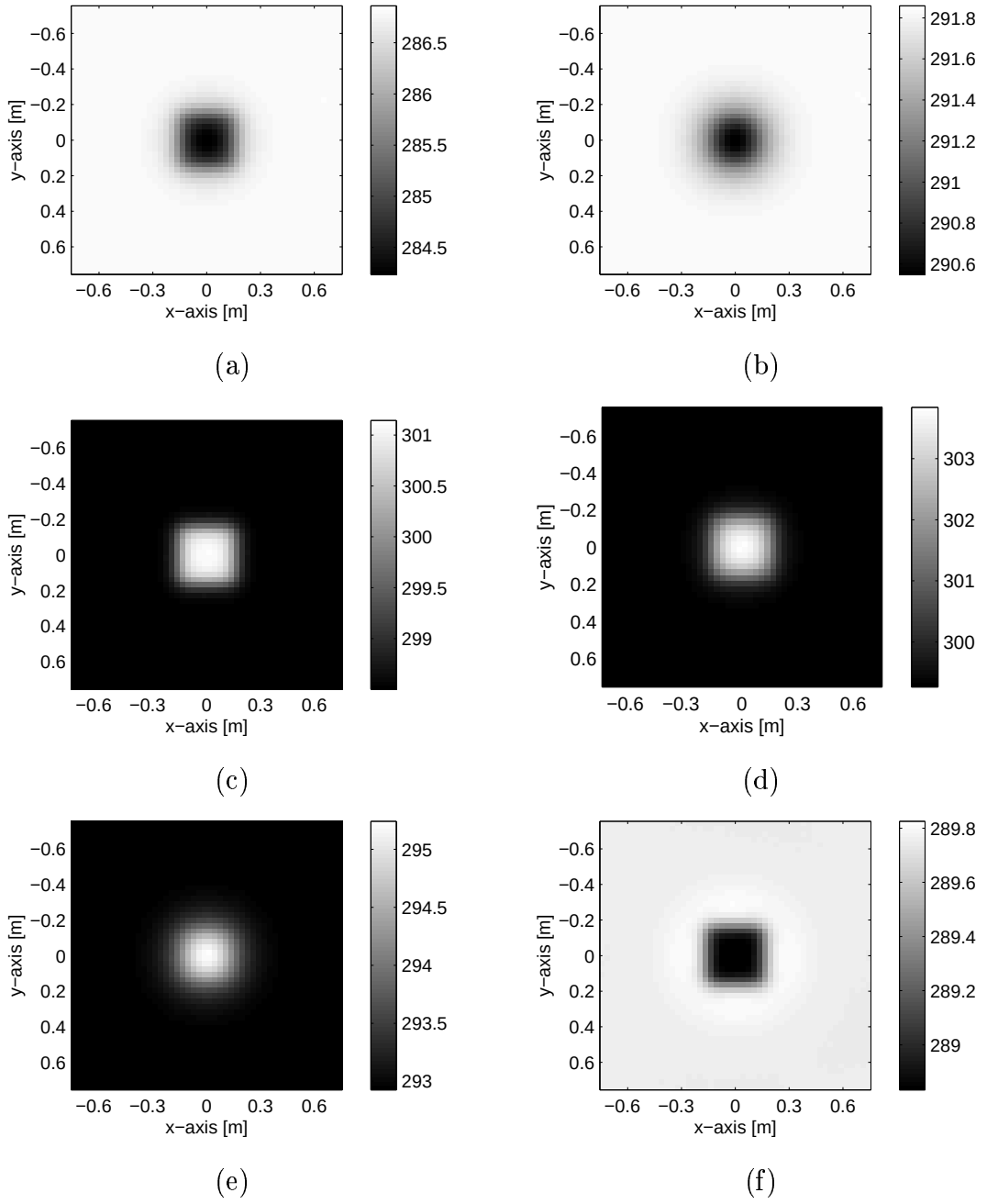


Fig. 6. Simulation of the soil surface temperature difference over a rectangular AT mine buried 3.74 cm under a smooth soil surface. The temperature distribution is evaluated (a) at dawn, (b) 3 hours after sunrise, (c) at noon, (d) 3 hours after noon, (e) at sunset, and (f) 3 hours after sunset.

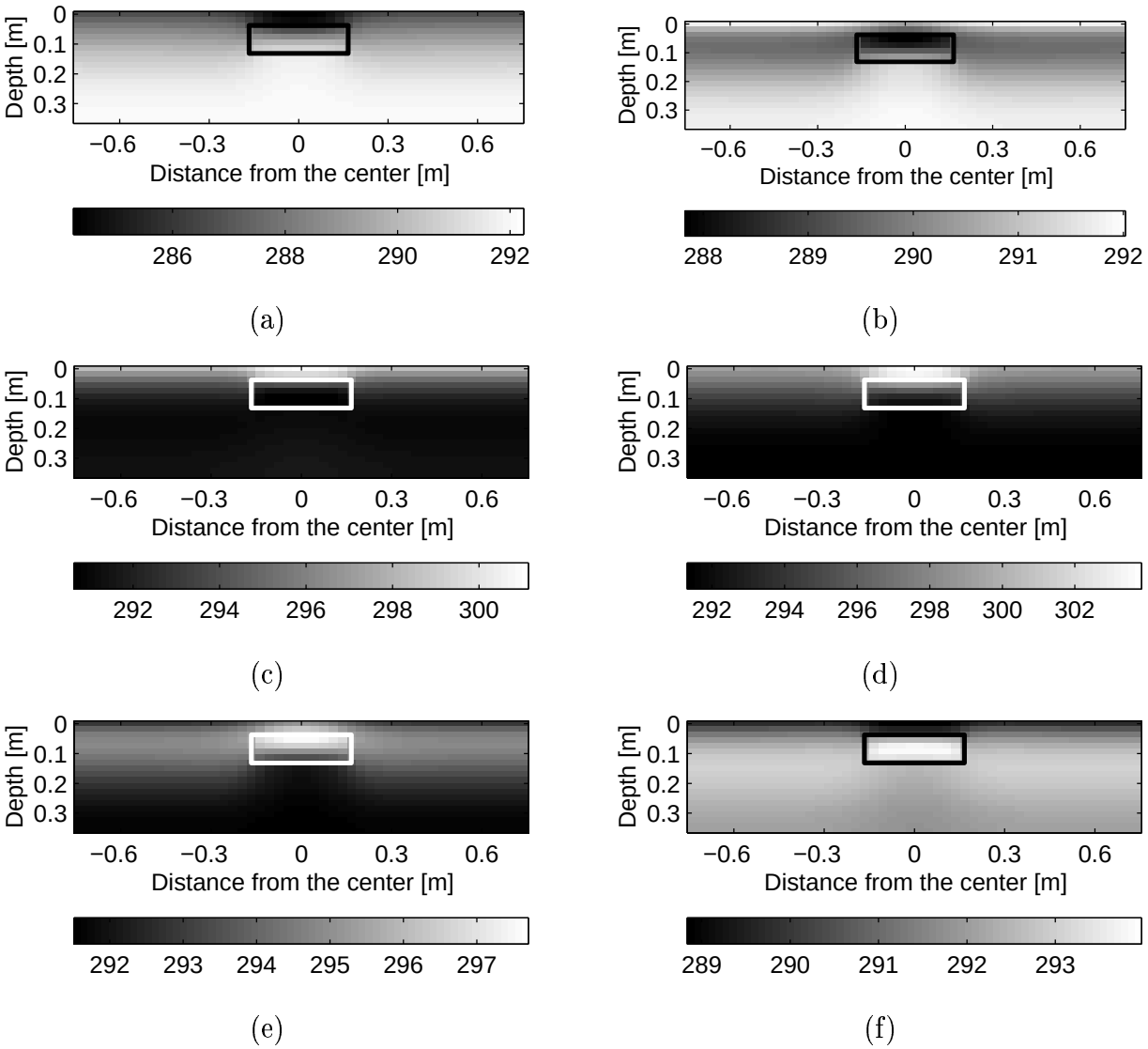


Fig. 7. Simulation of the temperature distribution at depth taken through the middle of the rectangular AT mine buried 3.74 cm under a smooth soil surface. The temperature distribution is evaluated (a) at dawn, (b) 3 hours after sunrise, (c) at noon, (d) 3 hours after noon, (e) at sunset, and (f) 3 hours after sunset.

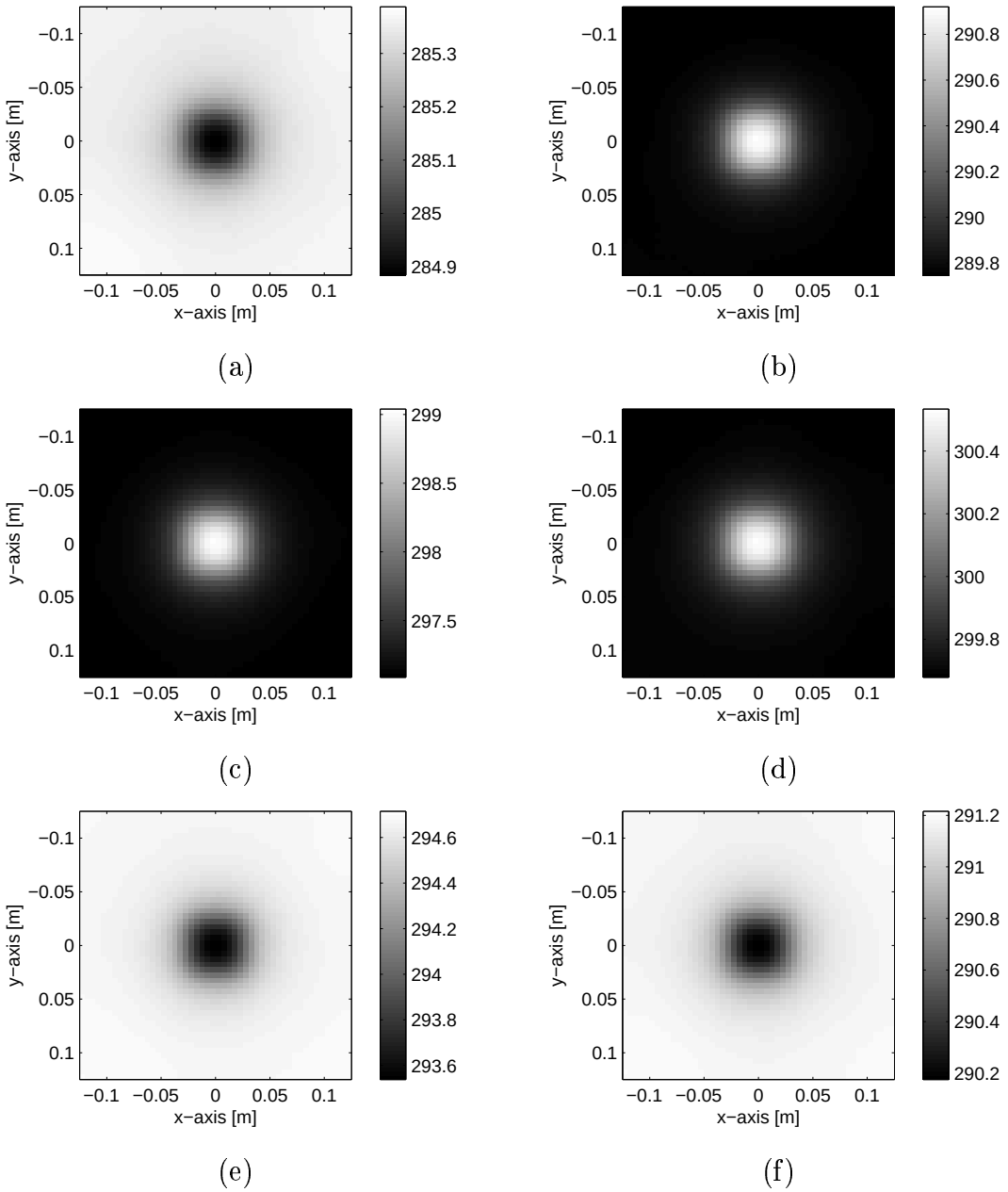


Fig. 8. Simulation of the soil surface temperature difference over a rectangular AP mine buried 1 cm under a smooth soil surface. The temperature distribution is evaluated (a) at dawn, (b) 3 hours after sunrise, (c) at noon, (d) 3 hours after noon, (e) at sunset, and (f) 3 hours after sunset.

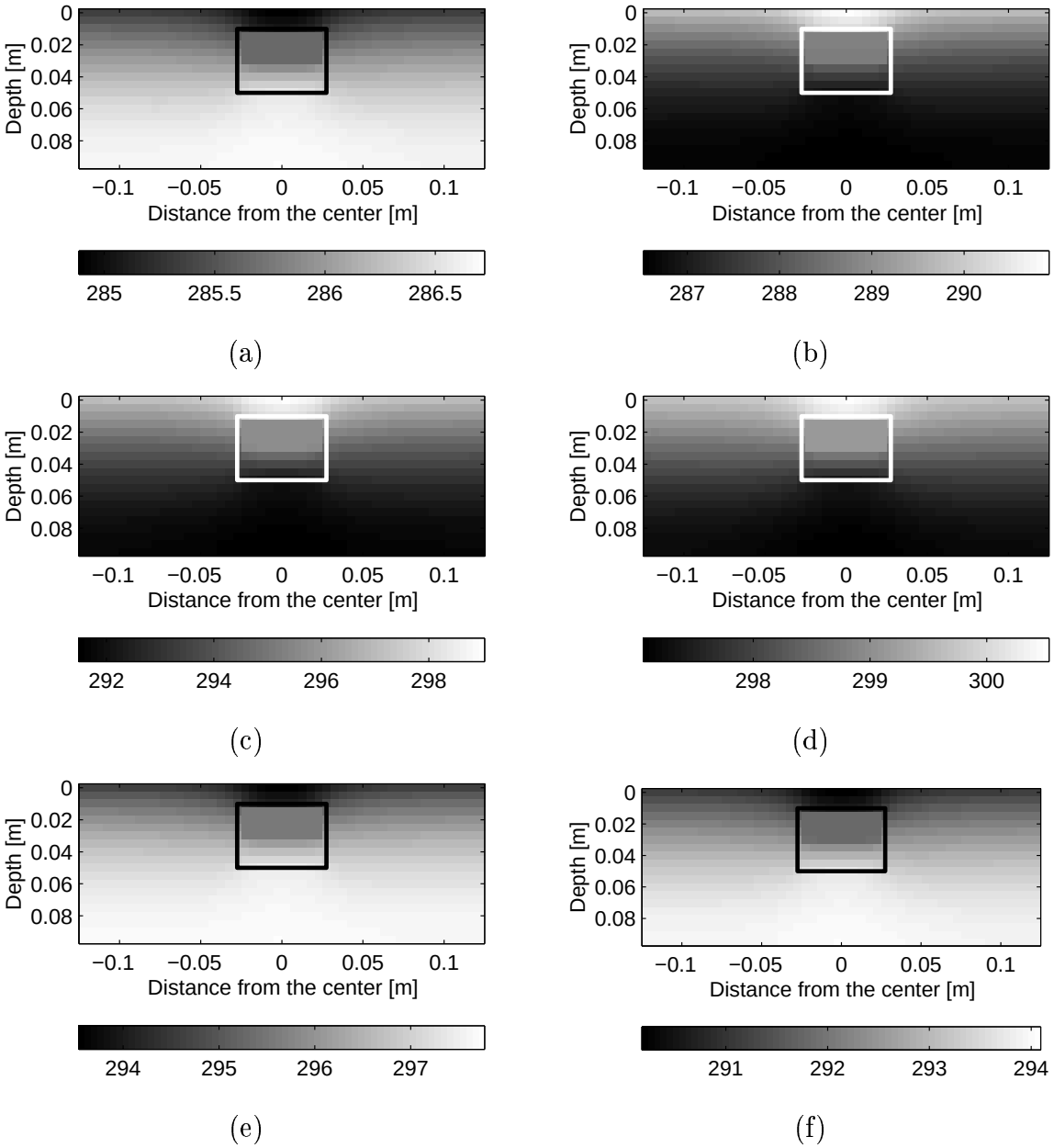


Fig. 9. Simulation of the temperature distribution at depth taken through the middle of the rectangular AP mine buried 1 cm under a smooth soil surface. The presence of the air gap is evidenced by the different temperatures in the top and bottom halves of the mine. The temperature distribution is evaluated (a) at dawn, (b) 3 hours after sunrise, (c) at noon, (d) 3 hours after noon, (e) at sunset, and (f) 3 hours after sunset.

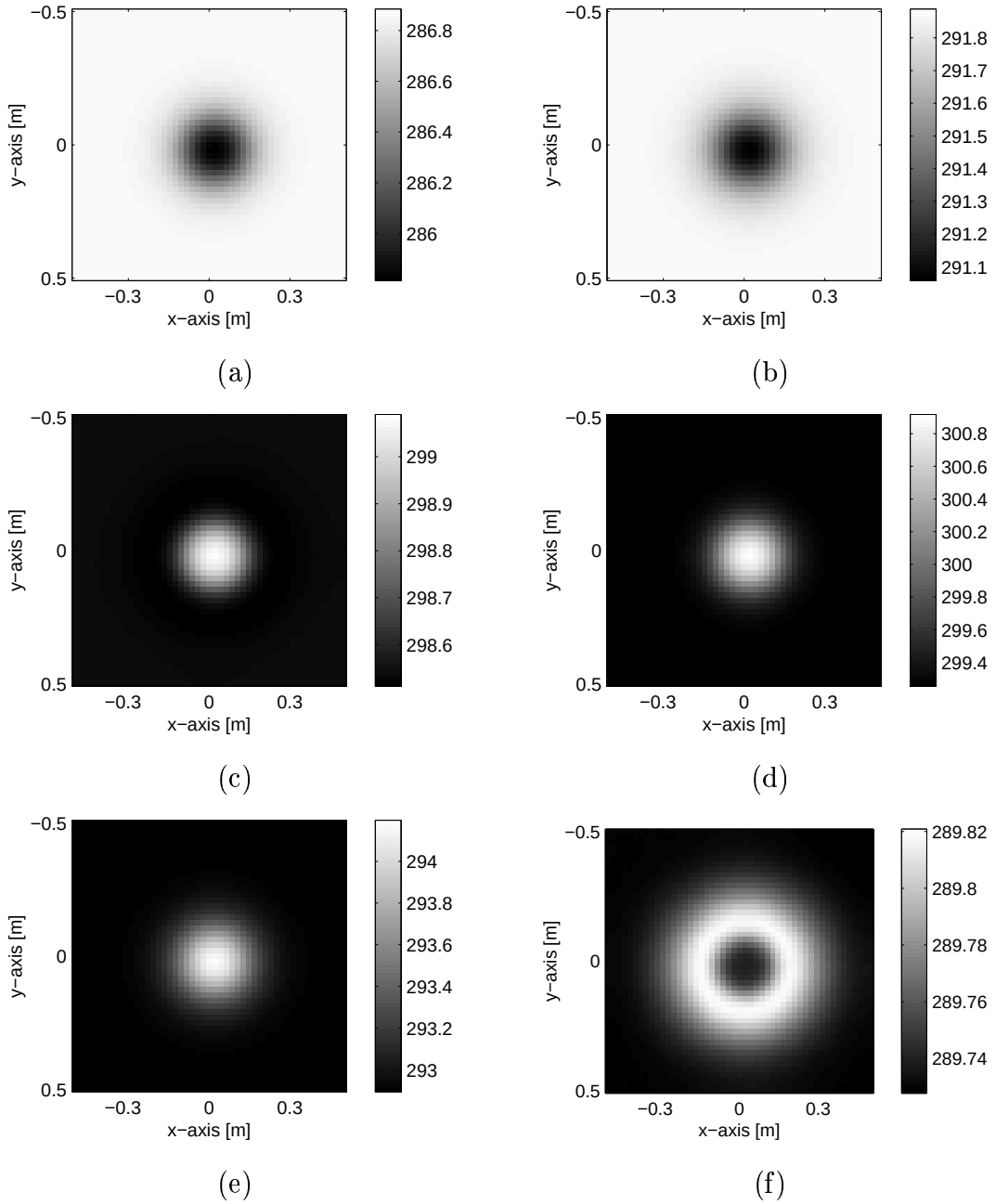


Fig. 10. Simulation of the soil surface temperature difference over a cylindrical anti-tank mine buried 6.66 cm under a smooth soil surface. The temperature distribution is evaluated (a) at dawn, (b) 3 hours after sunrise, (c) at noon, (d) 3 hours after noon, (e) at sunset, and (f) 3 hours after sunset.

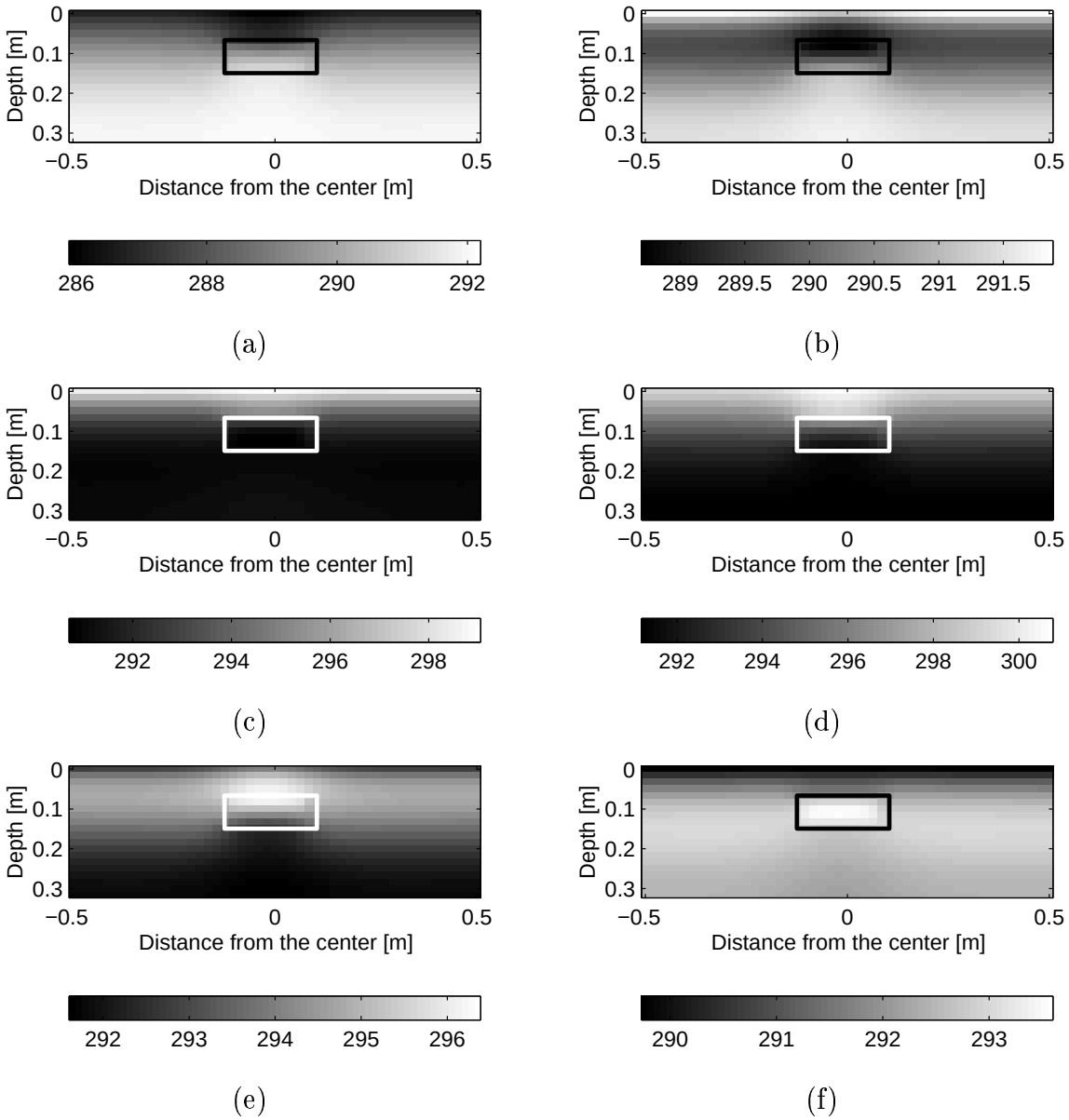


Fig. 11. Simulation of the temperature distribution at depth taken through the middle of the cylindrical anti-tank mine surrogate buried 6.66 cm under a smooth soil surface. The temperature distribution is evaluated (a) at dawn, (b) 3 hours after sunrise, (c) at noon, (d) 3 hours after noon, (e) at sunset, and (f) 3 hours after sunset.

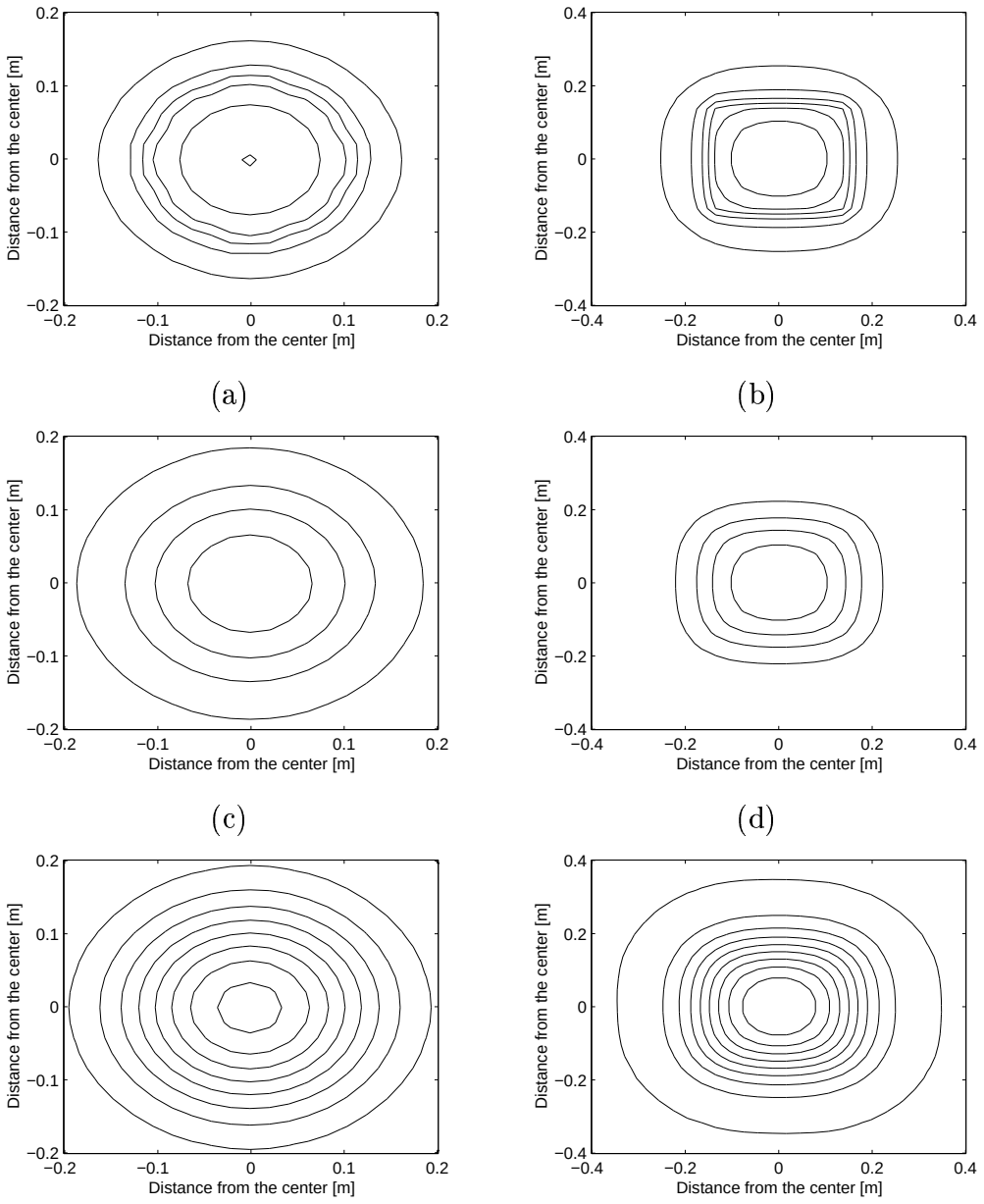


Fig. 12. Isothermal contours for cylindrical and rectangular anti-tank mines at 16:00 plotted at different depths (a) On the cylindrical mine surface. (b) On the rectangular mine surface. (c) At a depth of half of the mine burial depth for cylindrical mine. (d) At a depth of half of the mine burial depth for rectangular mine. (e) On the soil surface over the cylindrical mine. (f) On the soil surface over the rectangular mine.

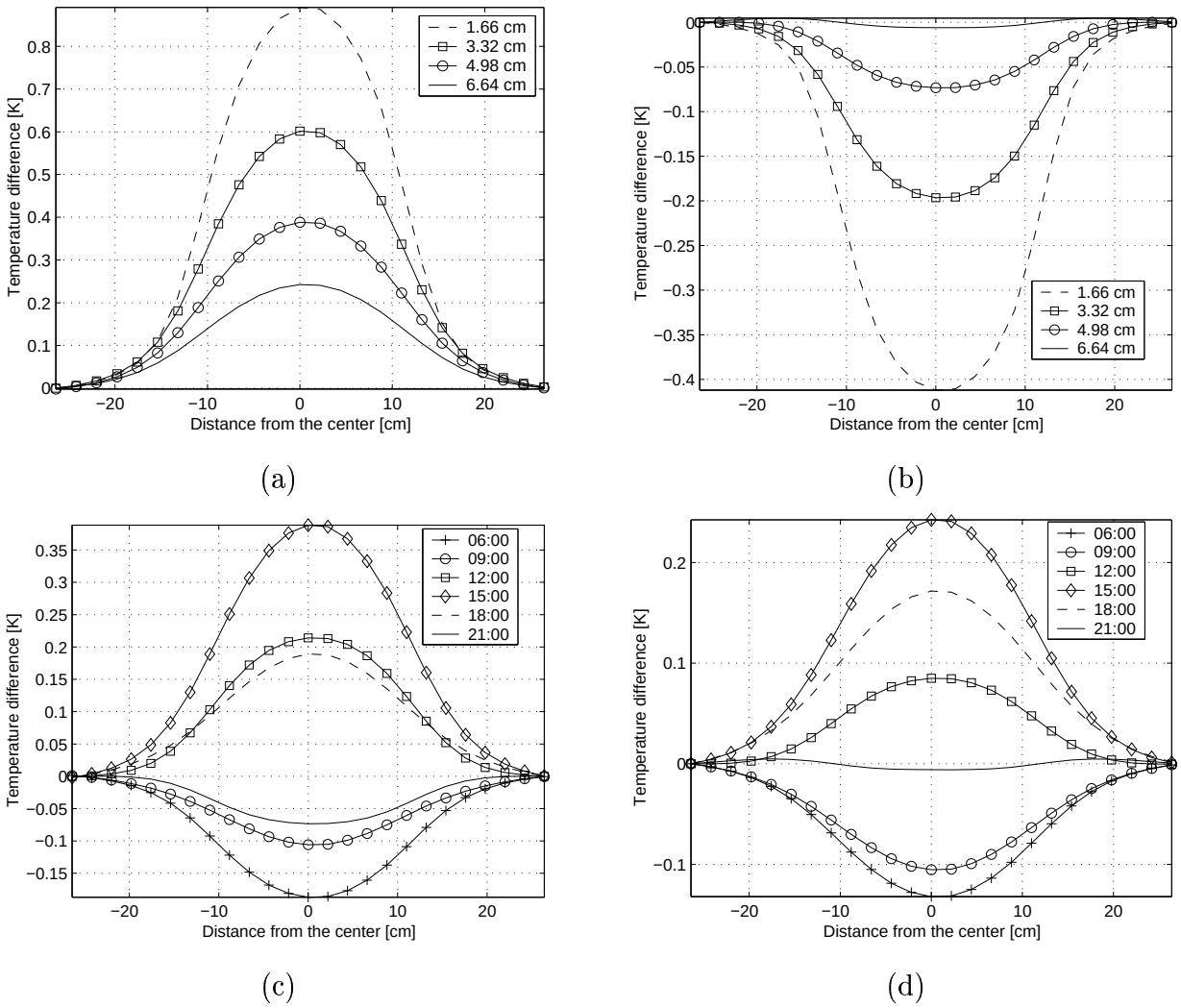


Fig. 13. The mine signatures as a function of distance from the center (a) Dependence of mine burial depth at 15:00. (b) Dependence of mine burial depth at 21:00. (c) Dependence of time of day for a mine buried at 3.32 cm. (d) Dependence of time of day for a mine buried at 6.64 cm.

Fusion of Data From Forward-Looking Demining Sensors

B. A. Baertlein*, W.-J. Liao and D.-H. Chen

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

Phone: (614) 292-0076

Fax: (614) 292-7297

E-mail: baertlein.1@osu.edu

Category: Session 38

ABSTRACT

We present fusion algorithms and detection performance for two sensors of buried land mines capable of forward-looking operation. The first is a laser Doppler vibrometer (LDV) sensor of acoustically induced soil vibrations developed by the University of Mississippi. The second is a wideband, forward-looking, synthetic aperture ground-penetrating radar developed by SRI. We demonstrate the performance of the individual sensors and fusion algorithms over a data set acquired at Yuma Proving Grounds, AZ. During that collection the LDV sensor was operating in a down-looking mode, but its ability to collect forward-looking data has been confirmed in other tests. The amount of fusible data available for these sensors is small, which limits the sophistication and performance of fusion algorithms used with these data. We observed that the GPR and LDV sensors are complementary to a degree. The GPR sensor is significantly better at detecting metal mines, while the LDV sensor is somewhat better at detecting plastic mines. A classical decision-level "OR" fusion algorithm is shown to be effective for this sensor suite.

1. INTRODUCTION

Physical limitations force many mine-detecting sensors to operate in close proximity to the soil, but for safety reasons greater sensor standoff is preferable. In response to this need, there is increased interest in forward-looking sensors, particularly for buried mine detection.

Several technologies have shown an ability to detect buried mines at significant standoff distances. It has long been recognized that passive infrared (IR) sensors will detect buried mines under certain conditions, but IR sensors have several weaknesses. Perhaps the most serious limitation of those sensors is that the signature of a buried mine is detectable only at certain times of the day and under favorable environmental conditions. This fact makes IR sensors unattractive in some roles.

Ground penetrating radar (GPR) is a second sensor capable of detecting buried mines from a significant standoff distance. When the radar is operated in a forward-looking geometry, the strong specular surface reflection that plagues down-looking GPRs is greatly reduced. A number of systems of this type have been demonstrated over the years. Recently, SRI has produced a forward-looking ground-penetrating radar (FLGPR) for demining.

Another sensor that has recently shown promise for buried mine detection is remote sensing of acoustically stimulated ground motion. An implementation of this concept using a laser Doppler vibrometer (LDV) for detecting surface motion has recently been demonstrated by the University of Mississippi.

The stressing performance requirements of demining motivate an investigation of multi-sensor data fusion. Fusion offers several potential benefits including an increase in detection rate, a reduction in false alarm rate, and an ability to operate in more diverse environments.

In this paper we describe the application of sensor fusion to the FLGPR and LDV sensors. We begin in Section 2 with a brief description of the sensors and the data they produce. In Section 3 we describe algorithms for processing data from these sensors. Recently, both sensors were used to acquire data at Yuma Proving Grounds (YPG), AZ. Single sensor performance, fusion algorithms, and fused sensor performance are described in Section 4. The resulting fusible data set is small, which has serious implications for fusion (described in Section 4.1). Moreover, sensor performance characteristics estimated from a small data set do not necessarily represent true sensor performance.

Nonetheless, we show that some relatively simple fusion techniques are effective. Concluding remarks appear in Section 5.

2. TEST SITE AND SENSOR SUITE

2.1. Test Site: Yuma Proving Grounds

During March 2000 the FLGPR and LDV sensors were operated at the YPG mine lanes. Those lanes comprise four 300 meter by 5 meter regions, the content of which is publicaly available. A fourth region of equal size is reserved for blind tests. The lanes contain solely AT mines of types TM62M, TM62P, and VS2.2. Lanes 1 and 2 each contain 15 TM62Ps, 15 TM62Ms, 15 VS2.2s, and 9 filled holes. Lane 3 contains 20 mines of each type and no filled holes. In total, the three lanes contain 50 of each mine type, of which 11 are buried at each of the depths zero (flush), 5 cm, 10 cm and 15 cm. The remaining 9 mines of each type are placed on the surface.

2.2. SRI Forward Looking Ground-Penetrating Radar (FLGPR)

The SRI FLGPR has been described previously.[1,2] Key features of the system as configured for the 2000 YPG data collection are a 5 cm resolution in range and cross-range. This resolution is achieved in the range dimension through a 2.7 GHz bandwidth (0.3-3.0 GHz). For the cross-range dimension, a pair of quad-ridged horns (one for transmit and another for receive) are mechanically scanned through a 4 meter synthetic aperture. SAR processing is used to image a rectangular region eight meters wide and 23 meters long (seven to 30 meters downrange from the radar). Polarimetric data (HH, VV, and HV) are acquired. The radar is mounted on a vehicle (a delivery van), which is normally moved down-range in increments of a few meters between data acquisitions. For each data collection the van location (i.e., its `x`, `y`, and `z` coordinates) and orientation (roll, pitch, and yaw) were recorded and used to derive ground coordinates from image coordinates. An empirically derived amplitude correction was used to boost the received signal with increasing range so as to generate a constant return amplitude with range.

As noted in Section 1, a forward-looking radar offers potential benefits both in operator safety and in reduced surface clutter. The latter limits the ability of down-looking radars to detect near-surface targets. The forward-looking geometry, however, has some drawbacks. In particular, there is a decrease in the amount of radar energy that impinges on a buried target. In addition, scattering from surface clutter is distributed in range for a forward looking sensor, which can cause a surface clutter return to be received at the same time as the return from a buried target.

The FLGPR acquired 541 images of Lanes 1 through 4 at YPG, which include multiple views of all 150 known mines. In using those data we restricted our attention to images in which the mine of interest was viewed at a range of approximately 15 meters. Both the VV and HH returns were considered in our analysis. To date we have not examined the cross-polarized response.

2.3. U. Mississippi Acoustically Driven Laser Doppler Vibrometer

The UM LDV system comprises a pair of acoustical speakers directed at the earth.[3–5] The speakers are driven by pseudo-random noise, producing a broadband acoustic excitation at the earth's surface. Coupling of this acoustic energy into seismic vibrations is detected by a raster-scanned LDV. The data produced by this sensor comprise a square array 16 pixels on a side acquired on a uniform linear spatial grid with nominal overall dimensions of 1 meter square. At each spatial position a Doppler frequency response is measured, producing a spatial estimate of the surface velocity as a function of frequency. The measured frequency band ranges from a few tens of Hz to a few hundred Hz. The UM collection yielded 223 LDV images, which comprise samples of 96 known mines, 29 unique clutter sites (of which eight were filled holes at known locations) and 29 sites in the blind test area (Lane 4). During the Yuma collection the LDV was operated in a down-looking geometry, but its ability to collect data in a forward-looking geometry has been demonstrated previously in tests at other sites.

3. SINGLE-SENSOR ALGORITHMS

Before it is possible to fuse sensor data having different formats, the data must be reduced to a common form. In our work the sensor data are reduced to a small number of features. In this section we describe that processing for each sensor.

3.1. FLGPR

In this preliminary evaluation, we used as features the VV and HH signal-to-clutter ratios estimated by SRI. These values are determined as the ratio of (1) the brightest pixel within a specified distance from the expected target location and (2) the average of 36 clutter pixels taken from a region offset in cross-range from the true mine position. Since the resulting data are a ratio of a peak to an average, they are necessarily greater than or equal to unity.

The FLGPR collected multiple images of each mine, and in this work we have winnowed those images using the following approach: Preference was given to data from a range of 15 meters, since the FLGPR signal-to-clutter ratio is range dependent, and the most effective detection typically occurs at that range. When multiple S/C values were available for a given mine at a given range, the average value was used. When data from a 15 meter range were not available, data from 20 meters or 10 meters were used instead.

3.2. LDV

In processing the LDV data, it is first necessary to estimate and remove background trends observed in the images. Such trends are thought to arise from the acoustic wave sweeping over the region being interrogated. A low-order polynomial model is fitted to the data from each frequency band, and the unknown coefficients are determined via least squares. The procedure is described and illustrated in a companion paper,[6] and it will not be repeated here.

The residual image is then processed using a CFAR algorithm[7]: a spatial shape for the mine signature is assumed, and a spatially matched filter is applied at each point in the image. The unknown spectral amplitude of the mine signature is computed using a maximum likelihood procedure which, for the assumption of Gaussian clutter, is equivalent to the solution of a set of linear equations. The residual error variance in the filter output is compared to a local estimate of the image clutter variance. An estimate of the log-likelihood ratio is obtained, which is used for detection.

In addition to the log-likelihood ratio, we also found that the image variance was a good discriminator of mines. High variance was found to correlate well with the presence of a mine.

4. SINGLE-SENSOR AND FUSED PERFORMANCE

The sensor data described above were used to generate ROC curves for single-sensor and fused detections. In this section we review the performance of the individual sensors, describe the fusion algorithm, and demonstrate the performance benefits of fusion.

4.1. Data Set Limitations

Although a significant amount of data was collected with both sensors, the data suitable for fusion are surprisingly limited. Only data acquired at the same location can be combined through fusion. We refer to those data as "fusible." The coordinates of the FLGPR data are known from the position and orientation of the vehicle. The radar positions can also be checked by examining the positions of known fiducials in the imagery. During the LDV collection, the mine positions were known and imagery collected over known mines can be combined with the relevant FLGPR data for fusion. The fusible target set comprises 33 TM62M mines, 33 TM62P mines, and 29 VS2.2 mines. The locations of the LDV clutter data, however, were not recorded. As a result, the fusible clutter data comprise the responses of eight filled holes for which the locations were known. This small set of non-targets greatly limits our ability to train and test classifiers and fusion algorithms, but it is sufficient to make some qualitative tests of fusion and detection concepts. It is also important to note that because the size of the clutter set is small, performance estimates shown here are not necessarily representative of the sensor's true performance.

The size of the data set also limits the techniques available for classifier design and fusion. With two sensors having two features each, it is not feasible to train feature-level fusion algorithms for the resulting four-dimensional space. We adopt a soft decision-level fusion approach, in which the outputs of the LDV and GPR detectors (prior to thresholding) are input to the fusion algorithm.

4.2. FLGPR Performance

The performance of the FLGPR for all fusible mines in the YPG test site is given in Figure 1. We present separate ROC curves for the VV and HH signal-to-clutter ratios (SCRs). A study of these data reveals that the VV and HH SCRs are correlated, and fusion of them provides little benefit. In what follows, we use the HH output as the FLGPR detector.

It is common in theoretical analyses of detection and classification to assume that data from a given sensor comprise identically distributed random variables. That assumption is violated for the FLGPR. To see this, consider the ROC curves shown in Figures 2 and 3 for metal (TM62M) and plastic (TM62P and VS2.2) mines respectively. For this small fusible data set the FLGPR has limited success in detecting mines in general (cf. Figure 1) and very poor performance for plastic mines (cf. Figure 3), but its performance on metal mines is excellent (Figure 2). The distribution of S/C values for metal mines differs significantly from that of plastic mines, and this fact can be used to advantage in what follows.

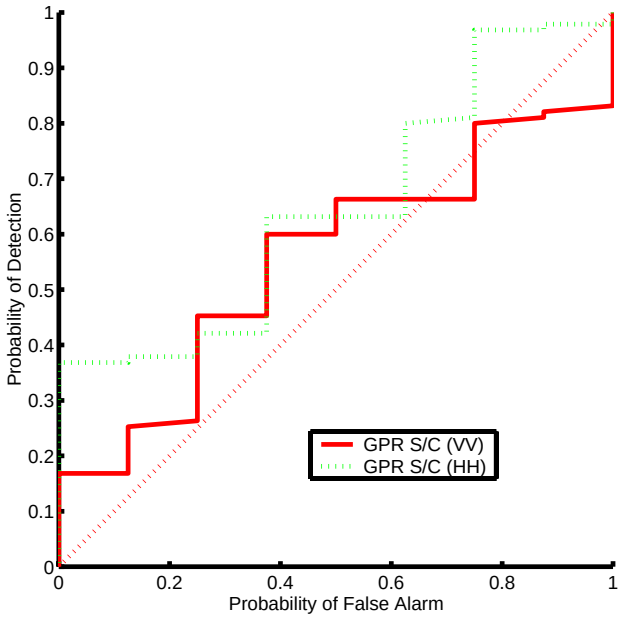


Figure 1. ROC curve for the FLGPR on metal and plastic mines.

4.3. LDV Performance

The performance of the LDV on all samples at YPG is shown in Figure 4. The figure shows ROC curves derived from the individual features (CFAR output and image variance) and also a classifier output that uses both features. The classifier comprises a weighted sum of nonlinearly remapped features. Specifically, both the CFAR value (feature f_1) and the image variance (feature f_2) are scaled to the range [0,1]. The fused detector feature (f_F) is given by

$$f_F = w_1 f_1 + w_2 f_2^p \tag{1}$$

where weights w_1 and w_2 and power law p are selected to maximize performance. The weights w_1 and w_2 account for the different performance of these two features, while the power-law remapping (on the normalized interval [0,1]) permits us to define simple nonlinear boundaries in the (f_1, f_2) feature space. Alternatively, we can think of the power-law mapping as a method for rescaling the features to comparable threshold levels.

The responses of the LDV sensor for metal and plastic mines are shown in Figures 5 and 6. It can be seen that the response of this sensor is somewhat better for plastic mines. The variance features is significantly more effective for plastic mines and small false alarm rates.

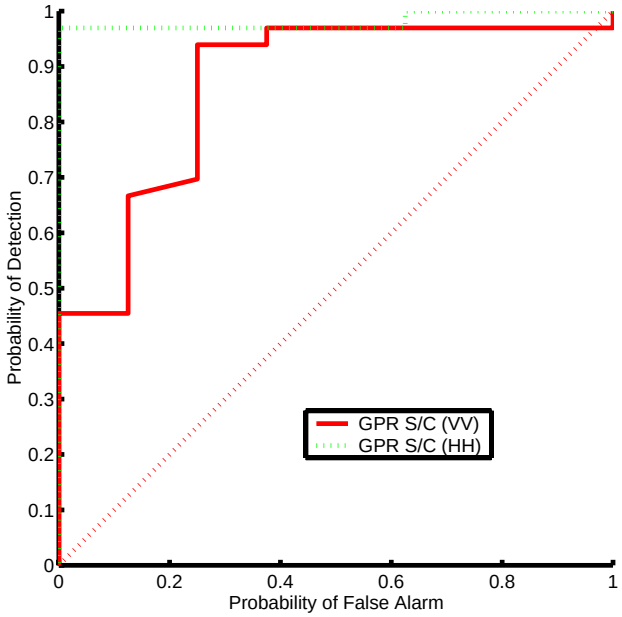


Figure 2. ROC curve for the FLGPR on metal mines.

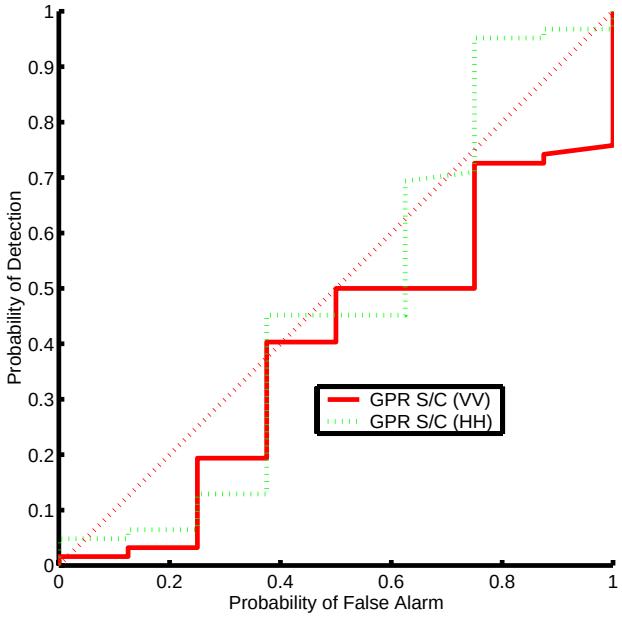


Figure 3. ROC curve for the FLGPR on plastic mines.

4.4. Fusion Performance

The foregoing results show that the FLGPR and LDV sensors are complementary to a degree. The FLGPR works markedly better for metallic mines, which the LDV works somewhat better for plastic mines. Fusion allows us to exploit this behavior by using simple "OR" fusion of hard (binary) decision-level outputs. The procedure used is as follows:

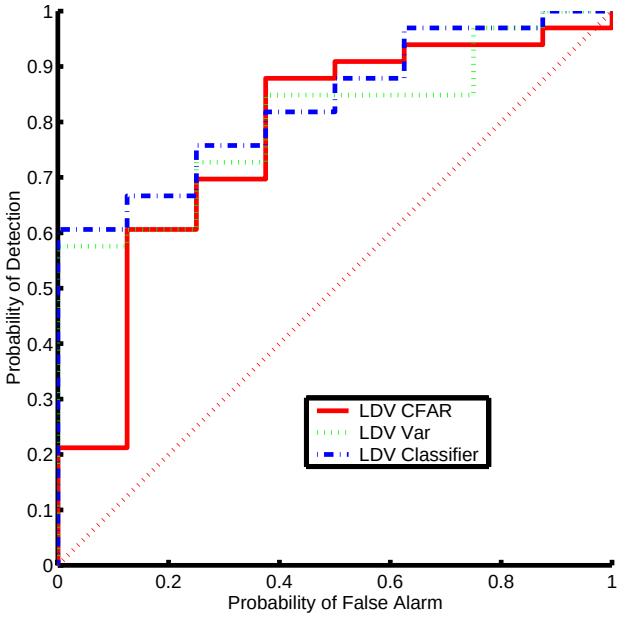


Figure 4. ROC curves for LDV features and the LDV classifier on metal and plastic mines.

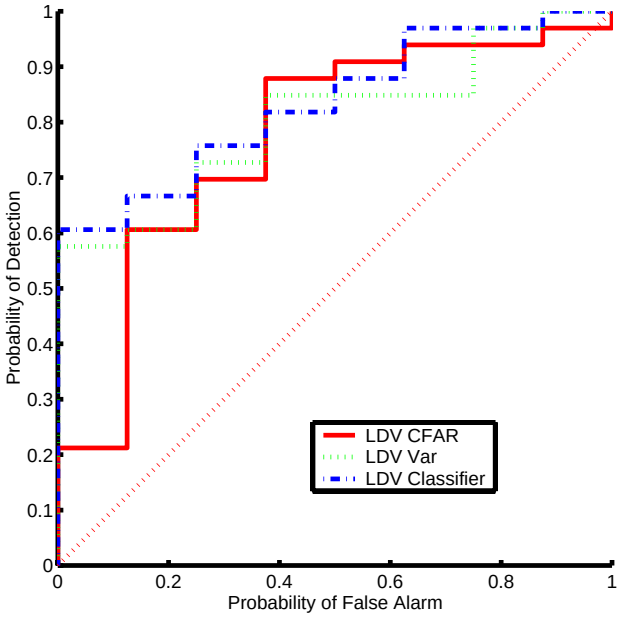


Figure 5. ROC curve for the LDV on metal mines.

The soft-decision outputs from each sensor are scaled to the interval [0,1]. A single threshold is applied to those outputs, resulting in two binary decision values. The decision values are logically "OR"ed together, producing a single fused decision.

An ROC curve computed from this fusion algorithm is shown in Figure 7. The responses of the two sensors individually to the fusible data sets are also shown. We observe that fusion has produce a gain in P_d throughout

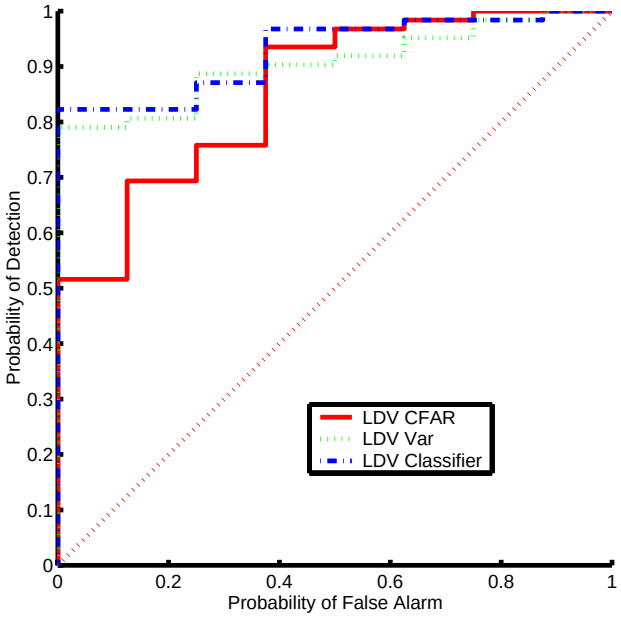


Figure 6. ROC curve for the LDV on plastic mines.

most of the range of P_{fa} as a result of the metal mines being detected by the GPR. The gain, however, is not large since the metal mines are only one third of the total targets and the LDV already has some capability to detect those targets. Nonetheless, it is notable that although the GPR performs poorly for the complete fusible data set, its excellent performance for metal mines can provide a benefit via fusion.

5. CONCLUDING REMARKS

Forward-looking demining sensors are attractive for the added operator safety they provide, and fusion of multi-sensor data is an effective method of improving the performance of a forward-looking system. In this paper we have demonstrated this concept.

During March 2000, data were acquired at Yuma Proving Grounds using both a forward-looking GPR sensor and a forward-looking laser Doppler vibrometer that sensed acoustically driven seismic soil motion. In fusing these data we found that the GPR is a very effective sensor of metal mines, but it is less effective on plastic mines. Conversely, the LDV sensor is somewhat more effective on plastic mines than on metal mines. The amount of fusible data is small (in particular, the amount of fusible clutter data is very limited), which limits the number and sophistication of fusion techniques that are suitable to this problem. We found that a simple "OR" fusion approach can exploit the complementarity of these sensors and provides an improvement in performance.

6. ACKNOWLEDGMENTS

The authors gratefully acknowledge the assistance of SRI in interpreting the FLGPR data and in providing the S/C values used for this study. Similarly, the assistance of the University of Mississippi in providing and interpreting the LDV data is gratefully acknowledged.

This project was supported by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.

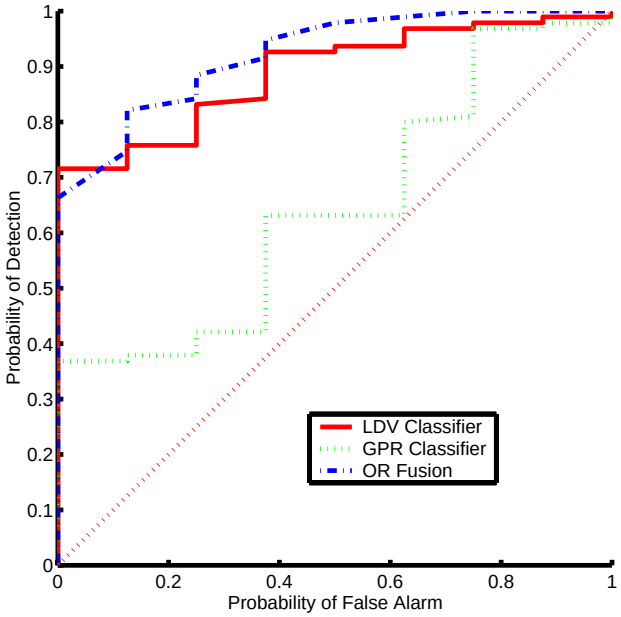


Figure 7. ROC curve for fusion of FLGPR and LDV data for metal and plastic mines.

REFERENCES

1. J. Kositsky and P. Milanfar, "A forward-looking high-resolution GPR system," in Detection and Remediation Technologies for Mines and Minelike Targets IV, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. Dugan, eds., Proc. SPIE 3710, pp. 1052–1062, 1999.
2. J. Kositsky, "Results from a forward-looking GPR mine detection system," in Detection and Remediation Technologies for Mines and Minelike Targets V, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. Dugan, eds., Proc. SPIE 4038, pp. 1077–1087, 2000.
3. J. M. Sabatier and N. Xiang, "Laser-Doppler based acoustic-to-seismic detection of buried mines," in Detection and Remediation Technologies for Mines and Minelike Targets IV, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. Dugan, eds., Proc. SPIE 3710, pp. 215–222, 1999.
4. T. R. Witten, K. Sherbondy, J. Habersat, and J. M. Sabatier, "Acoustic technology for land mine detection - Past test, present requirements and future concepts," in Detection and Remediation Technologies for Mines and Minelike Targets V, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. Dugan, eds., Proc. SPIE 4038, pp. 781–789, 2000.
5. N. Xiang and J. M. Sabatier, "Land mine detection measurements using acoustic-to-seismic coupling," in Detection and Remediation Technologies for Mines and Minelike Targets V, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. Dugan, eds., Proc. SPIE 4038, pp. 645–655, 2000.
6. B. A. Baertlein, W.-J. Liao, and D.-H. Chen, "Comparison of detection performance for stand-alone and fused mine sensors," in Proceedings of the UXO/Countermine Forum, 2001.
7. I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," IEEE Trans. Acoustics, Speech and Signal Proc. 38, pp. 1760–1770, October 1990.

Detection of Land Mines in Multi-Spectral and Multi-Temporal IR Imagery

W.-J. Liao, D.-H. Chen and B. A. Baertlein*

The Ohio State University ElectroScience Laboratory

1320 Kinnear Road, Columbus, OH 43212

Phone: (614) 292-0076

Fax: (614) 292-7297

E-mail: baertlein.1@osu.edu

Category: Session 15, Detection

ABSTRACT

In this paper we present techniques for detecting mines using data acquired at multiple wavelengths (multi-spectral) or at multiple times (multi-temporal). Both types of data employ vector-valued pixels, and therefore they have a common mathematical structure. For the multi-spectral data we show that the performance of standard algorithms can be improved significantly through spatial whitening when there is spatial correlation in the scene. For multi-temporal data we employ a classifier of the vector-valued pixels to identify mine-like and clutter-like regions. Both multi-spectral and multi-temporal algorithms are demonstrated using experimental data collected at Fort A.P. Hill, VA under a variety of conditions.

1. INTRODUCTION

It has long been known that passive IR signatures of both land mines and natural clutter depend on wavelength. Differences in the spectral responses of mines and clutter have led to the investigation of both multi-spectral and hyper-spectral systems for mine detection. Multi-spectral mine detection has been investigated extensively in the context of military mine-detection systems, and it continues to be an area of active interest.

It is also widely recognized that the responses of both mines and clutter are time dependent. Changes in the illumination geometry, surface temperature, and the spectral content of the illumination all produce differences in mine and clutter signatures. Multi-temporal information is not available in many military operations, but it is attractive for humanitarian demining, in which the observer is able to acquire multiple looks at the region of interest.

In this work we describe algorithms for processing multi-spectral and multi-temporal data. We first treat detection in multi-spectral imagery in Section 2. Many existing detection algorithms rely on the shape of the signature and its contrast with the background. We have found that the performance of standard algorithms can be improved by taking the spatial correlation of the clutter into account. Processed results are presented for multi-spectral data acquired in a recent collection at Ft. A. P. Hill. Algorithms for multi-temporal signals are described in Section 3. A classifier (a support vector machine) is trained to recognize surface and buried mines based on their thermal history. This approach is demonstrated for data acquired at Ft. Huachuca and Ft. A. P. Hill. Concluding remarks appear in Section 4.

2. MULTI-SPECTRAL DETECTION

2.1. Prior Work

Many authors have explored target detection in multi- and hyper-spectral imagery. A widely referenced approach to this problem is the RX algorithm, named for its authors, Reed and (Xiaoli) Yu.[1] This algorithm, which we describe in more detail in the next section, has been extended to include spectrally correlated (but spatially uncorrelated) clutter, which is estimated with an adaptive algorithm and used to design a spectral pre-whitening filter.[2] It was found that a priori knowledge of the target spectral distribution often makes little difference in performance.[2] Those authors and their co-workers have proposed several forms of the algorithm.[3–5] The RX algorithm has previously

been used in the COBRA system.[6] A application of the RX algorithm for hyper-spectral detection of military and cultural features was described by Stellman et al.[7]

Other approaches to multi-spectral have been examined. Schweizer and Moura[8] developed a CFAR algorithm based on a first-order Gauss-Markov random field (GMRF) model for the clutter. In that algorithm a maximum likelihood (ML) technique was used to estimate the clutter parameters, which are then used in a GLRT detector.

Several authors have attempted to exploit target spectral characterstics. Ashton[9] used clustering algorithms to find sub-pixel anomalies in multi-spectral IR terrain imagery. Ashton and Schaum[10] used RX to search for anomalies in background-suppressed spectral signatures.

A very different approach to mine detection in multi-spectral imagery was explored by Banerji and Goutsias,[11] who used mathematical morphology to detect mines in individual bands followed by fusion of the band information. Correlation among the bands was addressed by using a maximum noise fraction transform to generate independent bands.

2.2. Generalized Matched Filter Processing

For a target with a known signature in spatially white clutter, a matched filter is the optimal linear detector. When the target signature is known with the exception of its amplitude, a generalized likelihood ratio test may be constructed in which an ML estimator is used to determine the signature amplitude. For the special case of Gaussian noise, the amplitude estimator reduces to a linear system of equations.

The RX algorithm embodies this approach. A detailed description of the algorithm has appeared previously,[1] which the reader may consult for more details. That algorithm is a locally adaptive CFAR detector, which assumes spatially uncorrelated Gaussian clutter and a known target spatial signature with unknown spectral distribution. A local estimate of the clutter variance is made at each position within the scene and a local likelihood ratio is formed from the output of a matched filter and the clutter variance.

A key limiting assumption of the RX algorithm is that the clutter is spatially uncorrelated. In this paper we extend the RX algorithm to treat the problem of spatial correlation. In a prior work it was suggested that the spatially uncorrelated assumption is valid for low-resolution imagery,[2] but in the images described below significant correlation is evident. Figure 1 shows imagery and power spectra for mines and clutter. It is evident that the spectra overlap, which has a strong effect on detection algorithms.

The optimal detector of a known signature in colored noise involves a modification of the matched filter concept. Instead of a simple correlation of the measured data $\mathbf{X}$ with the known template (signature) $\mathbf{S}$ in the form $\mathbf{S}^\mathsf{T} \cdot \mathbf{X}$, both the data and template are pre-whitened as follows

$$(\mathbf{V} \cdot \mathbf{S})^\mathsf{T} \cdot (\mathbf{V} \cdot \mathbf{X}) \tag{1}$$

where $\mathbf{V}$ is a factorization of the clutter covariance matrix Σ.

$$\mathbf{\Sigma} = \mathbf{V}\mathbf{V}^\mathsf{T} \tag{2}$$

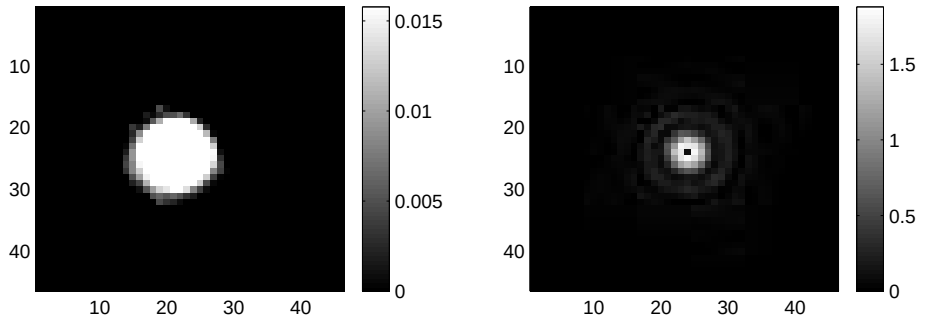
The factorization matrix $\mathbf{V}$ is readily found from an eigenvector decomposition of Σ.

It is difficult to implement this optimal approach because of both computational and practical limitations. For an image chip of size $J \times K$ the dimensions of $\mathbf{\Sigma}$ are $(JK) \times (JK)$. The local covariance matrix must be estimated from a set of clutter chips, the number of which is nominally much larger than JK. For chips of reasonable size, this computation becomes impractical. Furthermore, the clutter statistics of soil vary spatially. As a result, it may be impossible to find the large number of image samples required to estimate the covariance matrix within an image region having the desired clutter characteristics.

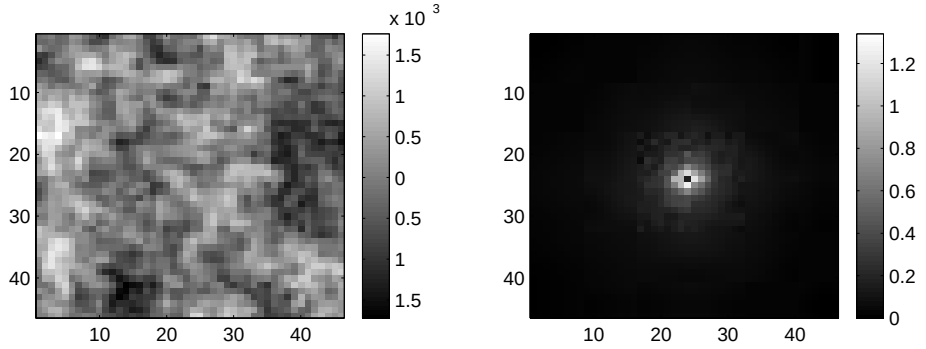
Over the spatial scales of interest in mine detection the correlation of these images is approximately first-order Markov, which leads to an exponential correlation function, which can be inverted. For isotropic clutter having correlation coefficient ρ, one can express the whitening operation as a convolution with a simple 3×3 kernel[12]

$$\begin{bmatrix} \rho^2 & -\rho(1+\rho^2) & \rho^2 \\ -\rho(1+\rho^2) & (1+\rho^2)^2 & -\rho(1+\rho^2) \\ \rho^2 & -\rho(1+\rho^2) & \rho^2 \end{bmatrix} \tag{3}$$

Using a locally derived estimate of ρ, this operator can be used to produce spatially whitened imagery at a modest computational cost.



(a) Surface mine image and power spectral density.



(b) Clutter image and power spectral density.

Figure 1. Some mine chips and their power spectra.

2.3. A Multi-Spectral Data Collection

Multi-spectral data was collected at Fort A.P. Hill, VA in May 2000. The data comprise anti-tank mines including TM62, VS1.6 and RAAM laid on the surface and buried. Several ground covers were imaged including sand, dirt, short grass and tall grass. Registered, multi-spectral images were collected using a visible broadband CCD camera and a MWIR camera. Multiple filters were used to extract specific spectral bands. The visible camera produces 768 × 512 data samples per frame and a 20 × 30 degree FOV, while the MWIR camera provides 256 × 256 data samples per frame and a 22 × 22 degree FOV. Both cameras were mounted on a cherry picker lifted 60 feet above the ground. The resulting images cover approximately a 3.5 m × 5.2 m area in the visible CCD and a 4 m × 4 m area in the MWIR camera.

To efficiently explore the response of our algorithms to a variety of ground covers and environmental conditions while keeping the computational burden manageable, we extracted a set of mine and clutter image chips. A total of 64 mine and clutter chips were extracted. Chips of two mine types on the surface (TM62 and VS1.6) were selected from three background types (sand, short grass and tall grass). Figure 2 shows sample multi-spectral target images in different environments. The test set also included clutter chips that show homogeneous backgrounds, inhomogeneous backgrounds (without emplaced non-mine objects), and clutter with emplaced objects that have a mine-like appearance. Those objects included crushed paint cans, pizza boxes, and wood sticks. Figure 3 shows some background and clutter chips.

Each image chip comprises six channels of data: five visible/IR bands (blue, green, yellow, red, and NIR) and one MWIR band. The chip size is 46 by 46 pixels. Because different cameras were involved the visible/NIR and MWIR chips have different resolutions, which is evident in Figure 2.

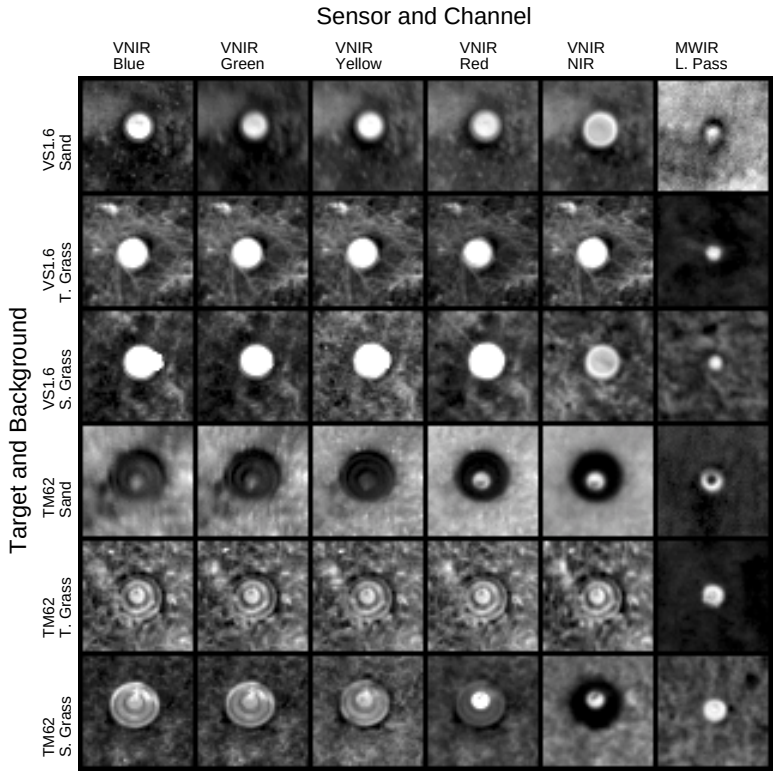


Figure 2. Examples of multi-spectral target chips.

2.4. Example Results

The foregoing algorithms were applied to the multi-spectral image chips described above. ROC curves for the baseline RX algorithm and RX with a pre-whitening filter are shown in Figure 4. Prewhitening significantly increases P_{d} for all but the smallest false alarm rates. At the lowest P_{fa} levels, the small loss in signal intensity caused by whitening is sufficient to cause some weak targets to be missed. In general, however, there is an improvement in false alarm rate for a fixed level of P_{d}. The improvement is most pronounced at the highest levels. For $P_{\mathsf{d}} = 0.95$ prewhitening decreases P_{fa} roughly five-fold.

3. MULTI-TEMPORAL DETECTION

Although it is well known that the temperature of surface mines and of soil over buried mines are time dependent, this fact does not appear to have been widely used in mine detection. We have identified only one group[13] that has previously explored this concept.

The problems of detection in multi-spectral and multi-temporal data are mathematically similar: vector-valued pixels (spectral responses or temporal responses) at each image location are examined for the presence of anomalies with mine-like characteristics. In practice, however, we find they are quite different. For multi-spectral data, the pixel vectors (spectral responses) at the mine are nearly constant over a mine or target, but for reasons which are as yet unclear, they are essentially random from target to target. Thus, the spectral information appears to contain little discriminatory power. Conversely, the temporal dependence of a mine signature has considerable discriminative power, which can be exploited by a classifier. In the sections that follow we describe the application of this procedure to an example data set.

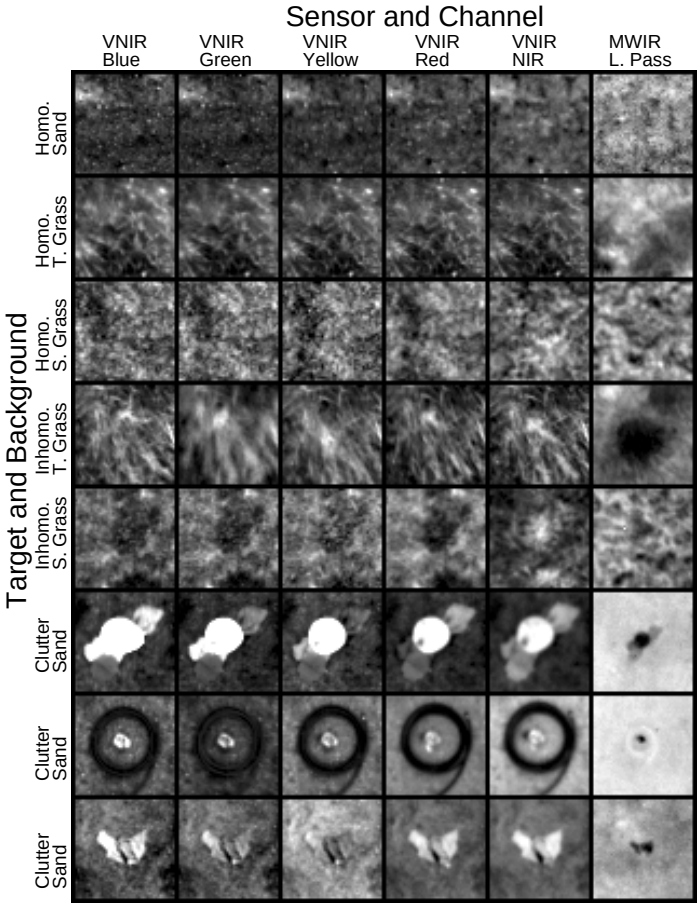


Figure 3. Sample multi-spectral clutter chips

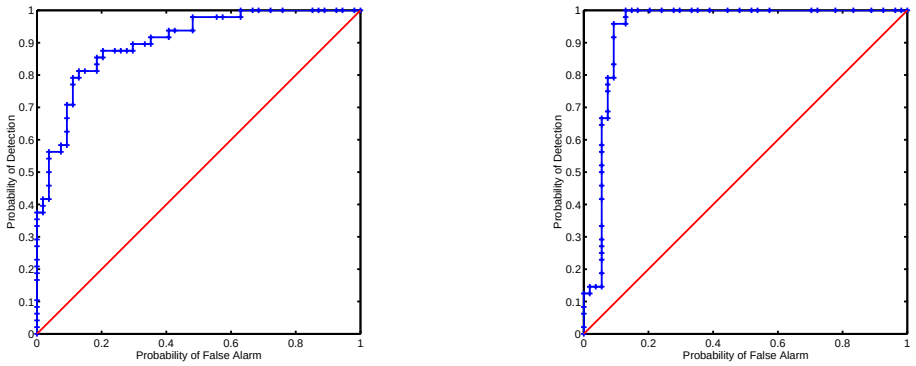
3.1. A Multi-Temporal Data Set

During 1998 a stationary IR camera was used to image surface-laid and buried targets at two sites, Fort A.P. Hill, VA and Fort Huachuca, AZ. The layout of the minefields were similar for these two collections, and a description of the A.P. Hill data set is shown in Figure 5. Data were acquired at 15 minute intervals for a duration of roughly 24 hours. Both Agema and Amber cameras were used. An example image from the Agema camera (8-12 μm) is shown in Figure 6.

A study was made of the temporal history of these targets. Sixty chips (thirty mines and thirty clutter) were extracted from this imagery. Figure 7 presents temperature histories for surface mines and buried mines in relation to clutter. It is evident that the surface mines all vary temporally in a similar manner, and are quite distinct from the background. (Compare the mine and clutter responses during the evening and during mid-day.) The temperature history of the buried mines are also distinct from the clutter, but much less so.

3.2. Detection Performance

Mine detection in multi-temporal imagery was investigated using two techniques. In the first approach we searched for spatial temperature anomalies. The time histories were sampled at six points taken roughly four hours apart starting at approximately 4:00 AM, and the resulting data were used with the RX algorithm. A spatially uniform mine signature was used. We next used a classifier to detect mines based on their temperature history. A support vector machine was trained using the foregoing data. The ROC curves generated by RX and by leave-one-out testing



(a) The baseline RX algorithm.

(b) The RX algorithm with clutter whitening.

Figure 4. ROC curves for the RX algorithm applied to multi-spectral imagery without and with clutter whitening.

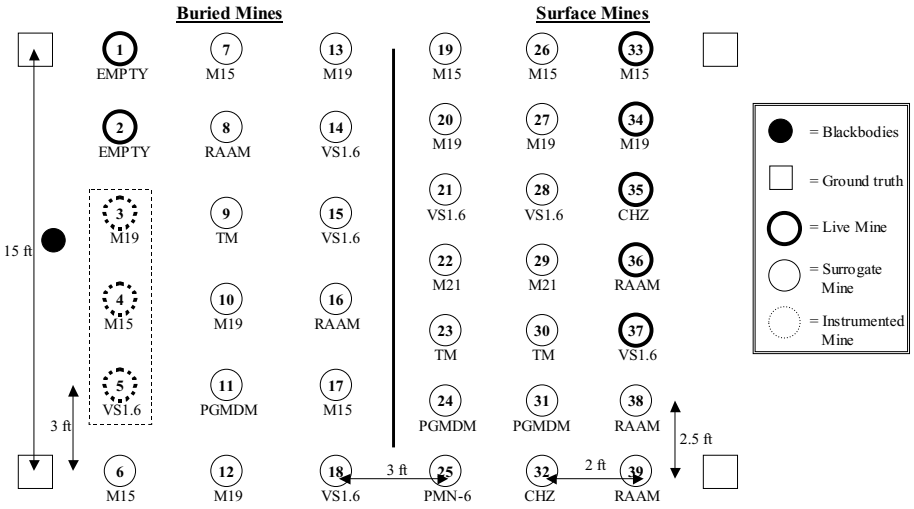


Figure 5. Mine layout for the AP Hill collection.

of the SVM are shown in Figure 8 for surface mines and buried mines. We see that the classifier approach performs significantly better than the RX algorithm for these data. The results suggest that the time histories contain more information than the spatial imagery. As expected, buried mines are much more difficult to detect than surface mines, but even in this case the classifier approach outperforms the RX algorithm.

4. CONCLUDING REMARKS

Processing techniques have been demonstrated for detection of land mines in multi-spectral and multi-temporal imagery. For the multi-spectral data we exploited primarily the spatial information in the signature. We used a whitening filter to eliminate spatial correlation in the imagery. A first-order Gauss-Markov random field approximation for the clutter permits use of a computationally efficient method for whitening prior to processing with other algorithms. It was shown that pre-whitening can greatly improve the performance of the the widely used RX algorithm. For multi-temporal detection we exploited the discrimative power in the time-histories of individual pixels. We found that applying a classifier to pixel time histories produced good detection.

Both the multi-spectral and multi-temporal algorithms are now being applied to images (as opposed to image chips). Both methods segment the imagery into mine-like and non-mine-like regions. Simple morphological operations

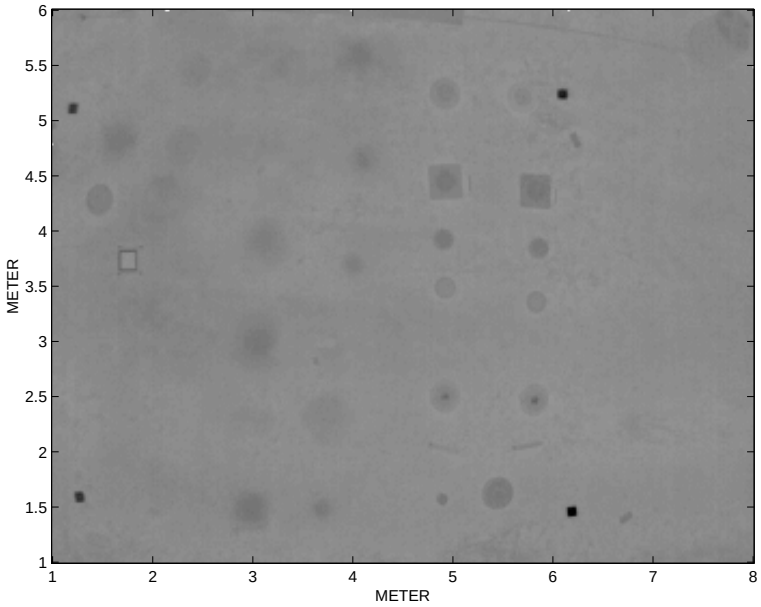
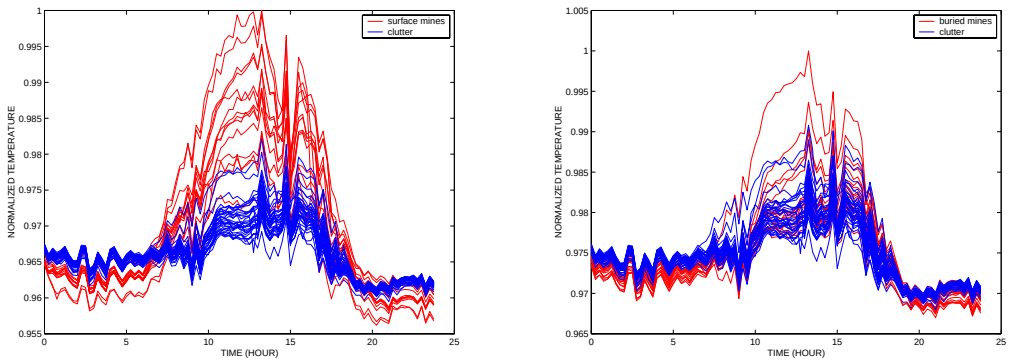


Figure 6. Example imagery from the AP Hill collection.



(a) Surface mines versus clutter.

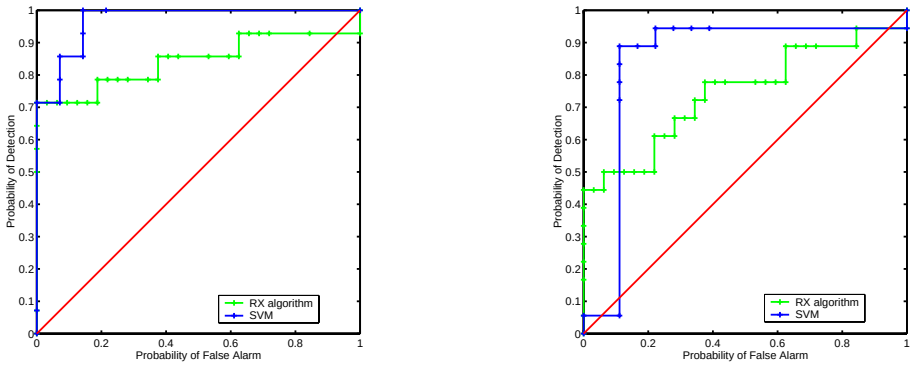
(b) Buried mines versus clutter.

Figure 7. Time histories for buried and surface mines.

on those regions have yielded good detection performance, compared to a straightforward application of the RX algorithm. Future work will address greater use of the spectral information in the multi-spectral imagery, and greater use of the spatial information in the multi-temporal imagery.

5. ACKNOWLEDGMENTS

This project was supported in part by funds from Duke University under an award from the ARO (the OSD MURI program). The findings, opinions and recommendations expressed therein are those of the author and are not necessarily those of Duke University or the ARO.



(a) Surface mines.

(b) Buried mines.

Figure 8. ROC curve for multi-temporal data using RX and SVM algorithms.

REFERENCES

1. I. S. Reed and X. Yu, "Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution," IEEE Trans. Acoustics, Speech and Signal Proc. 38, pp. 1760–1770, October 1990.
2. X. Yu and I. S. Reed and A. D. Stocker, "Comparative performance analysis of adaptive multiplespectral detectors," IEEE Trans. Acoustics, Speech and Signal Proc. 41, pp. 2639–2654, August 1993.
3. J. Y. Chen and I. S. Reed, "A detection algorithm for optical targets in clutter," IEEE Trans. Aerosp. Electron. Syst. AES-23, pp. 46–59, February 1987.
4. A. Margalit, I. S. Reed, and R. M. Gagliardi, "Adaptive optical target detection using correlated images," IEEE Trans. Aerosp. Electron. Syst. AES-21, pp. 394–405, June 1985.
5. X. Yu, L. E. Hoff, I. S. Reed, A. M. Chen, and L. B. Stotts, "Automatic target detection and recognition in multiband imagery: A unified ml detection and estimation approach," IEEE Trans. Image Processing 6, pp. 143–156, January 1997.
6. Q. A. Holmes, C. R. Schwartz, J. H. Seldin, J. A. Wright, and L. W. Witter, "Adaptive multispectral cfar detection of land mines," in Detection Technologies for Mines and Minelike Targets, A. C. Dubey, I. Cindrich, J. M. Ralston, and K. Rigano, eds., Proc. SPIE 2496, pp. 421–432, 1995.
7. C. M. Stellman, G. G. Hazel, F. Bucholtz, J. V. Michalowsicz, A. Stocker, and W. Schaaf, "Real-time hyperspectral detection and cuing," Opt. Eng. 39, pp. 1928–1935, July 2000.
8. S. M. Schweizer and J. M. F. Moura, "Hyperspectral imagery: Clutter adaptation in anomaly detection," IEEE Trans. Information Theory 46, pp. 1855–1871, August 2000.
9. E. A. Ashton, "Detection of subpixel anomalies in multispectral infrared imagery using an adaptive bayesian classifier," IEEE Trans. Geoscience and Remote Sensing 36, pp. 506–517, March 1998.
10. E. A. Ashton and A. Schaum, "Algorithms for the detection of sub-pixel targets in multispectral imagery," Photogrammetric Engineering and Remote Sensing 64, pp. 723–731, July 1998.
11. A. Banerji and J. Goutsias, "A morphological approach to automatic mine detection problems," IEEE Trans. Aerospace and Electronic Systems 34, pp. 1085–1096, October 1998.
12. A. Arcess, P. H. Mengaert, and E. W. Trombini, "Image detection through bipolar correlation," IEEE Trans. Information Theory IT-16, pp. 534–541, September 1970.
13. M. Lundberg and I. Y. H. Gu, "A 3-d matched filter for detection of land mines using spatio-temporal thermal modeling," in Detection Technologies for Mines and Minelike Targets V, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, eds., Proc. SPIE 4038, pp. 179–188, 2000.

Thermal Emission from a Layered Medium Bounded by a Slightly Rough Interface

Joel T. Johnson, *Member, IEEE*

***Abstract*—The small perturbation method (SPM) is applied to study thermal emission from a layered medium bounded by a slightly rough interface. Brightness temperatures are calculated to second order in surface height, including both specular reflection coefficient corrections and incoherent Bragg scatter terms. Unlike the homogeneous medium case, in which the SPM applied for emission predictions produces an expansion in surface slope, the theory remains a small height expansion, and convergence of the series is shown to depend on properties of the layered medium. Results from this theory can be applied in studies of soil moisture, sea ice, or sea surface remote sensing and buried object detection with microwave radiometers.**

***Index Terms*—Microwave radiometry, remote sensing, rough surfaces.**

I. Introduction

MODELS for microwave thermal emission from deterministic or statistically described rough surfaces are of interest in passive remote sensing of soil moisture, sea ice, and the ocean surface. Several models for thermal emission from a rough surface bounding a homogeneous medium have been developed previously [1]–[6], primarily through application of standard surface scattering approximate methods to calculate surface emissivity using Kirchhoff's law. Models based on both the small perturbation method (SPM) and the physical optics (PO) approximation have been presented. A recent work [7] has further revealed that use of the SPM for emission from a rough surface bounding a homogeneous medium results in a small slope (rather than small height) emission approximation identical to that which would be obtained from the small slope approximation of [8]. The SPM can thus provide accurate emission predictions even for surfaces with large heights in terms of the electromagnetic wavelength. Numerical tests of the SPM for a set of canonical periodic surfaces have confirmed this statement [9]. These results motivate use of the SPM/small slope approximation (SPM/SSA) for the study of homogeneous medium thermal emission. However, the SPM for emission from a rough surface bounding a layered medium has apparently not previously been considered. This problem is potentially more relevant to soil moisture remote sensing studies given the variations in soil moisture content with depth which typically occur. A model for emission from a layered medium bounded by a rough interface could also be applied to model sea water covered with foam, ice, or other materials, or used to estimate surface clutter influences on microwave radiometry systems for buried object detection [10], [11].

The SPM has previously been applied to study backscattering from a layered medium bounded by a slightly rough interface [12]–[16] and results demonstrated that the presence of a layered medium can cause significant changes in first order surface scattered fields. An approximation for modifying two layer medium surface scattering predictions to the case of a finite object buried beneath a rough interface was also suggested in [12] for application to ground penetrating radar problems.

In this paper, the results of [12], [13] are extended to enable calculation of thermal emission from a slightly rough interface bounding a layered medium. Note that studies of thermal emission require both the first order SPM terms as in [12], [13] and also second order corrections to the specular reflection coefficient [2]. These quantities are derived and presented in Sections II and III. As in the homogeneous medium case, the resulting expression for the rough surface-induced correction to flat surface brightness temperatures is expressed in terms of an integral over the surface directional spectrum multiplied by a "weighting" function [17]. Studies of the weighting functions allow properties of the emission physics to be inferred independent of the surface statistics considered. Consideration of the weighting functions for isotropic (i.e., azimuthally symmetric) surfaces in Section IV reveals that a small height and not small slope expansion is obtained in the layered medium case, and convergence properties of this series are discussed. Studies of the weighting functions for azimuthally asymmetric surfaces in Section V again show a small height expansion for the second and higher azimuthal harmonics of all polarimetric brightness temperatures. Sample results applying the theory are presented in Section VI, and implications of the study considered in Section VII.

II. Formulation

A systematic solution of SPM equations for scattering from a rough surface bounding a homogeneous medium has recently been developed in [18]. This procedure applies the Rayleigh hypothesis to study scattered and transmitted plane wave amplitudes from a periodic surface excited by an incident plane wave, as in the original SPM formulation of [19]. SPM results for a nonperiodic surface are obtained in the limit as the surface periods become large following [19], [20]. The same procedure is applied in this paper for a rough surface bounding a layered

Manuscript received July 23, 1999; revised June 15, 2000. This work was sponsored by the Office of Naval Research, Washington, DC, Contract N00014-97-1-0541, National Science Foundation, Washington, DC, Project ECS-9701678, and by a Grant from Duke University, Durham, NC, as part of the OSD MURI on Humanitarian Demining.

The author is with the Department of Electrical Engineering and Electro-Science Laboratory, The Ohio State University, Columbus, OH 43210.

Publisher Item Identifier S 0196-2892(01)01157-3.

0196–2892/01$10.00 © 2001 IEEE

medium. The details of the procedure are very similar to [18], so only the basic formulation and results from the method are summarized here.

Consider a zero mean periodic surface profile $z = f(x, y)$ with periods P_x and P_y in the x and y directions, respectively, which separates free space in region zero (permittivity ϵ_0, permeability μ_0) for $z > f(x, y)$ from region one, a homogeneous nonmagnetic dielectric medium with permittivity $\epsilon_d = \epsilon\epsilon_0$ for $-d < z < f(x, y)$. Initially, a two layer configuration is considered, with Region 2 for $z < -d$ a homogeneous nonmagnetic dielectric medium with permittivity $\epsilon_{d2} = \epsilon_2\epsilon_0$ as illustrated in Fig. 1. Extension to a multilayer medium is straightforward and will be described below. The periodic surface $f(x, y)$ can also be expressed in terms of its Fourier series coefficients

$$f(x, y) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \exp\left(i\frac{2\pi nx}{P_x}\right)\exp\left(i\frac{2\pi my}{P_y}\right) h_{n,m} \tag{1}$$

$$h_{n,m} = \frac{1}{P_x}\frac{1}{P_y}\int_0^{P_x} dx \int_0^{P_y} dy \exp\left(-i\frac{2\pi nx}{P_x}\right) \cdot \exp\left(-i\frac{2\pi my}{P_y}\right) f(x, y). \tag{2}$$

Consider an incident electromagnetic plane wave that illuminates this periodic surface from the free space region, with electric and magnetic fields given by

$$\overline{E}^i = \hat{e}_i \exp\left(i\overline{k}_i \cdot \overline{r}\right) \tag{3}$$

$$\overline{H}^i = \frac{\hat{k}_i \times \hat{e}_i}{\eta_0} \exp\left(i\overline{k}_i \cdot \overline{r}\right) \tag{4}$$

where $\hat{e}_i$ represents the polarization vector of the incident electric field.

$$\overline{k}_i = k_0\hat{k}_i = \hat{x}k_{xi} + \hat{y}k_{yi} - \hat{z}k_{zi} \tag{5}$$

represents the propagation vector of the incident plane wave with wavenumber $k_0 = 2\pi/\lambda$.

$$\overline{r} = \hat{x}x + \hat{y}y + \hat{z}z \tag{6}$$

is a position vector in Cartesian space, and $\eta_0 = \sqrt{(\mu_0/\epsilon_0)}$ is the impedance of free space. Note an $\exp(-i\omega t)$ time convention is assumed.

Under the Rayleigh hypothesis, the scattered field in region zero consists of a sum of upgoing plane waves (or "Floquet modes"), which can be written as

$$\overline{E}^s = \sum_m \sum_n \left[\hat{h}_s^{n,m}\alpha_{n,m} + \hat{v}_s^{n,m}\beta_{n,m}\right] \cdot \exp\left(i\overline{k}_s^{n,m} \cdot \overline{r}\right) \tag{7}$$

$$\overline{H}^s = \frac{1}{\eta_0}\sum_m \sum_n \left[-\hat{v}_s^{n,m}\alpha_{n,m} + \hat{h}_s^{n,m}\beta_{n,m}\right] \cdot \exp\left(i\overline{k}_s^{n,m} \cdot \overline{r}\right) \tag{8}$$

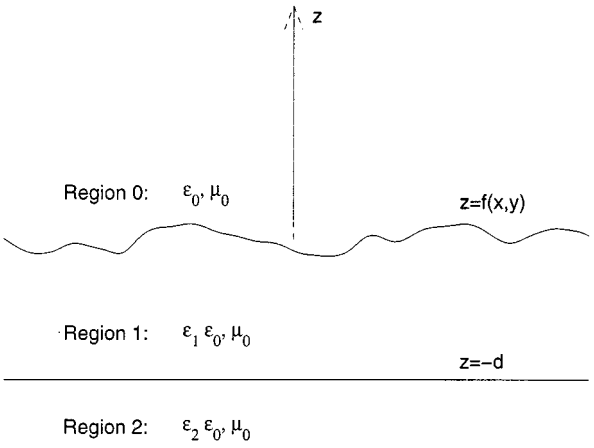


Fig. 1. Geometry of two layer medium bounded by a slightly rough interface.

while fields in region one consist of both upgoing and downgoing plane, waves which can be written as

$$\overline{E}^t = \sum_m \sum_n \left[\hat{h}_s^{n,m}\gamma_{n,m} + \hat{v}_t^{n,m}\delta_{n,m}\right] \cdot \exp\left(i\overline{k}_t^{n,m} \cdot \overline{r}\right) + \sum_m \sum_n \left[\hat{h}_s^{n,m}\Gamma_{n,m} + \hat{v}_u^{n,m}\Delta_{n,m}\right] \cdot \exp\left(i\overline{k}_u^{n,m} \cdot \overline{r}\right) \tag{9}$$

$$\overline{H}^t = \frac{1}{\eta_1}\sum_m \sum_n \left[-\hat{v}_t^{n,m}\gamma_{n,m} + \hat{h}_s^{n,m}\delta_{n,m}\right] \cdot \exp\left(i\overline{k}_t^{n,m} \cdot \overline{r}\right) + \frac{1}{\eta_1}\sum_m \sum_n \left[-\hat{v}_u^{n,m}\Gamma_{n,m} + \hat{h}_s^{n,m}\Delta_{n,m}\right] \cdot \exp\left(i\overline{k}_u^{n,m} \cdot \overline{r}\right) \tag{10}$$

where $\eta_1 = \eta_0/\sqrt{\epsilon}$ is the impedance of the lower medium. Note that all sums are assumed to be from $-\infty$ to ∞ unless otherwise notated. Fields in region two consist only of downgoing plane waves, written as

$$\overline{E}^l = \sum_m \sum_n \left[\hat{h}_s^{n,m}P_{n,m} + \hat{v}_l^{n,m}Q_{n,m}\right] \cdot \exp\left(i\overline{k}_l^{n,m} \cdot \overline{r}\right) \tag{11}$$

$$\overline{H}^l = \frac{1}{\eta_2}\sum_m \sum_n \left[-\hat{v}_l^{n,m}P_{n,m} + \hat{h}_s^{n,m}Q_{n,m}\right] \cdot \exp\left(i\overline{k}_l^{n,m} \cdot \overline{r}\right) \tag{12}$$

where $\eta_2 = \eta_0/\sqrt{\epsilon_2}$ is the impedance of the lower medium. In the above equation, α, β, γ, δ, Γ, Δ, P, and Q are the unknown complex amplitudes of the Floquet modes in each region. Plane wave propagation vectors are defined by the Floquet theorem as

$$\overline{k}_s^{n,m} = \hat{x}k_{xn} + \hat{y}k_{ym} + \hat{z}k_{znm} \tag{13}$$

$$\overline{k}_t^{n,m} = \hat{x}k_{xn} + \hat{y}k_{ym} - \hat{z}k_{z1nm} \tag{14}$$

$$\overline{k}_u^{n,m} = \hat{x}k_{xn} + \hat{y}k_{ym} + \hat{z}k_{z1nm} \tag{15}$$

$$\overline{k}_l^{n,m} = \hat{x}k_{xn} + \hat{y}k_{ym} - \hat{z}k_{z2nm} \tag{16}$$

where

$$k_{xn} = k_{xi} + \frac{2\pi n}{P_x} \tag{17}$$

$$k_{ym} = k_{yi} + \frac{2\pi m}{P_y} \tag{18}$$

$$k_{\rho nm} = \sqrt{k_{xn}^2 + k_{ym}^2} \tag{19}$$

$$k_{znm} = \sqrt{k_0^2 - k_{\rho nm}^2} \tag{20}$$

$$k_{z1nm} = \sqrt{k_0^2 \epsilon - k_{\rho nm}^2} \tag{21}$$

$$k_{z2nm} = \sqrt{k_0^2 \epsilon_2 - k_{\rho nm}^2}. \tag{22}$$

Modes for which $k_{\rho nm}$ becomes greater than k_0, k_1, or k_2 have k_{znm}, k_{z1nm}, and k_{z2nm}, respectively, defined so that attenuation occurs as upgoing fields propagate away from the rough surface in region zero and as downgoing fields propagate away from the layered medium boundaries in regions one and two. Orthogonal horizontal and vertical polarization vectors for these plane waves are defined as

$$\hat{h}_i = \hat{x}\frac{k_{yi}}{k_{\rho i}} - \hat{y}\frac{k_{xi}}{k_{\rho i}} \tag{23}$$

$$\hat{h}_s^{n,m} = \hat{x}\frac{k_{ym}}{k_{\rho nm}} - \hat{y}\frac{k_{xn}}{k_{\rho nm}} \tag{24}$$

$$\hat{v}_i = \hat{x}\frac{k_{xi}k_{zi}}{k_0 k_{\rho i}} + \hat{y}\frac{k_{yi}k_{zi}}{k_0 k_{\rho i}} + \hat{z}\frac{k_{\rho i}}{k_0} \tag{25}$$

$$\hat{v}_s^{n,m} = -\hat{x}\frac{k_{xn}k_{znm}}{k_0 k_{\rho nm}} - \hat{y}\frac{k_{ym}k_{znm}}{k_0 k_{\rho nm}} + \hat{z}\frac{k_{\rho nm}}{k_0} \tag{26}$$

$$\hat{v}_t^{n,m} = \hat{x}\frac{k_{xn}k_{z1nm}}{k_1 k_{\rho nm}} + \hat{y}\frac{k_{ym}k_{z1nm}}{k_1 k_{\rho nm}} + \hat{z}\frac{k_{\rho nm}}{k_1} \tag{27}$$

$$\hat{v}_u^{n,m} = -\hat{x}\frac{k_{xn}k_{z1nm}}{k_1 k_{\rho nm}} - \hat{y}\frac{k_{ym}k_{z1nm}}{k_1 k_{\rho nm}} + \hat{z}\frac{k_{\rho nm}}{k_1} \tag{28}$$

$$\hat{v}_l^{n,m} = \hat{x}\frac{k_{xn}k_{z2nm}}{k_2 k_{\rho nm}} + \hat{y}\frac{k_{ym}k_{z2nm}}{k_2 k_{\rho nm}} + \hat{z}\frac{k_{\rho nm}}{k_2} \tag{29}$$

where $k_1 = k_0\sqrt{\epsilon}$ and $k_2 = k_0\sqrt{\epsilon_2}$ are the wavenumbers in regions one and two, respectively.

Boundary conditions on the layered medium interfaces specify that tangential electric and magnetic fields must be continuous. At $z = f(x, y)$, this becomes

$$(\hat{z} - \overline{\partial f}) \times (\overline{E}^i + \overline{E}^s) = (\hat{z} - \overline{\partial f}) \times \overline{E}^t \tag{30}$$

$$(\hat{z} - \overline{\partial f}) \times (\overline{H}^i + \overline{H}^s) = (\hat{z} - \overline{\partial f}) \times \overline{H}^t \tag{31}$$

since a vector normal to the surface can be written as $\hat{z} - \overline{\partial f}$, where $\overline{\partial f} = \hat{x}(\partial f/\partial x) + \hat{y}(\partial f/\partial y)$. At $z = -d$, the boundary conditions specify

$$\hat{z} \times \overline{E}^t = \hat{z} \times \overline{E}^l \tag{32}$$

$$\hat{z} \times \overline{H}^t = \hat{z} \times \overline{H}^l. \tag{33}$$

Substituting the Rayleigh hypothesis fields from (9)–(12) into (32), (33), the following relationships can be derived:

$$P_{n,m} = \gamma_{n,m} \exp(i(k_{z1nm} - k_{z2nm})\, d) \cdot \left(\frac{2k_{z1nm}}{k_{z1nm} + k_{z2nm}}\right) \tag{34}$$

$$Q_{n,m} = \delta_{n,m} \exp(i(k_{z1nm} - k_{z2nm})\, d) \cdot \left(\frac{2k_{z1nm}\sqrt{\epsilon\epsilon_2}}{\epsilon_2 k_{z1nm} + \epsilon k_{z2nm}}\right) \tag{35}$$

$$\Gamma_{n,m} = \gamma_{n,m} \exp(2ik_{z1nm}d) \left(\frac{k_{z1nm} - k_{z2nm}}{k_{z1nm} + k_{z2nm}}\right) \tag{36}$$

$$= \gamma_{n,m} \exp(2ik_{z1nm}d) R_h(k_{z1nm}, k_{z2nm}) \tag{37}$$

$$\Delta_{n,m} = \delta_{n,m} \exp(2ik_{z1nm}d) \left(\frac{\epsilon_2 k_{z1nm} - \epsilon k_{z2nm}}{\epsilon_2 k_{z1nm} + \epsilon k_{z2nm}}\right) \tag{38}$$

$$= \delta_{n,m} \exp(2ik_{z1nm}d) R_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}) \tag{39}$$

which defines the lower layer reflection coefficients $R_h(k_{z1nm}, k_{z2nm})$ and $R_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm})$, respectively. These definitions make the generalization to a layered medium below the rough interface clear. $R_h(k_{z1nm}, k_{z2nm})$ and $R_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm})$ are simply replaced by the reflection coefficients of the layered medium below $z = -d$. Expressions for $P_{n,m}$ and $Q_{n,m}$ are no longer correct in this case, but since thermal emission can be computed from the reflectivity in the region zero (i.e., in terms of α and β only), it is still possible to compute brightness temperatures for the layered medium by modifying only these reflection coefficients.

Substituting the above relationships and the Rayleigh hypothesis fields from (7)–(10) into (30), (31), the following equations result if only x and y components are considered:

$$\begin{aligned}
&\sum_m \sum_n \exp\left(i\frac{2\pi n x}{P_x}\right) \exp\left(i\frac{2\pi m y}{P_y}\right) \\
&\quad \cdot \Big[(\hat{z} \times \hat{h}_s^{n,m})(\alpha_{n,m} \exp(ik_{znm}z) - \gamma_{n,m} \\
&\qquad \cdot (\exp(-ik_{z1nm}z) + R_h(k_{z1nm}, k_{z2nm}) \\
&\qquad\quad \cdot \exp(2ik_{z1nm}d) \exp(ik_{z1nm}z))) + (\hat{z} \times \hat{v}_s^{n,m}) \\
&\qquad \cdot \Big(\beta_{n,m} \exp(ik_{znm}z) + \frac{k_0 k_{z1nm}}{k_1 k_{znm}} \delta_{n,m} \\
&\qquad\quad \cdot (\exp(-ik_{z1nm}z) - R_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}) \\
&\qquad\quad \cdot \exp(2ik_{z1nm}d) \exp(ik_{z1nm}z))\Big)\Big] \\
&= -(\hat{z} \times \hat{e}_i) \exp(-ik_{zi}z) + (\overline{\partial f} \times \hat{e}_i) \exp(-ik_{zi}z) \\
&\quad + \sum_m \sum_n \exp\left(i\frac{2\pi n x}{P_x}\right) \exp\left(i\frac{2\pi m y}{P_y}\right) \\
&\quad \cdot \Big[(\overline{\partial f} \times \hat{v}_s^{n,m}) \Big(\beta_{n,m} \exp(ik_{znm}z) - \frac{k_0}{k_1} \\
&\qquad \cdot \delta_{n,m}(\exp(-ik_{z1nm}z) + R_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}) \\
&\qquad\quad \cdot \exp(2ik_{z1nm}d) \exp(ik_{z1nm}z))\Big)\Big]
\end{aligned} \tag{40}$$

$$\begin{aligned}
&\sum_m \sum_n \exp\left(i\frac{2\pi n x}{P_x}\right) \exp\left(i\frac{2\pi m y}{P_y}\right) \\
&\quad \cdot \Big[(-\hat{z} \times \hat{v}_s^{n,m}) \Big(\alpha_{n,m} \exp(ik_{znm}z) + \frac{k_{z1nm}}{k_{znm}} \\
&\qquad \cdot \gamma_{n,m}(\exp(-ik_{z1nm}z) - R_h(k_{z1nm}, k_{z2nm}) \\
&\qquad\quad \cdot \exp(2ik_{z1nm}d) \exp(ik_{z1nm}z))\Big) \\
&\qquad + (\hat{z} \times \hat{h}_s^{n,m}) \Big(\beta_{n,m} \exp(ik_{znm}z) - \frac{k_1}{k_0}
\end{aligned}$$

$$\cdot \delta_{n,m}(\exp(-ik_{z1nm}z) + R_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}) \cdot \exp(2ik_{z1nm}d)\exp(ik_{z1nm}z))\Big)\Big]$$
$$= -(\hat{z}\times\hat{k}_i\times\hat{e}_i)\exp(-ik_{zi}z) + \Big(\overline{\partial f}\times\hat{k}_i\times\hat{e}_i\Big)\cdot\exp(-ik_{zi}z) + \sum_m\sum_n \exp\Big(i\frac{2\pi nx}{P_x}\Big)\exp\Big(i\frac{2\pi my}{P_y}\Big)\cdot\Big[\Big(-\overline{\partial f}\times\hat{v}_s^{n,m}\Big)(\alpha_{n,m}\exp(ik_{znm}z) - \gamma_{n,m}(\exp(-ik_{z1nm}z) + R_h(k_{z1nm}, k_{z2nm}) \cdot \exp(2ik_{z1nm}d)\exp(ik_{z1nm}z)))\Big]. \quad (41)$$

Note that the above equations (which have two components each) provide four scalar equations for the four scalar unknown functions α, β, γ, and δ, with unknowns Γ, Δ, P, and Q determined from (34)–(39).

At this point, a small height expansion is used by expanding the exponentials in the above equations in power series

$$\exp(\pm ik_z z) = \sum_{q=0}^{\infty}\frac{(\pm ik_z z)^q}{q!} \quad (42)$$

and by substituting a perturbation series for the unknowns

$$\alpha_{n,m} = \alpha_{n,m}^{(0)} + \alpha_{n,m}^{(1)} + \cdots = \sum_{l=0}^{\infty}\alpha_{n,m}^{(l)} \quad (43)$$

with similar definitions for β, γ, and δ. Perturbation series terms are defined so that the lth term is of order f^l or equivalent combinations of f and its derivatives since $(\partial f/\partial x)$ and $(\partial f/\partial y)$ are assumed to be the same order as f.

III. Solution of SPM Equations

The systematic procedure for solving SPM equations described in [18] is applied to (40) and (41) once the perturbation series is substituted. Solutions at the zeroth order are found to consist only of the specular plane waves in each region (represented by $n = m = 0$), with fields in region zero given by

$$\alpha_{0,0}^{(0)} = \frac{k_{zi} - k_{z1i}Z_h(k_{z1i}, k_{z2i}, d)}{k_{zi} + k_{z1i}Z_h(k_{z1i}, k_{z2i}, d)} = R_{Hl} \quad (44)$$
$$\beta_{0,0}^{(0)} = 0 \quad (45)$$

for a horizontally polarized incident field (i.e., $\hat{e}_i = \hat{h}_i$), while

$$\alpha_{0,0}^{(0)} = 0 \quad (46)$$
$$\beta_{0,0}^{(0)} = \frac{\epsilon k_{zi} - k_{z1i}Z_v(\epsilon_2 k_{z1i}, \epsilon_1 k_{z2i}, d)}{\epsilon k_{zi} + k_{z1i}Z_v(\epsilon_2 k_{z1i}, \epsilon_1 k_{z2i}, d)} = R_{Vl} \quad (47)$$

for a vertically polarized incident field. The above equations are the total reflection coefficients at $z = 0$ of the layered medium with a flat surface R_{Hl} and R_{Vl} and also involve

$$Z_h(k_{z1nm}, k_{z2nm}, d) = \frac{1 - R_h(k_{z1nm}, k_{z2nm})\exp(2ik_{z1nm}d)}{1 + R_h(k_{z1nm}, k_{z2nm})\exp(2ik_{z1nm}d)} = \frac{1 - G_h(k_{z1nm}, k_{z2nm}, d)}{1 + G_h(k_{z1nm}, k_{z2nm}, d)}$$
$$Z_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}, d) = \frac{1 - R_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm})\exp(2ik_{z1nm}d)}{1 + R_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm})\exp(2ik_{z1nm}d)} = \frac{1 - G_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}, d)}{1 + G_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}, d)}. \quad (48)$$

A general form for first order solutions is

$$\zeta_{n,m}^{(1)} = h_{n,m}g_\zeta^{(1)}(k_{xn}, k_{ym}) \quad (49)$$

where $\zeta = \alpha$, β, γ, or δ, and $g_\zeta^{(1)}$ is a corresponding function. Solutions in region zero for a horizontally polarized incident field are

$$g_\alpha^{(1)} = \frac{-2ik_{zi}(k_0^2 - k_1^2)}{(k_{znm} + k_{z1nm})(k_{zi} + k_{z1i})}c_{i,n} \cdot\left[\frac{1 + G_h(k_{z1i}, k_{z2i}, d)}{1 + G_h(k_{z1i}, k_{z2i}, d)R_h(k_{zi}, k_{z1i})}\right]\left[\frac{1 + G_h(k_{z1nm}, k_{z2nm}, d)}{1 + G_h(k_{z1nm}, k_{z2nm}, d)R_h(k_{znm}, k_{z1nm})}\right] \quad (50)$$

$$g_\beta^{(1)} = \frac{-2ik_{zi}(k_0^2 - k_1^2)}{(\epsilon k_{znm} + k_{z1nm})(k_{zi} + k_{z1i})}\left(\frac{k_{z1nm}}{k_0}\right)s_{i,n} \cdot\left[\frac{1 + G_h(k_{z1i}, k_{z2i}, d)}{1 + G_h(k_{z1i}, k_{z2i}, d)R_h(k_{zi}, k_{z1i})}\right] \cdot\left[\frac{1 - G_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}, d)}{1 + G_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}, d)R_v(\epsilon k_{znm}, k_{z1nm})}\right] \quad (51)$$

while

$$g_\alpha^{(1)} = \frac{-2ik_{zi}(k_0^2 - k_1^2)}{(k_{znm} + k_{z1nm})(\epsilon k_{zi} + k_{z1i})}\left(\frac{k_{z1i}}{k_0}\right)s_{i,n} \cdot\left[\frac{1 + G_h(k_{z1nm}, k_{z2nm}, d)}{1 + G_h(k_{z1nm}, k_{z2nm}, d)R_h(k_{znm}, k_{z1nm})}\right] \cdot\left[\frac{1 - G_v(\epsilon_2 k_{z1i}, \epsilon k_{z2i}, d)}{1 + G_v(\epsilon_2 k_{z1i}, \epsilon k_{z2i}, d)R_v(\epsilon k_{zi}, k_{z1i})}\right] \quad (52)$$

$$g_\beta^{(1)} = \frac{-2ik_{zi}(k_0^2 - k_1^2)}{(\epsilon k_{znm} + k_{z1nm})(\epsilon k_{zi} + k_{z1i})} \cdot\left[\frac{1 + G_v(\epsilon_2 k_{z1i}, \epsilon k_{z2i}, d)}{1 + G_v(\epsilon_2 k_{z1i}, \epsilon k_{z2i}, d)R_v(\epsilon k_{zi}, k_{z1i})}\right] \cdot\left[\frac{1 + G_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}, d)}{1 + G_v(\epsilon_2 k_{z1nm}, \epsilon k_{z2nm}, d)R_v(\epsilon k_{znm}, k_{z1nm})}\right]$$

$$\cdot\left(\frac{\epsilon k_{\rho i}k_{\rho nm}}{k_0^2}-\frac{k_{z1i}k_{z1nm}}{k_0^2}c_{i,n}\cdot\left[\frac{1-G_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)}{1+G_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)}\right]\cdot\left[\frac{1-G_v(\epsilon_2 k_{z1i},\,\epsilon k_{z2i},\,d)}{1+G_v(\epsilon_2 k_{z1i},\,\epsilon k_{z2i},\,d)}\right]\right) \tag{53}$$

for a vertically polarized incident field. In the above equations, $c_{i,n}$ and $s_{i,n}$ represent the cosine and sine functions

$$c_{i,n}=\frac{k_{xi}k_{xn}+k_{yi}k_{ym}}{k_{\rho i}k_{\rho nm}} \tag{54}$$

$$s_{i,n}=\frac{k_{xi}k_{ym}-k_{yi}k_{xn}}{k_{\rho i}k_{\rho nm}} \tag{55}$$

while the terms in brackets are due to the presence of the layered medium. Note these terms become unity as the reflection at $z=-d$ vanishes (i.e., G_h and G_v approach zero), and first order results for a homogeneous medium are obtained. These results illustrate the "Bragg scatter" phenomenon of first order perturbation theory, since scattered fields at a particular angle [i.e., $(n,\,m)$] are directly proportional to the amplitude of a particular surface Fourier component. Equations (50)–(53) are identical to those in [12] when backscattering is considered but apply for general bistatic scattering angles.

A second order correction to the specularly reflected fields in region zero can also be derived as

$$\zeta_{0,0}^{(2)}=\sum_m\sum_n|h_{n,m}|^2 g_\zeta^{(2)}(k_{xi},\,k_{yi},\,k_{xn},\,k_{ym}) \tag{56}$$

where $\zeta=\alpha$ or β, and $g_\zeta^{(2)}$ is a corresponding function. For a horizontally polarized incident field, see (57) and (58), shown at the bottom of the page. For a vertically polarized incident field, see (59), shown at the bottom of the page, and $g_\alpha^{(2)}$ is -1 times $g_\beta^{(2)}$ for horizontal incidence in (58). The above results reduce to the second order specular reflection coefficient corrections described in [2] when the reflections at $z=-d$ vanish, except for a minus sign difference in cross polarized terms due to differing coordinate systems.

$$g_\alpha^{(2)}=\frac{-2k_{zi}(k_0^2-k_1^2)}{(k_{zi}+k_{z1i}Z_h(k_{z1i},\,k_{z2i},\,d))^2}\left\{k_{z1i}Z_h(k_{z1i},\,k_{z2i},\,d)+\frac{k_{znm}k_{z1nm}(1-\epsilon)Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)}{\epsilon k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)}+c_{i,n}^2\cdot\left(\frac{k_0^2-k_1^2}{k_{znm}+k_{z1nm}Z_h(k_{z1nm},\,k_{z2nm},\,d)}-\frac{k_{znm}k_{z1nm}(1-\epsilon)Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)}{\epsilon k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)}\right)\right\} \tag{57}$$

$$g_\beta^{(2)}=\frac{-2k_{zi}(k_0^2-k_1^2)}{(\epsilon k_{zi}+k_{z1i}Z_v(\epsilon_2 k_{z1i},\,\epsilon_1 k_{z2i},\,d))(k_{zi}+k_{z1i}Z_h(k_{z1i},\,k_{z2i},\,d))}\cdot\left\{-s_{i,n}c_{i,n}\frac{k_{z1i}(\epsilon-1)Z_v(\epsilon_2 k_{z1i},\,\epsilon_1 k_{z2i},\,d)}{k_{znm}+k_{z1nm}Z_h(k_{z1nm},\,k_{z2nm},\,d)}\cdot\left(-k_0+\frac{k_{znm}k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)(k_{znm}+k_{z1nm}Z_h(k_{z1nm},\,k_{z2nm},\,d))}{k_0(\epsilon k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d))}\right)-s_{i,n}\frac{\epsilon k_{\rho i}k_{\rho nm}(k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d))}{k_0(\epsilon k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d))}\right\}. \tag{58}$$

$$g_\beta^{(2)}=\frac{-2k_{zi}(k_0^2-k_1^2)}{(\epsilon k_{zi}+k_{z1i}Z_v(\epsilon_2 k_{z1i},\,\epsilon_1 k_{z2i},\,d))^2}\left\{\frac{-\epsilon k_{\rho i}^2k_{\rho nm}^2(\epsilon-1)}{k^2(\epsilon k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d))}-\epsilon k_{z1i}Z_v(\epsilon_2 k_{z1i},\,\epsilon_1 k_{z2i},\,d)-\frac{k_{z1i}^2(Z_v(\epsilon_2 k_{z1i},\,\epsilon_1 k_{z2i},\,d))^2(1-\epsilon)}{k_{znm}+k_{z1nm}Z_h(k_{z1nm},\,k_{z2nm},\,d)}+c_{i,n}\left(\frac{2\epsilon k_{\rho i}k_{\rho nm}k_{z1i}Z_v(\epsilon_2 k_{z1i},\,\epsilon_1 k_{z2i},\,d)}{k^2}\right)\left(\frac{k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)}{\epsilon k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)}\right)+c_{i,n}^2\left(\frac{k_{z1i}^2(Z_v(\epsilon_2 k_{z1i},\,\epsilon_1 k_{z2i},\,d))^2(1-\epsilon)}{k_{znm}+k_{z1nm}Z_h(k_{z1nm},\,k_{z2nm},\,d)}-\frac{k_{z1i}^2k_{znm}k_{z1nm}(1-\epsilon)Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d)(Z_v(\epsilon_2 k_{z1i},\,\epsilon_1 k_{z2i},\,d))^2}{k^2(\epsilon k_{znm}+k_{z1nm}Z_v(\epsilon_2 k_{z1nm},\,\epsilon k_{z2nm},\,d))}\right)\right\}. \tag{59}$$

Polarimetric brightness temperatures of a periodic surface can be calculated through the application of Kirchhoff's Law

$$\overline{T}_B = \begin{bmatrix} T_{Bh} \\ T_{Bv} \\ T_U \\ T_V \end{bmatrix} = T_s \left(\begin{bmatrix} 1 - r_h \\ 1 - r_v \\ -r_U \\ -r_V \end{bmatrix} \right) \tag{60}$$

where T_{Bh} and T_{Bv} are the brightness temperatures measured by horizontally and vertically polarized antennas, respectively, T_U and T_V are proportional to the real and imaginary parts of the correlation between fields in horizontal and vertical polarizations, respectively ([2]), and T_s refers to the layered medium physical temperature in Kelvin (assumed constant throughout the layered medium). Total reflectivities for the periodic surface to second order in surface height are shown in (61), at the bottom of the next page, where the first term is the reflectivity of the layered medium with a flat interface, and the following two terms are the Bragg scattering and reflection coefficient correction contributions, respectively. In the aforementioned equations, the subscripts hh and hv refer to g_α functions with horizontal incidence and vertical incidence, respectively, while vh and vv refer to g_β functions with horizontal and vertical incidence, respectively. Note that both of the summations are in terms of sums over the periodic surface power spectral density $|h_{n,m}|^2$ and can be combined into a single term that expresses the correction to brightness temperatures caused by surface roughness. In the limit that surface periods become large compared to both the electromagnetic wavelength and any roughness features, the sums can be replaced with integrals over the continuous power spectral density $W(k_{xn} - k_{xi}, k_{ym} - k_{yi}) = (|h_{n,m}|^2/\delta k_x \delta k_y)$, where $\delta k_x = (2\pi/P_x)$ and $\delta k_y = (2\pi/P_y)$. The final result for continuous surface brightness temperatures is shown in (62), at the bottom of the page, where the new $\tilde{g}$ functions include both the Bragg scatter and reflection coefficient correction terms described previously and are functions of the radiometer polar observation angle θ_i, the radiometer azimuthal observation angle ϕ_i, the layered medium properties ϵ, ϵ_2, and d, and the integration variables k'_ρ and ϕ'. Note also that a coordinate shift has been made in the above integration: $\tilde{g}$ functions are evaluated as in (61), except that $k_{xn} = k_{xi} + k'_\rho \cos\phi'$ and $k_{ym} = k_{yi} + k'_\rho \sin\phi'$. The dependence of the $\tilde{g}$ functions on $\phi' - \phi_i$ is found from (61), and the dependence on ϵ_2 is replaced by a dependence on R_h and R_v at $z = -d$ in the case of a general layered medium. Equation (62) expresses the brightness temperature of a layered medium bounded by a slightly rough interface in terms of the brightness temperature of the layered medium with a flat interface and a roughness correction. The roughness correction is obtained through an integration of the surface power spectral density W weighted by the $\tilde{g}$ weighting functions, which are distinct for each polarimetric quantity. Studies of these weighting functions therefore allow the physics of rough surface thermal emission to be examined independent of the particular surface power spectral density and are considered in the next section.

IV. Study of Weighting Functions for Isotropic Surfaces

To simplify the analysis, assume that an isotropic surface is considered (i.e., one with no directional properties) so that U and V brightnesses are zero [2] and so that the roughness correction becomes

$$\begin{aligned} \Delta T_{B\gamma} &= -T_s \left(\int_0^\infty dk'_\rho k'_\rho W(k'_\rho) \int_0^{2\pi} d\phi' \right. \\ &\qquad \left. \tilde{g}_\gamma \cdot (f, \theta_i, \epsilon, \epsilon_2, d, k'_\rho, \phi' - \phi_i) \right) \\ &= -T_s \left(\int_0^\infty dk'_\rho k'_\rho W(k'_\rho) \tilde{g}_{\gamma,0}(f, \theta_i, \epsilon, \epsilon_2, d, k'_\rho) \right) \end{aligned} \tag{63}$$

$$\begin{aligned} \begin{bmatrix} r_h \\ r_v \\ r_U \\ r_V \end{bmatrix} &= \begin{bmatrix} |R_{Hl}|^2 \\ |R_{Vl}|^2 \\ 0 \\ 0 \end{bmatrix} + \sum_m \sum_n |h_{n,m}|^2 \frac{\mathrm{Re}\{k_{znm}\}}{k_{zi}} \begin{bmatrix} |g_{hh}^{(1)}(k_{znm})|^2 + |g_{vh}^{(1)}(k_{znm})|^2 \\ |g_{hv}^{(1)}(k_{znm})|^2 + |g_{vv}^{(1)}(k_{znm})|^2 \\ 2\mathrm{Re}\{g_{hv}^{(1)}(k_{znm}) g_{hh}^{(1)*}(k_{znm}) + g_{vv}^{(1)}(k_{znm}) g_{vh}^{(1)*}(k_{znm})\} \\ 2\mathrm{Im}\{g_{hv}^{(1)}(k_{znm}) g_{hh}^{(1)*}(k_{znm}) + g_{vv}^{(1)}(k_{znm}) g_{vh}^{(1)*}(k_{znm})\} \end{bmatrix} \\ &\quad + \sum_m \sum_n |h_{n,m}|^2 \begin{bmatrix} 2\mathrm{Re}\{R_{Hl}^* g_{hh}^{(2)}(k_{xi}, k_{yi}, k_{xn}, k_{ym})\} \\ 2\mathrm{Re}\{R_{Vl}^* g_{vv}^{(2)}(k_{xi}, k_{yi}, k_{xn}, k_{ym})\} \\ 2\mathrm{Re}\{(R_{Hl}^* - R_{Vl}^*) g_{hv}^{(2)}(k_{xi}, k_{yi}, k_{xn}, k_{ym})\} \\ 2\mathrm{Im}\{(R_{Hl}^* + R_{Vl}^*) g_{hv}^{(2)}(k_{xi}, k_{yi}, k_{xn}, k_{ym})\} \end{bmatrix} \end{aligned} \tag{61}$$

$$\begin{bmatrix} T_{Bh} \\ T_{Bv} \\ T_U \\ T_V \end{bmatrix} = T_s \left(\begin{bmatrix} 1 - |R_{Hl}|^2 \\ 1 - |R_{Vl}|^2 \\ 0 \\ 0 \end{bmatrix} - \int_0^\infty dk'_\rho k'_\rho \int_0^{2\pi} d\phi' W(k'_\rho, \phi') \begin{bmatrix} \tilde{g}_h(f, \theta_i, \epsilon, \epsilon_2, d, k'_\rho, \phi' - \phi_i) \\ \tilde{g}_v(f, \theta_i, \epsilon, \epsilon_2, d, k'_\rho, \phi' - \phi_i) \\ \tilde{g}_U(f, \theta_i, \epsilon, \epsilon_2, d, k'_\rho, \phi' - \phi_i) \\ \tilde{g}_V(f, \theta_i, \epsilon, \epsilon_2, d, k'_\rho, \phi' - \phi_i) \end{bmatrix} \right) \tag{62}$$

where $\gamma = h$ or v. Studies of the $\tilde{g}_{\gamma,0}$ functions reveal further simplification if a factor of k_0^2 is removed

$$\Delta T_{B\gamma} = -T_s k_0^2 \cdot \left(\int_0^\infty dk'_\rho k'_\rho W(k'_\rho) \tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon_2, k_0 d, k'_\rho/k_0) \right) \tag{64}$$

illustrating that the $\tilde{\tilde{g}}$ weighting functions do not explicitly depend on frequency if length scales relative to the electromagnetic wavelength (i.e., $k_0 d$ and k'_ρ/k_0) are considered.

Fig. 2(a) plots $\tilde{\tilde{g}}_{h,0}$ and $\tilde{\tilde{g}}_{v,0}$ versus k'_ρ/k_0 for a two layer medium with $d = 0.1\lambda$, $\epsilon = 7.5 + i0.67$, $\epsilon_2 = 3 + i\ 0.08$ (intended to model a clay medium at 3 GHz with upper and lower layer moisture contents of approximately 20 and 5%) and for $\theta_i = 30°$. Logarithmic scales are used for both the horizontal and vertical axes to enable a large range of scales to be observed. Signs of these functions (defined as $+1$ for positive values and -1 for negative values) are displayed in Fig. 2 plot (b), with the $\tilde{\tilde{g}}_{h,0}$ and $\tilde{\tilde{g}}_{v,0}$ sign curves shifted by $+2$ and -2, respectively, to enable them to be distinguished. Note the constant valued weighting functions obtained as k'_ρ/k_0 becomes small (large length scales in the spectrum relative to λ). Notating the value of this constant as $\tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon_2, k_0 d, 0)$, Fig. 2 plots (c) and (d) illustrate the scaled difference functions

$$\tilde{\tilde{g}}_{\gamma,s} = \frac{\tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon_2, k_0 d, k'_\rho/k_0) - \tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon_2, k_0 d, 0)}{(k'_\rho/k_0)^2} \tag{65}$$

and their signs, respectively. Roughness induced changes in brightnesses can then be rewritten as

$$\begin{aligned}\Delta T_{B\gamma} &= -T_s k_0^2 \left(\tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon_2, k_0 d, 0) \int_0^\infty dk'_\rho k'_\rho W(k'_\rho) + \frac{1}{k_0^2} \int_0^\infty dk'_\rho k'^3_\rho W(k'_\rho) \tilde{\tilde{g}}_{\gamma,s}(\theta_i, \epsilon, \epsilon_2, k_0 d, k'_\rho/k_0) \right) \\ &= -T_s \left(\frac{k_0^2 h^2}{2\pi} \tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon_2, k_0 d, 0) + \int_0^\infty dk'_\rho k'^3_\rho W(k'_\rho) \tilde{\tilde{g}}_{\gamma,s}(\theta_i, \epsilon, \epsilon_2, k_0 d, k'_\rho/k_0) \right) \end{aligned} \tag{66}$$

where h^2 is the surface height variance. The above equation shows that $\tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon_2, k_0 d, 0)$ indicates a dependence on the surface variance, while $\tilde{\tilde{g}}_{\gamma,s}(\theta_i, \epsilon, \epsilon_2, k_0 d, k'_\rho/k_0)$ represents a function that weights the spectrum in computing the surface slope variance.

Fig. 2 (a) and (c) both show the "critical phenomena" effects [17] observed in the homogeneous medium case, and the vertical lines included in plot (c) mark the boundaries of the region within which critical phenomena can occur. As in the homogeneous medium case, Fig. 2 (c) demonstrates that length scales both much larger than or comparable to the electromagnetic wavelength can contribute to the roughness-induced correction through $\tilde{\tilde{g}}_{\gamma,s}$. The importance of $\tilde{\tilde{g}}_{\gamma,s}$ contributions, however, depends strongly on the magnitude of the $(k_0^2 h^2/2\pi)\tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon_2, k_0 d, 0)$ product.

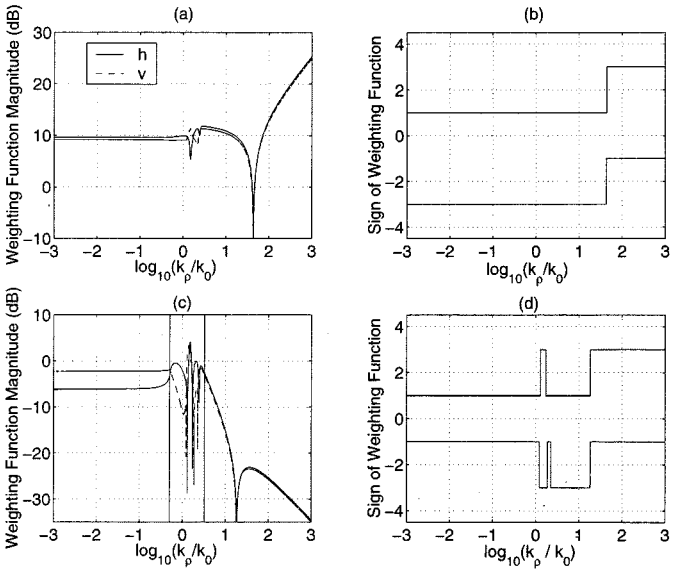


Fig. 2. Weighting functions for a two layer medium with $\theta_i = 30°$, $d = 0.1\lambda$, $\epsilon = 7.5 + i\ 0.67$ and $\epsilon_2 = 3 + i\ 0.08$. (a) Magnitude of weighting functions $\bar{g}_{\gamma,0}$. (b) Sign of weighting functions $\bar{g}_{\gamma,0}$. (c) Magnitude of weighting functions $\bar{g}_{\gamma,s}$. (d) Magnitude of weighting functions $\bar{g}_{\gamma,s}$. Note the sign curves are shifted by $+2$ and -2 for h and v polarizations, respectively.

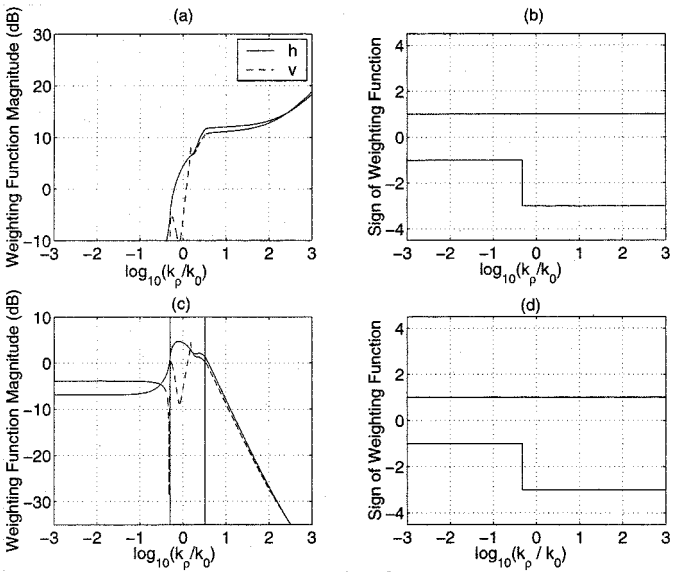


Fig. 3. Same as Fig. 2, but for $\epsilon_2 = 7.5 + i\ 0.67$.

To clarify the relationship between emission for a rough surface bounding, a two layer medium versus a homogeneous medium, Fig. 3 (a)–(d) illustrate the same functions as in Fig. 2, except that ϵ_2 is modified to equal ϵ, so that a homogeneous medium exists below the rough surface. Note the dramatic change in plot (a), as the constant-valued weighting functions for small k'_ρ/k_0 no longer occur since $\tilde{\tilde{g}}_{\gamma,0}(\theta_i, \epsilon, \epsilon, k_0 d, 0) = 0$. Thus, with a homogeneous medium, a surface height variance-dependent term is not obtained, resulting in a small slope approximation [7]. As is evident from Fig. 2 (a) however, height-dependent terms do not vanish in the layered medium case, and the theory remains a small height expansion.

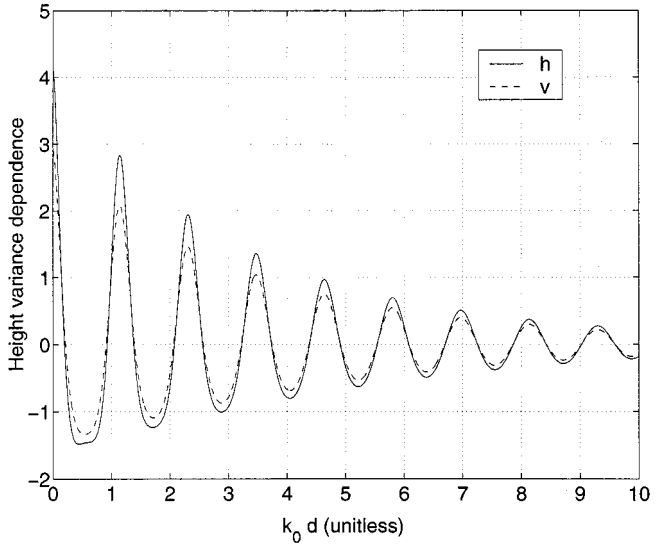


Fig. 4. Surface height dependent term $(1/2\pi)\bar{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)$ for the layered medium of Fig. 2 versus $k_0 d$.

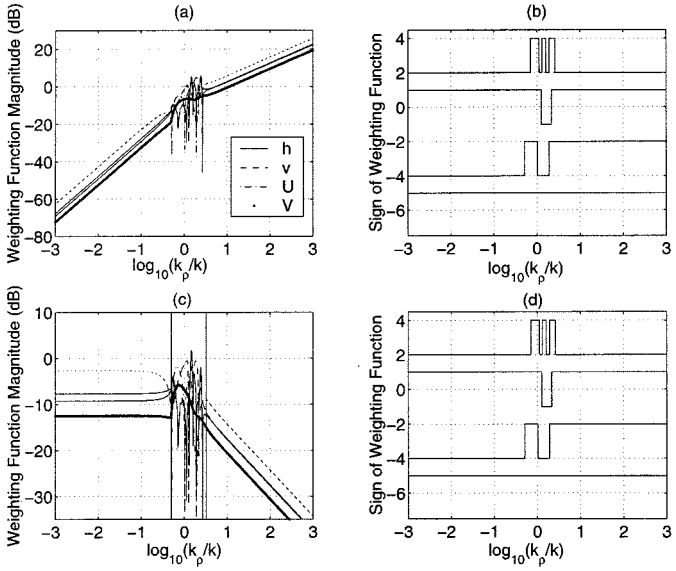


Fig. 5. Second harmonic weighting functions for a two layer medium with $\theta_i = 30°$, $d = 0.1\lambda$, $\epsilon = 7.5 + i\ 0.67$, and $\epsilon_2 = 3 + i\ 0.08$. (a) Magnitude of weighting functions $\bar{g}_{\gamma,2}$. (b) Sign of weighting functions $\bar{g}_{\gamma,2}$. (c) Magnitude of weighting functions $\bar{g}_{\gamma,2}/(k'_\rho/k_0)^2$. (d) Magnitude of weighting functions $\bar{g}_{\gamma,2}/(k'_\rho/k_0)^2$. Note the sign curves are shifted by $+3$, 0, -3, and -6 for h, v, U, and V polarizations, respectively.

Note also that the magnitude of the height variance-dependent correction is determined by $\tilde{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)$, which remains a function of the layered medium properties. Thus, convergence of the series is not determined by rough surface parameters alone. An estimate of the accuracy of the second order correction for surfaces with small slopes can be obtained by examining the magnitude of the $(k_0^2 h^2/2\pi)\tilde{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)$ product. Small values of this product ($\leq$0.025) should indicate that the second order theory provides reasonable answers, while larger values may require a higher order theory in order to obtain accurate predictions. Fig. 4 plots $(1/2\pi)\tilde{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)$ for the two layer medium above as a function of $k_0 d$. Note the periodic increases that occur. A reasonable prediction of roughness-induced corrections can still be obtained for depths at which $(1/2\pi)\tilde{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)$ becomes large, but a correspondingly smaller surface height variance would be required. A verification of the suggested convergence test was performed through comparison with a numerical solution of the SPM equations as described in [18], which enabled the fourth order correction to be determined. Choosing $|(1/2\pi)k_0^2 h^2 \tilde{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)| \leq 0.025$ [or equivalently, $kh \leq (0.4/\sqrt{|\tilde{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)|})$] was found in several tests to provide fourth order corrections significantly smaller than second order, although cases in which $\tilde{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)$ vanishes remain problematic and require careful consideration. Tests of the higher order theory indicate that a small slope expansion is not necessarily achieved when $\tilde{\tilde{g}}_{\gamma,0}(\theta_i,\ \epsilon,\ \epsilon_2,\ k_0 d,\ 0)$ vanishes unless the physical configuration approaches that of the homogeneous medium (i.e., region one becomes sufficiently lossy to obscure reflections from region two.)

V. Study of Weighting Functions for Asymmetric Surfaces

For surfaces that are not symmetric in azimuth, brightness temperatures become a function of the radiometer azimuthal observation angle ϕ_i and U, and V polarimetric brightnesses become nonzero. In this case, the roughness-induced correction is

$$\Delta T_{B\gamma} = -T_s \left(\int_0^\infty dk'_\rho k'_\rho \int_0^{2\pi} d\phi' W(k'_\rho,\ \phi') \tilde{g}_\gamma(f,\ \theta_i,\ \epsilon,\ \epsilon_2,\ d,\ k'_\rho,\ \phi' - \phi_i) \right) \tag{67}$$

where $\gamma = h,\ v,\ U$, or V. The above expression can be simplified as described in [17] through an expansion of both $W(k'_\rho,\ \phi')$ and $\tilde{g}_\gamma$ into Fourier series in azimuth. It is also assumed that the surface spectrum contains only even cosine harmonics. The resulting expression for the roughness-induced correction is shown in (68), at the bottom of the page, where

$$\Delta T_{B\gamma} = -T_s k_0^2 \left(\begin{bmatrix} \int_0^\infty dk'_\rho k'_\rho \tilde{\tilde{g}}_{\gamma,0}(k'_\rho/k_0) W_0(k'_\rho) \\ 0 \end{bmatrix} + \sum_{n=1}^{\infty} \begin{bmatrix} 2\cos(n\phi_i) \int_0^\infty dk'_\rho k'_\rho \operatorname{Re}\{\tilde{\tilde{g}}_{\gamma,n}(k'_\rho/k_0)\} W_n(k'_\rho) \\ -2\sin(n\phi_i) \int_0^\infty dk'_\rho k'_\rho \operatorname{Im}\{\tilde{\tilde{g}}_{\gamma,n}(k'_\rho/k_0)\} W_n(k'_\rho) \end{bmatrix} \right) \tag{68}$$

the upper row applies for h and v brightnesses, while the lower row applies for U and V. In the above equation

$$\tilde{\tilde{g}}_{\gamma,n}(k'_\rho/k_0) = \frac{1}{k_0^2}\int_0^{2\pi} d\phi' e^{in\phi'}\tilde{g}_\gamma(f,\theta_i,\epsilon,\epsilon_2,d,k'_\rho,\phi') \tag{69}$$

and is also a function of θ_i, ϵ, ϵ_2, and $k_0 d$, while

$$W_n(k'_\rho) = \frac{1}{2\pi}\int_0^{2\pi} d\phi' e^{-in\phi'} W(k'_\rho,\phi'). \tag{70}$$

Equation (68) demonstrates that particular azimuthal harmonics of polarimetric brightness temperatures [i.e., the $\cos(n\phi_i)$ and $\sin(n\phi_i)$ terms] are given by integrals of a distinct weighting function for each azimuthal harmonic ($\tilde{\tilde{g}}_{\gamma,n}(k'_\rho/k_0)$), multiplied with the corresponding surface spectrum harmonic $W_n(k'_\rho)$. Under these definitions, the zeroth harmonic weighting functions are identical to those considered in Figs. 2 and 3.

Fig. 5 plots the second harmonic weighting functions $\mathrm{Re}\{\tilde{\tilde{g}}_{\gamma,2}\}$ for h and v and $\mathrm{Im}\{\tilde{\tilde{g}}_{\gamma,2}\}$ for U and V versus k'_ρ/k_0 for the case considered in Fig. 2. Plot (a) in this figure illustrates the magnitudes of the second harmonic weighting functions in dB, while plot (b) illustrates their signs. Note again the four sign functions are shifted in steps of 3 to allow the curves to be more easily distinguished. The constant-valued weighting functions for small k'_ρ/k_0 observed in Fig. 2 are not obtained in this case, indicating that the second order SPM produces a surface slope-dependent term for second azimuthal harmonics. Plots (c) and (d) of Fig. 5 illustrate the weighting functions divided by $(k'_\rho/k_0)^2$, as in plots (c) and (d) of Fig. 2 and confirm the slope dependence. A slope rather than height dependence for azimuthal harmonics with $n > 0$ would be advantageous, particularly in studies of sea surface emission. However, further tests with the numerical SPM solution (again, implemented following the procedure in [18]) showed that azimuthal harmonics with $n > 0$ depend directly on the surface height variance at fourth and higher orders. Thus, the theory remains an expansion in surface height for all azimuthal harmonics of surface brightness temperatures, even though the second order prediction of azimuthal harmonics with $n > 0$ does not directly involve the surface height variance. Predictions of the second order theory for second harmonic coefficients were found adequate in cases for which the convergence requirement specified in Section IV was satisfied.

VI. Sample Results

To illustrate the influence of layered media on roughness-induced emission corrections, sample brightness temperatures are considered in this section. Fig. 6 illustrates results for surfaces with $T_s = 283$ K and with an isotropic, Gaussian roughness spectrum, completely characterized by the root mean squared (rms) surface height h and correlation length l parameters. Note a Gaussian roughness spectrum is not necessarily realistic for soil surfaces, but is commonly applied in theoretical studies due to its simplicity. Roughness induced brightness temperature corrections for the two layer medium considered in Fig. 2

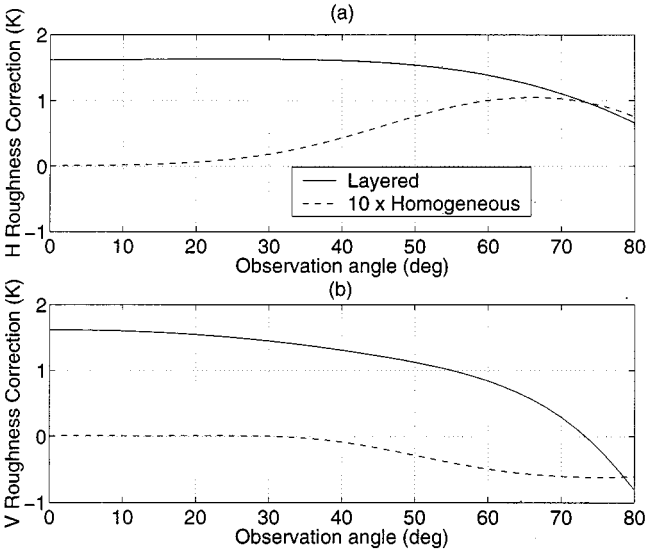


Fig. 6. Comparison of roughness induced brightness corrections for a two layer medium with $d = 0.1\lambda$, $\epsilon = 7.5 + i\ 0.67$, and $\epsilon_2 = 3 + i\ 0.08$ with a homogeneous medium $\epsilon_2 = 7.5 + i\ 0.67$ versus radiometer polar observation angle. The surface has an isotropic Gaussian roughness spectrum with $h = 0.01\lambda$ and $l = \lambda$. (a) Horizontal polarization and (b) vertical polarization. Note that homogeneous medium results are multiplied by ten in these plots.

and for $h = 0.01\lambda$, $l = \lambda$ are plotted versus observation angle in Fig. 6 and compared with those for a homogeneous medium with $\epsilon_2 = \epsilon$. Note the significant differences observed, both in the amplitude of roughness corrections and in their variations with observation angle. Homogeneous medium roughness corrections plotted in Fig. 6 are multiplied by ten to make their variations more clear. Height-dependent factors for all the cases shown were found to be <0.006, so that the second order correction should be accurate. Clearly, the presence of a layered medium below a rough interface can cause large changes in the influence of surface roughness on the medium boundary. Brightness temperatures for the flat surface medium are of course significantly different in the homogeneous and layered medium cases as well.

Fig. 7 illustrates the dependence of roughness induced emission corrections on layer depth d. The configuration is the same as that of Fig. 6 and for polar observation angle 30°. Note the oscillatory pattern observed versus depth, indicating the presence of coherent effects that are not disrupted by the small roughness used. Homogeneous medium roughness corrections for this case are less than 0.02 K in both polarizations.

Fig. 8 plots second harmonic coefficients of brightness temperatures versus observation angle for surfaces with an anisotropic Gaussian roughness spectrum with $h = 0.02\lambda$, $l_x = 0.5\lambda$, and $l_y = 1\lambda$, where l_x and l_y represent the correlation lengths in the x and y directions, respectively. Parameters of the layered medium are the same as those of previous examples, and again layered medium results are compared with those of a homogeneous medium. In this rougher surface case, height dependent factors remain <0.025 for all cases illustrated, so that second order theory predictions should be reasonable. Again the results show that the presence of a layered medium can cause significant changes in azimuthal variations of brightness temperatures, although the differences are smaller than those

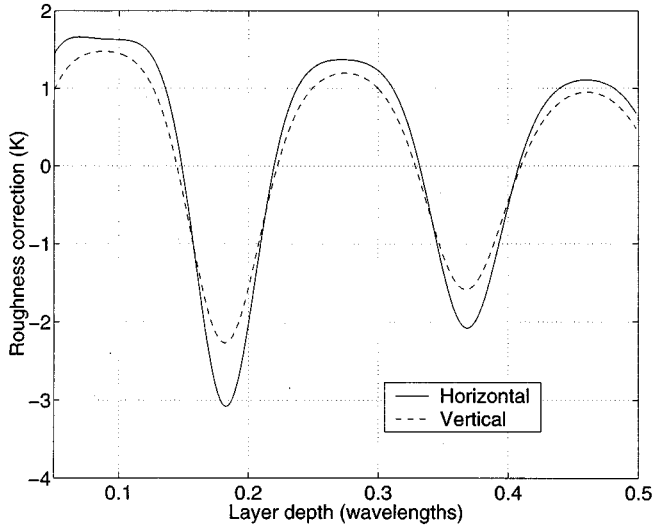


Fig. 7. Layer depth dependence of roughness induced brightness corrections for a two layer medium with $\epsilon = 7.5 + i\ 0.67$ and $\epsilon_2 = 3 + i\ 0.08$ at 30° polar observation angle. The surface has an isotropic Gaussian roughness spectrum with $h = 0.01\lambda$ and $l = \lambda$.

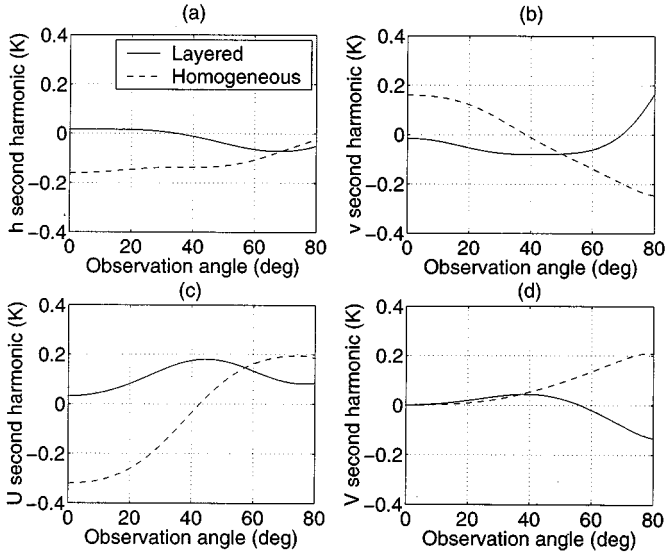


Fig. 8. Comparison of second harmonic brightness temperatures for a two layer medium with $d = 0.1\lambda$, $\epsilon = 7.5 + i\ 0.67$, and $\epsilon_2 = 3 + i\ 0.08$ with a homogeneous medium $\epsilon_2 = 7.5 + i\ 0.67$ versus radiometer polar observation angle. The surface has a Gaussian roughness spectrum with $h = 0.2\lambda$, $l_x = 0.5\lambda$, and $l_y = \lambda$. (a) Horizontal polarization, (b) vertical polarization, (c) U brightness, and (d) V brightness.

of Fig. 6 due to the absence of a height variance dependence in the second order prediction of second harmonic coefficients.

VII. Conclusions

Expressions for slightly rough surface-induced corrections to layered medium thermal emission have been derived through the small perturbation method in this paper. Results show that an expansion in surface height, as opposed to surface slope, is obtained, but accurate predictions can still be computed from the theory as long as the specified convergence rules are followed. The theory can be applied to problems in the remote sensing of soil moisture, sea ice, or sea surfaces to assess the effect of surface roughness on layered medium brightness temperatures. Typically, surface roughness corrections are expected to be small in microwave passive remote sensing, but current sensors are sufficiently accurate to observe these variations in many cases. Understanding and including rough surface effects in retrieval models can therefore potentially lead to more accurate sensing of layered medium parameters. The modification proposed in [12] for modeling scattering from a finite size object buried beneath an interface can also be applied in the emission equations to produce an approximate emission theory for a finite object buried beneath the ground. Studies of microwave radiometry for buried object detection like those of [10], [11], but including rough interface effects can therefore also be performed with the theory developed.

Finally, note that the change from a small slope to small height theory in the layered medium case raises some interesting issues, in particular regarding the magnitude of a change in permittivity in the medium required to introduce a significant height dependence. This question can be important in studies of sea surface remote sensing, since the height variance of sea surfaces is typically very large with respect to the wavelength at higher microwave frequencies, making a small height theory of sea surface emission impractical. The convergence expressions proposed can be applied to address these questions for a specified layered medium, but a more accurate emission theory will be needed in cases for which the convergence requirements are not satisfied.

References

[1] L. Tsang and R. W. Newton, "Microwave emissions from soils with rough surfaces," *J. Geophys. Res.*, vol. 87, pp. 9017–9024, 1982.

[2] S. H. Yueh, R. Kwok, F. K. Li, S. V. Nghiem, and W. J. Wilson, "Polarimetric passive remote sensing of ocean wind vectors," *Radio Sci.*, vol. 29, pp. 799–814, 1994.

[3] S. H. Yueh, "Modeling of wind direction signals in polarimetric sea surface brightness temperatures," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 1400–1418, 1997.

[4] D. B. Kunkee and A. J. Gasiewski, "Simulation of passive microwave wind direction signatures over the ocean using an asymmetric-wave geometrical optics model," *Radio Sci.*, vol. 32, p. 59, 1997.

[5] V. G. Irisov, "Microwave radiation from a weakly non-Gaussian surface," in *Int. Geoscience and Remote Sensing Symp. '98 Conf. Proc.*, vol. 5, 1998, pp. 2329–2332.

[6] A. K. Fung and M. F. Chen, "Emission from an inhomogeneous layer with irregular surfaces," *Radio Sci.*, vol. 16, pp. 289–298, 1981.

[7] V. G. Irisov, "Small-slope expansion for thermal and reflected radiation from a rough surface," *Waves Random Media*, vol. 7, pp. 1–10, 1997.

[8] A. G. Voronovich, *Wave Scattering from Rough Surfaces*. Berlin, Germany: Springer-Verlag, 1994.

[9] M. Zhang and J. T. Johnson, "Theoretical studies of ocean polarimetric brightness signatures," in *Int. Geoscience and Remote Sensing Symp. '98 Conf. Proc.*, vol. 5, Seattle, WA, 1998, pp. 2333–2335.

[10] G. De Amici, B. Hauss, and L. Yujiri, "Detection of landmines via a passive microwave radiometer," in *Proc. SPIE*, vol. 3710, 1999.

[11] J. T. Johnson, "Theoretical study of microwave radiometry for buried object detection," in *Proc. SPIE*, vol. 4098, 2000.

[12] I. M. Fuks, "Radar contrast polarization dependence on subsurface sensing," in *Int. Geoscience and Remote Sensing Symp. '98 Conf. Proc.*, vol. 3, 1998, pp. 1455–1459.

[13] I. M. Fuks and A. G. Voronovich, "Wave diffraction by rough interfaces in an arbitrary plane layered medium," *Waves Random Media*, vol. 10, pp. 253–272, 2000.

[14] P. I. Arseev, "Perturbation theory for the Green's function of an electromagnetic field on a rough surface," *Sov. Phys. JETP*, vol. 65, pp. 262–267, 1987.

[15] G. V. Rozhnov, "Electromagnetic wave diffraction by multilayer media with rough interfaces," *Sov. Phys. JETP*, vol. 69, pp. 646–651, 1989.

[16] N. P. Zhuk, A. V. Frankov, and A. G. Yarovoy, "Backward scattering of electromagnetic waves by the rough surface of a medium whose dielectric permittivity has an exponential profile," *Sov. J. Commun. Technol. Electron.*, vol. 38, pp. 108–111, 1993.

[17] J. T. Johnson and M. Zhang, "Theoretical study of the small slope approximation for ocean polarimetric thermal emission," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2305–2736, Sept. 1999.

[18] J. T. Johnson, "Third order small perturbation method for scattering from dielectric rough surfaces," *J. Opt. Soc. Amer. A*, vol. 16, no. 11, pp. 2720–2736, 1999.

[19] S. O. Rice, "Reflection of electromagnetic waves from slightly rough surfaces," *Commun. Pure Appl. Math*, vol. 4, pp. 361–378, 1951.

[20] G. R. Valenzuela, "Depolarization of EM waves by slightly rough surfaces," *IEEE Trans. Antennas Propagat.*, vol. AP-15, pp. 552–557, 1967.

Joel T. Johnson (M'96) received the B.E.E. degree from the Georgia Institute of Technology, Atlanta, in 1991, and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1993 and 1996, respectively.

He is currently an Associate Professor in the Department of Electrical Engineering and ElectroScience Laboratory, The Ohio State University, Columbus. His research interests are in the areas of microwave remote sensing, propagation, and electromagnetic wave theory.

Dr. Johnson is an Associate Member of commissions B and F of the International Union of Radio Science (URSI) and a member of Tau Beta Pi, Eta Kappa Nu, and Phi Kappa Phi. He received the 1993 best paper award from the IEEE Geoscience and Remote Sensing Society, and was named an Office of Naval Research Young Investigator, National Science Foundation Career awardee, and PECASE award recipient in 1997.

An Analytical Model for Studies of Soil Modification Effects on Ground Penetrating Radar

Jatupum Jenwatanavet and Joel T. Johnson, *Member, IEEE*

***Abstract*—Due to the similar dielectric constants of buried nonmetallic targets and dry soils, it is often difficult to detect and identify nonmetallic targets with ground penetrating radar. The addition of properly chosen chemical agents to modify soil properties can potentially provide improved detection. Previous studies using waveguide experiments have shown that the addition of water improves dielectric contrasts but also increases loss so that target detectability is not necessarily improved. The addition of liquid nitrogen to wet soils can reduce background medium loss and restore target visibility, and waveguide studies of target detection through controlled depth of nitrogen penetration have shown that scattering can be significantly enhanced if an optimal amount of nitrogen is added. In this paper, a simple physical optics (PO) model for scattering from a three-dimensional target buried below a half space is presented, and it is shown that the radar cross section of the target depends on the dielectric contrast with and attenuation in the background medium. The model is validated through comparison with a Method of Moments code and found to yield accurate predictions for near normal incidence geometries. Analytical studies of target detection with two concepts of soil modification are then described: obtaining an "optimal" homogeneous soil water content and the addition of a large quantity of water along with an optimal amount of liquid nitrogen. Finally, initial tests of these soil modification techniques with a dielectric rod antenna ground penetrating radar are performed and demonstrate that the addition of liquid nitrogen to excessively wet soils can reduce loss and enhance target visibility.**

***Index Terms*—Ground penetrating radar, landmine detection, radar cross section, subsurface sensing.**

I. Introduction

THE detection and identification of nonmetallic anti-personnel landmines remains a challenging problem for all current technologies [1]. Ground penetrating radar (GPR) systems are currently only of limited utility in this area because of the often low dielectric (i.e., complex permittivity) contrast between plastic mines and the surrounding soil. For example, the relative permittivity ϵ_r of most plastic materials has a very small imaginary part and a real part in the range of 2–4, while that of most dry soils is similar, making it difficult for a sensor which relies on scattering from dielectric contrasts to distinguish these two materials. Previous S-band waveguide studies [2] have shown that the homogeneous addition of water to dry sand can increase target/background medium dielectric contrasts so that larger scattered fields can be obtained. However, attenuation in the background medium is also increased and reduces scattered fields. For homogeneous water content soils an optimal water content exists to provide maximum target visibility, but for deeper targets even this "maximum" response can be quite small due to attenuation. For the nonhomogeneous water content soils created when water is poured onto a soil surface and allowed to penetrate over a moderate time period, previous waveguide experiments [3] have shown that moderate depth target visibility is not improved because a substantial quantity of water is required to obtain penetration to target depths and results in excessive losses near the surface for GPR systems operating at microwave frequencies. Thus the addition of water alone may not be sufficient to enhance target detection in many environments. Although many soil and target physical parameters can also influence GPR detection of buried objects [4], this study focuses on the influence of dielectric contrast on microwave frequency GPR systems for detection of small nonmagnetic, nonmetallic objects buried at relatively shallow depths (such as anti-personnel mines).

To address problems with excessive loss in wet soils, the addition of chemicals to the soil medium has been proposed [2], [3]. Although such a use of chemicals will clearly have many important practical issues which must be resolved before use in field tests is possible, the electromagnetic effects must first be considered in order to determine appropriate chemical choices. One chemical which has been proposed for modifying soil properties is liquid nitrogen [2], [3], which potentially could be generated on site for use in the field. Due to the much smaller loss tangents in ice than in water [5], target visibility can potentially be regained in excessively lossy soil situations. However, the dielectric constant of frozen soil also becomes more similar to that of nonmetallic targets, so again small dielectric contrasts become a problem in completely frozen soils. Waveguide experimental results [3] suggested that an "optimal" amount of liquid nitrogen should exist to enhance target visibility; this amount of nitrogen would be sufficient to freeze most of the soil above the target so that propagation loss is reduced, but would leave a small layer of wet soil around the target to retain dielectric contrast so that large scattered fields could be obtained. The importance of flow effects for water and liquid nitrogen have also been discussed [3]; improvements in target visibility in waveguide experiments were determined by subtracting measurements made with and without a surrogate mine target. Differences in the flow patterns of water and liquid nitrogen with and without a target can also be measured, but careful control of these chemicals was exercised to obtain repeatable measurements. Further waveguide analytical studies of soil modification through the addition of

Manuscript received October 8, 1999; revised December 11, 2000. This work was supported by NSF under Project ECS-9701678 and by a grant from Duke University as part of the Office of the Secretary of Defense MURI on Humanitarian Deminig.

The authors are with the Department of Electrical Engineering, ElectroScience Laboratory, Ohio State University, Columbus, OH 43212 USA.

Publisher Item Identifier S 0018-926X(01)05251-6.

0018–926X/01$10.00 © 2001 IEEE

water and liquid nitrogen [6] suggested a procedure for locating low-contrast targets at unknown depths through a gradual addition of increasing amounts of liquid nitrogen. Waveguide experiments confirmed this concept, even with the complication of water and liquid nitrogen flow effects not considered in the analytical studies. The resulting procedure allows unknown depth targets which have *exactly* the same dielectric constant as the original soil medium to be detected in principle.

In this paper, previous results are extended beyond the waveguide configuration through an analytical study of the physical optics approximation for scattering from a three dimensional target buried below a half space. Although many numerical codes [7]–[12] have been developed for this problem, the analytical equations obtained in this paper allow insight into the effects of soil modification. Results show that the radar cross section (RCS) of a buried target is proportional to the reflection coefficient between the target and the surrounding medium, confirming the importance of dielectric contrast. The dependence on background medium attenuation and transmission through the soil/air interface is also clarified. The model is validated by comparison with a method of moments (MoM) code and found to yield reasonable predictions for near normal incidence angles. Analytical studies of target detection with two concepts of soil modification are also described: adding water to obtain an "optimal" homogeneous soil water content and the addition of a large quantity of water along with an optimal amount of liquid nitrogen. Results from the analytical study of the optimal homogeneous soil water content show that optimal water contents decrease when target depths increase since loss even in the lower water content soils is significant for very deep targets. The analytical study of the addition of a large quantity of water along with an optimal amount of liquid nitrogen demonstrates the potential of liquid nitrogen to reduce loss and restore target responses. Finally, the soil modification techniques suggested by these analytical studies are applied in an initial measurement with a dielectric rod 2–6-GHz GPR system developed at The Ohio State University ElectroScience Laboratory [13], and results suggest that the techniques proposed should be applicable to general microwave GPR systems for detecting nonmetallic anti-personnel mines.

II. Physical Optics Model for the RCS of a Buried Target

In this study, it is assumed that a rectangular target with dimensions $\Delta x \times \Delta y \times h$, permeability μ_0, and complex permittivity ϵ_2 is buried at depth d inside a soil background medium with permeability μ_0 and complex permittivity ϵ_1. The geometry for a perpendicular polarized plane wave incident at angle θ_i with respect to the z axis is shown in Fig. 1. The goal of this study is to determine the RCS of the target by using the PO approximation on the top surface of the target. The target RCS will also be computed when the target is in free space and the ratio of buried/nonburied RCS considered to clarify soil medium effects. In both cases, total electric and magnetic fields at the top surface of the target must be determined to compute the equivalent electric and magnetic current densities of the target, and

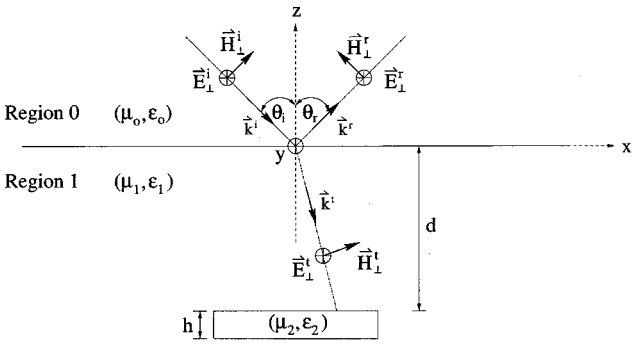


Fig. 1. Geometry for oblique incidence reflection: Perpendicular polarization.

these equivalent current densities are used to calculate the electric and magnetic vector potentials. The scattered electric field in the far zone can then be computed from these vector potentials. Multiple reflections between the ground and target surface are neglected in this analysis.

According to Fig. 1, a perpendicular polarized incident electric and magnetic field in region 0 can be written as

$$\vec{E}^i_\perp = \hat{y} E_0 e^{-j(k^i_x x - k^i_z z)} \tag{1}$$

$$\vec{H}^i_\perp = \left(\frac{\hat{x}k^i_z + \hat{z}k^i_x}{\omega\mu_0}\right) E_0 e^{-j(k^i_x x - k^i_z z)} \tag{2}$$

where E_0 is the amplitude of the incident electric field and an $e^{j\omega t}$ time convention is used. The total electric and magnetic fields in the soil medium (region 1) are modeled as a combination of the transmitted fields and fields reflected from the target. In the PO approximation, reflections from the top surface of the target are modeled as reflections from an infinite layer, so that total fields in region 1 are written as

$$\vec{E}_\perp = \hat{y} T_\perp E_0 e^{-j(k^i_x x - k^t_z z)} + \hat{y}\tilde{\Gamma}'_\perp T_\perp E_0 e^{-j(k^i_x x + k^t_z z)} \tag{3}$$

$$\vec{H}_\perp = \left(\frac{\hat{x}k^t_z + \hat{z}k^i_x}{\omega\mu_0}\right) T_\perp E_0 e^{-j(k^i_x x - k^t_z z)} + \left(\frac{-\hat{x}k^t_z + \hat{z}k^i_x}{\omega\mu_0}\right) \tilde{\Gamma}'_\perp T_\perp E_0 e^{-j(k^i_x x + k^t_z z)} \tag{4}$$

where $T_\perp$ is the transmission coefficient from region 0 to region 1

$$T_\perp = \frac{2k^i_z}{k^i_z + k^t_z}. \tag{5}$$

$\tilde{\Gamma}'_\perp$ is the reflection coefficient at the top surface of the target which can be derived by using the usual multilayer boundary condition procedure for the top surface of a dielectric slab with thickness h located at distance d below the ground surface

$$\tilde{\Gamma}'_\perp = \tilde{\Gamma}_\perp e^{-j2k^t_z d} \tag{6}$$

$$\tilde{\Gamma}_\perp = \frac{\left(\tilde{k}^{t2}_z - k^{t\,2}_z\right)\left(1 - e^{-j2\tilde{k}^t_z h}\right)}{\left(\tilde{k}^t_z - k^t_z\right)^2 e^{-j2\tilde{k}^t_z h} - \left(\tilde{k}^t_z + k^t_z\right)^2}. \tag{7}$$

Note that the propagation constants in each region are defined as

$$k_x^i = \omega\sqrt{\mu_0\epsilon_0}\sin\theta_i \tag{8}$$

$$k_z^i = \omega\sqrt{\mu_0\epsilon_0}\cos\theta_i \tag{9}$$

$$k_z^t = \sqrt{\omega^2\mu_0\epsilon_1 - \omega^2\mu_0\epsilon_0\sin^2\theta_i} \tag{10}$$

$$\tilde{k}_z^t = \sqrt{\omega^2\mu_0\epsilon_2 - \omega^2\mu_0\epsilon_0\sin^2\theta_i}. \tag{11}$$

Under the PO approximation, the equivalent electric and magnetic current densities at the top surface of the target are

$$\vec{J}_p = \hat{y}\frac{k_z^t T_\perp E_0}{\omega\mu_0}[1-\tilde{\Gamma}_\perp]e^{-j(k_x^i x + k_z^t d)} \tag{12}$$

$$\vec{M}_p = \hat{x}T_\perp E_0[1+\tilde{\Gamma}_\perp]e^{-j(k_x^i x + k_z^t d)} \tag{13}$$

and backscattered electric and magnetic vector potentials in the far field are approximately

$$\vec{A} \simeq \frac{\mu_0 e^{-jk_0 r}e^{-jk_z^t d}\tilde{T}_\perp}{4\pi r}\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}}\int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}}\vec{J}_p e^{-jk_x^i x'}\,dx'\,dy' \tag{14}$$

$$\vec{F} \simeq \frac{\epsilon_1 e^{-jk_0 r}e^{-jk_z^t d}\tilde{T}_\perp}{4\pi r}\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}}\int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}}\vec{M}_p e^{-jk_x^i x'}\,dx'\,dy'. \tag{15}$$

Here r is the distance between the origin and the observation point, and k_0 is the propagation constant in free space equal to $\omega\sqrt{\mu_0\epsilon_0}$. Also $\tilde{T}_\perp = 2k_z^t/(k_z^t + k_z^i)$ is the transmission coefficient from region 1 to 0. Equations (14) and (15) are constructed from the basic idea that the target is replaced by the equivalent electric and magnetic current sources under the PO approximation. The electric and magnetic fields radiated by these sources in the soil medium then propagate until meeting the interface between the soil medium and free space, at which point the transmission coefficient from region 1 to 0 is included. Substituting $\vec{J}_p$ and $\vec{M}_p$ from (12) and (13) into (14) and (15) and computing the integral, the vector potentials are shown in (16) and (17) at the bottom of the page. Scattered electric fields corresponding to these electric and magnetic vector potentials are shown in (18) and (19) at the bottom of the page, where $\eta_1 = \sqrt{\mu_0/\epsilon_1}$ is the intrinsic impedance of the soil medium. Therefore, the total scattered electric field can be written as

$$\vec{E}_\perp^s = \vec{E}_A^s + \vec{E}_F^s \tag{20}$$

$$= \hat{y}\frac{k_z^t T_\perp E_0 \tilde{T}_\perp e^{-jk_0 r}e^{-j2k_z^t d}\Delta x\Delta y\sin\left(k_x^i\Delta x\right)2\tilde{\Gamma}_\perp}{4\pi r k_x^i\Delta x}. \tag{21}$$

For normal incidence, (21) becomes

$$\vec{E}_\perp^s = \hat{y}\frac{k_1 T_\perp E_0 \tilde{T}_\perp e^{-jk_0 r}e^{-j2k_1 d}\Delta x\Delta y 2\tilde{\Gamma}_\perp}{4\pi r}. \tag{22}$$

The backscattering RCS is

$$\sigma = \lim_{r\to\infty}\left[\frac{4\pi r^2|\vec{E}^s|^2}{|\vec{E}^i|^2}\right]. \tag{23}$$

Substituting $\vec{E}_\perp^i$ and $\vec{E}_\perp^s$ from (1) and (22) into (23), the normal incidence RCS becomes

$$\sigma_\perp = \left(|k_1 T_\perp \tilde{T}_\perp \tilde{\Gamma}_\perp|^2 e^{-4\Im[k_1 d]}\right)\left(\frac{\Delta x^2\Delta y^2}{\pi}\right)|Q^2| \tag{24}$$

where k_1 is the propagation constant in the soil medium equal to $\omega\sqrt{\mu_0\epsilon_1}$ and $\Im$ indicates the imaginary part. An additional factor Q^2 has been introduced in (24) to account for modification of the power density per unit angle of the scattered spherical wave due to refraction through the interface. Consideration of the plane wave representation of a spherical wave shows that the factor $Q = k_z^i/k_z^t$. Note that the effect of spherical wave propagation between the two boundaries will cancel out if the incident wave from the source in the air region is a spherical wave and not a plane wave. Q can thus be written as

$$Q = \frac{k_z^i}{k_z^t} \quad \text{for plane wave incidence} \tag{25}$$

$$= 1 \quad \text{for spherical wave incidence.} \tag{26}$$

For a target in free space, $T_\perp = \tilde{T}_\perp = 1$, and the reflection coefficient at the top surface of the target $(\tilde{\Gamma}_\perp)$ from (7) is written as $\tilde{\Gamma}_\perp^0$. The normal incidence backscattering RCS of the target in free space can then be written as

$$\sigma_\perp^0 = \left|k_0\tilde{\Gamma}_\perp^0\right|^2\left(\frac{\Delta x^2\Delta y^2}{\pi}\right). \tag{27}$$

$$\vec{A} = \hat{y}\frac{k_z^t T_\perp E_0[1-\tilde{\Gamma}_\perp]\tilde{T}_\perp e^{-jk_0 r}e^{-j2k_z^t d}\Delta x\Delta y\sin\left(k_x^i\Delta x\right)}{4\pi r\omega k_x^i\Delta x} = \hat{y}A \tag{16}$$

$$\vec{F} = \hat{x}\frac{\epsilon_1 T_\perp E_0[1+\tilde{\Gamma}_\perp]\tilde{T}_\perp e^{-jk_0 r}e^{-j2k_z^t d}\Delta x\Delta y\sin\left(k_x^i\Delta x\right)}{4\pi r k_x^i\Delta x} = \hat{x}F. \tag{17}$$

$$\vec{E}_A^s \simeq -\hat{y}j\omega A \simeq -\hat{y}\frac{j\omega k_z^t T_\perp E_0[1-\tilde{\Gamma}_\perp]\tilde{T}_\perp e^{-jk_0 r}e^{-j2k_z^t d}\Delta x\Delta y\sin\left(k_x^i\Delta x\right)}{4\pi r\omega k_x^i\Delta x} \tag{18}$$

$$\vec{E}_F^s \simeq \hat{y}\eta_1 j\omega F \simeq \hat{y}\frac{j\omega\eta_1 k_z^t\epsilon_1 T_\perp E_0[1+\tilde{\Gamma}_\perp]\tilde{T}_\perp e^{-jk_0 r}e^{-j2k_z^t d}\Delta x\Delta y\sin\left(k_x^i\Delta x\right)}{4\pi r\omega\sqrt{\mu_0\epsilon_2}k_x^i\Delta x} \tag{19}$$

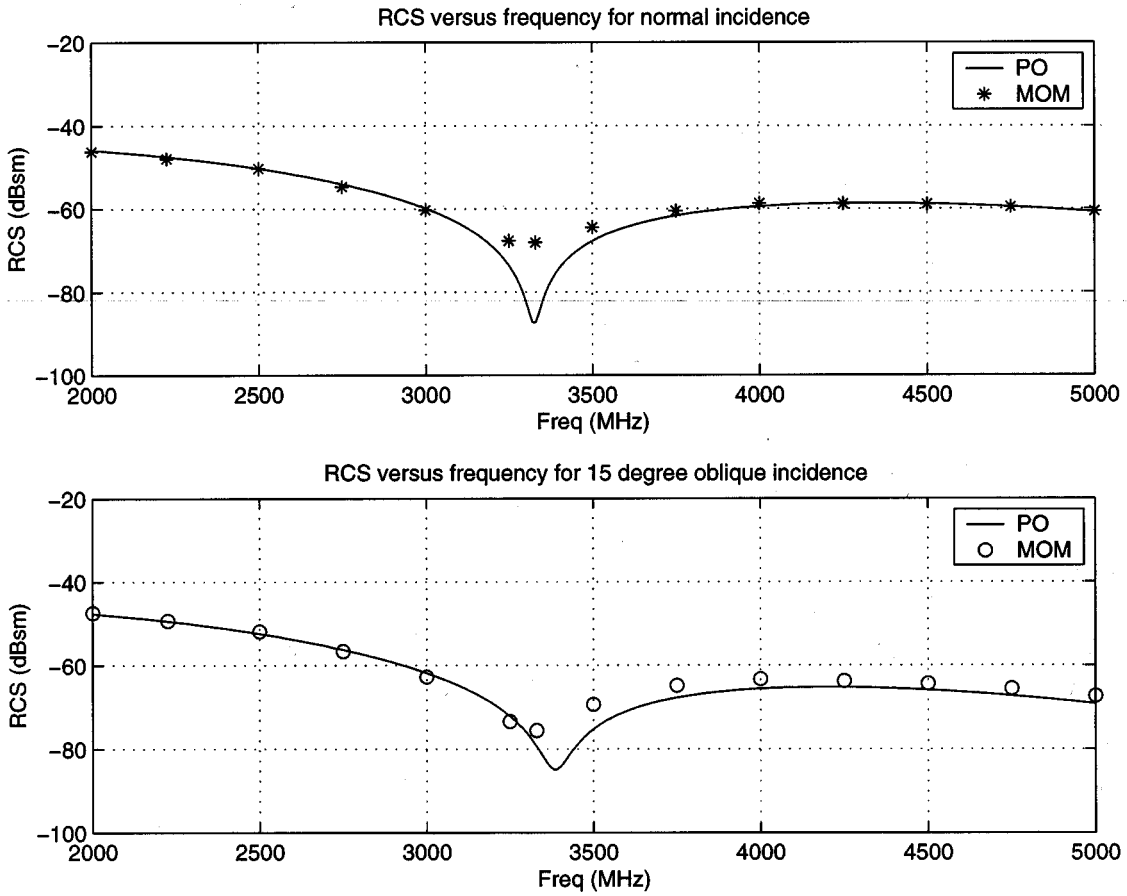


Fig. 2. Comparison of analytically calculated RCS from PO and numerically calculated RCS from MoM versus frequency for normal and 15 degree oblique incidence.

Finally, the ratio between the normal incidence target RCS in the soil medium and in free space can be calculated by dividing (24) by (27) to obtain

$$\frac{\sigma_\perp}{\sigma_\perp^0} = \left| \frac{k_1 \tilde{T}_\perp T_\perp \tilde{\Gamma}_\perp}{k_0 \tilde{\Gamma}_\perp^0} \right|^2 e^{-4\Im[k_1 d]} |Q^2|. \tag{28}$$

Results for parallel polarization can be derived in a similar manner and are not detailed here; the emphasis of the current study on near normal incidence should result in only slight differences between polarizations.

Obviously from (28), the ratio of normal incidence backscattered RCS of a target in the soil medium with respect to that for a a target in free space is dependent on reflection and transmission coefficients and on attenuation in the background medium. The reflection coefficient term $|\tilde{\Gamma}_\perp|^2$ clearly will increase as the dielectric contrast between the target and soil medium becomes larger, although the transmission coefficients $|\tilde{T}_\perp T_\perp|^2$ will tend to decrease as background medium contrast with free space increases.

To validate this simple PO model, the backscattered RCS of a target for 0 degree and 15 degree plane wave incidence on a soil medium was calculated and compared with numerical results from a MoM code [12]. In this validation a nylon target ($\epsilon_r \approx 3.15 - j0.03$ from S-band waveguide measurements described in [2]) with cross section 7.62 cm by 7.62 cm and thickness of 2.54 cm was assumed to be buried at 7.62 cm below the ground surface. Soil with dielectric constant $5 - j1.25$ (from S-band waveguide measurements of homogeneous sand with 7.5% water content [2]) was used as the background medium. The frequency range was chosen to be 2–5 GHz. Fig. 2 shows the comparison of the analytically calculated RCS from the PO model and the numerically calculated RCS from the MoM for both normal and 15 degree oblique incidence (using the fields of (21) in the PO model at oblique incidence) in perpendicular polarization. The analytical data were computed in 0.01 GHz steps whereas the numerical data used 0.25 GHz steps, due to the much larger computational requirements of the MOM code (for example 1060 unknowns used at 4 GHz, requiring 2.5 h CPU time on a 200-MHz Pentium Pro processor). From the plot a null is observed in PO predictions at frequency 3330 MHz due to a resonance in the dielectric target reflection coefficient. Clearly the analytical and numerical data match reasonably well. The average level of error excluding the null frequency is about 0.77 and 1.43 dB for normal and 15 degree oblique incidence, respectively. Fig. 3 demonstrates another comparison of analytically calculated RCS from the PO model and numerically calculated RCS from the MoM versus incidence angle at 4 GHz for perpendicular and parallel polarizations. Again, both analytical and numerical data show reasonable agreement, although the differences increase at more oblique angles as expected for a PO model. These results demonstrate

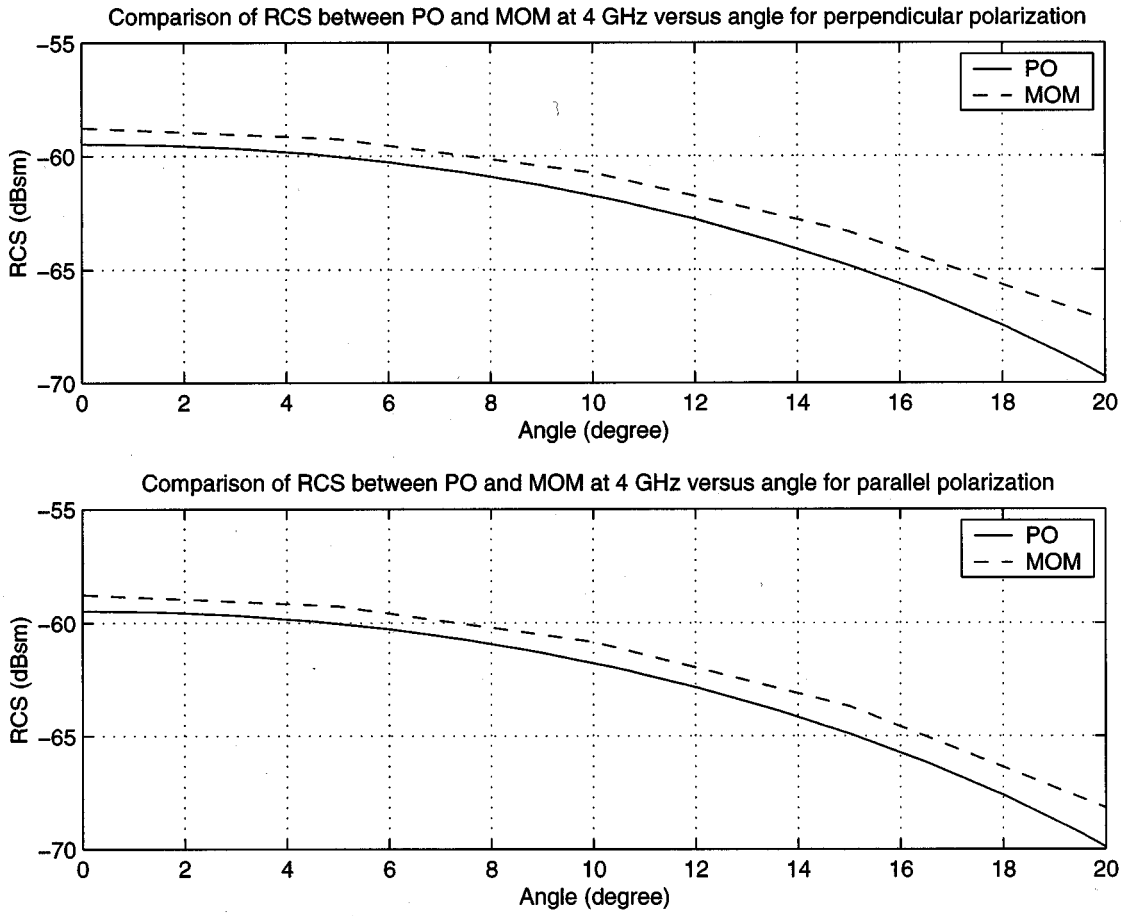


Fig. 3. Comparison of analytically calculated RCS from PO and numerically calculated RCS from MoM versus angle at 4 GHz for perpendicular and parallel polarizations.

that a simple PO model can be valid for RCS studies of a buried target at near normal incidence angles.

A comparison between the PO model and a GPR measurement is next considered. The experimental result was obtained using a dielectric rod antenna GPR system [13], depicted in Fig. 4. A network analyzer was set to transmit continuous wave (CW) electromagnetic signals with IF bandwidth 1000 Hz and power 10 dBm in 0.1 GHz steps from 2 to 6 GHz. The dielectric rod antenna was located at 1 in above the ground surface in order to obtain a very small illuminated area, and attached to a horizontal movement controller so that measurements could be taken as a function of position. The system has a basic calibration procedure involving two small cylindrical wire scatterers, but the near field nature of this system makes a complete calibration difficult so only qualitative image responses are compared. The target considered in this measurement was a 7.62-cm diameter nylon disk of 2.54 cm thickness buried at 6.35 cm depth. The measured result is presented in Fig. 5(a) while the analytical result is presented in Fig. 5(b). The theoretical result was generated using (21) (with $Q = 1$ for spherical wave incidence) to calculate the scattered electric field assuming that the dielectric constants of soil and nylon are the same as in the previous case except that the target was buried at 2.5-in depth. In this study, clutter returns from the ground surface which occur in measured early time images (which in practice require more advanced suppression methods [14]) are not included in the analytical model. To model the rod antenna movement, additional amplitude and phase factors must be included in (21). According to Fig. 6 which illustrates the geometry for the rod antenna movement, the factor $e^{-j2dk_z^t}$ in (21) should be replaced by $[e^{-|2\Im[k_z^t]d|}e^{-j2(k_x^i a+\Re[k_z^t]d)}e^{-j2k_0 s_1}]/[s_1+s_2]$ to take into account attenuation in the soil medium, spherical wave propagation, and phase shifts in both media; here $\Re$ indicates real part. Here h is the vertical distance from the tip of the rod antenna to the ground surface, x is the horizontal distance from the tip of the rod antenna to the center of the target and d is the depth of the target, all of which are known. The parameters a, s_1, and s_2 are determined by

$$s_1 = \sqrt{h^2 + b^2} \tag{29}$$

$$s_2 = \sqrt{d^2 + a^2} \tag{30}$$

$$b = h\tan\theta_i \tag{31}$$

$$a = x - b \tag{32}$$

$$x = h\tan\theta_i + d\tan\theta_t = h\tan\theta_i + \frac{d\sin\theta_i}{\Re\left[\sqrt{\frac{\epsilon_1}{\epsilon_0} - \sin^2\theta_i}\right]}. \tag{33}$$

The parameter θ_t in (33) is not the transmitted angle, which is a complex number, but the angle at which the constant phase

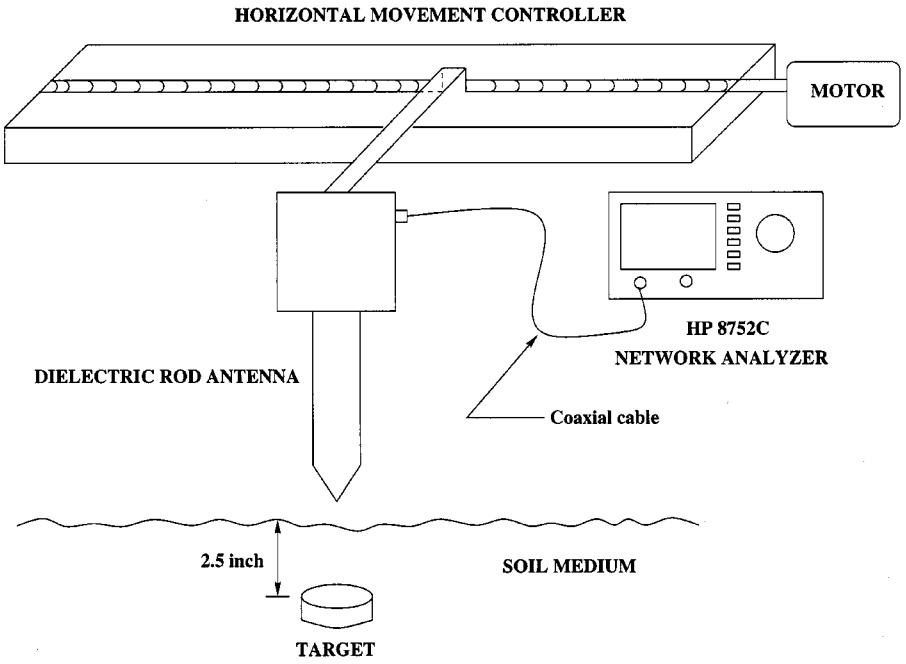
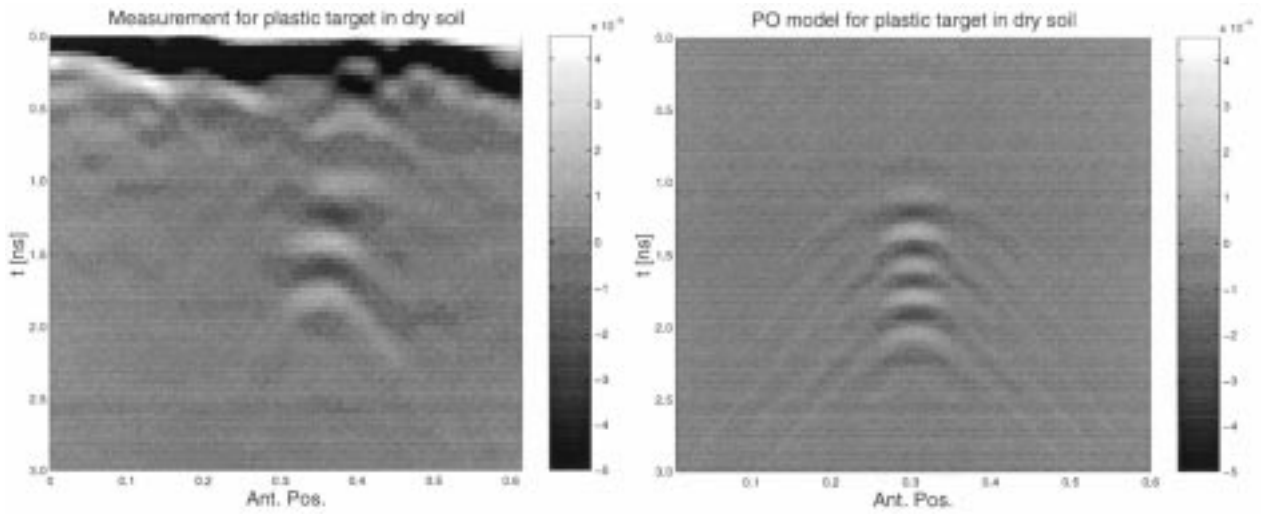


Fig. 4. Experimental configuration for dielectric rod antenna GPR system.



(a) (b)

Fig. 5. Comparison of time domain image scattered fields from PO model and GPR measurement: plastic target $(\epsilon_r = 3.15 - j0.03)$ buried at 2.5-in depth in soil $(\epsilon_r = 5 - j1.25)$; field units are defined such that their magnitude squared in the frequency domain is the radar cross section in square meters. (a) Measurement. (b) PO model.

planes of the transmitted plane wave are inclined at with respect to the z axis. The parameter θ_i is found by numerically solving the nonlinear equation (33). Both images in Fig. 5 are produced by plotting time domain scattered fields as a function of antenna horizontal position, and results are calibrated so that time 0 nS corresponds to the top surface of the ground. Note that both images are plotted with the same amplitude scale. As observed from the plot, features of both results are similar, with target responses occuring around the same time; however more oscillations are observed at later times in theoretical model results. In the theoretical model, the permittivity of the soil medium is assumed to be constant but the natural soil medium in the measurement is not homogeneous, so that it is possible to obtain for example more loss for deeper soil. However, since these two

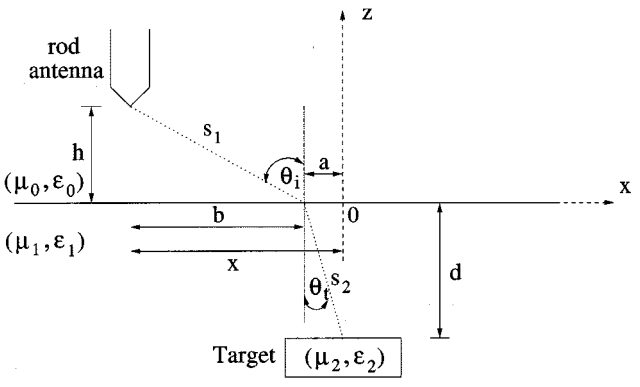


Fig. 6. Geometry for modeling dielectric rod antenna movement.

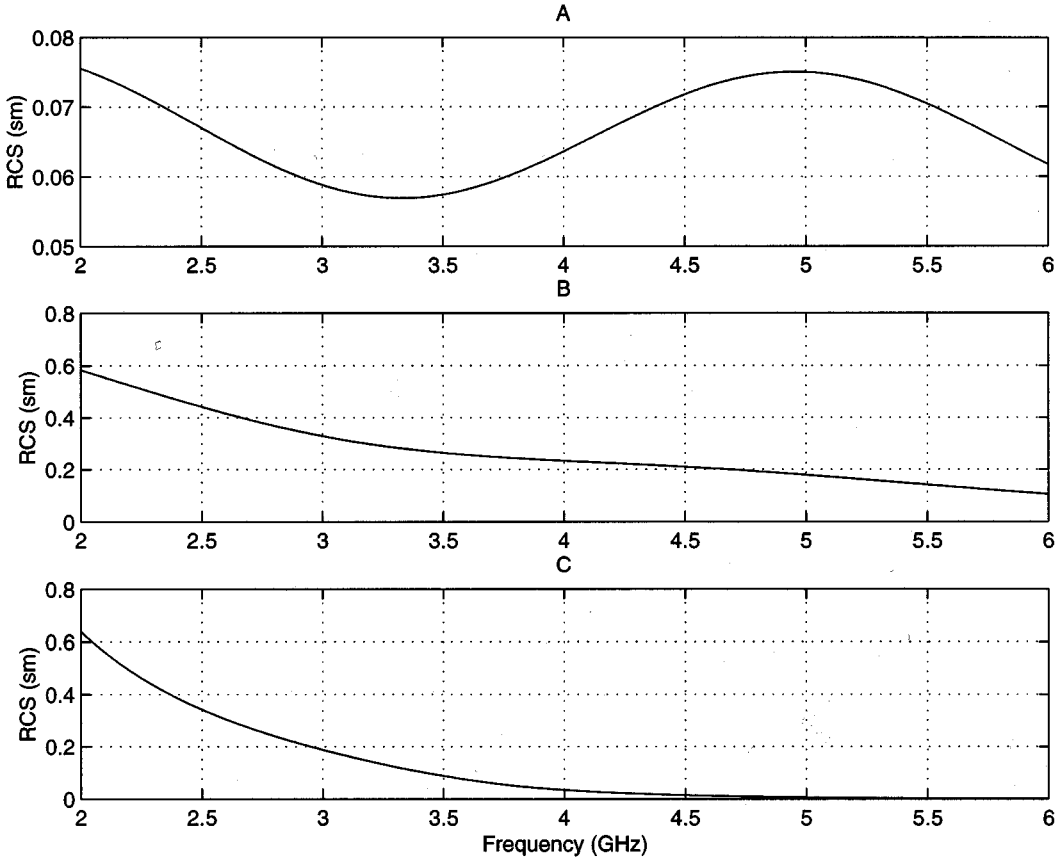


Fig. 7. Ratio of backscattering RCS of nylon target versus frequency for normal incidence in soil medium with respect to free space when soil medium is (A) dry sand (B) sand with 7.5% water content (C) sand with 30% water content.

images are in good qualitative agreement in general, this comparison also supports the validation of the PO model. The following section will concentrate on an analytical study of target detection using the PO model.

III. Analytical Study of Target Detection

To demonstrate that the RCS of a buried target depends on the dielectric contrast between the target and background medium, (28) was used to generate the ratio of backscattering RCS of a target in a homogeneous water content soil medium with respect to free space for normally incident spherical waves. In this study the nylon target $(\epsilon_r \approx 3.15 - j0.03)$ with thickness of 2.54 cm was assumed to be buried at 7.62 cm below the top ground surface. Sand with three different homogeneous water contents was used as the background medium: dry sand $(\epsilon_r \approx 2.6 - j0.005)$, sand with 7.5% water content $(\epsilon_r \approx 5.8 - j0.3)$ and sand with 30% water content $(\epsilon_r \approx 17.8 - j2)$; these permittivity values again were obtained from S-band waveguide measurements [2]. Dry sand was used in the case of no modification and expected to produce a small RCS. Sand with 7.5% water content was used in the case of "optimal" water modification and expected to produce larger RCS. Sand with 30% water content was used in the case of excessive water modification and expected to produce a small RCS again. Fig. 7 illustrates the simulated ratio of backscattering RCS with respect to free space of the nylon target versus frequency for normal incidence in the three soil media. The frequency range was chosen to be 2–6 GHz, and analytical data were calculated in 0.01-GHz steps. Plot (A) shows that the RCS of the plastic target in dry sand is small because of the low dielectric contrast. In plot (B), the RCS of the plastic target is larger than that in plot (A) since there is more dielectric contrast in this case; note the scale in plot (B) has been increased by a factor of ten. The RCS of plot (B) is around 4.5 times that of plot (A) at the center frequency (4 GHz). In plot (C) the RCS of the plastic target becomes obscured at high frequencies because electromagnetic wave absorption is significant at high frequencies. Even though the dielectric contrast in this case is larger than that in plot (B), loss in the background medium prohibits the incident wave from reaching the target so that the RCS is still low. However, at low frequencies it could be possible to detect the target since the loss is small compared to the loss at high frequencies. Clearly the RCS of the target in this plot is close to that in plot (A) at the center frequency. These results again demonstrate the dependence of buried target RCS on the dielectric contrast with and attenuation in the background medium.

The concept of soil modification through the addition of water to obtain an "optimal" homogeneous soil water content is next considered in more detail. The optimal homogeneous soil water content is defined here as the value which gives the largest peak backscattered field response in the time domain. An empirical model [15] is used to estimate the relative permittivity of soils for a given homogeneous water content and frequency and requires soil physical parameters such as volumetric moisture and

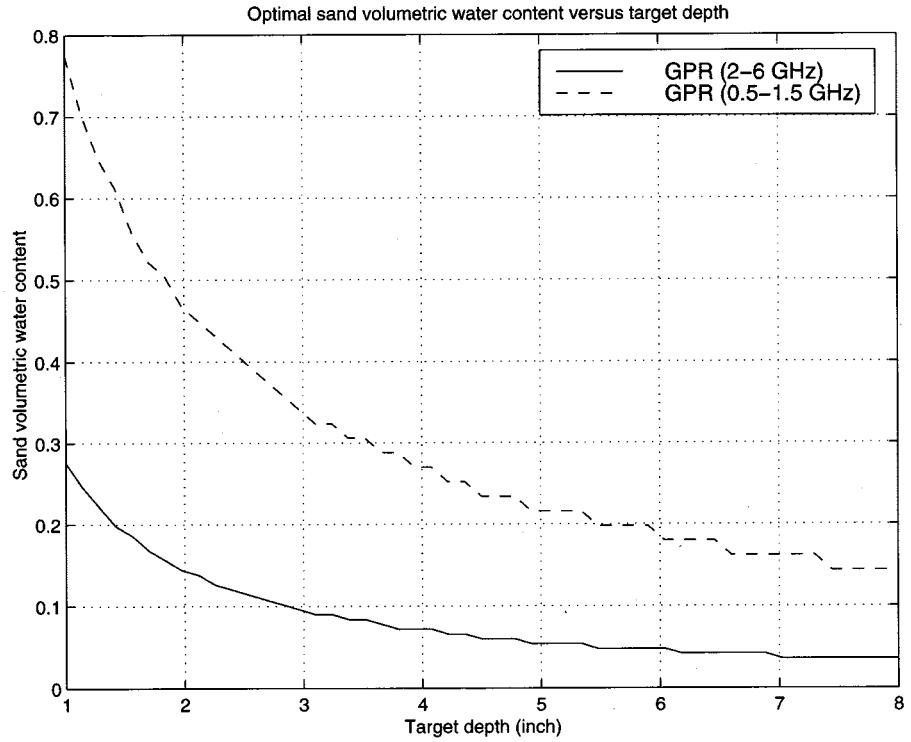


Fig. 8. Optimal sand volumetric water content obtained in analytical study versus target depth.

soil textural composition as input data. The analytical PO model is used along with the dielectric constant from this empirical model to perform a theoretical study of the maximum target response for each water content. Fig. 8 illustrates optimal homogeneous sand (bulk density $= 1.55$ g/cm^3) water contents obtained as a function of target depth. In this study, the target depths were varied from 1 to 8 in (0.0254 to 0.2032 m), and the relative permittivity of the nylon target was assumed to be $3 - j0.08$. Two frequency ranges (2–6 GHz and 0.5–1.5 GHz) were chosen to generate this plot. As observed from the plot, optimal water contents decrease when frequency increases because loss is more significant at higher frequencies. Optimal homogeneous soil water contents also decrease when target depths increase since loss even in the lower water content soil will become significant for very deep targets, and in many cases it may be difficult to obtain the optimal water content. However, the addition of liquid nitrogen onto a soil medium can reduce loss and restore target visibility due to the much smaller loss tangents in ice than water.

Fig. 9 illustrates soil modification through the addition of water and liquid nitrogen procedure proposed to locate targets at a known depth. In plot (A) of Fig. 9, a low contrast target is buried at a known depth d in dry sand. Scattered returns in plot (A) would be small due to the small dielectric contrast between the plastic target and dry sand medium. The second figure of plot (A) shows the configuration after a large amount of water is added onto the top sand surface. Scattered returns in this case would again be small due to an excessive attenuation in the very wet sand above the target. Fig. 9 plots (B) through (D) now illustrate the configuration as increasing amounts of liquid nitrogen are added, so that the frozen sand region above the target extends

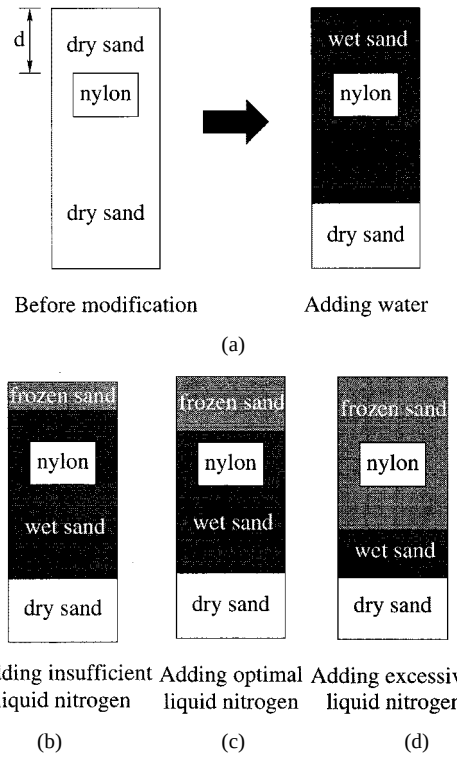


Fig. 9. Soil modification through the addition of water and liquid nitrogen.

to greater and greater depths. In plot (B) only a small quantity of nitrogen has been added so that loss in the wet sand region above the target still causes excessive attenuation and the target remains obscured. In plot (C), an "optimal" amount of liquid nitrogen has been added so that loss above the target is mini-

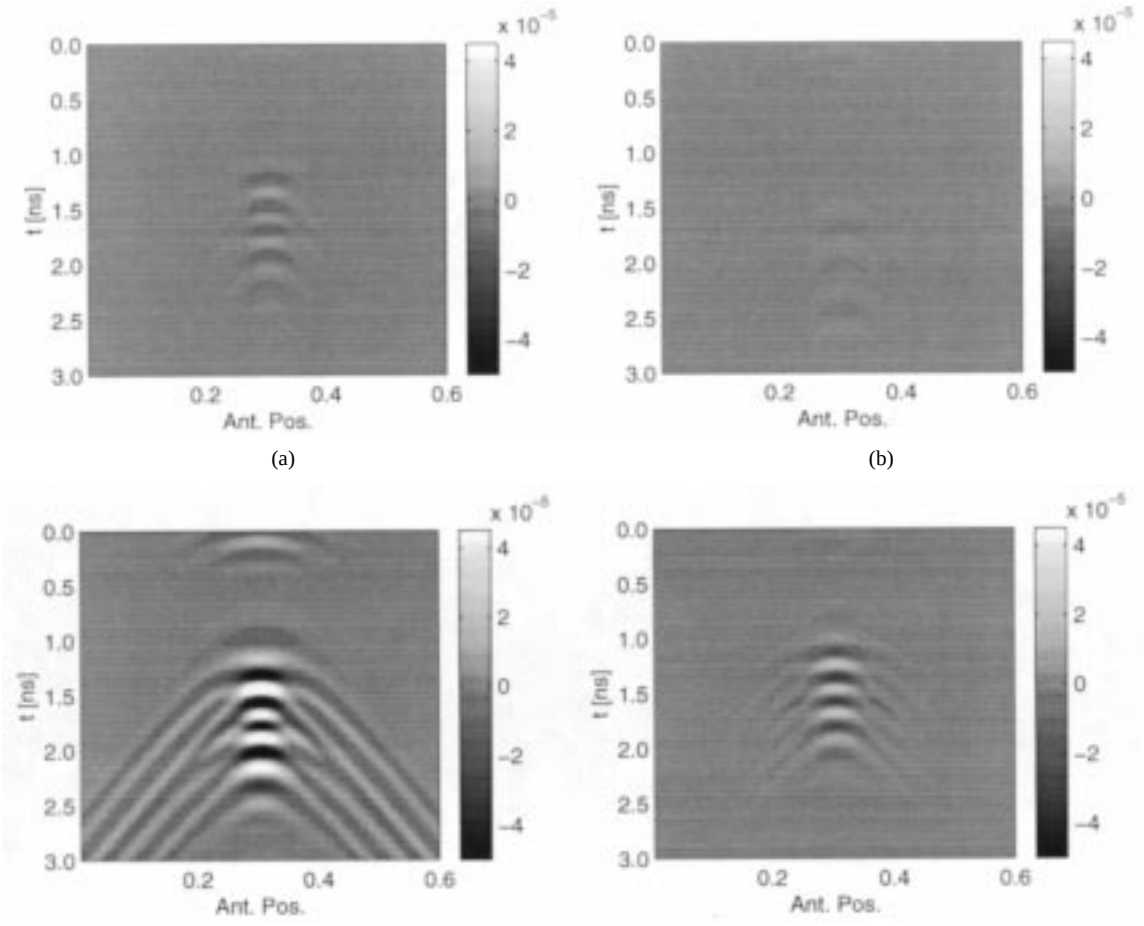


Fig. 10. Image results of time domain simulated scattered fields from PO model for the configurations of Fig. 8 with 6.35 cm depth nylon target. (a) Target in dry soil. (b) After the addition of water and a small amount of liquid nitrogen. (c) After the addition of an "optimal" amount of liquid nitrogen. (d) After the addition of an excessive amount of liquid nitrogen; field units are defined as in Fig. 5.

mized, but a large dielectric contrast in the region surrounding the target remains to provide scattering. In plot (D), an excessive amount of nitrogen has been added and again the target is obscured due to the low contrast between the nylon and frozen sand media. If the target depth is assumed known, the optimal amount of nitrogen to be added can be determined and cataloged for future measurements of similar depth targets. However, it is also clear from Fig. 9 that the procedure should be applicable to unknown depth targets as well, since nonoptimal amounts of liquid nitrogen do not produce large target returns. Thus, a procedure in which measurements are taken as gradually increasing amounts of liquid nitrogen are added should be sufficient to improve visibility of unknown depth targets as well.

Image results of time domain simulated scattered fields from the PO model based on the configurations of Fig. 9, plots (A)–(D), are shown in Fig. 10, plots (A)–(D) for a target at depth 6.35 cm. The model was extended to include multi-layer effects by generalizing the procedure of Section II, and reflections from boundaries between frozen and wet soils are removed to highlight target scattering effects. In this simulation, the dielectric constant of "dry" soil is approximated as $5 - j1.25$, the dielectric constant of frozen soil is approximated as $4 - j0.5$, and the dielectric constant of wet soil is approximated as $17.8 - j10$. The nylon target ($\epsilon_r \approx 3.15 - j0.03$) was assumed to be buried at 6.35 cm below soil surface. The thickness of frozen soil is assumed to be 2.54 cm for configuration (B), 5.08 cm for configuration (C), and 6.35 cm for configuration (D). Fig. 10, plot (A) confirms the small scattered returns obtained from the low contrast target in dry sand. Fig. 10 plot (B) illustrates the return for the configuration of Fig. 9, plot (B); here the small amount of nitrogen added is insufficient to reduce excessive background loss and the target remains obscured. Fig. 10 plot (C) shows the greatly enhanced target response obtained in the "optimal" configuration, in which 1.27 cm of wet sand remain above the target. Finally Fig. 10, plot (D) shows smaller responses as the entire background medium becomes frozen.

To confirm the soil modification procedure through the addition of water and liquid nitrogen suggested by the analytical models, measurements with the dielectric rod antenna GPR system were performed using natural soil in the backyard of the ElectroScience Laboratory. The experimental configuration was the same as described previously, and both metal and nylon disks with the same dimensions as discussed in Section II buried at 6.35 cm depth were considered in these measurements. In this test, substantial quantities of water and liquid nitrogen were applied to ensure that soil properties were modified. Measurements of the targets were first made with the soil in its ini-

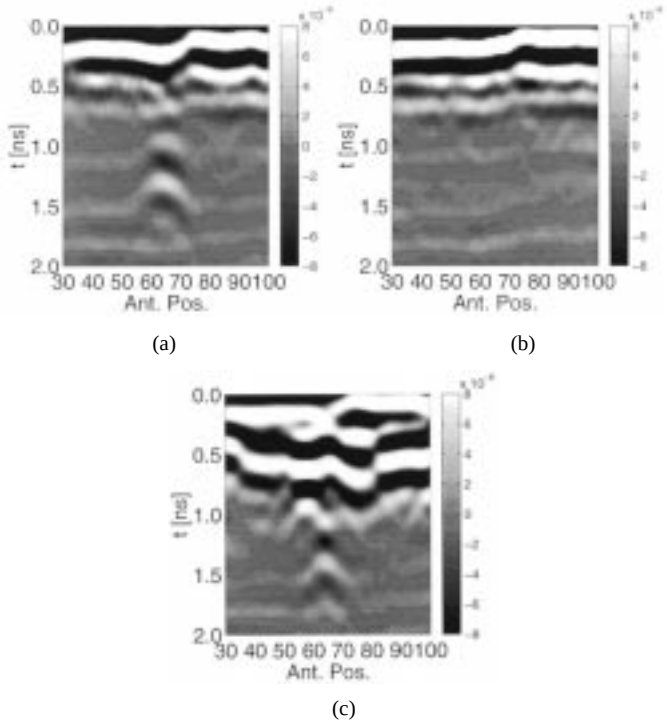


Fig. 11. Image results of time domain measured fields for dielectric rod antenna with 6.35 cm depth metal target. (a) Target in dry soil. (b) After the addition of 2 gallons of water. (c) After the addition of 15 liters of liquid nitrogen; field units are defined as in Fig. 5.

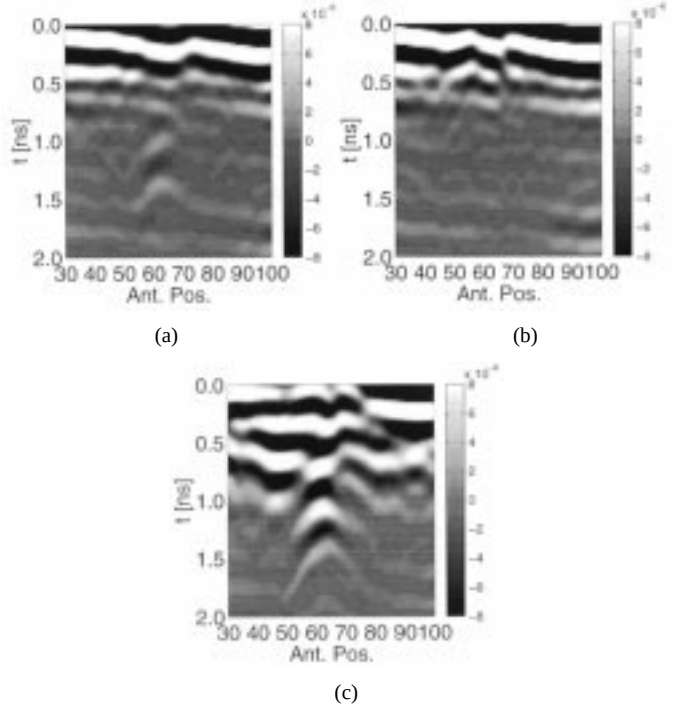


Fig. 12. Image results of time domain measured fields for dielectric rod antenna with 6.35 cm depth nylon target. (a) Target in dry soil. (b) After the addition of 2 gallons of water. (c) After the addition of 15 liters of liquid nitrogen; field units are defined as in Fig. 5.

tial state, and then two gallons of water were poured over an area of approximately 0.5-m square which contained the target. Measurements in this "excessively wet" configuration were then taken, and finally approximately 15 of liquid nitrogen were applied over the same area and the final measurements made. Tests of the nitrogen penetration depth were made by digging up the soil after the measurements were completed, and it was estimated that penetration to approximately 8 cm was obtained with this quantity of nitrogen. Although the chemical amounts used in this demonstration were substantial, it is expected that further studies can develop more efficient application procedures that will reduce the required quantities.

Measured results for the metallic and nylon targets are presented in Figs. 11 and 12, respectively as images. These images are produced by plotting time domain measured fields as a function of antenna horizontal position; the target location was known to be between antenna positions 60 and 70, and results are calibrated so that time 0 nS corresponds to the top surface of the ground. Returns observed from 0 to 0.6 ns correspond to clutter and sidelobe returns from the strong surface reflections in the measurement; clutter reduction signal processing methods can be applied to reduce these contributions [14]. Figs. 11 and 12 both contain three plots: plot (a) is the image before any water or nitrogen has been added, plot (b) is the image after the addition of water, and plot (c) is image after the addition of water and liquid nitrogen. Fig. 11 plots (b) and (c) clearly demonstrate the improved visibility of a metallic target in excessively wet soil obtained after the addition of liquid nitrogen. Even though the target is visible also in the un-modified soil, liquid nitrogen addition would be useful in situations where the un-modified soil had properties closer to the wet soil considered here. Results in Fig. 12 are also interesting: a much smaller response than the metallic target is obtained in plot (a) with the nylon target, and again the response is completely obscured in excessively wet soil. The enhanced response in part (c) again shows the effectiveness of liquid nitrogen addition. The relative contributions of reduced loss in the background and increased dielectric contrast are difficult to assess from this single measurement, but these initial qualitative results demonstrate that the soil modification techniques suggested by the analytical model should have applicability to general microwave GPR systems for detecting anti-personnel mines.

IV. Summary and Conclusion

The detection of nonmetallic landmines with ground penetrating radar through soil modification has been presented. The physical optics approximation for the RCS of a buried target has been derived, and results demonstrate that the RCS of a buried target is proportional to reflection and transmission coefficients and to attenuation in the background medium. The comparison between analytical results from this simple PO model and numerical results from the method of moments validates the model for near normal incidence studies. Moreover, the analytical data is similar to measured data from a dielectric rod antenna GPR system. Soil modification through an optimal homogeneous soil water content is shown to improve target detection, but optimal water contents decrease when the target depths increase since loss even in the lower water content soil will become significant for very deep targets. In many situations, it may not be possible to obtain effective soil modification using water alone. Soil modification through the addition of water and liquid nitrogen

was also described. Simulated results from the PO model confirm that an "optimal" amount of liquid nitrogen should exist to enhance target visibility. Initial measurements with a dielectric rod antenna GPR system also demonstrate that the addition of liquid nitrogen to excessively wet soils can reduce loss and enhance target visibility. Results from this study show potential for soil modification techniques to improve target detection. Clearly the soil modification procedures described can produce improvement to microwave GPR systems in two limiting cases: targets with strong scattering responses (i.e., metallic) obscured by excessive loss in the background medium and targets whose dielectric constant is identical to that of the background medium. In addition, the soil modification techniques described can potentially produce advantages not only in target detection but also in target identification, since scattered target signatures should be enhanced under larger dielectric contrasts.

REFERENCES

[1] A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, "Detection and remediation technologies for mines and minelike targets V," in Proc. SPIE, vol. 4038, 2000.

[2] J. T. Johnson, J. Jenwatanavet, N. Wang, R. Caldecott, and J. Young, "Waveguide studies of soil modification techniques for enhanced mine detection with ground penetrating radar," in *Detection and Remediation Technologies for Mines and Minelike Targets III, Proc. SPIE*, vol. 3392, A. C. Dubey, J. F. Harvey, and J. T. Broach, Eds., 1998, pp. 716–724.

[3] J. Jenwatanavet, J. T. Johnson, N. Wang, and R. Caldecott, "Development of soil modification techniques for enhanced detection/discrimination," in *IGARSS'98 Conf. Proc.*, vol. 1, 1998, pp. 210–212.

[4] M. H. Powers and G. R. Olhoeft, "Computer modeling to transfer GPR UXO detectability knowledge between sites," in *UXO Forum 1996, Conf. Proc.*, 1996, pp. 347–356.

[5] H. Frohlich, *Theory of Dielectrics; Dielectric Constant and Dielectric Loss*. New York, NY: Oxford Univ. Press, 1958.

[6] J. T. Johnson, J. Jenwatanavet, and N. Wang, "Soil modification studies for enhanced mine detection with ground penetrating radar," in *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proc. SPIE*, vol. 3710, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, Eds., 1999, pp. 739–748.

[7] N. Geng, D. Jackson, and L. Carin, "On the resonances of a dielectric BOR buried in a dispersive layered medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 1305–1313, Aug. 1999.

[8] N. Geng and L. Carin, "Wide-band electromagnetic scattering from a dielectric BOR buried in a layered lossy dispersive medium," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 610–619, Apr. 1999.

[9] T. Dogaru and L. Carin, "Time-domain sensing of targets buried under a rough air-ground interface," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 360–372, Mar. 1998.

[10] N. Geng and L. Carin, "Short-pulse electromagnetic scattering from arbitrarily oriented subsurface ordnance," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2111–2113, July 1999.

[11] N. Geng, C. Baum, and L. Carin, "On the low-frequency natural response of conducting and permeable targets," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 347–359, Jan. 1999.

[12] E. Newman, "A User's Manual for the Electromagnetic Surface Patch Code: Preliminary Version ESP5.0," ElectroScience Laboratory, The Ohio State University, unpublished report, 1997.

[13] S. Nag, J. L. Peters, I. J. Gupta, and C.-C. Chen, "Ramp response for the detection of anti-personnel mines," in *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proc. SPIE*, vol. 3710, A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan, Eds., 1999, pp. 1313–1322.

[14] A. V. der Merwe and I. J. Gupta, "A novel signal processing technique for clutter reduction in GPR measurements of small, shallow land mines," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 2627–2637, 2000.

[15] F. Ulaby, R. Moore, and A. Fung, *Microwave Remote Sensing: Active and Passive*. New York: Addison-Wesley, 1981, vol. 3.



Jatupum Jenwatanavet obtained the B.E. degree in electrical engineering from Chulalongkorn University, Bangkok, Thailand, and the M.S. degree in electrical engineering from The Ohio State University, Columbus. From 1997 to 1999, he was a graduate research associate at The ElectroScience Laboratory, The Ohio State University. His research interests include theoretical and experimental studies in remote sensing, radar, scattering, and antennas. He is currently an electrical engineer with Motorola, Inc. in Piscataway, NJ.

Joel T. Johnson (M'96) received the B.E.E. degree in electrical engineering from the Georgia Institute of Technology in 1991 and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology in 1993 and 1996, respectively.

He is currently an Associate Professor in the Department of Electrical Engineering and ElectroScience Laboratory of The Ohio State University. His research interests are in the areas of microwave remote sensing, propagation, and electromagnetic wave theory.

Dr. Johnson is an associate member of commissions B and F of the International Union of Radio Science (URSI), and a member of Tau Beta Pi, Eta Kappa Nu, and Phi Kappa Phi. He received the 1993 best paper award from the IEEE Geoscience and Remote Sensing Society, and was named an Office of Naval Research Young Investigator, National Science Foundation Career awardee, and PECASE award recipient in 1997.

Coupled Canonical Grid/Discrete Dipole Approach for Computing Scattering from Objects Above or Below a Rough Interface

Joel T. Johnson, *Member, IEEE,* and Robert J. Burkholder, *Senior Member, IEEE,*

***Abstract*—A numerical model for computing scattering from a three-dimensional (3-D) dielectric object above or below a rough interface is described. The model is based on an iterative method of moments solution for equivalent electric and magnetic surface current densities on the rough interface and equivalent volumetric electric currents in the penetrable object. To improve computational efficiency, the canonical grid method and the discrete dipole approach (DDA) are used to compute surface to surface and object to object point couplings, respectively, in $O(N \log N)$, where N is the number of surface or object sampling points. Two distinct iterative approaches and a preconditioning method for the resulting matrix equation are discussed, and the solution is verified through comparison with a Sommerfeld integral-based solution in the flat surface limit. Results are illustrated for a sample landmine detection problem and show that a slight surface roughness can modify object backscattering returns.**

***Index Terms*—Electromagnetic scattering, ground penetrating radar (GPR), radar cross section, rough surface scattering.**

I. Introduction

ELECTROMAGNETIC scattering from objects is affected by the surrounding medium. Many realistic geometries involve objects in the presence of the ground, which is often modeled as a planar dielectric boundary. However, roughness on the ground surface can potentially modify object scattering returns from those with a flat surface, particularly in cases where the roughness size becomes larger than a fraction of the electromagnetic wavelength. Analysis of these problems is complicated by the many possible scattering interactions between the rough surface and object; at present, approximate analytical solutions exist only in the small roughness limit [1]–[4]. The development of numerical models is therefore of interest since these models avoid any approximations and thus can clarify important scattering effects when analytical methods fail. The fact that a combined object/rough surface scattering model can be applied to studies of landmine detection [5], foliage and surface remote sensing [2], or target detection and signal processing problems [6] provides further motivation for model development.

Several previous numerical studies of scattering from objects in the presence of the ground have been reported. For geometries with a flat surface model, a Sommerfeld Green's function can be derived for frequency domain methods, and discretization is required only on the object [7], [8]. Time domain studies of objects in the presence of flat boundaries have also been reported [9]. Previous studies including objects and surface roughness have primarily been limited to one-dimensional (1-D) surface geometries (i.e., surfaces having roughness only in one horizontal direction) [10]–[13] since surface unknowns must be included and computational requirements are increased. Problems involving two-dimensional (2-D) surfaces and 3-D objects have been considered only in a small number of studies [6], [14] due to the large computational requirements, and have been limited in most cases to perfectly conducting objects.

In this paper, an efficient numerical model is described which makes studies of scattering from a 3-D dielectric object in the presence of a 2-D rough surface possible. The model is similar to that used in the perfectly conducting object study of [6], and is based on an iterative method of moments solution. The integral equation formulation and resulting matrix equation are presented in Section II. Iterative solutions of the matrix equation based on a standard nonstationary algorithm [15] or based on the "multiple interaction" iteration of [6] are then discussed in Section III. Efficiency of the iterative solution is improved through use of the "canonical grid" (CAG) method [16]–[22] and the discrete-dipole approximation (DDA) [23], [24] for computing surface to surface and object to object point couplings, respectively, as described in Section IV. A sample application of the model to studies of scattering from a subsurface object is then illustrated in Section V, and final conclusions presented in Section VI.

II. Formulation

Fig. 1 illustrates the geometry considered: a dielectric object with relative permittivity ϵ_3 is located in the presence (in this case, below) of a rough interface $z = f(x, y)$ between Regions 1 (relative permittivity ϵ_1) and 2 (relative permittivity ϵ_2). Integral equations to determine time harmonic fields scattered in this problem can be formulated in terms of equivalent electric and magnetic surface current densities on the interface and equivalent electric volumetric currents in the object.

Manuscript received September 30, 2000; revised December 15, 2000. This work was supported by ONR Contracts N00014-97-1-0541 and N00014-00-1-0399, NSF Project ECS-9701678, and by a grant from Duke University as part of the OSD MURI on Humanitarian Demining. Use of the IBM SP system at the Maui High Performance Computing Center is acknowledged and was sponsored by the Air Force Research Laboratory, Air Force Material Command under cooperative Agreement F29601-93-2-0001. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the U.S. Air Force, Air Force Research Laboratory, or the U.S. Government.

The authors are with the Department of Electrical Engineering and ElectroScience Laboratory, Ohio State University, Columbus, OH 43210 USA.

Publisher Item Identifier S 0196-2892(01)04841-0.

0196–2892/01$10.00 © 2001 IEEE

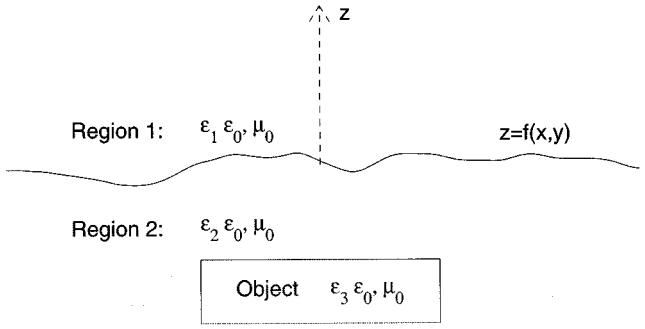


Fig. 1. Geometry of problem: object may be above or below boundary.

A. Surface Integral Equations

For convenience, electric and magnetic surface current densities on the interface are defined here as $\overline{F} = \eta_1 \overline{N} \times \overline{H}_1$ and $\overline{G} = -\overline{N} \times \overline{E}_1$, respectively, where $\overline{E}_1$ and $\overline{H}_1$ are the total electric and magnetic fields in Region 1 on the interface, η_1 is the characteristic impedance of Region 1, and $\overline{N}$ is an upward pointing normal vector $\hat{z} - \hat{x}(\partial f/\partial x) - \hat{y}(\partial f/\partial y)$. Following the formulation of [25] and assuming an $e^{-i\omega t}$ time dependence, integral equations for $\overline{F}$ and $\overline{G}$ are

$$\frac{\overline{F}}{2} = \overline{N} \times \iint_A dx'\, dy' \cdot \left\{ ik_1 g_1 \overline{G} + \nabla g_1 \times \overline{F} - \left(\frac{\nabla g_1}{ik_1}\right) |\overline{N}'| \left[\nabla' \cdot \left(\frac{\overline{G}}{|\overline{N}'|}\right)\right]\right\} + \eta_1 \overline{N} \times \overline{H}_{inc,1} \quad (1)$$

$$-\frac{\overline{G}}{2} = -\overline{N} \times \iint_A dx'\, dy' \cdot \left\{ ik_1 g_2 \overline{F} - \nabla g_2 \times \overline{G} - \left(\frac{\nabla g_2}{ik_1}\right) |\overline{N}'| \left[\frac{\epsilon_1}{\epsilon_2} \nabla' \cdot \left(\frac{\overline{F}}{|\overline{N}'|}\right)\right]\right\} + \overline{N} \times \overline{E}_{inc,2} \quad (2)$$

where $k_j = k_0\sqrt{\epsilon_j}$ for $j = 1 \cdots 3$ is the wavenumber in Region j; $k_0 = \omega\sqrt{\mu_0 \epsilon_0} = (2\pi/\lambda)$ is the free-space wavenumber for electromagnetic wavelength λ and radian frequency ω; and g_j is the free-space Green's function in Region j

$$g_j(\overline{r}, \overline{r}') = \frac{e^{ik_j|\overline{r}-\overline{r}'|}}{4\pi|\overline{r}-\overline{r}'|} \quad (3)$$

where ∇ and ∇' denote the gradient operator in terms of unprimed (observation) and primed (source) coordinates, respectively, and the integration is over the horizontal projected area A of surface $z = f(x, y)$. The incident fields $\overline{H}_{inc,1}$ and $\overline{E}_{inc,2}$ describe fields which impinge from above or below the boundary, respectively, and will include contributions from the fields $\overline{E}_{inc}$ and $\overline{H}_{inc}$ which excite the problem (assumed to approach from Region 1) as well as the dielectric object. Use of the Stratton–Chu form and surface divergence operators in (1) and (2) results in four scalar unknown functions (i.e., the x and y components of $\overline{F}$ and $\overline{G}$, from which the z components can be determined) to be found, as opposed to the six unknown scalar function formulations which have been used previously [6], [18]. A standard point matching discretization (i.e., surface currents assumed constant over a surface patch and integral equations "tested" at a set of points) is applied to these equations as described in [16]–[22], [25], so that surface current divergences $\nabla' \cdot (\overline{F}/|\overline{N}'|)$ and $\nabla' \cdot (\overline{G}/|\overline{N}'|)$ can be computed numerically using a centered difference method. Examination of self terms in (1) and (2) (i.e., when source and observation points overlap) shows cancellation of all terms involving ∇g_j for surface patches modeled locally as tilted planes due to symmetries in the integration. Singularities in the g_j terms are integrable, and calculated efficiently through standard analytical subtraction/addition methods.

B. Object Integral Equation

A dielectric object in Region j is replaced by an equivalent electric volumetric current density $\overline{J}$, for which the resulting integral equation is

$$\overline{E} = \overline{E}_{inc,v} + \iiint_{\text{object}} dV' i\omega\mu_0 g_j \cdot \left[\overline{\overline{I}} - \hat{R}\hat{R} + \frac{ik_j R - 1}{k_j^2 R^2}\left(\overline{\overline{I}} - 3\hat{R}\hat{R}\right)\right] \cdot \overline{J} \quad (4)$$

where

R $|\overline{r} - \overline{r}'|$;
$\hat{R}$ $((\overline{r} - \overline{r}'/(|\overline{r} - \overline{r}'|)$;
$\overline{\overline{I}}$ unit dyad;
$\overline{E}_{inc,v}$ field which impinges upon the object including contributions from electric and magnetic surface current densities on the interface $z = f(x, y)$ and the field which excites the problem if the object is located in Region 1.

In a point discretization of $\overline{J}$ with uniform volume ΔV, (4) can also be expressed in terms of a set of "dipole moments" $\overline{P} = \Delta V \overline{J}(i/4\pi\epsilon_j\epsilon_0\omega)$, leading to a "discrete dipole" representation of equivalent currents [23]. Equation (4) involves three scalar unknown functions: the $\hat{x}$, $\hat{y}$, and $\hat{z}$ components of $\overline{J}$. Self terms in this formulation are described in [23], and replace the electric field on the left-hand side of (4) with $\overline{P}/\alpha$ where α is given by

$$\alpha = \frac{3\Delta V}{4\pi}\left(\frac{\frac{\epsilon_3}{\epsilon_j} - 1}{\frac{\epsilon_3}{\epsilon_j} + 2}\right). \quad (5)$$

Note that inhomogeneous objects can also be included in the present formulation simply by varying values of ϵ_3 and α for points in the object; homogeneous objects are considered in the results shown for simplicity.

C. Surface and Object Coupling

Coupling between surface and object is computed using formulations similar to (1), (2), and (4). For an object in Region 2, coupling from surface sources to object observation points is calculated from

$$\overline{E}_{inc,v} = -\iint_A dx'\, dy' ik_1 g_2 \cdot \left[\overline{\overline{I}} - \hat{R}\hat{R} + \frac{ik_2 R - 1}{k_2^2 R^2}\left(\overline{\overline{I}} - 3\hat{R}\hat{R}\right)\right] \cdot \overline{F} + \iint_A dx'\, dy' g_2 \frac{ik_2 R - 1}{R}(\hat{R} \times \overline{G}) \quad (6)$$

and coupling from object sources to surface observation points is found by evaluating $\overline{E}_{inc,2}$ from the integral term in (4) with $k_j = k_2$. For an object in Region 1, surface source to object coupling is obtained from

$$\overline{E}_{inc,v} = \overline{E}_{inc} + \iint_A dx'\,dy' ik_1 g_1 \cdot \left[\overline{\overline{I}} - \hat{R}\hat{R} + \frac{ik_1R-1}{k_1^2R^2}\left(\overline{\overline{I}} - 3\hat{R}\hat{R}\right)\right]\cdot\overline{F} - \iint_A dx'\,dy' g_1 \frac{ik_1R-1}{R}(\hat{R}\times\overline{G}) \quad (7)$$

where $\overline{E}_{inc}$ is the electric field which excites the problem. Object source to surface coupling is found from

$$\overline{H}_{inc,1} = \overline{H}_{inc} + \iiint_{\text{object}} dV' g_1 \frac{ik_1R-1}{R}(\hat{R}\times\overline{J}) \quad (8)$$

where $\overline{H}_{inc}$ is the magnetic field which excites the problem corresponding to $\overline{E}_{inc}$.

In the results to be shown, object and surface points do not overlap, so that no singular terms in the coupling equations are encountered. For geometries which do contain overlapping points, singularities can be extracted and integrated by the methods described for (1), (2), and (4), although this has not been performed in the current implementation.

D. Combined Matrix Equation

Combining the surface, object, and coupling equations results in a matrix equation in terms of $\overline{F}$, $\overline{G}$, and $\overline{J}$. For an object in Region 2, the matrix equation can symbolically be written as

$$\begin{bmatrix} \text{Surface : Surface} & -\text{Object : Surface} \\ -\text{Surface : Object} & \text{Object : Object} \end{bmatrix} \cdot \begin{bmatrix} \overline{F}, \overline{G} \\ \overline{J} \end{bmatrix} = \begin{bmatrix} \overline{Q}_{inc} \\ 0 \end{bmatrix} \quad (9)$$

where "Surface : Surface" block represents minus the integral terms on the right-hand sides (RHSs) of (1) and (2). The "Object : Object" block represents minus the integral term on the RHS of (4), and "Object : Surface" and "Surface : Object" blocks are derived from integral terms of the equations described in Section II-C [written as negative since they originally appear on the RHSs of (1), (2), and (4)]. The case for an object in Region 1 is similar and will not be separated in the following discussions.

The "Surface : Surface" and "Object : Surface" blocks of the matrix equation are ordered so that the $\hat{x}$ and $\hat{y}$ components of (1) for a single observation point are following by the $\hat{x}$ and $\hat{y}$ components of (2); this arrangement places the $\pm 1/2$ self contributions on the diagonal of the "Surface : Surface" component when the $\overline{F}$ and $\overline{G}$ unknowns are ordered as F_x, F_y, G_x, and G_y for a given source point in the unknown vector. The RHS $\overline{Q}_{inc}$ thus is ordered as $\hat{x}\cdot\eta_1\overline{N}\times\overline{H}_{inc}$, $\hat{y}\cdot\eta_1\overline{N}\times\overline{H}_{inc}$, 0, 0 for each surface observation point. The "Surface : Object" and "Object : Object" blocks are arranged so that $\hat{x}$, $\hat{y}$, and $\hat{z}$ components of (4) appear in succession for a given object observation point, while the unknown vector contains successive J_x, J_y, and J_z values for a given object source point. These choices locate $1/\alpha$ terms on the diagonal of the "Object : Object" matrix.

In many problems of interest, scattering from the rough interface can be much larger than that from the dielectric object (e.g., the case of a low contrast landmine buried in a lossy medium). In this case, a matrix equation in terms of object minus no-object difference fields can be formulated to clarify effects of the object. If the no-object problem is first solved through

$$[\text{Surface : Surface}][\overline{F}_0, \overline{G}_0] = [\overline{Q}_{inc}] \quad (10)$$

then the object minus no-object currents can be obtained from

$$\begin{bmatrix} \text{Surface : Surface} & -\text{Object : Surface} \\ -\text{Surface : Object} & \text{Object : Object} \end{bmatrix} \cdot \begin{bmatrix} \overline{F}-\overline{F}_0, \overline{G}-\overline{G}_0 \\ \overline{J} \end{bmatrix} = \begin{bmatrix} 0 \\ \text{Surface : Object}\{\overline{F}_0, \overline{G}_0\} \end{bmatrix}. \quad (11)$$

This differencing procedure is useful with an iterative solution of the matrix equation, since it insures that accuracy is retained in object scattering effects. The surface only solution is typically substantially less expensive than the combined surface/object problem (particularly with the canonical grid method described in Section IV) so that the additional computations are tolerable.

Consideration of object minus no object scattered fields also demonstrates an important issue in interpreting results for combined surface/object problems. Since object scattering is typically discussed in terms of radar cross sections while surface scattering is more appropriately described as a cross section per unit area, surface/object scattering problems include both standard and area extensive geometries making the definition of a radar cross section ambiguous. The results of this paper treat fields radiated in Region 1 by the difference currents $\Delta\overline{F} = \overline{F}-\overline{F}_0$, $\Delta\overline{G} = \overline{G}-\overline{G}_0$ as equivalent to those radiated by an object in defining a radar cross section. However, since these difference currents contain object scattering but also interactions between the surface and object, a standard cross section is not obtained and results remain sensitive to the particular surface geometry and incident field considered.

III. Iterative Methods

Since the matrix equation (9) for 2-D surfaces and 3-D dielectric objects is likely to contain a large number of unknowns, direct inversion of (9) is computationally inefficient. Iterative solutions can be developed based on a multiple interaction procedure [6] or through an application of standard nonstationary iterative methods [15]. The former is accomplished for object minus no-object difference currents by iterating

$$[\text{Object : Object}]\left[\overline{J}^{(n+1)}\right] = \left[\text{Surface : Object}\left\{\overline{F}_0 + \Delta\overline{F}^{(n)}, \overline{G}_0 + \Delta\overline{G}^{(n)}\right\}\right] \quad (12)$$

$$[\text{Surface : Surface}]\left[\Delta\overline{F}^{(n+1)}, \Delta\overline{G}^{(n+1)}\right] = \left[\text{Object : Surface}\left\{\overline{J}^{(n+1)}\right\}\right] \quad (13)$$

with the iteration initialized through $\Delta\overline{F}^{(0)} = 0$, $\Delta\overline{G}^{(0)} = 0$. The "Surface : Surface" and "Object : Object" matrix equations are solved with the nonstationary algorithm described below, with an initial guess for the $(n+1)$th solution taken as the nth solution to improve convergence at each surface to object iteration. Since each iteration corresponds to an additional interaction between the surface and object, the method should be expected to converge rapidly for cases lacking strong surface/object interactions. Convergence can be monitored in terms of a "pseudo-residual" by examining the norms of $\Delta\overline{F}^{(n+1)} - \Delta\overline{F}^{(n)}$ and $\Delta\overline{G}^{(n+1)} - \Delta\overline{G}^{(n)}$.

Nonstationary methods applied to the combined object/surface matrix equation (11) can potentially provide improved convergence properties for cases with strong object/surface interactions and also converge in terms of the absolute residual of the matrix equation. The biconjugate gradient-stabilized algorithm (Bi-CGSTAB) was chosen for this paper, since it has been found effective in previous studies of rough surface scattering. Preconditioning can be beneficial with nonstationary methods if an approximation to the original matrix can be found which is easily inverted. This is accomplished for the matrix of (11) by neglecting surface/object coupling and employing an approximate solution of the "Surface : Surface" block based on a "quasi-physical optics" procedure as described in [21]. The "Object : Object" block of the preconditioner is solved through another Bi-CGSTAB routine set with only a moderate convergence accuracy. The effectiveness of this preconditioner is determined by the ratio of the number of surface and object unknowns; for cases with a much larger number of surface unknowns, the "Object : Object" solution is not expensive and the preconditioner is warranted. The matrix multiply acceleration methods described in the next Section further make the preconditioner effective. The results illustrated in Section V were computed using the combined object/surface matrix preconditioned Bi-CGSTAB algorithm to insure that high accuracy was retained; limited tests with the surface/object iteration procedure of (12) and (13) showed good performance as well.

IV. Canonical Grid and Discrete Dipole Algorithms

To improve computational efficiency, the CAG [16]–[22] and DDA [24] methods are used to perform "Surface : Surface" and "Object : Object" block multiplies, respectively. Both algorithms are based on use of a uniform Cartesian grid for description of the surface profile (2-D grid) and dielectric object (three dimensional grid), and a Toeplitz representation for "Surface : Surface" and "Object : Object" matrices so that multiplies can be performed using the fast Fourier transform (FFT). Toeplitz matrices are obtained for the "Surface : Surface" block by expanding matrix elements in a series under the assumption of small height difference relative to the horizontal separation between points on the surface. The number of series terms retained is a parameter of the method, along with a "strong distance" between points inside of which the series expansion is not used. The method is most effective for small to moderate surface slopes and small to moderate surface heights in terms of λ, for which a small number of series terms and a small strong distance are needed. Other surface scattering acceleration techniques [14], [26] are likely to be more effective for large surface slopes and heights, but are unlikely to obtain the efficiency of the canonical grid method for small to moderate roughness surfaces. For the "Object : Object" block, the object is placed inside a uniform Cartesian grid and volumetric electric current sources set to zero in grid locations not occupied by the object. Three-dimensional FFT [27] computations involve the entire object grid, however, so the acceleration routine is most effective if the object occupies a significant fraction of the grid volume; elongated Cartesian grids can be used if necessary to insure that this occurs. Use of a uniform Cartesian grid can introduce "stair-stepping" errors in non-Cartesian boundaries, which are alleviated only through increased sampling rates. Reference [23] considers these errors for a spherically shaped object and shows that sampling rates from 10 to 20 points per wavelength in the object are acceptable. Both the canonical grid and discrete dipole methods provide $O(N_s \log N_s)$ and $O(N_o \log N_o)$ multiplies for "Surface : Surface" and "Object : Object" calculations, respectively, where N_s is the number of surface unknowns and N_o is the number of object unknowns.

Note that "Surface : Object" and "Object : Surface" coupling matrices are not accelerated in the current method, and therefore require $O(N_s N_o)$ computations. Development of FFT-based acceleration methods for these terms is complicated by the differing sampling requirements on $\overline{J}$ and $(\overline{F}, \overline{G})$ since differing dielectric media are involved. Use of FFT-based methods for sources and observation points on differing sampling rate grids would require careful consideration. Other matrix multiply acceleration methods can again avoid these problems. However, for problems in which $N_s \gg N_o$ acceleration of coupling matrix multiplies will provide only moderate gains in overall routine efficiency.

V. Sample Application: Scattering from a Buried "Landmine"

Fig. 2 illustrates a sample combined object/surface problem: a rectangular box dielectric object with dimensions $7.62\,\text{cm} \times 7.62\,\text{cm} \times 2.54\,\text{cm}$ and relative permittivity $\epsilon_3 = 3 + i0.03$ is located 7.62 cm below an interface between free-space and a medium with relative permittivity $\epsilon_2 = 5 + i1.25$. Scattering from this object is to be determined for a field incident at 15° from normal incidence at frequencies from 2 to 5.1 GHz. Results will be shown for both flat and rough interfaces between Regions 1 and 2. This geometry is intended to model ground penetrating radar observation of a low contrast (i.e., plastic) target in moderately attenuating soil; it is clear that object returns may be quite small when compared to potential interface scattering effects.

Fig. 2 also shows that a finite size interface (here, $1.28\,\text{m} \times 1.28\,\text{m}$) between Regions 1 and 2 is used in the model. To avoid artificial edge scattering effects due to this truncation, the problem is excited by the "tapered wave" incident field described in [17]. This field is designed to mimic an incident plane wave but provides attenuation of incident fields as surface edges are approached. For tapering parameter

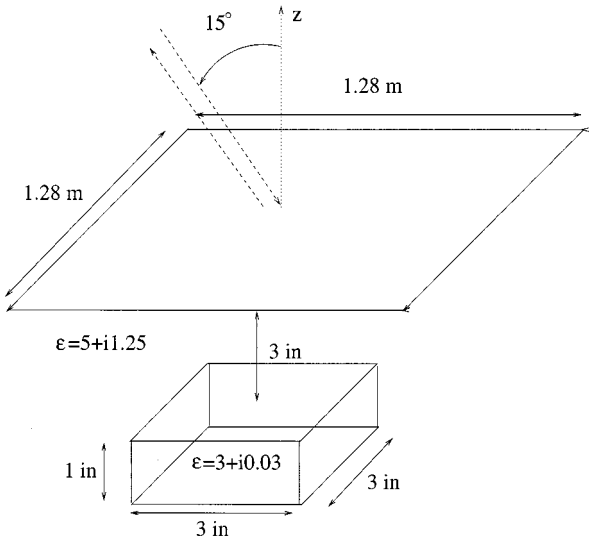


Fig. 2. "Landmine" detection geometry considered.

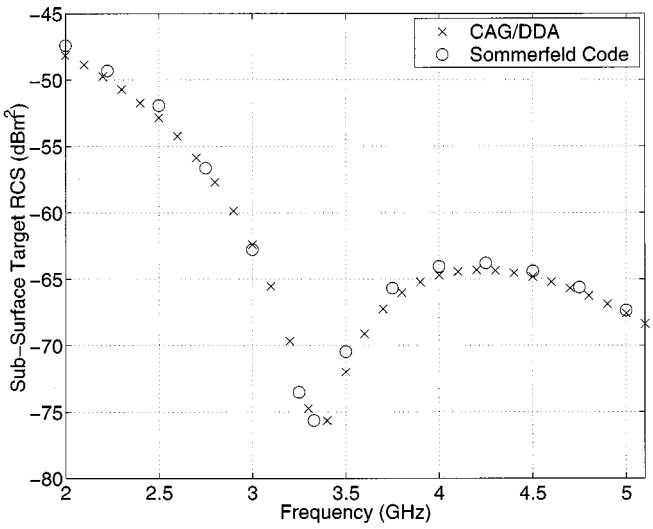


Fig. 3. Object minus no object radar cross sections for geometry of Fig. 2 versus frequency, HH polarization. Comparison of CAG/DDA method with a Sommerfeld Green's function-based solution.

$g = 5.333$, the field incident on surface edges is approximately 60 dB lower than that at the center, and the object horizontal cross section projected onto the flat interface is well within the 3-dB incident spot size. A test of tapered wave influence will be described below through comparison with a plane wave incidence Sommerfeld Green's function code in the flat surface limit.

A. Computational Issues

Since the $1.28\,\text{m} \times 1.28\,\text{m}$ interface ranges between 8.5 and 21.76 free-space wavelengths as the frequency is varied from 2 to 5.1 GHz, the interface is sampled into 128×128 points in the results shown. Although this is somewhat small for the higher frequencies given that the lower medium wavelength is approximately 2.25 times shorter than that in free-space, a set of comparisons using 256×256 points showed only slight changes in scattered cross sections. While a smaller number of points could be used for the lower frequencies, a constant number of points sampling the interface as frequency is varied was chosen for convenience in describing the rough surface profile. The resulting number of field unknowns on the interface is 65 536. The object is sampled on a $32 \times 32 \times 32$ point grid with step size 3.175 mm (ranging from approximately 1/27 to 1/11 of the wavelength in the object as frequency varies), so that the object occupies 14% of the grid volume; again this choice was made for simplicity in initial tests but more effective grids (e.g., a $32 \times 32 \times 8$ point grid) could be used. The total number of unknowns for the object is then 13 824, and the combined problem contains 79 360 unknowns.

Although the problem considered can be solved on a PC level platform, total computing times for the multiple cases considered in this paper were further reduced through use of IBM SP parallel computing resources at the Maui High Performance Computing Center [28]. Since results as a function of frequency were of interest, single frequency calculations were performed on individual nodes of the parallel computer (comparable to PC platforms) to obtain 32 frequencies between 2 and 5.1 GHz. Single frequency computing times on a single node ranged from approximately one to eight hours depending on convergence parameters and the iterative method used; attempts to optimize computing times have at present not been performed extensively.

B. Validation

To validate the CAG/DDA model, backscattered radar cross sections from the subsurface object were compared with those obtained by a Sommerfeld Green's function-based code [29] in the flat surface limit. Fig. 3 illustrates the comparison versus frequency for HH polarization, and shows good agreement between the CAG/DDA and Sommerfeld codes even through a relatively large variation in cross sections. Small cross section values observed between 3 and 3.5 GHz can be explained due to interference between reflections from the top and bottom interfaces of the target, as predicted through a simple multilayer reflection coefficient model. Some discrepancies within 1 dB are observed, but are difficult to resolve due to computational limitations of the Sommerfeld code which is based on direct matrix equation solution. The good agreement also confirms that the tapered wave incident field used in the CAG/DDA model is effectively modeling a plane wave incident field while reducing surface edge scattering effects.

C. Effect of a Slight Surface Roughness

Fig. 4 repeats the results of Fig. 3 but includes a slight surface roughness on the interface between Regions 1 and 2. A Gaussian random process description of $f(x, y)$ was used with an isotropic Gaussian correlation function. A surface height standard deviation of 3.58 mm and correlation length of 3.58 cm were chosen; this height standard deviation ranges from approximately 1/42 to 1/17 of the free-space wavelength so that the roughness exceeds limitations of the analytical small perturbation method (SPM) [30] for rough surface scattering only at the highest frequencies. Two terms in the canonical grid series and a "strong distance" of 15 points were used to insure accurate inclusion of roughness effects in "Surface : Surface"

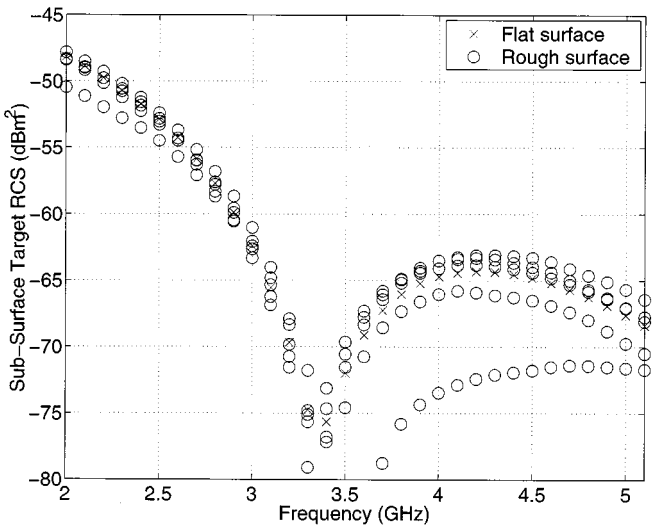


Fig. 4. Effect of surface roughness on object minus no object radar cross sections versus frequency, HH polarization. Results are illustrated for five surface realizations.

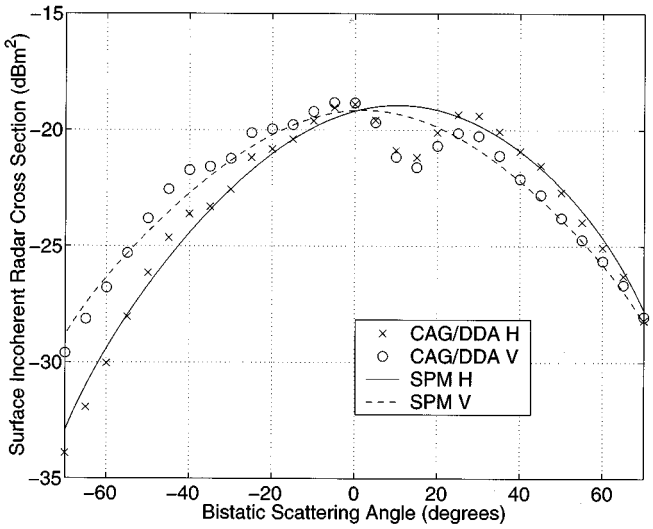


Fig. 5. Average in-plane incoherent surface bistatic scattering cross sections. SPM results are included for comparison.

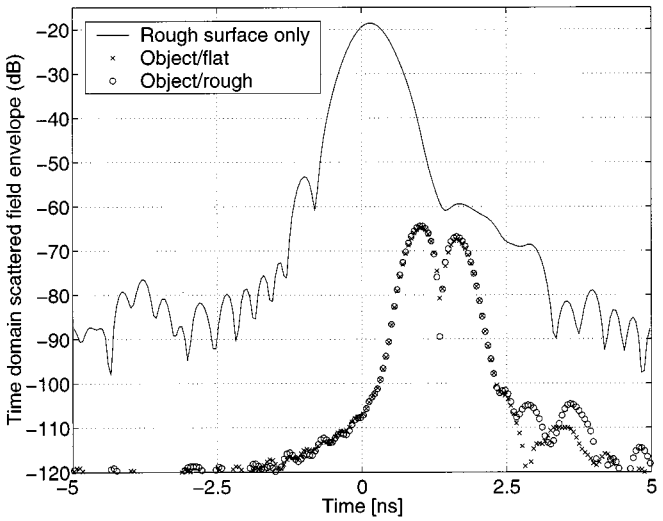


Fig. 6. Envelope of time domain HH backscattered fields for geometry of Fig. 2.

coupling. Results were calculated for five distinct realizations of the surface random process and object minus no object "radar cross sections" are compared with the flat surface case in Fig. 4. The presence of surface roughness is observed to have only a moderate influence on object backscattering versus frequency, with larger effects obtained at the higher frequencies where roughness is larger in terms of the wavelength. Although further calculations could be performed to obtain a Monte Carlo averaged object RCS in the presence of the specified random process, the single realization results illustrated are also of interest since they also give some idea as to the level of variations that may be observed for a given surface profile.

An additional validation of surface scattering contributions can be performed for the problem considered through comparison of surface only scattering with the SPM. Fig. 5 shows the comparison of in-plane HH and VV incoherent bistatic radar cross sections (scattering angle $-15°$ indicates backscattering) at 3 GHz averaged over 32 surface realizations. Results are plotted in terms of absolute cross sections rather than cross sections per unit area: SPM predictions are simply scaled by the area illuminated by the tapered incident field to obtain the curves shown. Note the much larger surface scattering cross sections compared to those of the object in Fig. 4; the low contrast target and moderate attenuation in Region 2 produce this effect, even with the slight surface roughness considered. Some discrepancies between CAG/DDA and SPM predictions near the forward scattering region are observed due to inaccuracies in subtraction of the coherent forward scattered tapered wave that exists for this slight surface roughness. Overall, however, the good agreement obtained validates the CAG/DDA method developed for calculation of both object and surface scattering.

A final illustration of combined object/surface scattering effects is provided in Fig. 6, which plots the envelope of time domain backscattered fields obtained from 2 to 5.1 GHz data for the flat surface case and one of the rough surfaces. Object scattering cross sections are observed to return later in time, as expected, and again to have significantly lower amplitudes than rough surface scattering contributions. Surface roughness effects are observed to have only a moderate effect on time domain object cross sections; results with other surface realizations show similar levels of differences. The comparison of Fig. 6 demonstrates the typical problem of small object scattering in the presence of much larger surface clutter for ground penetrating radar systems. Effective signal processing algorithms are thus required to detect these objects.

VI. Conclusions

The CAG/DDA model presented makes numerical calculation of combined surface/dielectric object geometries relatively efficient, and can be applied to many areas of current interest. Further improvements of the method can be obtained through acceleration of "Surface : Object" and "Object : Surface" multiplies with FFT-based methods or other techniques, and also through further tests to optimize parameter choices for the algorithm. The results shown illustrate the importance of including

both object and surface scattering contributions. Evaluations of analytical models for surface/object interactions [1]–[4] are also possible with the method, and insights obtained from further studies should enable improved analytical models to be created.

REFERENCES

[1] Y. Zhang, Y. E. Yang, H. Braunisch, and J. A. Kong, "Electromagnetic wave interaction of conducting object with rough surface by hybrid SPM/MOM technique," *PIER 22: Progr. Electromagn. Res.*, vol. 22, pp. 315–335, 1999.

[2] T. Chiu and K. Sarabandi, "Electromagnetic scattering interaction between a dielectric cylinder and a slightly rough surface," *IEEE Trans. Antennas. Propagat.*, vol. 47, pp. 902–913, May 1999.

[3] J. T. Johnson, "Thermal emission from a layered medium bounded by a slightly rough interface," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 368–378, Feb. 2001.

[4] I. M. Fuks and A. G. Voronovich, "Wave diffraction by rough interfaces in an arbitrary plane-layered medium," *Waves Random Media*, vol. 10, pp. 253–272, 2000.

[5] *Proc. SPIE 4098: Detection and Remediation Technologies for Mines and Minelike Targets V*, A. C. Dubey *et al.*, Eds., 2000.

[6] G. F. Zhang, L. Tsang, and K. Pak, "Angular correlation function and scattering coefficient of electromagnetic waves scattered by a buried object under a two-dimensional rough surface," *J. Opt. Soc. Amer. A*, vol. 15, pp. 2995–3002, 1998.

[7] J. Q. He, T. J. Yu, N. Geng, and L. Carin, "Method of moments analysis of electromagnetic scattering from a general three-dimensional dielectric target embedded in a multilayered medium," *Radio Sci.*, vol. 32, pp. 305–313, 2000.

[8] T. J. Cui and W. C. Chew, "Fast algorithm for electromagnetic scattering by buried 3-D dielectric objects of large size," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2597–2608, Sept. 1999.

[9] S. Tjuatja, A. K. Fung, and J. Bredow, "Radar imaging of buried objects," in *Proc. Int. Geoscience and Remote Sensing Soc.*, 1998, pp. 524–526.

[10] K. O'Neill, R. F. Lussky, and K. D. Paulsen, "Scattering from a metallic object embedded near the randomly rough surface of a lossy dielectric," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 367–376, 1996.

[11] T. Dogaru and L. Carin, "Time-domain sensing of targets buried under a rough air-ground interface," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 360–372, Mar. 1998.

[12] A. V. der Merwe and I. J. Gupta, "A novel signal processing technique for clutter reduction in GPR measurements of small, shallow land mines," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 2627–2637, Nov. 2000.

[13] M. R. Pino, L. Landesa, J. L. Rodriguez, F. Obelleiro, and R. J. Burkholder, "The generalized forward–backward method for analyzing the scattering from targets on ocean-like rough surfaces," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 961–969, 1999.

[14] M. El-Shenawee and C. Rappaport, "Modeling clutter from random rough ground for GPR subsurface sensing applications," in *Proc. PIERS*, Cambridge, MA, 2000, p. 39.

[15] R. Barrett, M. Berry, T. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. Philadelphia, PA: SIAM, 1993.

[16] L. Tsang, C. H. Chan, K. Pak, and H. Sangani, "Monte Carlo simulations of large scale problems of random rough surface scattering and applications to grazing incidence with the BMIA/canonical grid method," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 851–859, Aug. 1995.

[17] K. Pak, L. Tsang, C. H. Chan, and J. T. Johnson, "Backscattering enhancement of electromagnetic waves from two dimensional perfectly conducting random rough surfaces based on Monte Carlo simulations," *J. Opt. Soc. Amer. A*, vol. 12, pp. 2491–2499, 1995.

[18] K. Pak, L. Tsang, and J. T. Johnson, "Numerical simulations and backscattering enhancement of electromagnetic waves from two dimensional dielectric random rough surfaces with sparse matrix canonical grid method," *J. Opt. Soc. Amer. A*, vol. 14, pp. 1515–1529, 1997.

[19] J. T. Johnson, R. T. Shin, J. A. Kong, L. Tsang, and K. Pak, "A numerical study of the composite surface model for ocean scattering," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 72–83, Jan. 1998.

[20] J. T. Johnson and H. T. Chou, "Numerical studies of low grazing angle backscatter from 1-D and 2-D impedance surfaces," in *Proc. Int. Geoscience and Remote Sensing Soc.*, vol. 4, 1998, pp. 2295–2297.

[21] J. T. Johnson, "A numerical study of low grazing angle backscatter from ocean-like impedance surfaces with the canonical grid method," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 114–120, 1998.

[22] ——, "On the canonical grid method for two dimensional scattering problems," *IEEE Trans. Antennas Propagat.*, vol. 46, pp. 297–302, 1998.

[23] B. T. Draine and P. J. Flatau, "Discrete-dipole approximation for scattering calculations," *J. Opt. Soc. Amer. A*, vol. 11, pp. 1491–1499, 1994.

[24] P. J. Flatau, "Improvements in the discrete-dipole approximation method of computing scattering and absorption," *Opt. Lett.*, vol. 22, pp. 1205–1207, 1997.

[25] J. T. Johnson, R. T. Shin, J. A. Kong, L. Tsang, and K. Pak, "A numerical study of ocean polarimetric thermal emission," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 8–20, Jan. 1999.

[26] H. T. Chou and J. T. Johnson, "A novel acceleration algorithm for the computation of scattering from rough surfaces with the forward–backward method," *Radio Sci.*, vol. 33, no. 5, pp. 1277–1287, 1998.

[27] M. Frigo and S. G. Johnson. FFTW 2.1.3 Manual. [Online]. Available: www.fftw.org.

[28] Maui high performance computing center world wide web site. [Online]. Available: www.mhpcc.edu

[29] E. Newman, "A user's manual for the electromagnetic surface patch code: Preliminary version ESP5.0," ElectroScience Lab., The Ohio State Univ., Columbus, 1997.

[30] S. O. Rice, "Reflection of electromagnetic waves from slightly rough surfaces," *Commun. Pure Appl. Math.*, vol. 4, pp. 361–378, 1951.

Joel T. Johnson (M'96) received the B.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1991 and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, in 1993 and 1996, respectively.

He is currently an Associate Professor in the Department of Electrical Engineering and ElectroScience Laboratory, The Ohio State University (OSU), Columbus. His research interests are in the areas of microwave remote sensing, propagation, and electromagnetic wave theory.

Dr. Johnson is an Associate Member of Commissions B and F of the International Union of Radio Science (URSI), and a Member of Tau Beta Pi, Eta Kappa Nu, and Phi Kappa Phi. He received the 1993 Best Paper Award from the IEEE Geoscience and Remote Sensing Society, and was named an Office of Naval Research Young Investigator, National Science Foundation Career Awardee, and PECASE Award Recipient in 1997.

Robert J. Burkholder (S'85–M'89–SM'97) received the B.S., M.S., and Ph.D. degrees in electrical engineering from The Ohio State University (OSU), Columbus, in 1984, 1985, and 1989, respectively.

Since 1989, he has been with the OSU ElectroScience Laboratory, Department of Electrical Engineering, where he is currently Research Scientist and Adjunct Associate Professor. His research specialties are high-frequency asymptotic techniques and their hybrid combination with numerical techniques for solving large-scale electromagnetic radiation and scattering problems. He has contributed extensively to the EM analysis of large cavities, such as jet inlets/exhausts, and is currently working on the more general problem of EM radiation, propagation, and scattering in realistically complex environments.

Dr. Burkholder is a Full Member of URSI Commission B, and a Member of the Applied Computational Electromagnetics Society (ACES).

TABLE 1 Effect of the Counterclockwise Winding of the Parasitic QHA onthe Performance of the Axial Ratio and HPBW

Pitch Angle (α)	Length of Copper Wires (Normalized to λ)	With Printed Copper Strips	ResonantFrequency (GHz)	Bandwidth with AR < 3 dB (MHz)	HPBW
30°	0.936	Yes	3	81.3	54°
	0.936	No	3	40.5	38°
35°	0.75	Yes	2.67	60.3	41°
	0.75	No	2.67	20	41°
40°	0.5	Yes	3.17	80.2	39°
	0.5	No	3.17	39.9	33°
45°	0.5	Yes	3.64	170	68°
	0.5	No	3.64	40	52°
50°	0.7	Yes	2.67	55.8	79°
	0.7	No	2.67	20.8	50°

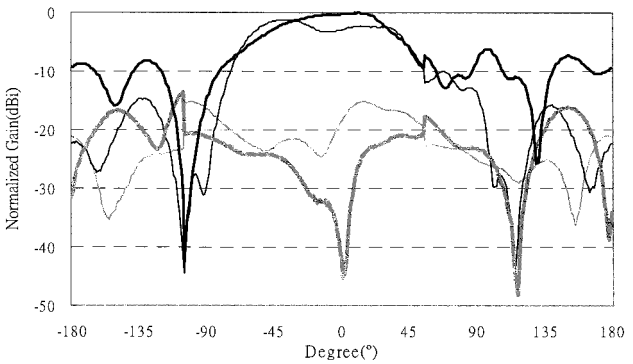


Figure 4 Radiation pattern of configurations in a counterclockwise winding at $l = 0.5\lambda$ and $\alpha = 45°$ at 3.64 GHz. ····· LHCP of the QHA with parasitic strips, ····· LHCP of the QHA without parasitic strips, —— RHCP of the QHA with parasitic strips, —— RHCP of the QHA without parasitic strips

use of a parasitic copper strip QHA is still under investigation.

4. CONCLUSIONS

A novel QHA with a second parasitic QHA is proposed and intensively studied. It is found that the return loss can be reduced from −28 to −47 dB, and the impedance bandwidth (with SWR < 2) can be broadened to 39%. The bandwidth with circular polarization can also be significantly increased. The HPBW of the radiation pattern can be increased from 53 to 68°. This compact antenna is ready for use in mobile satellite communications.

REFERENCES

1. A.T. Adams and C. Lumjiak, Optimization of the quadrifilar helix antenna, IEEE Trans Antennas Propagat (July 1971), 547–548.
2. T. Kamei, S. Takashi, H. Morishita, and T. Nagao, Quadrifilar helical antennas with a variable conical beam, 10th Int Conf Antennas Propagat, 1997, vol. 1, pp. 543–545.
3. K. Fujimoto and J.R. James, Mobile antenna systems handbook, Artech House, Norwood, MA, 1994, chap. 6.5.2, pp. 455–457, 535–536.
4. T. Hirota, A. Minakawa, and M. Muraguchi, Reduced-size branch-line and rat-race hybrids for uniplanar MMIC's, IEEE Trans Microwave Theory Tech 38 (1990), 270–275.

© 2001 John Wiley & Sons, Inc.

A STUDY OF THE FOUR-PATH MODEL FOR SCATTERING FROM AN OBJECT ABOVE A HALF SPACE

Joel T. Johnson[1]
[1] Department of Electrical Engineering and ElectroScience Laboratory
The Ohio State University
Columbus, Ohio 43210

Received 15 February 2001

ABSTRACT: *A study of scattering from a dielectric object located above a half space is performed to clarify the accuracy of a "four-path" model which includes single scattering effects only. An iterative method-of-moments solution is used for comparison, and both frequency- and time-domain examples are illustrated to clarify the scattering interactions which occur.* © 2001 John Wiley & Sons, Inc. Microwave Opt Technol Lett 30: 130–134, 2001.

Key words: *electromagnetic scattering; radar cross section*

1. INTRODUCTION

Many practical geometries for radar applications involve objects in the presence of the Earth surface. Although the ground surface typically has complex features such as vegetation, surface roughness, or dielectric constant variations, in many cases, these effects can be neglected, and the surface of the Earth can be modeled as a planar boundary between free space and a homogeneous dielectric medium. Electromagnetic scattering from objects above a half space has been studied extensively, but analytical solutions are not readily available due to the complexity of the half-space Green's function. A simple "four-path" approximate model which includes only single scattering effects has been widely applied [1–4], but the accuracy of this model has not been explored in detail. Numerical solutions which include all scattering interactions have also been developed [5–7], but insight into the scattering process is not directly provided by these techniques. An analysis of numerical model results to clarify the important scattering mechanisms would therefore be useful for improving understanding and for testing the simple four-path model.

Contract grant sponsor: ONR
Contract grant number: N00014-97-1-0541, N00014-00-1-0399
Contract grant sponsor: NSF
Contract grant number: Project ECS-9701678
Contract grant sponsor: Duke University as part of the OSD MURI on Humanitarian Demining

In this paper, an examination of the four-path model is performed through comparison with results from an efficient numerical model for scattering from a three-dimensional dielectric object in the presence of a dielectric half space. Results in both frequency and time domains are illustrated to allow distinct scattering interactions to be compared. Results show the four-path model to provide reasonable predictions, although higher order interactions are also observed in numerical model data. The following section describes the numerical model used, and Section 3 discusses the example problem considered. A description of the computational issues involved follows in Section 4, and results are presented in Section 5.

2. NUMERICAL MODEL

Figure 1 illustrates the basic geometry considered in this paper: a dielectric object with relative complex permittivity ϵ_3 is located above an interface between free space and a dielectric half space with relative complex permittivity ϵ_2. A method-of-moments formulation of this problem typically would be based on use of the half-space Green's function, so that unknowns would be required only on the dielectric object. However, the current study applies an algorithm [8] which places volumetric electric current unknowns inside the dielectric object and electric and magnetic surface current unknowns on the dielectric boundary, allowing use of the simpler unbounded medium Green's function. This model was developed for studies of scattering from objects in the presence of rough surfaces, for which the half-space Green's function formulation is invalid.

Although the current model will clearly have an increased number of unknowns when compared to a half-space Green's function method, the efficiency of the formulation still allows solutions to be obtained in a reasonable time. Computational efficiency is achieved through the use of an iterative solution of the moment-method matrix equation (using the bi-CGSTAB algorithm [9]) and through application of the canonical grid (CAG) [10–11] and discrete-dipole approximation (DDA) [12–13] methods for computing surface-to-surface and object-to-object point couplings, respectively, in order $N \log N$ computations, where N is the number of surface or object unknowns. A detailed description of the algorithm is provided in [8].

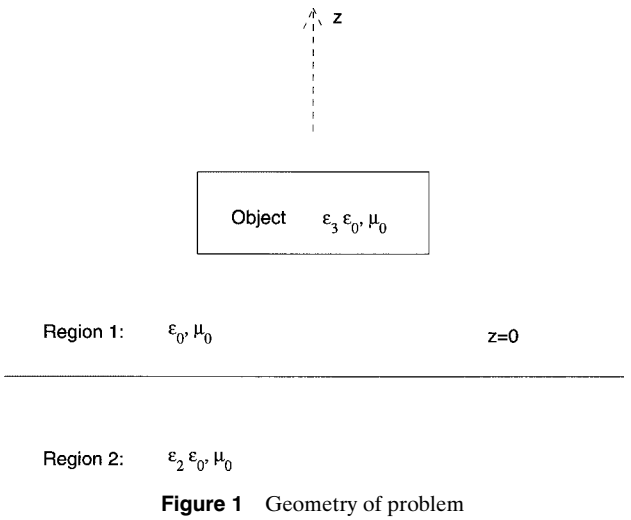


Figure 1 Geometry of problem

3. EXAMPLE PROBLEM

A dielectric rectangular box with dimensions 7.62 cm × 7.62 cm × 2.54 cm (thickness) and relative permittivity $\epsilon_3 = 3 + i0.03$ is used as the object in the results illustrated below. The center of the box is located 7.62 cm above an interface between free space and a medium with relative permittivity $\epsilon_2 = 5 + i1.25$. Scattering from this object is to be determined for a field incident at either 15 or 45° from normal incidence at frequencies from 2 to 5.1 GHz.

Because the model used places unknowns on the dielectric interface between regions one and two, the interface considered must be of finite size. A 1.28 m × 1.28 m interface is used, and artificial edge scattering effects are avoided through use of the "tapered-wave" incident field described in [14]. This field is designed to mimic an incident plane wave, but provides attenuation of incident fields as surface edges are approached. For tapering parameter $g = 3.333$, the field incident on the surface edges is approximately 24 dB lower than that at the center, and the object horizontal cross section projected onto the flat interface is well within the 3 dB incident spot size. A test of tapered-wave influence will be described below through comparison with a plane-wave incidence half-space Green's function code.

Figure 2 illustrates the scattering mechanisms considered in the approximate four-path model, which includes only a single scattering interaction with the dielectric object. Path 1 is a standard backscattering path, and is calculated as though the object were located in free space. Paths 2 and 3 involve a bistatic scatter from the object (again computed as if the object were in free space) and a reflection from the boundary, so an appropriate reflection coefficient must be included in the scattered field amplitude. Note that, since these two paths are identical under time reversal, their contributions are equal. Path 4 involves two reflections from the interface and a backscatter from the bottom surface of the object. Since the object considered in this example is symmetric in z, paths 1 and 4 involve the same object scattering coefficient. The four-path model represents scattering from an object in the presence of a dielectric boundary as the sum of these four mechanisms, including appropriate phase shifts due to the excess path lengths traveled in paths 2–4.

4. COMPUTATIONAL ISSUES

Since the 1.28 m × 1.28 m interface ranges between 8.5 and 21.76 free-space wavelengths as the frequency is varied from 2 to 5.1 GHz, the interface is sampled into 256 × 256 points in the results shown. Although this is somewhat small for the higher frequencies given that the lower medium wavelength is approximately 2.25 times shorter than that in free space, a set of comparisons using 512 × 512 points showed only slight changes in scattered cross sections. While a smaller number of points could be used for the lower frequencies, a constant number of points sampling the interface as the frequency is varied was chosen for convenience. The resulting number of field unknowns on the interface is 262,144. The object is sampled on a 32 × 32 × 8 point grid with step size 3.175 mm (ranging from approximately 1/27 to 1/11 of the wavelength in the object as the frequency varies), resulting in a total number of 13,824 object unknowns. The combined problem thus contains approximately 276,000 unknowns.

Although the problem considered can be solved on a PC level platform, total computing times for the multiple cases considered in this paper were further reduced through use of

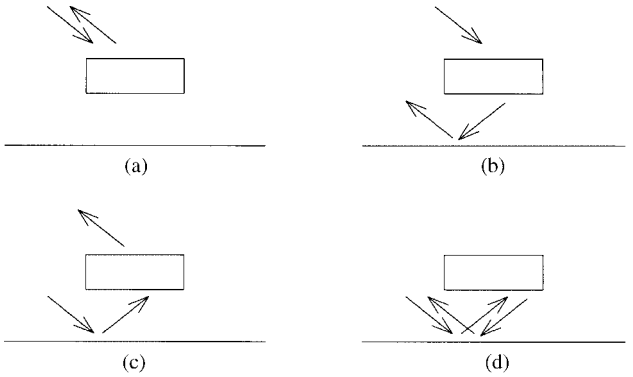


Figure 2 Scattering mechanisms of four-path model. (a) Path 1. (b) Path 2. (c) Path 3. (d) Path 4

IBM SP parallel computing resources at the Maui High Performance Computing Center [15]. Since results as a function of frequency were of interest, single-frequency calculations were performed on individual nodes of the parallel computer (comparable to PC platforms) to obtain 32 frequencies between 2 and 5.1 GHz. Single-frequency computing times on a single node were approximately 2 h; attempts to optimize computing times have, at present, not been performed extensively. Note that these times compare favorably with half-space Green's function codes due to the extensive computational time required for calculation of the half-space Green's function.

Four-path model contributions were calculated using an object in free-space DDA code [12–13] with the same grid as in the combined surface/object code. The smaller number of unknowns in the object-only problem allows a complete solution for all 32 frequencies to be obtained within 10 min on a 700 MHz Pentium III PC. The dramatic decrease in time for these computations indicates the motivation for use of the four-path model; development of approximate solutions for scattering from an object above a half space is also possible with the four-path model if an approximate theory for bistatic scattering from an object in free space is available.

5. RESULTS

Figure 3 illustrates *HH* polarized backscattering cross sections versus frequency for an incidence angle of 15°. For validation, predictions from the ESP5 half-space Green's-function-based code [16] are included for comparison at a subset of the computed frequencies; good agreement is observed, indicating that tapered-wave effects are not altering object cross sections significantly. Some discrepancies within 1 dB are observed, but are difficult to resolve due to computational limitations of the half-space code which is based on direct matrix equation solution. Results from the four-path model are also included, and show good agreement with numerical results. This comparison indicates that the four-path model captures the dominant scattering mechanisms of the problem, although discrepancies of up to 3.3 dB and an error standard deviation (in decibel values) of 1.2 dB indicate that multiple interaction effects not captured by the four-path model may be significant in some cases.

Figure 4 illustrates the same comparisons for 45° incidence and for both *HH* and *VV* cross sections. Polarization effects should be expected to become more important as the incidence angle increases due to polarized object scattering

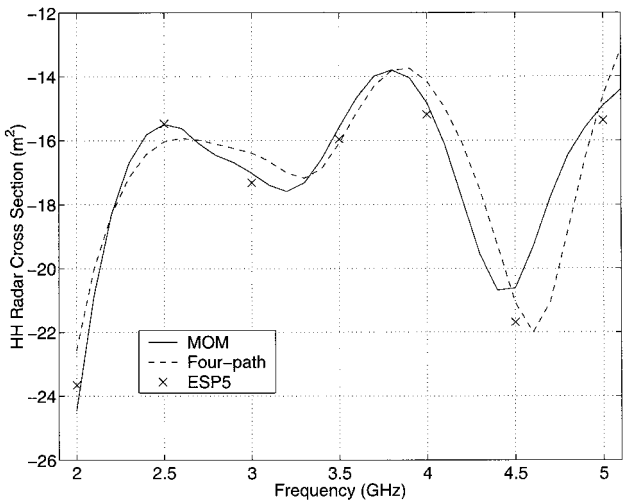


Figure 3 *HH* polarized radar cross sections versus frequency for 15° incidence. Comparison of method-of-moments solution with four-path model and ESP5 half-space Green's function solution

effects, and due to the difference between vertical and horizontal half-space reflection coefficients at larger incidence angles. Results indeed show significant differences between *HH* [plot (a)] and *VV* [plot (b)] results, and again the current model provides excellent agreement with the half-space Green's function code. The error of the four-path model increases compared to Figure 3, with maximum errors of approximately 6.5 dB for both *HH* and *VV* results, and with error standard deviations of approximately 2.1 and 2.6 dB, respectively. Note that the four-path model tends to overpredict numerical model *HH* cross sections; the mean error (again in decibel values) is 2.85 dB. These results continue to indicate that the four-path model captures dominant scattering effects, but neglects some mechanisms which may be significant,

To further explore the important scattering mechanisms, Figures 5 and 6 plot time-domain backscattered fields corre-

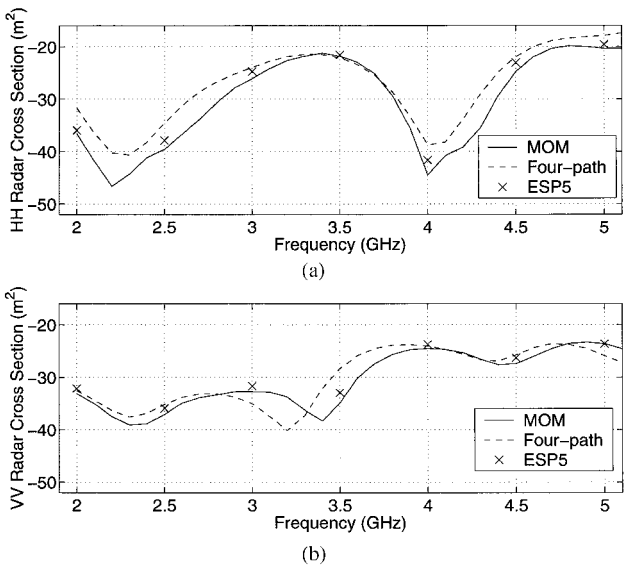


Figure 4 Radar cross sections versus frequency for 45° incidence. Comparison of method-of-moments solution with four-path model and ESP5 half-space Green's function solution. (a) *HH*. (b) *VV*

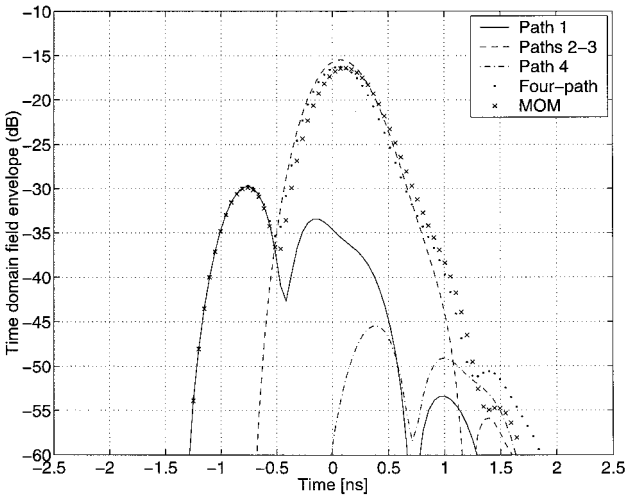


Figure 5 Envelope of *HH* time-domain backscattered fields for 15° incidence. Comparison of method of moments with four-path model

sponding to the results for Figures 3 (15° incidence) and 4 (45° incidence), respectively. The envelope of backscattered fields in decibels is illustrated in these figures, obtained from an inverse FFT after applying a Kaiser–Bessel window function with parameter $\beta = 6$. Time zero is defined at the origin of the half space, so object scattering returns appear at times less than zero. Four-path model results are included in these plots for each individual path as well as their combination; individual path results show the early time return from the object in path 1, the bistatic returns of paths 2 and 3 around time 0, and path 4 returns at longer time delays, as expected. Note in Figure 5 that the bistatic contribution of paths 2 and 3 is more significant than the direct object backscatter due to the relatively low contrast of the object, which results in the majority of energy remaining in the near-forward scattering region. In all cases, excellent agreement with numerical model results is obtained at early times where object only effects dominate scattering returns. However, some differences are observed around time zero, which become larger for the 45° cases, and particularly in vertical polarization. Longer time-delayed results show significant differences in all cases.

Several higher order object–surface interaction mechanisms can be proposed to explain these differences. In particular, a path in which the incident field is scattered by the target into a direction propagating normally onto the half space, reflected, and then rescattered by the object into the backscattering direction would produce contributions around time zero. Multiple reflections between the bottom surface of the target and the half space would then appear at later times, beginning with the approximate time of path 4. Each proposed higher order mechanism also has a corresponding set of four-paths, leading to further time-delayed contributions.

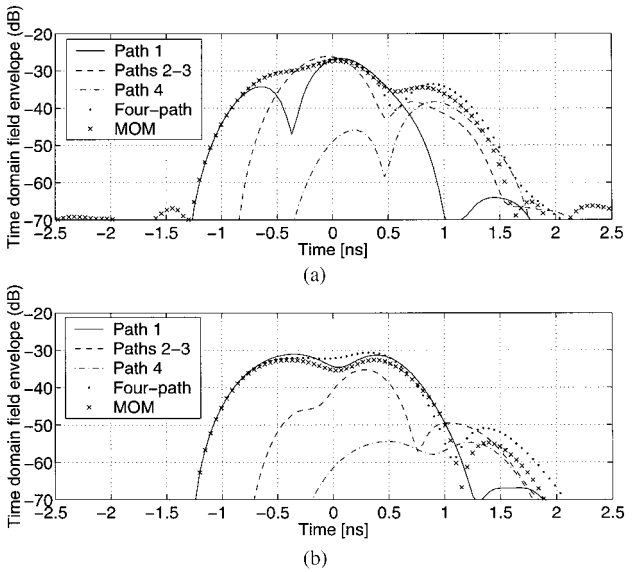


Figure 6 Envelope of time-domain backscattered fields for 45° incidence. Comparison of method of moments with four-path model. (a) *HH*. (b) *VV*

6. CONCLUSIONS

The results of this paper demonstrate that the basic premise of the four-path model for scattering from an object above a half space is valid, but that higher order effects can be significant in some cases which reduce four-path model accuracy. Clearly, the geometry of the target and its location relative to the half space will play an important role in the relevance of higher order mechanisms; the example illustrated here is one for which higher order interactions should be moderately strong. Time-domain backscattered returns allowed four-path model terms to be distinguished, and showed higher order mechanisms to become more important as time delays increase.

ACKNOWLEDGMENT

Discussions with Dr. R. J. Burkholder and Dr. P. H. Pathak are appreciated. Use of the IBM SP system at the Maui High Performance Computing Center is acknowledged, sponsored by the Air Force Research Laboratory, Air Force Material Command under cooperative agreement F29601-93-2-0001. Opinions, interpretations, conclusions, and recommendations are those of the authors, and are not necessarily endorsed by the United States Air Force, Air Force Research Laboratory, or the U.S. Government.

REFERENCES

1. L. Tsang, J.A. Kong, and R.T. Shin, Theory of microwave remote sensing, Wiley, New York, 1985.
2. E.A. Shtager, An estimation of sea surface influence on radar reflectivity of ships, IEEE Trans Antennas Propagat 47 (1999), 1623–1627.
3. M.A. Sletten, D.B. Trizna, and J.P. Hansen, Ultrawide-band radar observations of multipath propagation over the sea surface, IEEE Trans Antennas Propagat 44 (1996), 646–651.
4. T. Chiu and K. Sarabandi, Electromagnetic scattering interaction between a dielectric cylinder and a slightly rough surface, IEEE Trans Antennas Propagat 47 (1999), 902–913.
5. J.Q. He, T.J. Yu, N. Geng, and L. Carin, Method of moments analysis of electromagnetic scattering from a general three-dimensional dielectric target embedded in a multilayered medium, Radio Sci 32 (2000), 305–313.
6. N. Geng, M.A. Ressler, and L. Carin, Wide-band VHF scattering from a trihedral reflector situated above a lossy dispersive half-space, IEEE Trans Geosci Remote Sensing 37 (1999), 2609–2617.
7. T.J. Cui and W.C. Chew, Fast algorithm for electromagnetic scattering by buried 3-D dielectric objects of large size, IEEE Trans Geosci Remote Sensing 37 (1999), 2597–2608.
8. J.T. Johnson and R.J. Burkholder, Coupled canonical grid/discrete dipole approach for computing scattering from objects

above or below a rough interface, IEEE Trans Geosci Remote Sensing (2000, accepted).
9. R. Barrett, M. Berry, T. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst, Templates for the solution of linear systems: Building blocks for iterative methods, available by ftp from netlib2.cs.utk.edu, 1993.
10. L. Tsang, C.H. Chan, K. Pak, and H. Sangani, Monte Carlo simulations of large scale problems of random rough surface scattering and applications to grazing incidence with the BMIA/canonical grid method, IEEE Trans Antennas Propagat 43 (1995), 851–859.
11. J.T. Johnson, R.T. Shin, J.A. Kong, L. Tsang, and K. Pak, A numerical study of the composite surface model for ocean scattering, IEEE Trans Geosci Remote Sensing 36 (1998), 72–83.
12. B.T. Draine and P.J. Flatau, Discrete-dipole approximation for scattering calculations, J Opt Soc Amer A 11 (1994), 1491–1499.
13. P.J. Flatau, Improvements in the discrete-dipole approximation method of computing scattering and absorption, Opt Lett 22 (1997), 1205–1207.
14. K. Pak, L. Tsang, C.H. Chan, and J.T. Johnson, Backscattering enhancement of electromagnetic waves from two dimensional perfectly conducting random rough surfaces based on Monte Carlo simulations, J Opt Soc Amer 12 (1995), 2491–2499.
15. Maui High Performance Computing Center World Wide Web Site, on the World Wide Web at www.mhpcc.edu.
16. E. Newman, A user's manual for the electromagnetic surface patch code: Preliminary version ESP5.0, ElectroScience Laboratory, The Ohio State University, unpublished rep, 1997.

© 2001 John Wiley & Sons, Inc.

TWO INTEGRATED STACKED SHORTED PATCH ANTENNAS FOR DCS / WLAN DUAL-BAND OPERATIONS

Jian-Yi Wu[1] and Kin-Lu Wong[1]
[1] Department of Electrical Engineering
National Sun Yat-Sen University
Kaohsiung, Taiwan 804, R.O.C.

Received 7 February 2001

ABSTRACT: *A new design of two stacked shorted patch antennas integrated on the same microwave substrate for DCS and WLAN dual-band operations is presented. The two antennas together occupy a small area of 19 × 28.5 mm², and the heights for the DCS and WLAN antennas are 10 and 5.2 mm, respectively. Details of the measured antenna performances are presented.* © 2001 John Wiley & Sons, Inc. Microwave Opt Technol Lett 30: 134–136, 2001.

Key words: *stacked shorted patch antenna; dual-band operation; DCS operation; WLAN operation*

1. INTRODUCTION

Stacked shorted patch antennas can have enhanced impedance bandwidths compared to a regular single-layer shorted patch antenna, and several promising designs have been reported [1–3]. These available designs, however, are mainly for single-band operation. In this paper, we present a new design for integrating two stacked shorted patch antennas on the same microwave substrate to perform DCS (digital communication system, 1710–1880 MHz) and WLAN (wireless local-area network, 2400–2485 MHz) operations. The two antennas, having separate feeding ports, are integrated in a compact configuration, and meet the bandwidth requirements for the DCS and WLAN operations, respectively. Details of the antenna design are described, and experimental results of a constructed prototype are presented and discussed.

2. ANTENNA DESIGN

Figure 1 shows the proposed compact configuration of two stacked shorted patch antennas for DCS and WLAN operations. The side views of the DCS and WLAN antennas are, respectively, shown in Figure 1(a) and (b). Figure 1(c) depicts the top view of the integrated configuration of the two antennas, whose lower or driven patches are printed on the same microwave substrate (thickness 3.2 mm and relative permittivity 4.4 in this design) and short circuited to the same ground plane through the microwave substrate. Six and four shorting pins, symmetrically arranged close to the two ends of the shorted edge, are used for short circuiting the lower patches of the DCS and WLAN antennas to the ground plane, respectively. On the other hand, the upper or parasitic patches of the two antennas are short circuited to the same ground plane using a single shorting pin at the center of the shorted edge. The region between the shorted lower and upper patches is an air layer, and plastic posts (not shown in the figure) are used for supporting the upper patch above the lower one. Also note that separate coax feeds are used for the two antennas, and the feed position is selected along the center line of the lower patch.

In order to achieve a compact integrated configuration for the two antennas, the lower patch of the DCS antenna is transformed to have an L shape, which effectively reduces the required lower patch length to be 19 mm (without the patch shape transformation, the required lower patch is about 23 mm), and placed close to that of the WLAN antenna with a small gap of 2 mm. This arrangement makes the two antennas together occupy a small area of 19 × 28.5 mm². As

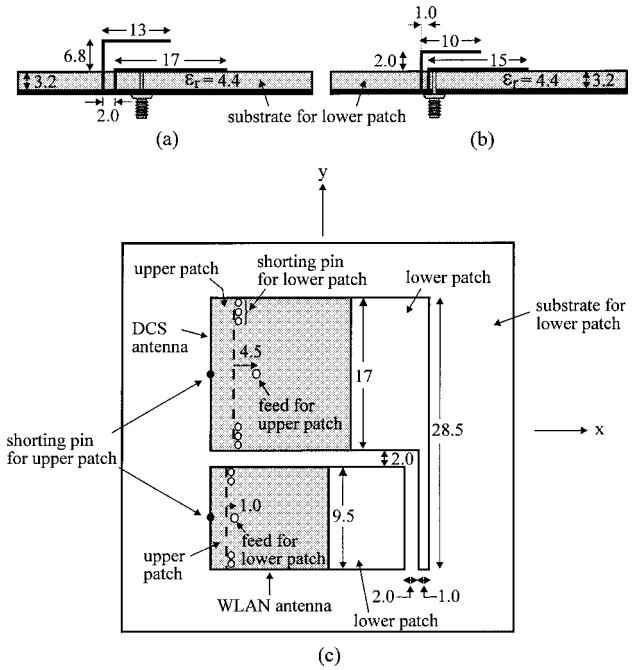


Figure 1 (a) Side view of the stacked shorted patch antenna for DCS operation. (b) Side view of the stacked shorted patch antenna for WLAN operation. (c) Top view of the DCS and WLAN antennas integrated on the same microwave substrate. Dimensions given in the figures are in millimeters, and the ground-plane size is 50 × 50 mm²

Radar images of rough surface scattering: Comparison of numerical and analytical models

Hyunjun Kim and Joel T. Johnson

The Ohio State University

Department of Electrical Engineering

ElectroScience Laboratory

1320 Kinnear Road, Columbus, Ohio 43212

hjk@esl.eng.ohio-state.edu

Abstract: Rough surface scattering theories are investigated through analysis of radar images. Backscatter results from $10 \sim 14$ GHz under tapered wave illumination are considered for 1-D random rough surface realizations which satisfy an impedance boundary condition. Back-projection tomography is applied to form 2-D synthetic aperture radar images from deterministic surface scattered field data at multiple incidence angles and frequencies. Numerical predictions of surface backscattered fields are obtained from an accelerated forward-backward method, and the resulting images are compared with those obtained from approximate scattering theories such as the physical optics (PO) approximation, the small slope approximation (SSA) and the non-local SSA (NLSSA). The resulting radar images illustrate scattering sources associated with single and multiple scattering on the boundary, and a ray tracing analysis confirms the locations of time-delayed image points due to double reflections. For single scattering effects, images demonstrate excellent agreement between analytical and numerical methods in both horizontal and vertical polarizations. For surfaces with rms height 2.0 cm and correlation length 7.5 cm at normal incidence, multiple scattering effects are observed and successfully captured when the lowest order NLSSA is employed.

1 Introduction

In recent years extensive studies of rough surface scattering have been conducted to develop accurate and efficient models valid over a wide range of incident and scattering angles. Several analytical approaches have shown promising results for specified ranges of surface statistics [1, 2]. No approximate solution, however, clearly explains all possible scattering mechanisms, and each approximation is limited to a particular range of surface roughness or electromagnetic parameters. Evaluation of analytical theories is typically based on results for average cross sections. However, data averaged over surface realizations does not provide a detailed description of surface scattering phenomena, so that detailed investigations of the scattering physics captured by the approximate theories have not been performed. Due to the lack of information on the physical behavior of scattering from rough surfaces, a more descriptive approach can be helpful to understand the existing theories.

Imaging techniques have been widely used in other areas of electromagnetics as a tool for analysis and understanding of scattering and propagation phenomena. For example, studies of natural target scattering have used high-resolution radar images to identify small foliage targets [3–6]. Imaging techniques to assist in separating buried targets from surface clutter are also of interest in subsurface target detection problems [7, 8]. The advantages of imaging techniques for improving understanding motivate the use of high resolution imaging for studies of rough surface scattering phenomena [9].

Although computational requirements in image analysis are compounded by the need for data at multiple frequencies and aspect angles, advances in computing facilities and the recent development of efficient numerical methods for backscatter predictions enable radar image formation with numerical scattering models. Images formed from numerical scattering models can be used as a reference solution to evaluate the performance of the existing analytical models such as the small slope approximation (SSA) and the non-local SSA (NLSSA). Recent studies have proven the success of these approximations in terms of average radar cross sections and for surfaces with moderate rms slopes [10–17].

In this paper, high spatial resolution radar images of a single realization of a Gaussian random process surface are formed using both approximate and numerical surface scattering models. In

Section 2, the scattering geometry is described, and a brief review of the analytical and numerical methods to be evaluated is presented, followed by a description of the 2-D synthetic aperture radar (SAR) image formation procedure. Images from the scattering models are compared in Section 3, including a discussion of the performance of the approximate theories. A summary and conclusions are presented in Section 4.

2 Scattering geometry and background theories

A. Problem geometry and incident field

Figure 1 illustrates the scattering geometry of a typical 1-D rough surface, $z = f(x)$, described as a realization of a Gaussian random process with a Gaussian correlation function. Surface statistics can be characterized by the rms height (σ) and correlation length (l_c). The rms slope (s) of the surface is proportional to the ratio of the two parameters ($s = \sqrt{2}\sigma/l_c$) for a Gaussian spectrum. The medium is assumed to be described by an impedance boundary condition with a relative permittivity of $(39.7 + i40.2)$.

A single surface realization is used in this paper, generated with rms height 1.25 cm and correlation length 7.5 cm. A larger amplitude version of this surface is also employed by multiplying the original surface amplitude by 1.6 to obtain an rms height of 2.0 cm. The resulting rms slopes for the two surfaces become 0.236 (13.28°) and 0.377 (20.66°), respectively.

The length of the surface profile studied is $L = 1.92$ meters (64λ at 10 GHz). To avoid scattering effects from surface edges, a Gaussian "tapered wave" [18] is used as the incident field. According to the criterion described in [19], a 64λ surface with tapering parameter $g = L/5$ is sufficient to provide accurate surface scattering calculations up to the maximum 55° incidence angle considered in this paper.

Scattering cross sections under tapered wave illumination are obtained for the analytical theories through a superposition of plane wave responses with a Gaussian weighting in the spectral domain:

$$\Psi_s(k_{ix}, k_{sx}) = \frac{g}{2\sqrt{\pi}} \int_{-\infty}^{\infty} \psi_s(k_{ix} + k_x, k_{sx}) e^{-(k_x g/2)^2} dk_x, \tag{1}$$

where ψ_s is the scattered field under uniform plane wave excitation, $g = L/5$ determines the taper

size of the incident field, and k_{ix} and k_{sx} denote the wave numbers of incident and scattered fields in the x direction, respectively.

B. Scattering theories

A numerical solution based on an iterative method of moments (MOM) called the forward-backward (FB) method is used as a reference backscatter result. To improve the efficiency of the computations, the novel spectral acceleration algorithm is employed [20]. A total of 1024 points were used in sampling the surface profile and surface fields for all scattering calculations of this paper. The physical optics (PO) approximation, the small slope approximation (SSA) and the non-local SSA (NLSSA) are also investigated in terms of radar images in this paper. Results from small perturbation method (SPM) are not included in this paper because SPM is valid only for small height surfaces [1, 21].

In PO predictions, the scattering phenomena involve only single scattering on the surface neglecting interactions between points of the surface [1]. It has been shown that a PO model should be adequate for surfaces with large radii of curvature and for near specular scattering regardless of surface height. PO results however show little polarization dependence even at oblique observation angles.

The small slope approximation (SSA), proposed by Voronovich, is based on a series expansion in generalized surface slope [10–15]. Though this series gradually improves results as additional terms are included, additional integrations in higher order terms make the computations more complicated. However, SSA results have shown good agreement with exact solutions when moderate incident angles are considered (less than 45°) and when rms slopes of surfaces are relatively small (less than 30°) so that only a few series terms are required [15]. In this paper, zeroth (first order in slope) and first (second order in slope) order SSA results are considered. Note that the first order SSA requires an additional integration for backscatter computation compared to PO calculations, but remains very efficient compared to the MOM/FB solution.

The non-local SSA (NLSSA) is an improvement of the SSA which attempts to include non-local interactions more accurately [16, 17]. To include the multiple interactions explicitly, a second kind integral equation (IE) is expanded to include double scattering terms, and then the expansion in

generalized surface slope is applied. A stationary phase analysis shows that the lowest order NLSSA can capture the double scattering mechanism [16]. In the references, expressions for kernels in the NLSSA integrations were developed for PEC surfaces. Since these kernels are based on combinations of SPM solutions, the NLSSA formula for an impedance surface can be easily achieved without loss of generality. Also expressions for vertical polarization can be obtained by duality principles. In this paper, the zeroth order NLSSA formulation is employed, which also requires an additional integration compared to PO calculations. However, the NLSSA double integration can be separated into a pair of 1-D Fourier transforms so that very efficient calculations are possible.

C. Construction of radar images

A 2-D SAR image of a deterministic surface can be constructed from a set of frequency and angular swept complex backscatter field data. This corresponds to a "spotlight" SAR image in which the incident beam is oriented to illuminate a fixed surface area. Tomographic processing using an inverse Fourier transform with back projection is employed to generate the images of this paper [22, 23].

For normal incidence, the down- and cross-range directions coincide with the $-z$ and $+x$ axes as shown in Figure 1, respectively (i.e., the incident field approaches the surface from above). Down- and cross-range resolutions of the image can be determined by the frequency and angular bandwidths, respectively. The down-range and cross-range resolutions, r_d and r_c, are given by

$$r_d = \frac{c}{2B}, \qquad r_c = \frac{c}{2f_o \sin\Theta}, \tag{2}$$

where c is the velocity of light, and B and Θ represent the frequency bandwidth centered on f_o and the angular rotation, respectively. To resolve surface variations on the order of a wavelength, backscatter data were collected over a 4 GHz frequency bandwidth (10 - 14 GHz) and a 20° angular bandwidth corresponding 3.75 cm down- and 3.65 cm cross-range resolution in the image domain, respectively.

The unambiguous down- and cross-ranges, D_d and D_c, can be obtained by the following equations:

$$D_d = \frac{c}{2\delta f}, \qquad D_c = \frac{c}{2f_o \delta\theta}, \tag{3}$$

where δf and $\delta\theta$ denote the steps in frequency and angle, respectively. Step sizes 50 MHz (D_d = 3 meters) and 0.2° (D_c = 3.6 meters at f_o = 12 GHz) are used so that image formation with large down range and cross range unambiguous regions are possible at normal incidence. Note that these unambiguous ranges should also take into account the possible ranges of time-delayed images due to multiple scattering.

To reduce the side-lobe level, a proper choice of windows is necessary. Since image formation is closely related to the Fourier transform, there is a tradeoff between side-lobe level and spatial resolution. Windows functions in both frequency and angle are chosen to set the relationship between these quantities. In this paper the well-known simple windows are used [24]. A rectangular window has optimum resolution, but the first side-lobe level is relatively high (-13 dB) so that minor scattering events other than strong single scattering can be completely obscured by the side-lobes. The Hamming widow has a low side-lobe level (-43 dB) with a wider main lobe. The disadvantage of the Hamming window is that the side-lobe level does not decrease significantly at wide ranges. Throughout this paper a Kaiser-Bessel window ($\alpha = 2$) is selected as an appropriate choice for the windowing function, resulting in a fast decaying side-lobe level at the expense of degrading image resolution.

3 Backscatter Results and 2-D SAR Images

A. Moderate height surface

First, SAR images of the moderate rms height surface with σ = 1.25 cm and l_c = 7.5 cm are investigated for aspect angles centered at normal incidence. The images in Figure 2 show the horizontal (HH) and vertical (VV) polarization images reconstructed from the exact numerical results (MOM/FB). Each image is expressed within the dynamic range of 60 dB ($-60 \sim 0$ dB) and composed of 200 × 200 pixels in a 2 m × 2 m range so that the pixel size is much smaller than the image resolution (note only a 2 m × 1.6 m range is shown in the figure). As mentioned in the previous section, the image resolution depends on the choice of windowing function as well as the angular and frequency bandwidths.

Images are observed to have a maximum scattering level at the center of the surface due to the

tapered wave illumination on the surface. The surface profile, $z = f(x)$, is also overlaid to match scattering centers to the corresponding surface points. The main scattering center distributions are from single scattering responses corresponding to the near specular points as shown in the figure. The maximum pixel amplitudes of the images are -1.18 dB and -1.23 dB for HH and VV, respectively.

Radar images obtained from the PO and SSA theories were found to be virtually identical to those in Figure 2, and thus are not plotted. PO results show maximum pixel amplitudes of -1.20 dB for HH and -1.19 dB for VV, showing very little polarization dependence as expected. Zeroth order and first order SSA images are also almost identical, and produce maximum pixel amplitudes of -1.26 and -1.30 dB for HH and VV, respectively. The maximum pixel amplitude differences between PO, first order SSA, and MOM images are thus less than 0.1 dB, demonstrating the accuracy of approximate solutions in predicting single scattering effects for this particular surface at near normal incidence.

It is also interesting to calculate image differences in the spatial domain between analytical and numerical results more quantitatively. To estimate the performance of the theories, the rms error is defined as

$$\epsilon = \frac{\sqrt{\sum_m \sum_n \left| \Phi_{m,n} - \tilde{\Phi}_{m,n} \right|^2}}{\sqrt{\sum_m \sum_n \left| \Phi_{m,n} \right|^2}}, \tag{4}$$

where $\Phi_{m,n}$ and $\tilde{\Phi}_{m,n}$ represent the pixel intensity of the exact solution and approximations, respectively. With the above definition, for both polarizations, the rms errors are 0.6% and 1.5% for the PO and first order SSA images, respectively. The error estimation shows that approximate theories such as PO and SSA for moderate height surfaces can predict the overall scattering events quite accurately.

B. Large height surface

For the moderate rms height case only single scattering contributions were observed. However as the surface rms height increases some additional effects can appear. Figure 3 shows radar images from the exact MOM solution for the scaled large rms height surface ($\sigma = 2.0$ cm). When compared

to the moderate rms height case, the maximum pixel amplitudes (-3.41 dB for HH and -3.47 dB for VV) are decreased by approximately 2.2 dB due to the fact that the moderate height surface has more specular points than the large height surface. Images for both polarizations show additional scattering points below the surface, possibly from multiple scattering effects.

To study the origin of these additional scattering points, a ray tracing analysis was carried out to predict the locations of points which occur due to double reflection. Figure 4 illustrates predicted locations corresponding to double bounces between two points. Overlaid by the images of Figure 3, the predicted points match the time-delayed images from multiple scattering effects. Another property of the multiple scattering images is that the spots are deformed into stripe-like images with an increasing angular bandwidth whereas single scattering spots become more focused. These effects were clearly observed when the angular bandwidth was doubled from 20° to 40°.

First order SSA predictions for the large rms height surface are presented in Figure 5 and fail to resolve the non-local interactions. However, good results for single scattered responses are obtained when compared to the MOM images of Figure 3. PO images (maximum pixel amplitude of -3.43 dB for HH and -3.42 dB for VV) are again virtually identical to the first order SSA (maximum pixel amplitude of -3.49 dB for HH and -3.53 dB for VV) for this case, and maximum pixel amplitudes from both theories are within 0.1 dB of the maximum pixel amplitude for numerical results in both polarizations. PO images produce an rms error of 4.4% for HH and 6.8% for VV while first order SSA images obtain an rms error of 4.9% for HH and 6.8% for VV in the image domain, respectively, which are larger than the moderate rms surface case. This is expected because both the PO and SSA predictions should be more accurate for surfaces with smaller slopes.

Non-local SSA (NLSSA) images are also presented in Figure 5. Excellent predictions for single scattering are acquired from NLSSA for both polarizations. More importantly, the NLSSA also captures multiple scattering contributions in VV polarization successfully. Again, the maximum pixel amplitude (-3.42 dB for HH and -3.46 dB for VV) shows less than a 0.1 dB difference from MOM/FB results, and the rms error is reduced to 4.1% and 2.1% for HH and VV, respectively. Unlike the first order SSA, NLSSA produces a smaller rms error for the VV case because NLSSA is able to capture multiple scattering effects comparable to the level of the MOM. For the HH case, however, the NLSSA underestimates multiple scattering effects. Changing the windowing function

to a lower side-lobe windowing such as the Nuttall window (3 cosines), multiple scattering points can also be observed in the NLSSA HH case. The underprediction of HH multiple scattering effects observed is consistent with the discussion of polarization effects in multiple scattering NLSSA contributions provided in [16].

C. Oblique incidence images

Figure 6 shows images from MOM and NLSSA results for aspect angles centered at 45° for the large rms height surface. With a 20° angular bandwidth, the highest incident angle becomes 55°. The incident field propagates in the $+x$ and $-z$ directions in the images shown. When compared to the normal incidence case, scattering centers have moved to points near the specular direction at 45°, and the maximum pixel amplitude is -7.58 dB for HH and -7.45 dB for VV. Images formed with the moderate height surface show similar features and a maximum HH and VV pixel amplitudes of -31.8 dB and -31.5 dB, respectively. Because rms slopes for the moderate and large height surfaces are 13.28° and 20.66°, respectively, the probability of obtaining specular features at 45° incidence angle is much greater for the large height surface, resulting in the much larger maximum pixel amplitudes compared to the moderate height case. Overall scattering cross sections for the moderate height surface are primarily non-specular in this angle range, and thus decrease rapidly as the angle increases. At high incidence angles multiple scattering events are no longer observable for the large height surface although the exact images show a very low level of contributions when the dynamic range is extended. This is because the chance to have near specular reflections for two points simultaneously reduces significantly for smooth Gaussian roughness surfaces and, thus dominant scattering mechanisms are from the specular points.

Performance of the approximate theories for the large height surface is as follows: PO (1.6% error, -7.45 dB maximum), SSA (1.6% error, -7.45 dB maximum), NLSSA (5.7% error, -8.1 dB maximum) in HH, while VV performance is: PO (0.5% error, -7.43 dB maximum), SSA (9.0% error, -6.66 dB maximum), NLSSA (23.9% error, -5.5 dB maximum). For the moderate height surface case, performance is: PO (2.8% error, -31.6 dB maximum), SSA (1.8% error, -31.7 dB maximum), NLSSA (4.9% error, -32.3 dB maximum), in HH, while VV performance is: PO (0.64% error, -31.5 dB maximum), SSA (13.9% error, -30.3 dB maximum), NLSSA (21.6% error, -29.96

dB maximum). Errors obtained in HH polarization for both surfaces are comparable to those at normal incidence, although the NLSSA shows slightly worse performance compared to PO and SSA results. Vertical polarization results however show greatly increased errors for the NLSSA and SSA theories, while PO continues to perform well. This somewhat surprising result suggests that the large height surface is sufficiently rough to begin to approach the geometrical optics (GO) limit, in which scattering is completely dominated by specular points. The similarity of HH and VV maximum pixel amplitudes from the MOM/FB model further indicates a GO scattering behavior. Since GO results are captured by the PO theory, but only approximately by the SSA and NLSSA theories as more terms are included, the low order SSA and NLSSA methods used here obtain somewhat larger errors. Of course, generalization of these results will require further image studies of a larger number of surface realizations.

4 Summary and Conclusion

Backscatter radar images of a deterministic rough surface with a Gaussian spectrum have been investigated. Observations in the image domain enabled interpretation of the major and secondary scattering events on the rough surfaces. Dominant scattering events were found at surface points related to the near specular directions.

For the moderate rms height surface ($\sigma = 1.25$ cm and $l_c = 7.5$ cm) single scattering returns are dominant and multiple scattering effects were negligible at normal incidence. In this case, PO and first order SSA predictions showed excellent agreement with numerical results for both horizontal and vertical polarizations. Estimated image domain errors between PO, first order SSA, and numerical images were less than 2%.

Multiple scattering effects at normal incidence were found to become more significant as the rms height increased, as depicted in the numerical images. For the large rms height surface with $\sigma = 2.0$ cm and $l_c = 7.5$ cm, time delayed spots due to multiple scattering clearly appeared on the image space underneath the surface. A simple ray tracing algorithm was able to specify the location of these spots due to double reflection between two points. Images revealed that multiple scattering effects were captured successfully from the lowest order NLSSA results, especially in

VV polarization, but not from the PO or the first order SSA results which emphasize only single scattering and local interactions.

The studies of this paper demonstrate that radar images provide a means for better understanding of rough surface scattering problems, and confirm that the NLSSA can capture some multiple scattering effects.

Acknowledgements:

Use of the RIDE imaging code provided by Dr. I. J. Gupta of the ElectroScience Laboratory is appreciated. This work was sponsored by ONR contracts N00014-97-1-0541 and N00014-00-1-0399, and NSF project ECS-9701678.

References

[1] L. Tsang, J. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing,* Artech House, 1985.

[2] J. T. Johnson, R. T. Shin, J. A. Kong, L. Tsang, and K. Pak, "A numerical study of the composite surface model for ocean backscattering," *IEEE Trans. Geosc. Rem. Sens.*, vol. 36, pp. 72–83, 1998.

[3] G. Zhang and L. Tsang, "Wave scattering and scene image of trees generated by Lindenmayer systems," *Proc. Int. Geosci. Remote Sensing Symp.*, Lincoln, Nebraska, pp. 728–729, 1996.

[4] H. Kim, J. T. Johnson, and B. Baertlein, "High resolution Ka-band backscatter images of a small tree: Measurements and models," *IEEE Trans. Geosc. Rem. Sens.*pp. 899–910, Mar. 2000.

[5] S. Brown and J. C. Bennett, "High-resolution microwave polarimetric imaging of small trees," *IEEE Trans. Geosc. Rem. Sens.*, vol. 37, pp. 48–53, Jan. 1999.

[6] J. Fortuny and A. J. Sieber, "Three-dimensional synthetic aperture radar imaging of a fir tree: first results," *IEEE Trans. Geosc. Rem. Sens.*, vol. 37, pp. 1006–1014, Mar. 1999.

[7] S. E. Shih, K. H. Ding, Y. Zhang, and J. A. Kong, "Subsurface detection based on enhanced SAR signatures using angular correlation function," *Proc. Int. Geosci. Remote Sensing Symp.*, Seattle, WA, pp. 530–532, 1998.

[8] G. Zhang and L. Tsang, "Application of angular correlation function of clutter scattering and correlation imaging in target detection," *IEEE Trans. Geosc. Rem. Sens.*, vol. 36, pp. 1485–1493, Sep. 1998.

[9] H. Kim and J. T. Johnson, "A comparison of rough surface scattering theories at low grazing angles," *Proc. URSI*, Toronto, Canada, p. 196, 1999.

[10] A. G. Voronovich, "Small-slope approximation in wave scattering by rough surfaces," *Sov. Phys.-JETP*, vol. 62, pp. 65–70, 1985.

[11] A. G. Voronovich, "Small-slope approximation for electromagnetic wave scattering at a rough interface of two dielectric half-spaces," *Waves Random Media*, vol. 4, pp. 337–367, 1994.

[12] A. G. Voronovich, *Wave Scattering from Rough Surfaces,* Springer Series on Wave Phenomena 17, 1995.

[13] S. L. Broschat, "Small slope approximation reflection coefficient for scattering from a "Pierson-Moskowitz" sea surface," *IEEE Trans. Geosc. Rem. Sens.*, vol. 31, pp. 1112–1114, Nov. 1993.

[14] E. I. Thorsos and S. L. Broschat, "An investigation of the small slope approximation for scattering from rough surfaces. Part I. Theory," *J. Acoust. Soc. Amer.*, vol. 97, pp. 2082–2093, 1995.

[15] S. L. Broschat and E. I. Thorsos, "An investigation of the small slope approximation for scattering from rough surfaces. Part II. Numerical studies," *J. Acoust. Soc. Amer.*, vol. 101, pp. 2615–2625, 1997.

[16] A. G. Voronovich,"Non-local small-slope approximation for wave scattering from rough surfaces," *Waves Random Media*, vol. 6, pp. 151–167, 1996.

[17] S. L. Broschat, "Reflection loss from a "Pierson-Moskowitz" sea surface using the nonlocal small slope approximation," *IEEE Trans. Geosc. Rem. Sens.*, vol. 37, pp. 632–634, Jan. 1999.

[18] E. I. Thorsos, "The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Amer.*, vol. 83, pp. 78–92

[19] D. A. Kapp, *A New Method to Calculate Wave Scattering from Rough Surfaces at Low Grazing Angles,* Ph.D. dissertation, Virginia Polytechnic Institute and State University, 1995.

[20] H.-T. Chou and J. T. Johnson,"A novel acceleration algorithm for the computation of scattering from rough surfaces with the forward-backward method,"*Radio Science*, vol. 33, pp. 1277–1287, 1998.

[21] S. O. Rice,"Reflection of electromagnetic waves from slightly rough surfaces,"*Commum. Pure Appl. Math*, vol. 4, pp. 361–378, 1951.

[22] I. J. Gupta and A. Gandhe, "Radar image editing and data extrapolation," Technical report no. 727723-9, The Ohio State University ElectroScience Laboratory, Apr. 1996.

[23] D. L. Mensa, *High Resolution Radar Cross Section Imaging,* Artech House, 1991.

[24] F. J. Harris,"On the use of windows for harmonic analysis with the discrete Fourier transform", *Proc. IEEE*, vol. 66, pp 51–83, 1978.

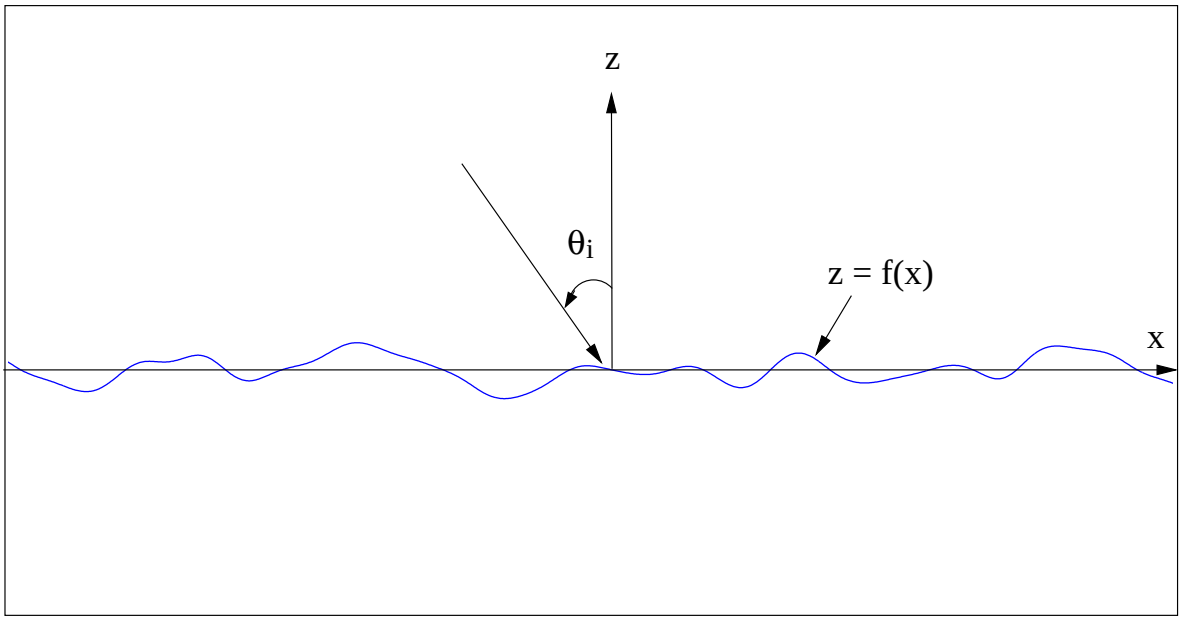


Figure 1: Scattering geometry of a rough surface

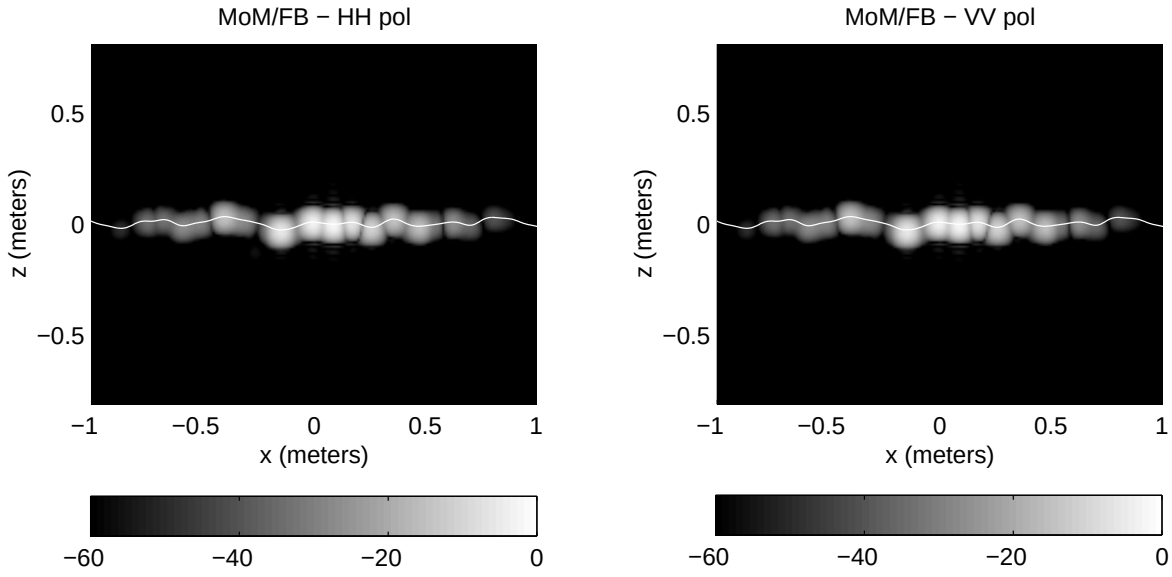


Figure 2: 2-D SAR images for a Gaussian rough surfaces with moderate rms height ($\sigma = 1.25$ cm and $l_c = 7.5$ cm): MOM/FB, $f = 10 \sim 14$ GHz, $\delta f = 50$ MHz, $\theta_i = -10^\circ \sim 10^\circ$ and $\delta\theta = 0.2^\circ$ (a) Horizontal polarization (b) Vertical polarization

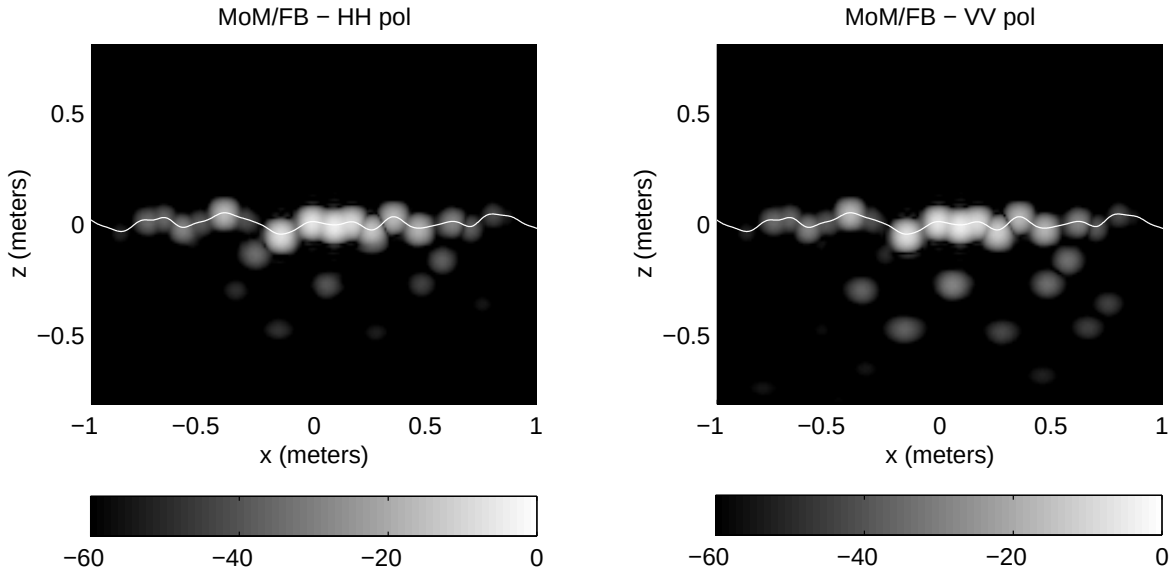


Figure 3: 2-D SAR images for a Gaussian rough surfaces with high rms height ($\sigma = 2.0$ cm and $l_c = 7.5$ cm): MOM/FB, $f = 10 \sim 14$ GHz, $\delta f = 50$ MHz, $\theta_i = -10^\circ \sim 10^\circ$ and $\delta\theta = 0.2^\circ$ (a) Horizontal polarization (b) Vertical polarization

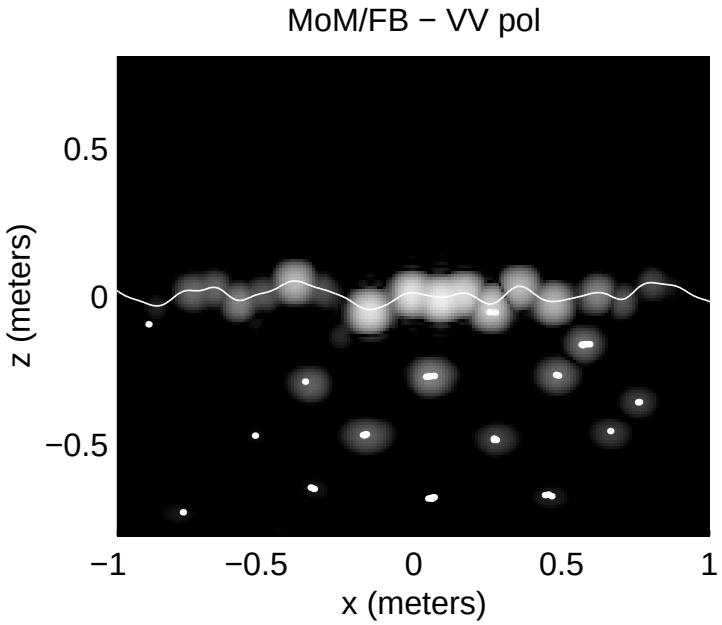


Figure 4: Ray tracing prediction for the location of image spots due to multiple scattering at normal incidence: Overlay of time-delayed scattering centers (marked as dots) and the same image as in Figure 3

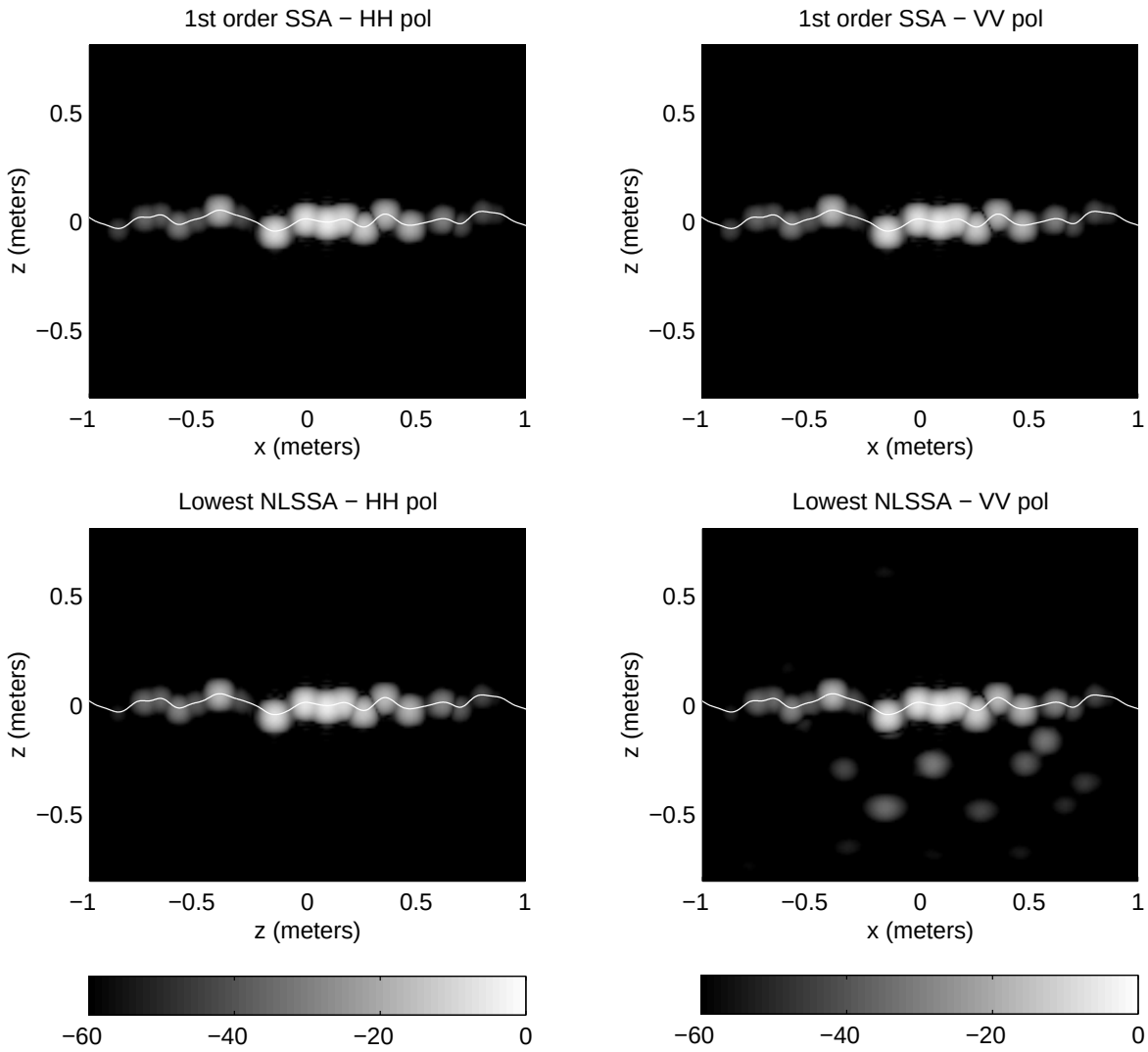


Figure 5: 2-D SAR images of a Gaussian rough surface with high rms height ($\sigma = 2.0$ cm and $l_c = 7.5$ cm): first order SSA and Non-local SSA, $f = 10 \sim 14$ GHz, $\delta f = 50$ MHz, $\theta_i = -10^\circ \sim 10^\circ$ and $\delta\theta = 0.2^\circ$ (a) Horizontal polarization (b) Vertical polarization

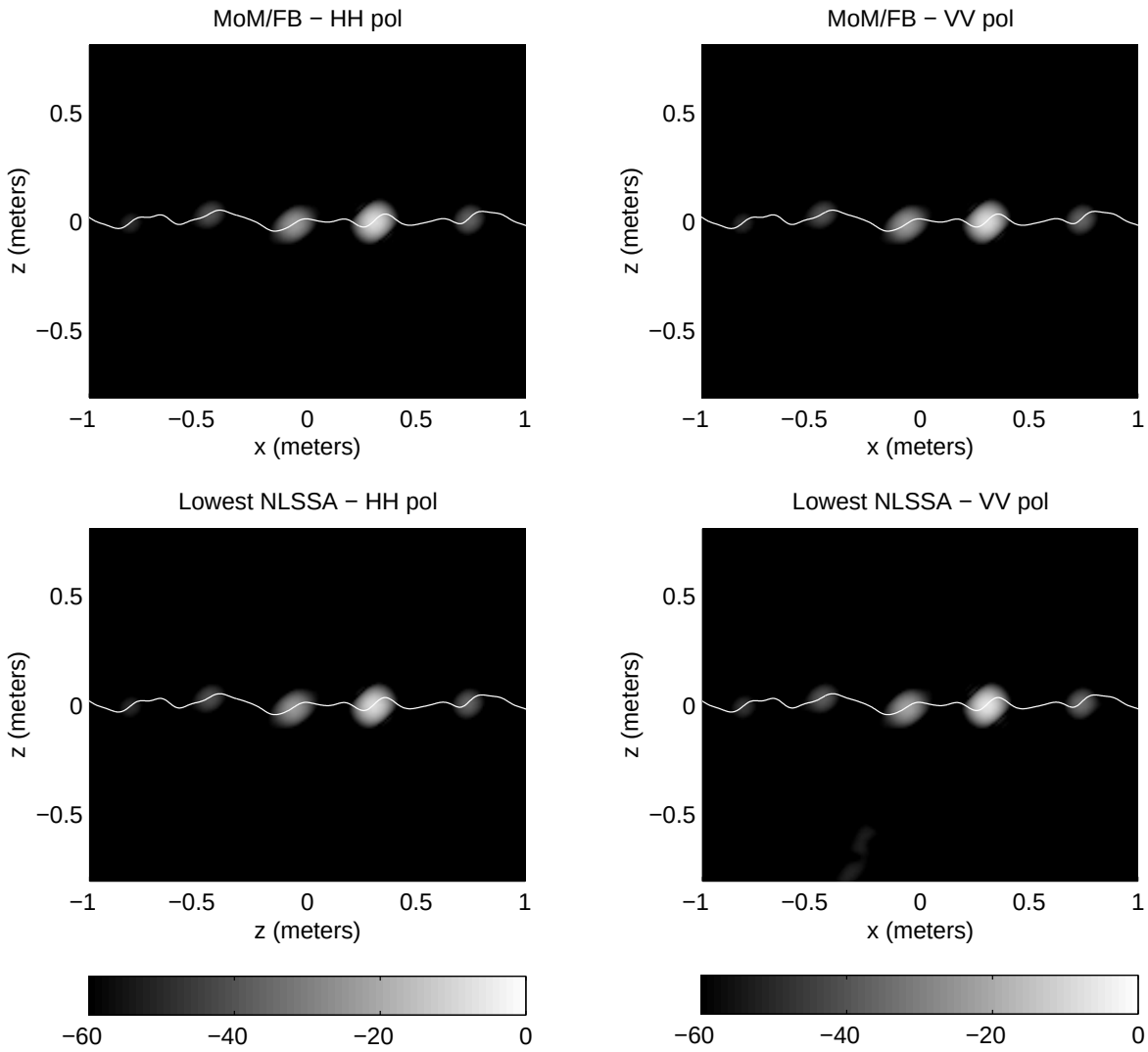


Figure 6: 2-D SAR images for a Gaussian rough surfaces with high rms height and large aspect angles ($\sigma = 2.0$ cm and $l_c = 7.5$ cm): MOM/FB and Non-local SSA, $f = 10 \sim 14$ GHz, $\delta f = 25$ MHz, $\theta_i = 35^\circ \sim 55^\circ$ and $\delta\theta = 0.2^\circ$ (a) Horizontal polarization (b) Vertical polarization

A Numerical Study of Scattering from an Object Above a Rough Surface

Joel T. Johnson

Department of Electrical Engineering and ElectroScience Laboratory

The Ohio State University

205 Dreese Laboratories

2015 Neil Ave

Columbus, OH 43210

(614) 292-1593, (614) 292-7297 FAX

johnson@ee.eng.ohio-state.edu

ABSTRACT

A numerical model is applied in a Monte Carlo study of scattering from a three dimensional penetrable object above a lossy dielectric rough interface. The model is based on an iterative method of moments solution for equivalent electric and magnetic surface current densities on the rough interface and equivalent volumetric electric currents in the penetrable object. Both time and frequency domain results are investigated to illustrate the relative importance of coherent and incoherent scattering effects. Results show that a four-path model using a reduced reflection coefficient can be reasonable for coherent scattering predictions, and that incoherent scattering in the combined object-surface problem can be significantly different than that obtained with the rough surface alone.

Keywords: rough surface scattering, radar cross section, electromagnetic scattering

1 Introduction

Electromagnetic scattering from objects is affected by the surrounding medium. Many realistic geometries involve objects in the presence of the Earth surface, which is often modeled as a planar dielectric boundary [1]-[4]. However, roughness on the Earth surface can potentially modify object scattering returns from those with a flat surface, particularly in cases where the roughness size becomes larger than a fraction of the electromagnetic wavelength. Analysis of these problems is complicated by the many possible scattering interactions between the rough surface and object; at present, approximate analytical solutions exist only in the small roughness limit [5]-[9].

Recent works have explored numerical solutions of the combined object/rough surface scattering problem [10]-[15], but have concentrated primarily on two dimensional scattering problems to reduce computational complexity. The majority of previous numerical studies have also been directed toward studies of scattering from objects beneath a rough surface for application to ground penetrating radar problems. Substantial motivation also exists, however, for studying problems in which objects are located above a rough surface, as demonstrated in [15]-[17] among other references.

In this paper, a numerical study of scattering from a three dimensional penetrable object located above a lossy dielectric rough interface is performed to provide an illustration of some of the coherent and incoherent scattering effects which can occur. A Monte Carlo simulation is used to obtain scattered field statistics as a function of frequency from 2 to 5 GHz, and results are illustrated in both the frequency and time domains to clarify the scattering physics. Results show that a "four-path" model [4] can remain reasonable for prediction of total coherent scattered fields if a rough surface reflection coefficient [18] is employed. An examination of incoherent scattered fields shows that significant differences can be obtained from those obtained in the presence of the rough surface alone.

The next section briefly reviews the numerical model employed in the study, and Section 3 describes the particular problem for which simulations are performed. Computational issues are discussed in Section 4, and results are presented in Section 5.

2 Numerical model

Figure 1 illustrates the basic geometry considered in this paper: a dielectric object with relative complex permittivity ϵ_3 is located above a rough interface $z = f(x, y)$ between free space and a dielectric medium with relative complex permittivity ϵ_2. The numerical model applied to solve this problem is an iterative method of moments solution for single frequency induced volumetric currents in the dielectric object and induced electric and magnetic surface currents on the rough interface. A point matching formulation is applied, and matrix multiply computations required in the iterative method are accelerated through use of the canonical grid method [19]-[20] and the discrete dipole approach [21]-[22] to compute surface to surface and object to object point couplings, respectively, in $O(N \log N)$, where N is the number of surface or object sampling points. A standard iterative method (the "Bi-conjugate gradient stabilized" (BiCG-stab) algorithm [23]) is used on the combined object/surface matrix equation, and the system is preconditioned through a "flat surface" approximation for surface to surface contributions and a low accuracy DDA solution for object to object contributions. The model is described in detail in [24], where an example of scattering from an object located below a rough interface is provided.

The rough surface profiles used in the study are realizations of a Gaussian random process and for simplicity are chosen to have an isotropic Gaussian correlation function. The resulting surface statistics are described completely by the surface rms height h and correlation length l. Due to the statistical nature of this problem, scattered field results obtained from an ensemble of surface realizations are considered. While a large number of realizations is desirable for more accurate estimates of scattered field statistics, computational issues described in Section 4 limit the current study to twenty realizations. Convergence tests with the obtained data show that average cross sections estimates should be accurate to within approximately 3 dB.

Because the rough interface modeled in the simulation is of finite size, a "tapered wave" incident field is used to avoid surface edge scattering effects. Incidence angles of 0 and 45 degrees from normal incidence are considered in this paper, and the respective tapered wave formulations are provided in [25] and [26]. The tapered waves used in the study are chosen so that the object is well within the

3 dB spot size of the incident field while approximately 60 dB incident field attenuation is obtained at surface edges. Tests of tapered wave influence in the flat surface limit have been performed in [4] through comparison with a plane wave incidence halfspace Green's function numerical solution [27]. Results of the comparison show only slight differences (within 1.5 dB) between tapered wave and plane wave radar cross sections obtained from object and object/surface interaction effects. However, note that direct surface backscattering at normal incidence is strongly influenced by tapered wave parameters; for this reason, only incoherent surface backscattering at normal incidence will be presented. Scattered fields in the study are calculated both for the combined object/rough surface problem and the rough surface only problem, so that rough surface scattering effects can be "separated" from object and object/surface interaction effects if desired.

Due to the presence of both object and distributed source (the rough surface) scatterers, total radar cross sections obtained are also dependent on the rough surface area illuminated by the incident tapered wave. For example, in the limit of a very large "spot size" incident field, rough surface scattering effects become more likely to dominate object scattering effects due to the larger surface area illuminated. Thus it should be noted that the results presented apply only for the particular incident field used. However, tests with larger tapered wave spot sizes showed only slight changes in object and object/surface interaction cross sections.

Consideration of the primary scattering effects of this problem suggests that coherent scattered fields (neglecting the direct surface reflection at normal incidence) should resemble those obtained for an object above a flat surface in the small roughness limit, and those obtained for an object in free space in the large roughness limit. A four-path model [4] based on image theory and a single scattering interaction with the object can be developed to describe this process by including the standard rough surface reflection coefficient modification [18] for fields which encounter the rough surface. Incoherent scattered fields should be caused both by direct surface backscattering and by object/surface interaction effects. The four-path model suggests that the latter are likely to be dominated by paths which involve a single bistatic scatter from the object combined with near specular scattering from the rough surface. However, since incoherent scattering from the rough

surface is distributed through a range of angles, incoherent object/surface interaction effects can be very complex and difficult to describe completely. Examination of frequency and time domain results in Section 5 however will give some indication as to the most important contributions.

3 Example problem

A dielectric rectangular box with dimensions 7.62 cm by 7.62 cm by 2.54 cm (thickness) and relative permittivity $\epsilon_3 = 3 + i0.03$ is used as the object in this study. The center of the box is located 8.89 cm above the rough interface between free space and a medium with relative permittivity $\epsilon_2 = 5 + i1.25$. Scattering for this geometry is to be determined for a field incident at either 0 or 45 degrees from normal incidence at sixteen frequencies from 2 to 5 GHz. A rough surface correlation length of 3.58 cm and surface rms heights of 3.58 mm or 1 cm are used, so that the surfaces range from slightly rough at the lowest frequency ($kh = 0.15$ or 0.42, respectively, where k is the electromagnetic wavenumber) to slightly to moderately rough at the highest frequency ($kh = 0.375$ or 1.05, respectively). However, rms slopes for these surfaces are approximately 8 degrees and 22 degrees, respectively, making the larger height surface exceed the limitations of standard perturbation theory [28]. Surface-only backscattered fields will be compared with results from the first two terms of the small slope approximation (SSA) [29] calculated in a separate SSA Monte Carlo simulation following the procedure described in [30].

Time domain scattered fields are obtained from frequency swept data through an FFT operation preceded by multiplication with a Kaiser-Bessel window (parameter $\beta = 3$) to reduce side-lobe levels. Time zero is defined to correspond to the center of the mean level of the rough surface ($z = 0$) in Figure 1, so that object scattering returns occur at negative times. A calculation of expected time delays shows object scattering contributions at approximate times of -0.60 ns and -0.42 ns for 0 and 45 degrees incidence respectively. Surface scattering returns within the 3 dB tapered wave spot at 45 degrees incidence are spread in time from approximately -0.7 ns to 0.7 ns. This time spreading of surface clutter at oblique observation angles and its effects on detection of objects has been previously described in [14], [31]. Time domain field statistics are calculated in

terms of the mean and variance of the field envelope as a function of time to clarify the time locations of various coherent and incoherent scattering effects. Of course, rough surface incoherent scattered fields should show no particular time location, but object/surface incoherent interaction effects do contain some time information which can help to indicate the important scattering mechanisms.

4 Computational issues

A 1.281 m by 1.281 m surface size is used which ranges from 8.5 to 21.35 free space wavelengths side dimension as the frequency varies from 2 to 5 GHz. The tapered wave 3 dB spot diameter with parameter $g = 5.333$ is then 28.3 cm so that the object is well within the tapered wave illumination pattern. The interface is sampled into 256 by 256 points, producing a sampling rate of 5.36 points per wavelength in the dielectric medium at the highest frequency; tests with 512 by 512 points in the flat surface limit showed negligible cross section variations. While a smaller number of surface points could be used for the lower frequencies, a constant number of points sampling the rough interface as frequency is varied was chosen for convenience. The resulting number of field unknowns on the interface is 262144. A "strong" bandwidth of 15 points and one canonical grid series term were used in rough surface matrix elements, as described in [24]; single realization tests confirmed that these parameters should provide accurate results for rough surface scattering contributions.

The object is sampled on a 32 by 32 by 8 point grid with step size 3.175 mm (ranging from approximately 1/27 to 1/11 of the wavelength in the object as frequency varies), resulting in a total number of 13824 object unknowns. The combined problem thus contains approximately 276000 unknowns. While this large number of unknowns would be prohibitive for many integral equation based methods, the efficient algorithm applied makes the current study possible.

Although the problem considered can be solved on a PC level platform for a single realization, total computing times for the multiple cases considered in this paper were further reduced through use of IBM SP parallel computing resources at the Maui High Performance Computing Center [32]. Since results as a function of frequency for multiple realizations were of interest, single frequency-single realization calculations were performed on individual nodes of the parallel com-

puter (comparable to PC platforms) to obtain twenty realizations with 16 frequencies between 2 and 5 GHz. Single frequency computing times on a single node ranged from approximately six to fourteen hours depending on frequency, incidence angle, and surface statistics; attempts to optimize computing times have at present not been performed extensively. Four-path model contributions were calculated using an object in free space DDA code [21]-[22] with the same grid as in the combined surface/object code, and synthesized following the procedure described in [4].

5 Results

A. Frequency domain

Figure 2 plots average coherent (plot (a)) and incoherent (plot (b)) backscattered co-polarized radar cross sections versus frequency for 0 degrees observation and for both the rms height 3.58 mm and 1 cm cases. Again, direct surface reflections are not included in plot (a). Also included are the corresponding cross sections for the object above a flat surface, as well as predictions for coherent cross sections using the reduced reflection coefficient four-path model. Coherent cross sections for the 3.58 mm rms height surface are very similar to those obtained with the object above a flat surface, while those for the rougher surface are significantly different and approach those for the object in free space. The four-path model is found to perform very well for this case, indicating that its approximations remain reasonable even in the presence of rough surfaces. The success of the four-path model indicates that terms involving more than one object scattering process can be neglected for normal incidence in this problem.

Incoherent cross sections in plot (b) are shown for surface-only contributions (obtained from the surface only numerical solution), combined surface/object total contributions, and object/surface interaction terms (obtained with total fields minus surface only fields.) All of these terms are found to increase with frequency, as expected for these surface statistics since more power from the coherent field is converted to incoherent power at higher frequencies. For the smaller rms height surface, total incoherent scattering contributions remain significantly smaller than coherent returns, while incoherent scattering is larger than coherent scattering at some frequencies for the rougher

surfaces. Surface only backscattering dominates object/surface interaction effects at all frequencies for the rougher surface, but object/surface interaction effects become comparable to surface only scattering at higher frequencies for the smaller height surface.

Figures 3 and 4 illustrate coherent and incoherent scattering returns, respectively, for 45 degrees incidence in HH (plot (a)) and VV (plot (b)) polarizations. Polarization differences should be observable in this problem for oblique incidence backscattering due to polarized object scattering and due to the polarization sensitivity of rough surface scattering at oblique angles. Coherent cross sections indeed show significant differences between HH and VV returns. Differences of rough surface coherent cross sections from those with a flat surface are less noticeable than in the 0 degrees case, due to the reduced Rayleigh parameters obtained at oblique incidence and smaller differences between the object in free space and object above a flat surface returns at 45 degrees. The accuracy of the four-path model (plotted only for the rougher surface case) is also reduced compared to Figure 2; similar levels of error are observed when comparing four-path and numerical model results with a flat surface. These discrepancies indicate that paths involving more than one object scattering process are more important for oblique paths, as discussed in [4].

Incoherent returns in Figure 4 show that incoherent scattering can be greater than coherent scattering even with the small height surface at some frequencies. Incoherent returns for the larger height surface dominate coherent returns at almost all frequencies for both HH and VV polarizations. Again surface incoherent scattering is dominant over object/surface interaction effects for almost all cases, but higher frequency HH results show a region where object/surface interaction contributions exceed those from the rough surface alone. As stated previously, the four-path model for incoherent scattering suggests that object/surface interaction effects can be dominated by paths involving a bistatic object scattering followed by a near-specular surface scattering. Studies of surface-only incoherent scattering cross sections show large increases in forward scattered fields with frequency at 45 degrees, while backscattered returns increase only slightly or decrease with frequency. Furthermore, the ratio of forward to backscattered incoherent cross sections is greatest for the HH, smaller rms height surface at the highest frequency, where it approaches 25

dB. Forward to backscattered cross section ratios for all other cases obtain maximum values of 15 dB. Thus, the results obtained support the four-path model idea that both surface backscattered and forward scattered incoherent cross sections are important in combined object/rough surface interaction problems.

A validation of rough surface-only incoherent cross sections is presented in Figure 5, where results at 0 degrees (plot (a)) and 45 degrees (plot (b)) are compared with predictions of the first two terms of the SSA. A Monte Carlo simulation using 100 surface realizations was used to obtain SSA results [30], so that the curves obtained show some residual variations due to the finite number of realizations. Numerical model results are in good general agreement with the SSA, although some differences within approximately 4 dB at the lower frequencies (where the tapered wave causes a larger degree of angular averaging) are observed. Overall the reasonable agreement obtained however validates both the numerical model and the SSA prediction for the surfaces considered.

B. Time domain

Figure 6 presents time domain backscattered field envelopes (in decibels) for 0 degrees incidence in the rms height 3.58 mm (plot (a)) and rms height 1 cm (plot (b)) cases. Both coherent and "incoherent" (i.e. the standard deviation of the field envelope as a function of time) returns are included, as well as returns with the object above a flat surface. Incoherent returns are plotted for total, surface only, and object/surface interaction contributions as well. Coherent returns in Figure 6 show general agreement with flat surface results for the lower rms height surface, but appreciable differences for the rougher surface. Note object scattering returns centered around time -0.6 ns show only minor deviations from the flat surface case since no surface scattering sources have been encountered (other than sidelobe contributions from later times.) Later time coherent returns in the rms height 1 cm begin to approach time domain returns with the object in free space (not plotted). Incoherent scattering contributions at zero degrees occur primarily at times after initial object returns, so that time domain object detection strategies would work well in this case. Incoherent returns show contributions from both surface only and object/surface

interactions, although surface only scattering dominates in the rougher case for most times. Initial object/surface incoherent interaction effects are found to be slightly time shifted from surface only scattering. This is consistent with the dominant four-path mechanism of a bistatic scattering from the object followed by a specular scattering from the rough surface, since a transmission through the object would result in a slight time delay. Object/surface interaction effects are observed to be more significant at later times, as would be expected for multiple object/surface interactions.

Figures 7 and 8 illustrate HH (plot (a)) and VV (plot (b)) time domain statistics for the rms height 3.58 mm and 1 cm cases, respectively. Similar observations regarding coherent fields are obtained in this case, with only slight differences from flat surface returns observed with rms height 3.58 mm while larger differences are observed in the rougher case as coherent fields approach those for an object in free space. Total incoherent returns show the time spreading associated with oblique observation of a rough surface, so that object scattering no longer occurs prior to rough surface returns, making target detection more difficult. As observed in Figures 4 and 5, incoherent scattering dominates coherent scattering for the rougher surface case, and surface only contributions generally are more significant than object/surface interactions even at later times. However, the HH polarized, low rms height surface shows object/surface interaction effects to exceed surface only contributions around time zero, consistent with the four-path model explanation.

6 Conclusions

The results of this study demonstrate some of the coherent and incoherent scattering effects which can occur in combined object/rough surface scattering problems. Coherent cross sections were found to resemble those for an object above a flat surface in the small roughness limit but to approach those for an object in free space as the roughness increased. A four-path model using a rough surface reduced reflection coefficient was found to match coherent cross sections well for normal incidence observation, although the accuracy was degraded at oblique observation where multiple object scattering effects can become more important. Incoherent scattered fields in both the time and frequency domains showed that both direct surface backscattering and object/surface

interaction terms can be important depending on the frequency, surface statistics, polarization, and scattering geometry. Incoherent object/surface interaction effects observed appear consistent with a four-path model interpretation in which the dominant contribution is from an object bistatic scattering followed or preceded by surface forward scattering. Thus both surface backscattering and forward scattering effects must be considered when analyzing returns from object above a rough surface. Further applications of these results and the iterative method of moments model include evaluation of approximate models for combined surface/object problems [5]-[9], as well as tests of target detection algorithms in the presence of rough surface clutter.

7 Acknowledgements

This work was sponsored by ONR contracts N00014-97-1-0541 and N00014-00-1-0399, NSF project ECS-9701678, and by a grant from Duke University as part of the OSD MURI on Humanitarian Demining. Use of the IBM SP system at the Maui High Performance Computing Center is acknowledged, sponsored by the Air Force Research Laboratory, Air Force Materiel Command under cooperative agreement F29601-93-2-0001. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Air Force, Air Force Research Laboratory, or the U.S. Government.

References

[1] He, J. Q., T. J. Yu, N. Geng, and L. Carin, "Method of moments analysis of electromagnetic scattering from a general three-dimensional dielectric target embedded in a multilayered medium," *Radio Science,* vol. 32, pp. 305–313, 2000.

[2] Geng, N., M. A. Ressler, and L. Carin, "Wide-band VHF scattering from a trihedral reflector situated above a lossy dispersive halfspace," *IEEE Trans. Geosc. Rem. Sens.,* vol. 37, pp. 2609–2617, 1999.

[3] Cui, T. J. and W. C. Chew, "Fast algorithm for electromagnetic scattering by buried 3-D dielectric objects of large size," *IEEE Trans. Geosc. Rem. Sens.,* vol. 37, pp. 2597–2608, 1999.

[4] Johnson, J. T., "A study of the four-path model for scattering from an object above a halfspace," submitted to *Microwave Opt. Tech. Lett.,* 2001.

[5] Zhang Y., Y. E. Yang, H. Braunisch, and J. A. Kong, "Electromagnetic wave interaction of conducting object with rough surface by hybrid SPM/MOM technique," *PIER 22: Progress in Electromagnetics Research,* vol. 22, pp. 315–335, 1999.

[6] Ishimaru, A., J. D. Rockway, and Y. Kuga, "Rough surface Green's function based on the first-order modified perturbation and smoothed diagram methods," *Waves in Random Media,* vol. 10, pp. 17–31, 2000.

[7] Chiu, T. and K. Sarabandi, "Electromagnetic scattering interaction between a dielectric cylinder and a slightly rough surface," *IEEE Trans. Ant. Prop.,* vol. 47, pp. 902–913, 1999.

[8] Johnson, J. T., "Thermal emission from a layered medium bounded by a slightly rough interface," accepted by *IEEE Trans. Geosc. Remote Sens.,* 2000.

[9] Fuks, I. M. and A. G. Voronovich, "Wave diffraction by rough interfaces in an arbitrary plane-layered medium," *Waves in Random Media,* vol. 10, pp. 253–272, 2000.

[10] O'Neill, K., R. F. Lussky, and K. D. Paulsen, "Scattering from a metallic object embedded near the randomly rough surface of a lossy dielectric," *IEEE Trans. Geosc. Rem. Sens.,* vol. 34, pp. 367–376, 1996.

[11] Dogaru, T. and L. Carin, "Time-domain sensing of targets buried under a rough air-ground interface," *IEEE Trans. Ant. Prop.,* vol. 46, pp. 360–372, 1998.

[12] Tjuatja, S., A. K. Fung, and J. Bredow, "Radar imaging of buried objects," *IGARSS'98,* conference proceedings, pp. 524–526, 1998.

[13] Zhang, G. F., L. Tsang, and K. Pak, "Angular correlation function and scattering coefficient of electromagnetic waves scattered by a buried object under a two-dimensional rough surface," *J. Opt. Soc. Am. A,* vol. 15, pp. 2995-3002, 1998.

[14] A. V. der Merwe and I. J. Gupta, "A novel signal processing technique for clutter reduction in GPR measurements of small, shallow land mines," *IEEE Trans. Geosci. Remote Sensing,* vol. 38, pp. 2627–2637, 2000.

[15] Pino. M. R., L. Landesa, J. Rodriguez, F. Obelleiro, and R. J. Burkholder, "The generalized forward-backward method for analyzing the scattering from targets on ocean-like rough surfaces," *IEEE Trans. Ant. Prop.,* vol. 47, pp. 961–969, 1999.

[16] Sletten, M. A., D. B. Trizna, and J. P. Hansen, "Ultrawide-band radar observations of multipath propagation over the sea surface," *IEEE Trans. Ant. Prop.,* vol. 44, pp. 646–651, 1996.

[17] Shtager, E. A., "An estimation of sea surface influence on radar reflectivity of ships," *IEEE Trans. Ant. Prop.* vol. 47, pp. 1623–1627, 1999.

[18] Tsang, L., J. A. Kong and R. T. Shin, *Theory of Microwave Remote Sensing*, John Wiley and Sons, New York, 1985.

[19] Tsang, L., C. H. Chan, K. Pak, H. Sangani, "Monte Carlo simulations of large scale problems of random rough surface scattering and applications to grazing incidence with the BMIA/canonical grid method," *IEEE Trans. Ant. Prop.,* vol. 43, pp. 851–859, 1995.

[20] Johnson, J. T., R. T. Shin, J. A. Kong, L. Tsang, and K. Pak, "A numerical study of the composite surface model for ocean scattering," *IEEE Trans. Geosc. Remote Sens.,* vol. 36, pp. 72-83, 1998.

[21] Draine, B. T. and P. J. Flatau, "Discrete-dipole approximation for scattering calculations," *J, Opt. Soc. Am. A,* vol. 11, pp. 1491–1499, 1994.

[22] Flatau, P. J., "Improvements in the discrete-dipole approximation method of computing scattering and absorption," *Optics Letters,* vol. 22, pp. 1205–1207, 1997.

[23] Barrett, R., M. Berry, T. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods,* available by *ftp* from *netlib2.cs.utk.edu*, 1993.

[24] Johnson, J. T. and R. J. Burkholder, "Coupled canonical grid/discrete dipole approach for computing scattering from objects above or below a rough interface," accepted by *IEEE Trans. Geosc. Rem. Sens.,* 2000.

[25] Pak, K., L. Tsang, C. H. Chan, and J. T. Johnson, "Backscattering enhancement of electromagnetic waves from two dimensional perfectly conducting random rough surfaces based on Monte Carlo simulations," *J. Optical Soc. of America*, vol. 12, pp. 2491–2499, 1995.

[26] Johnson, J. T., R. T. Shin, J. A. Kong, L. Tsang, and K. Pak, "A numerical study of ocean polarimetric thermal emission," *IEEE Trans. Geosc. Remote Sens.,* vol. 37, pp. 8–20, 1999.

[27] E. Newman, *A User's Manual for The Electromagnetic Surface Patch Code: Preliminary Version ESP5.0,* ElectroScience Laboratory, The Ohio State University, unpublished report, 1997.

[28] Rice, S. O., "Reflection of electromagnetic waves from slightly rough surfaces," *Commun. Pure Appl. Math,* vol. 4, pp. 361–378, 1951.

[29] Voronovich, A. G., *Wave Scattering from Rough Surfaces,* Berlin: Springer-Verlag, 1994.

[30] McDaniel, S. T., "Acoustic and radar scattering from directional seas," *Waves in Random Media,* vol. 9, no. 4, pp. 537–549, 1999.

[31] J. L. Salvati, C. C. Chen, and J. T. Johnson, "Theoretical study of a surface clutter reduction algorithm" *IEEE Geoscience and Remote Sensing Symposium,* conference proceedings, pp. 1460–1462, 1998.

[32] *Maui High Performance Computing Center World Wide Web Site,* on the World Wide Web at *www.mhpcc.edu.*

List of Figures

Figure 1: Geometry of problem

Figure 2: Average backscattered radar cross sections versus frequency for 0 degrees incidence (a) Coherent (b) Incoherent

Figure 3: Coherent backscattered radar cross sections versus frequency for 45 degrees incidence (a) HH (b) VV

Figure 4: Incoherent backscattered radar cross sections versus frequency for 45 degrees incidence (a) HH (b) VV

Figure 5: Comparison of rough surface incoherent backscattering with small slope approximation versus frequency (a) 0 degrees incidence (b) 45 degrees incidence

Figure 6: Envelope of time domain backscattered fields for 0 degrees incidence. (a) Surface rms height 3.58 mm (b) Surface rms height 1 cm

Figure 7: Envelope of time domain backscattered fields for 45 degrees incidence: surface rms height 3.58 mm (a) HH (b) VV

Figure 8: Envelope of time domain backscattered fields for 45 degrees incidence: surface rms height 1 cm (a) HH (b) VV

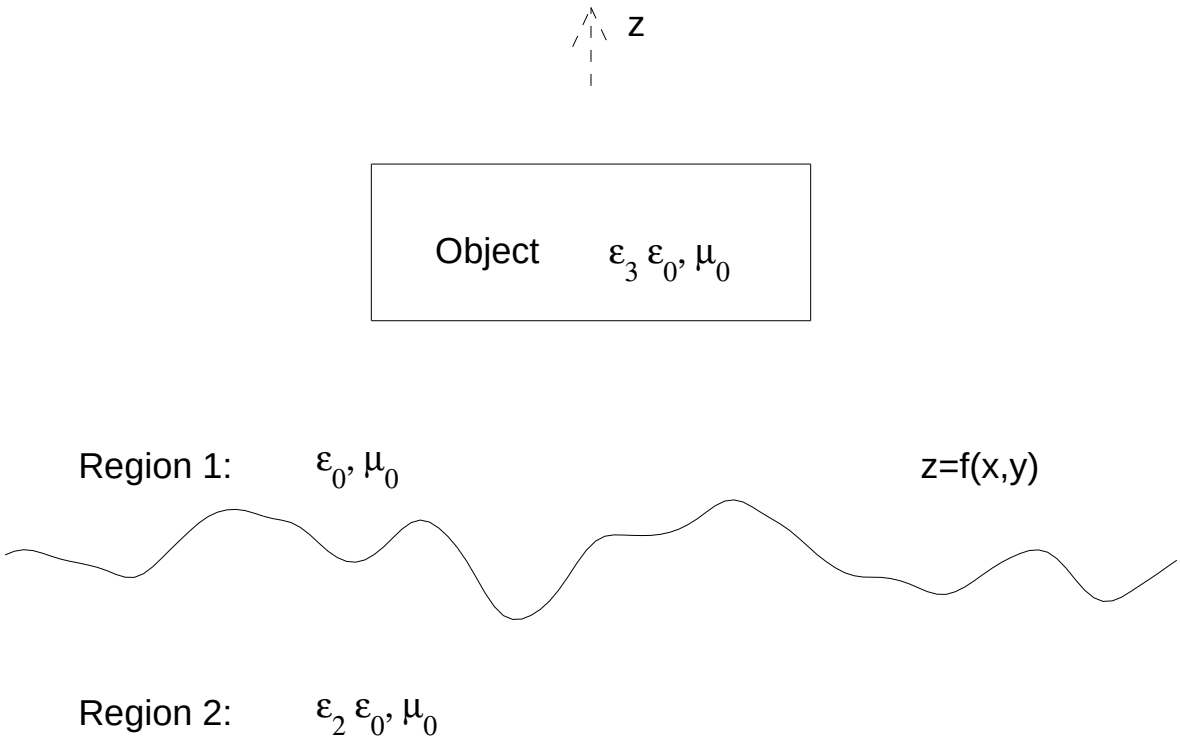


Figure 1: Geometry of problem

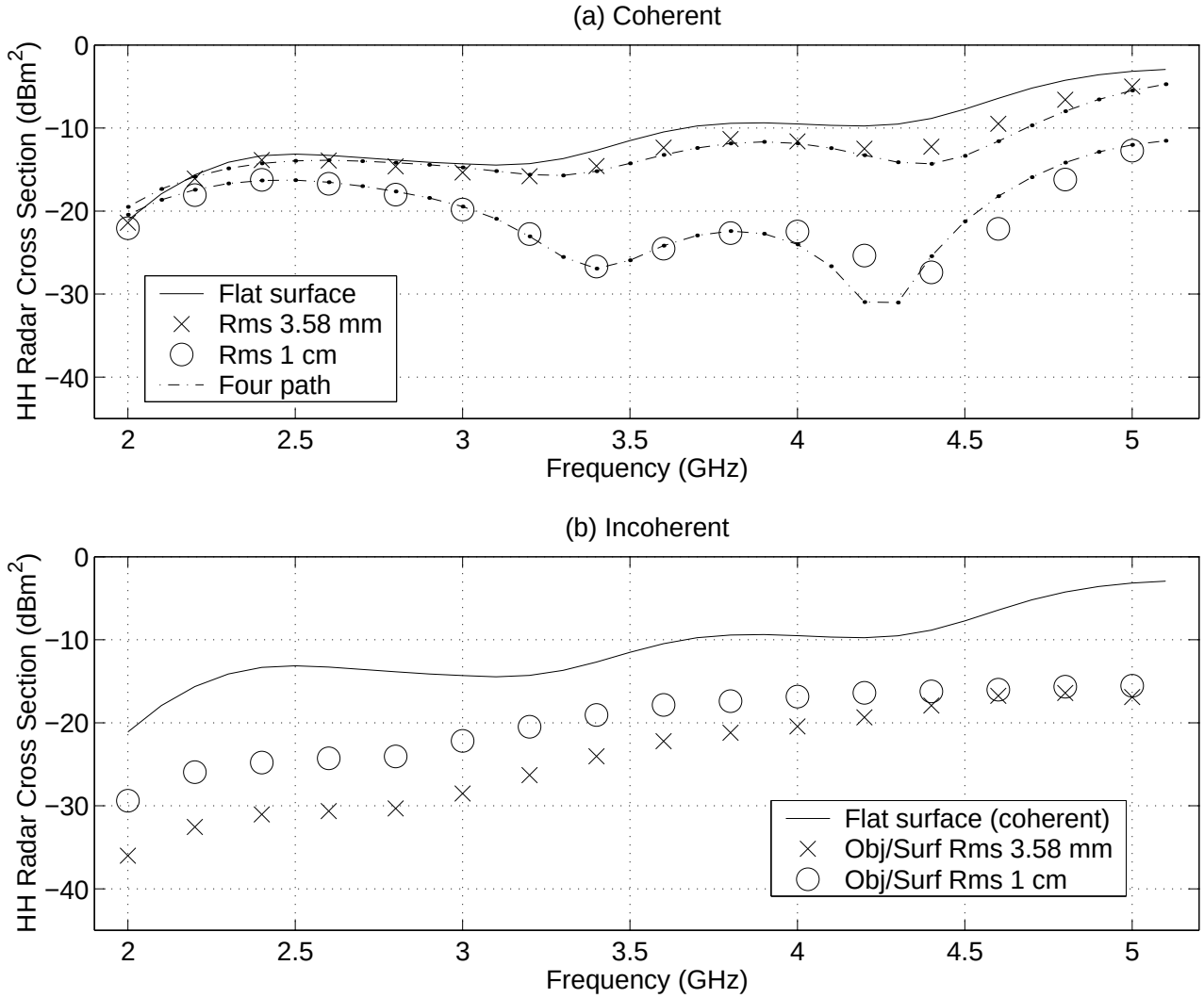


Figure 2: Average backscattered radar cross sections versus frequency for 0 degrees incidence (a) Coherent (b) Incoherent

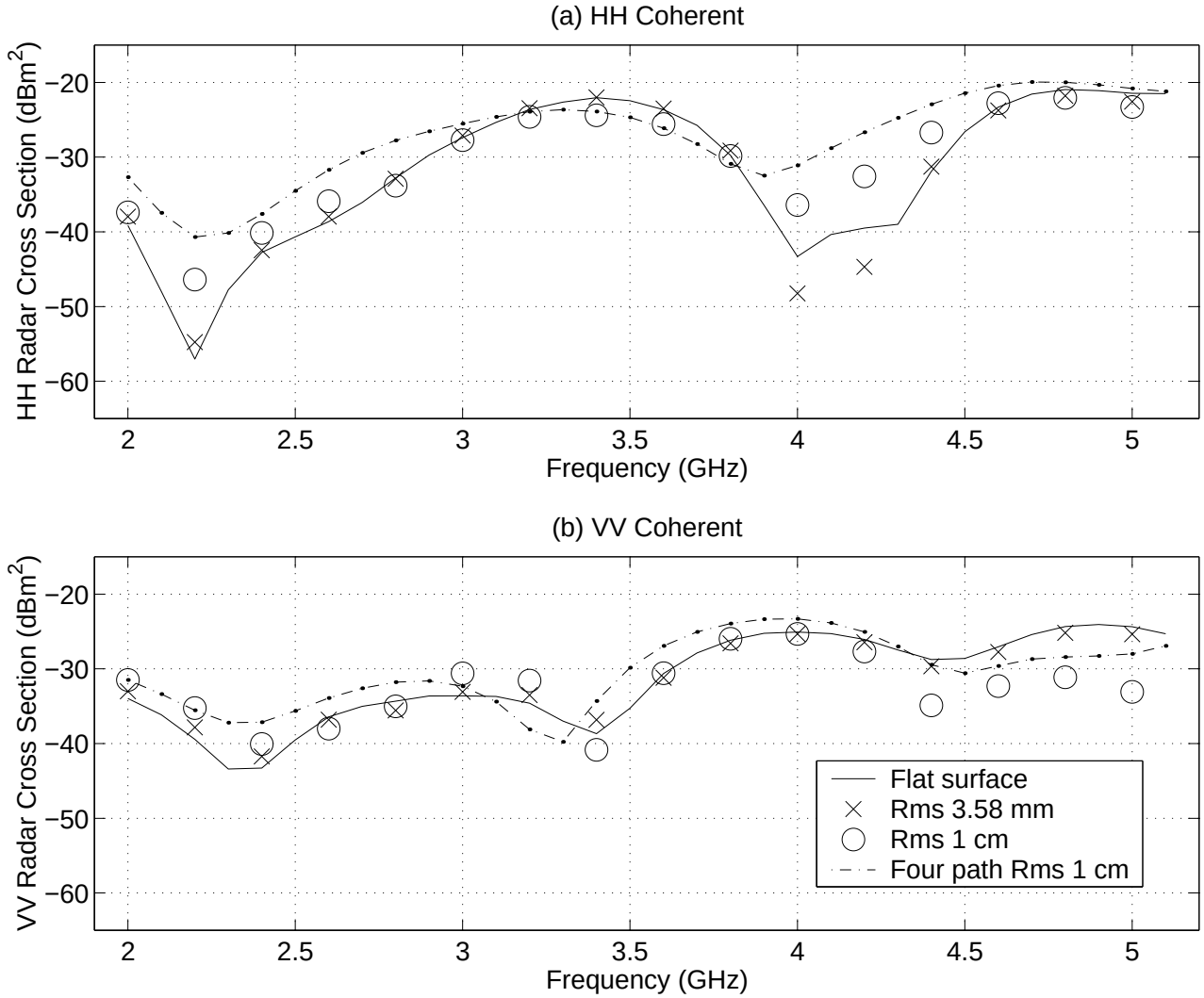


Figure 3: Coherent backscattered radar cross sections versus frequency for 45 degrees incidence (a) HH (b) VV

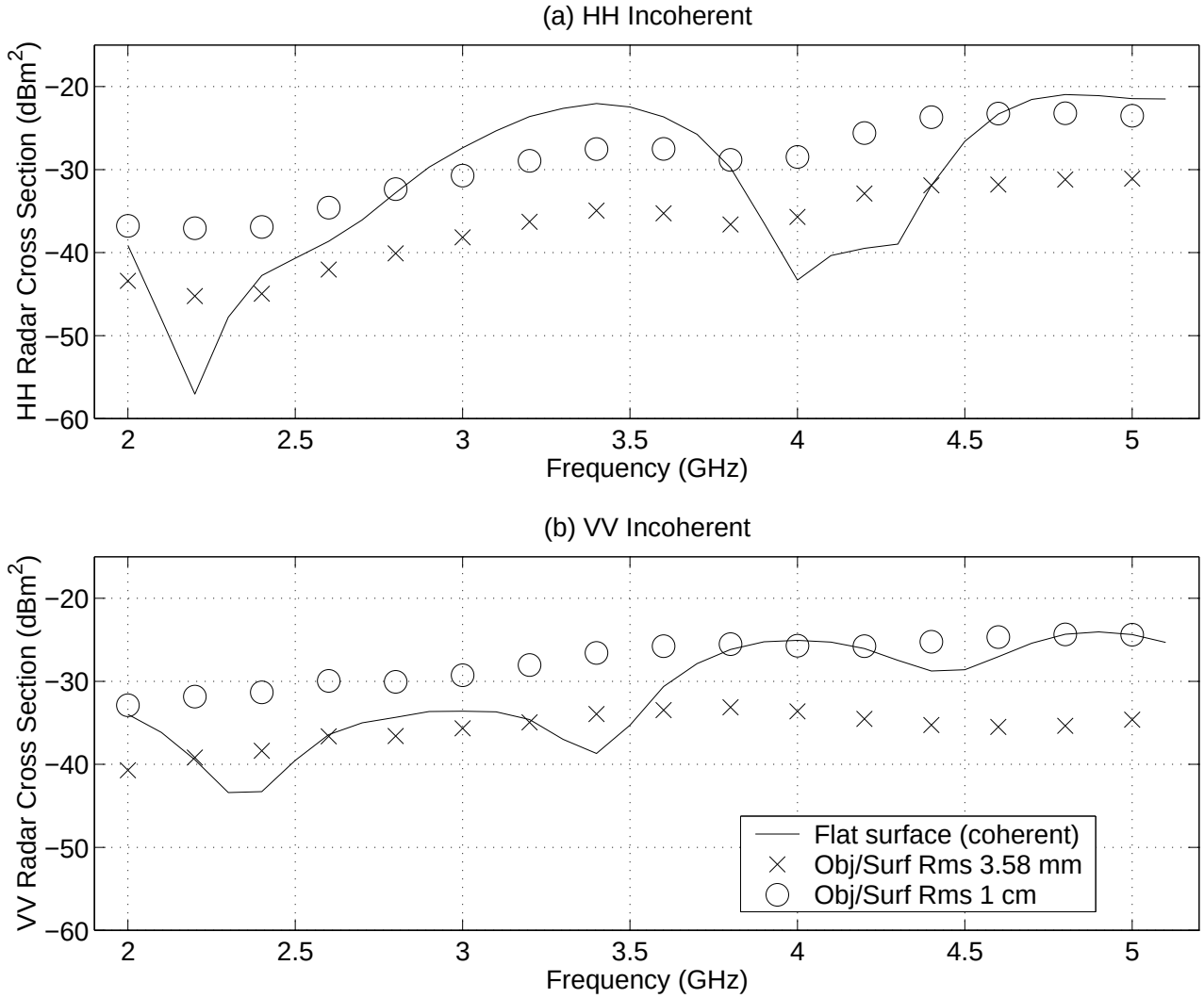


Figure 4: Incoherent backscattered radar cross sections versus frequency for 45 degrees incidence (a) HH (b) VV

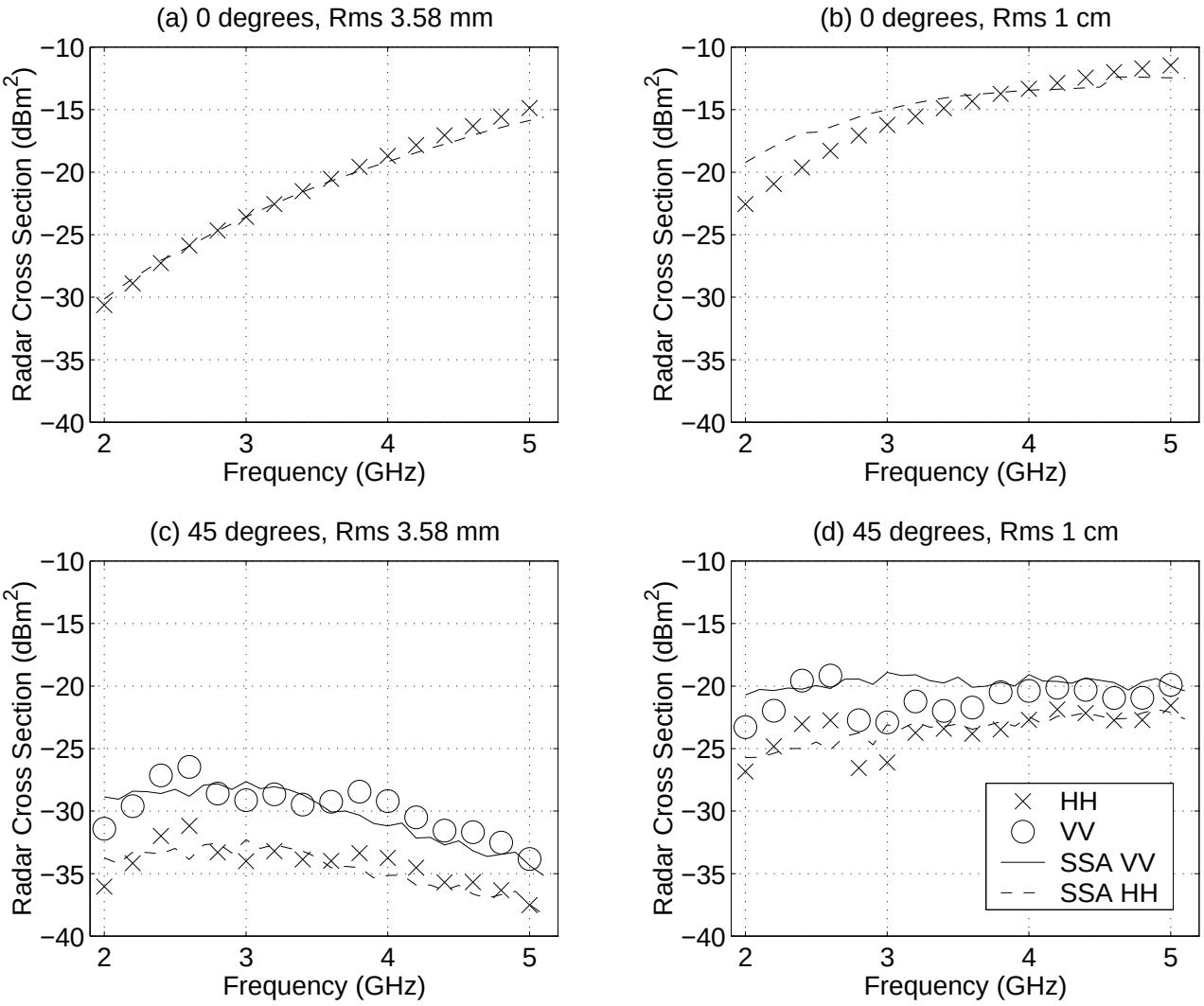


Figure 5: Comparison of rough surface incoherent backscattering with small slope approximation versus frequency (a) 0 degrees incidence (b) 45 degrees incidence

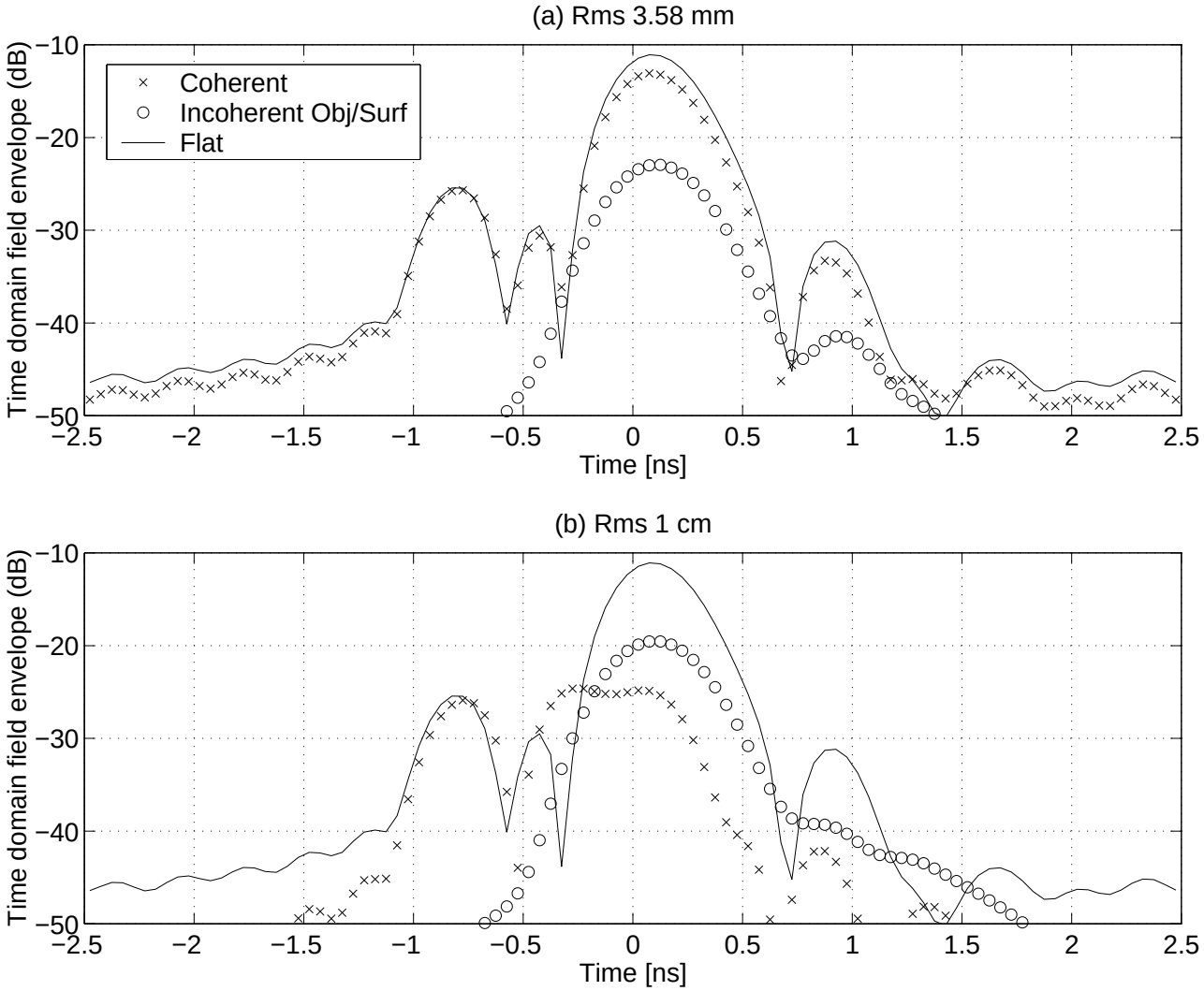


Figure 6: Envelope of time domain backscattered fields for 0 degrees incidence. (a) Surface rms height 3.58 mm (b) Surface rms height 1 cm

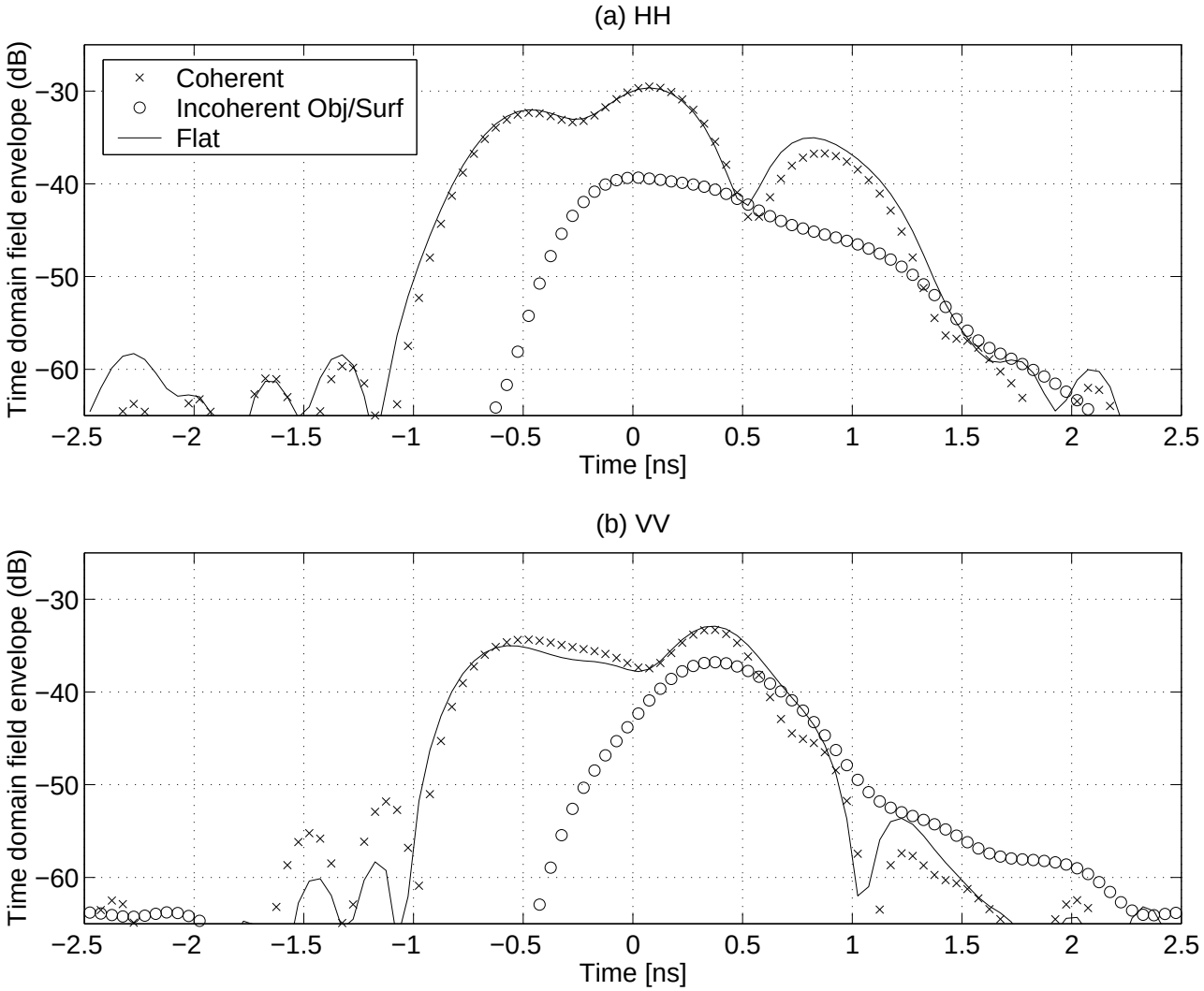


Figure 7: Envelope of time domain backscattered fields for 45 degrees incidence: surface rms height 3.58 mm (a) HH (b) VV

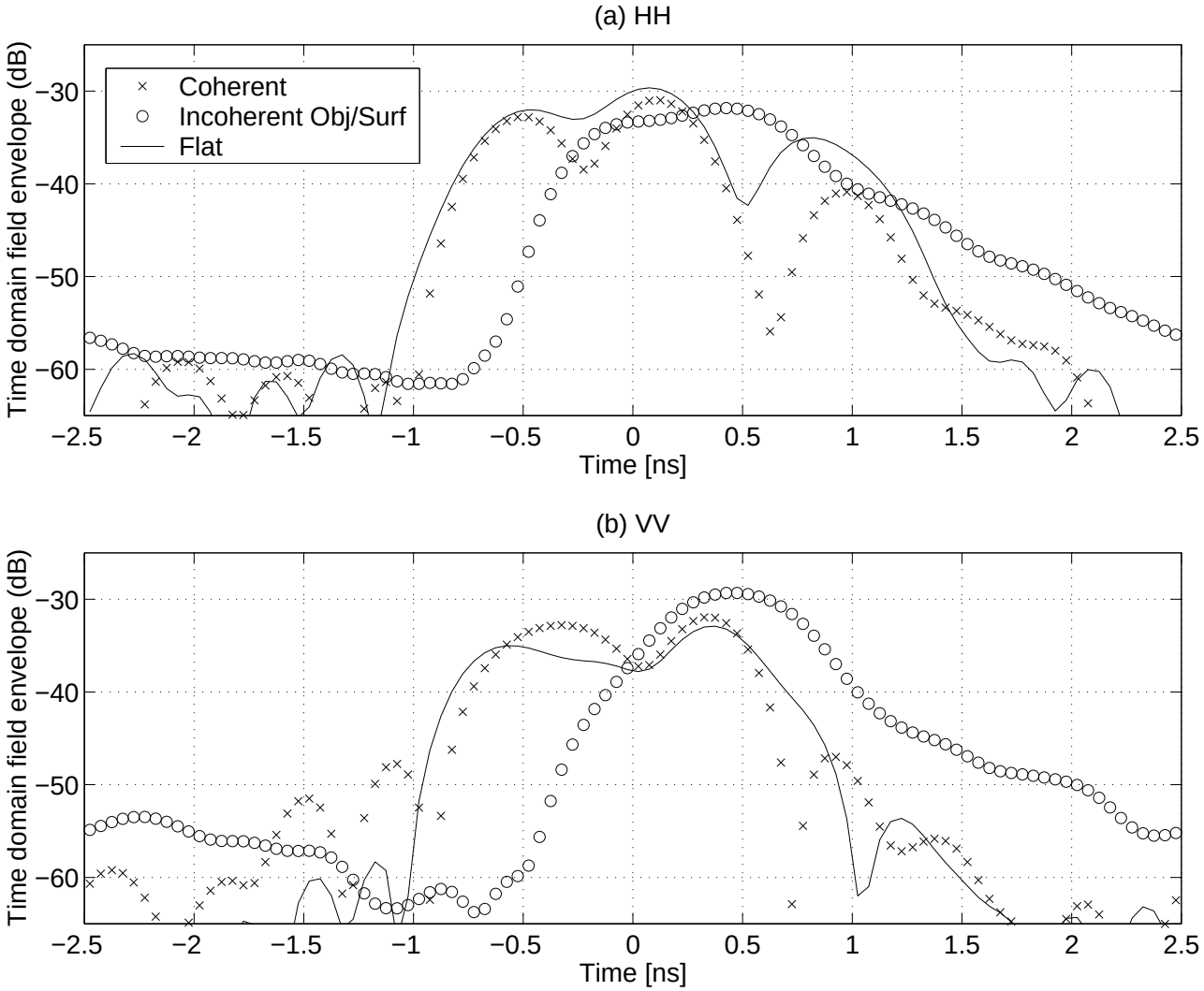


Figure 8: Envelope of time domain backscattered fields for 45 degrees incidence: surface rms height 1 cm (a) HH (b) VV

A Study of Microwave Thermal Emission from a Sub-surface Object

Baran U. Ungan and Joel T. Johnson

Department of Electrical Engineering and ElectroScience Laboratory

The Ohio State University

205 Dreese Laboratories

2015 Neil Ave

Columbus, OH 43210

(614) 292-1593

ABSTRACT

A numerical study of microwave thermal emission from a finite size buried object is described. Results show that emission from a finite size object can exhibit properties similar to emission from a layered (i.e., horizontally infinite) model of the object, including an oscillatory behavior versus frequency.

Keywords: mine detection, microwave radiometry, thermal emission, ground penetrating radar

1 Introduction

Microwave radiometers have been shown to be effective sensors for monitoring soil moisture and other geophysical data [1, 2]. Models for geophysical medium brightness temperatures often consider horizontally stratified geometries, because exact evaluation of observed brightness temperatures is possible for this case in terms of a sum involving layered medium reflection quantities [3]. For cases involving media at uniform physical temperature, the formulation reduces to "Kirchhoff's law" which relates emissivity to one minus the reflection coefficient of the layered medium.

Recent studies have begun to consider the use of microwave radiometers for detecting shallow, sub-surface objects such as anti-personnel landmines [4, 5]. Modeling studies [5] based on Kirchhoff's law and the three layer, horizontally stratified geometry shown in Figure 1, plot (a) (i.e.

objects are infinite layers) show that significant brightness temperature contrasts can be obtained in the presence of a sub-surface object if sufficient dielectric contrast is available and if soil medium attenuation is not excessive. Reference [5] further demonstrated that use of multi-frequency brightness measurements could potentially provide detection of objects even with low contrast or high attenuation due to the oscillatory behavior versus frequency observed in the presence of an object. Estimated environmental effects such as local surface temperature or soil moisture variations would not produce oscillatory frequency behavior (except in unusual circumstances [2]) so that detections would still be possible even in the presence of environmental "clutter". The possibility of directly detecting the dielectric contrast of a sub-surface object suggests an advantage for microwave passive sensors as compared to passive infrared [6] or millimeter wave sensors in which only surface temperature effects are detected.

However, previous buried object detection models have neglected effects of finite object size (Figure 1, plot (b)) since sub-surface objects were modeled as horizontally infinite layers. Numerically calculated emission results from a finite-size buried object are considered in this paper, through use of a numerical solution of the electromagnetic boundary value problem. The model applied is based on an iterative method of moments (including half-space Sommerfeld Green's functions) accelerated with the discrete dipole approximation [7]-[10]. Brightness temperatures and their variations with frequency are presented for a sample case, and amplitudes of brightness temperature oscillations in frequency are shown to vary according to the fraction of the observing antenna pattern occupied by the subsurface object. Finite size object solutions are also compared with results from the horizontally infinite layer model to determine the parameter space under which the layered model is applicable.

2 Formulation

A constant temperature medium is assumed so that brightness temperatures can be calculated using Kirchhoff's law:

$$T_B = T_s \left(1 - R\right) \tag{1}$$

where T_B is the brightness temperature measured by an observing radiometer, T_s is the physical temperature of the medium and sub-surface object (taken as 290 K in this paper), and R is the total reflectivity of the medium. The brightness temperature T_B is a function of the radiometer polar and azimuthal observation angles θ and ϕ (respectively), the frequency of observation f, the polarization of the radiometer, the dielectric properties of the medium in which the object is buried, and the geometric and dielectric properties of the sub-surface object. The total reflectivity R is equal to the total power scattered into the free space region when the subsurface object is illuminated by a plane wave field incident from angles θ and ϕ in the free space region with a polarization identical to that of the radiometer, and can therefore be determined by solving the corresponding scattering problem [3].

For an object modeled as an infinite layer, fields in the free space region consist only of the incident and reflected plane waves, with R determined by the amplitude of the reflected plane wave. However, in the presence of a finite size object, total fields in the free space region include both incident and reflected plane waves and spherical scattered wave components. The brightness temperature is then expressed as

$$T_B = T_s \left(1 - |\Gamma|^2 - \frac{P_{c0} + P_{s0}}{P_i} \right) \tag{2}$$

where Γ is the plane wave reflection coefficient of the half-space boundary, P_{s0} is the total scattered spherical wave power above the interface due to the sub-surface object, and P_{c0} is a cross-power term due to the interaction of the reflected plane wave and scattered spherical waves. Both these terms are normalized by the total power incident on the half-space medium P_i. The above equation is most conveniently derived by initially considering a "tapered" incident field [11] and evaluating the brightness temperature in the limit of a very large "spot size" on the boundary (i.e. the plane wave limit.)

If the half-space medium and object are assumed to be lossless, the same brightness temperature can also be calculated using the power radiated into the half-space medium as

$$T_B = T_s \left(|T|^2 \frac{\eta_0}{\eta_1} + \frac{P_{c1} + P_{s1}}{P_i} \right) \tag{3}$$

where T represents the plane wave transmission coefficient at the boundary, η_0 and η_1 are the characteristic impedances of free space and the half-space medium, respectively, and P_{s1} and P_{c1} are the scattered power and the cross-power terms, respectively, in the half-space medium. Comparison of the brightness temperatures obtained from equations (2) and (3) provides information on the level of power conservation obtained in the numerical solution; results will be plotted for brightnesses obtained from both methods.

For a plane wave incident field, the cross-power terms can be shown to reduce to

$$P_{c0} = \frac{2\pi}{\omega\mu}\mathrm{Im}\{\hat{e}_r \cdot \bar{F}_s\Gamma\} \tag{4}$$

and

$$P_{c1} = \frac{2\pi}{\omega\mu}\mathrm{Im}\{\hat{e}_t \cdot \bar{F}_s T\} \tag{5}$$

where $\omega = 2\pi f$ is the radian frequency of the radiometer, μ is the permeability of free space, Im denotes the imaginary part operator, $\hat{e}_r$ and $\hat{e}_t$ are unit vectors in the directions of the half-space reflected and the transmitted plane wave electric fields, respectively, and $\bar{F}_s$ is the scattered spherical wave amplitude in the reflected (for P_{c0}) or transmitted directions (for P_{c1}). The above equations are equivalent to an "optical theorem" [3] for an object in a lossless half-space.

Spherical wave scattered fields from a sub-surface object under plane wave illumination were numerically evaluated using an iterative method of moments algorithm. Because a half-space (Sommerfeld) Green's function was employed in the formulation, discretization was required only on the sub-surface object, which was sampled onto a three dimensional uniform Cartesian grid. Use of the volume equivalence principle and a point matching approach reduces the integral equations of the method of moments to a matrix equation for coupling between a set of "discrete dipoles" [7]. Computational efficiency is improved through the methods discribed in [8]-[10] so that electromagnetic coupling between all points on the object grid is computed in order $N \log N$, where N is the number of sampling points. In the brightness temperature calculations the cross-power term is computed using only specular scattered fields, whereas the scattered spherical wave power term requires a numerical integration of far-zone scattered powers over the upper or lower hemispheres.

Tests varying the number of points in this integration were performed to insure accurate total scattered power computations.

To clarify the influence of a sub-surface object, results will be presented in terms of the brightness temperature change caused by the presence of the object. Since the scattered power and cross-power terms vanish when no object is present, the change in the brightness temperatures caused by the object is

$$\Delta T_B = -T_s \left(\frac{P_{c0} + P_{s0}}{P_i} \right) = T_s \left(\frac{P_{c1} + P_{s1}}{P_i} \right) \tag{6}$$

A final issue involves evaluation of the P_i term when plane wave observations are considered. Since a plane wave would illuminate an infinite portion of the half-space in the corresponding scattering problem, P_i approaches infinity and the effect of the object becomes negligible. In fact, a radiometer observing a sub-surface object would be sensitive to only a finite sized portion of the half-space, so that P_i remains finite and approximately equal to

$$P_i \approx \frac{1}{2\eta_0} A \tag{7}$$

for nadiral observations ($\theta = 0$). Here A is the "spot-size" area of the observing antenna, which is assumed to have dimensions comparable to or larger than the electromagnetic wavelength so that the plane wave field model used in the numerical solution is valid. To make the results presented independent of the antenna spot size A, brightness temperature deviations are re-written as

$$\Delta T_B = -T_s \left(\frac{P_{c0} + P_{s0}}{\frac{1}{2\eta_0} A_{obj}} \right) \left(\frac{A_{obj}}{A} \right) \tag{8}$$

$$= -T_s \left(\frac{P_{c0} + P_{s0}}{\frac{1}{2\eta_0} A_{obj}} \right) F \tag{9}$$

where A_{obj} is the cross-sectional area of the object. The term F in the final equation thus represents a "beam fill factor" which describes the fraction of the observing antenna pattern occupied by the sub-surface object. The following plots use $F = 1$ for simplicity; it should be noted that the curves presented are to be multiplied by F for a specified antenna and object in order to predict the brightness changes observed by that antenna.

3 Results

Figure 2 illustrates ΔT_B as a function of frequency from the numerical solution (symbols) for a half-space relative permittivity of 4 and object relative permittivity of 3.15 (close to the values for dry soil and plastic, respectively.) The object is a cube with side length 2 cm centered 5 cm below the boundary, and was discretized into 16 by 16 by 16 points to insure accurate calculations over this range of frequencies. Nadiral observation $\theta = 0$ is considered, and the radiometer polarization is aligned with the cube edge. Numerical ΔT_B values plotted from both the reflected and the transmitted powers are in agreement to within approximately 2 K, confirming power conservation in the numerical solution to better than 1%. Results from the layered medium model (i.e. with the object modeled as an infinite layer) are also included as the solid curve in Figure 2, and show good agreement with finite size object brightness deviations even though the object is relatively small compared to the wavelengths considered. Only slight effects due to finite object size are observed, primarily at the higher frequencies, although again the beam fill factor for a specific antenna would reduce finite object results proportionately. Overall, results show that finite size object effects are not likely to modify the basic strategy of searching for oscillatory brightness features for detecting objects.

Figures 3 and 4 illustrate the individual contributions of the scattered spherical wave power and the cross-power terms to the ΔT_B results of Figure 2. The brightness temperature difference computed using the fields above the interface is presented in Figure 3. For this case the scattered power term is generally negligible as compared to the cross-power term. Unlike the case for the fields above the interface, the scattered power and the cross-power term both show important contributions with fields below the interface, as shown in Figure 4. However the two power terms still add up to give an accurate ΔT_B. This increase in the scattered power below the ground is possibly due to the fact that the object size is larger as compared to the wavelength in the medium.

Brightness temperature results for the slightly tilted object shown in Figure 5 are presented in Figure 6. Because the tilted object is represented on the original (untilted) grid, discretization error in the object boundary can potentially be significant. However, several tests of the results

were performed to insure that the results presented are accurate to within 2 K. Results show that only moderate changes in brightness temperatures are observed as the object is tilted to an angle of $\alpha = 20$ degrees.

4 Conclusions

Effects of finite object size on the detection of buried objects using microwave radiometry have been investigated. Thermal emission from buried objects has been shown to display an oscillatory behavior versus frequency, providing a means for detecting objects. While finite size object effects do affect the specific brightness temperature of the buried object, the basic strategy of searching for oscillatory features versus frequency for object detection appears to remain valid.

5 Acknowledgements

This work was sponsored by NSF project ECS-9701678 and by a grant from Duke University as part of the Office of the Secretary of Defense MURI on Humanitarian Demining.

References

[1] Ulaby, F. T., R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive,* Artech House: Norwood, MA, 1986.

[2] Schmugge, T. J., T. J. Jackson, P. E. O'Neill, and M. B. Parlange, "Observations of coherent emissions from soils", *Radio Science,* vol. 33, pp. 267–272, 1998.

[3] Tsang, L., J. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing,* Wiley: New York, 1985.

[4] De Amici, G., B. Hauss, and L. Yujiri, "Detection of landmines via a passive microwave radiometer," *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proceedings of SPIE*, vol. 3710, pp. 716–724, 1999.

[5] Johnson, J. T., "Theoretical study of microwave radiometry for buried object detection", *Detection and Remediation Technologies for Mines and Minelike Targets V, Proceedings of SPIE*, vol. 4038, pp. 286–297, 2000.

[6] Sendur, I. K. and B. A. Baertlein, "Techniques for improving buried mine detection in thermal IR imagery," *Detection and Remediation Technologies for Mines and Mine-Like Targets IV, Proceedings of SPIE,* vol. 3710, pp. 1272–1283, 1999.

[7] Draine, B. T. and P. J. Flatau, "Discrete-dipole approximation for scattering calculations", *J. Opt. Soc. Am. A,* vol. 11, pp. 1491–1499, 1994.

[8] Flatau, P. J., "Improvements in the discrete-dipole approximation method of computing scattering and absorption", *Optics Letters,* vol. 22, pp. 1205–1207, 1997.

[9] Cui, T. J. and W. C. Chew, "Fast algorithm for electromagnetic scattering by buried 3-D objects of large size", *IEEE Trans. Geosc. Rem. Sens.*, vol. 37, pp. 2597–2608, 1999.

[10] Johnson, J. T. and R. J. Burkholder, "Coupled canonical grid / discrete dipole approach for computing scattering from objects above or below a rough interface", *IEEE Trans. Geosc. Rem. Sens.*, vol. 39, pp. 1214–1220, 2001.

[11] Braunisch, H., Y. Zhang, C. O. Ao, S. E. Shih, Y. E. Yang, K. H. Ding, J. A. Kong, and L. Tsang, "Tapered wave with dominant polarization state for all angles of incidence", *IEEE Trans. Antennas Prop.*, vol. 48, pp. 1086–1095, 2000.

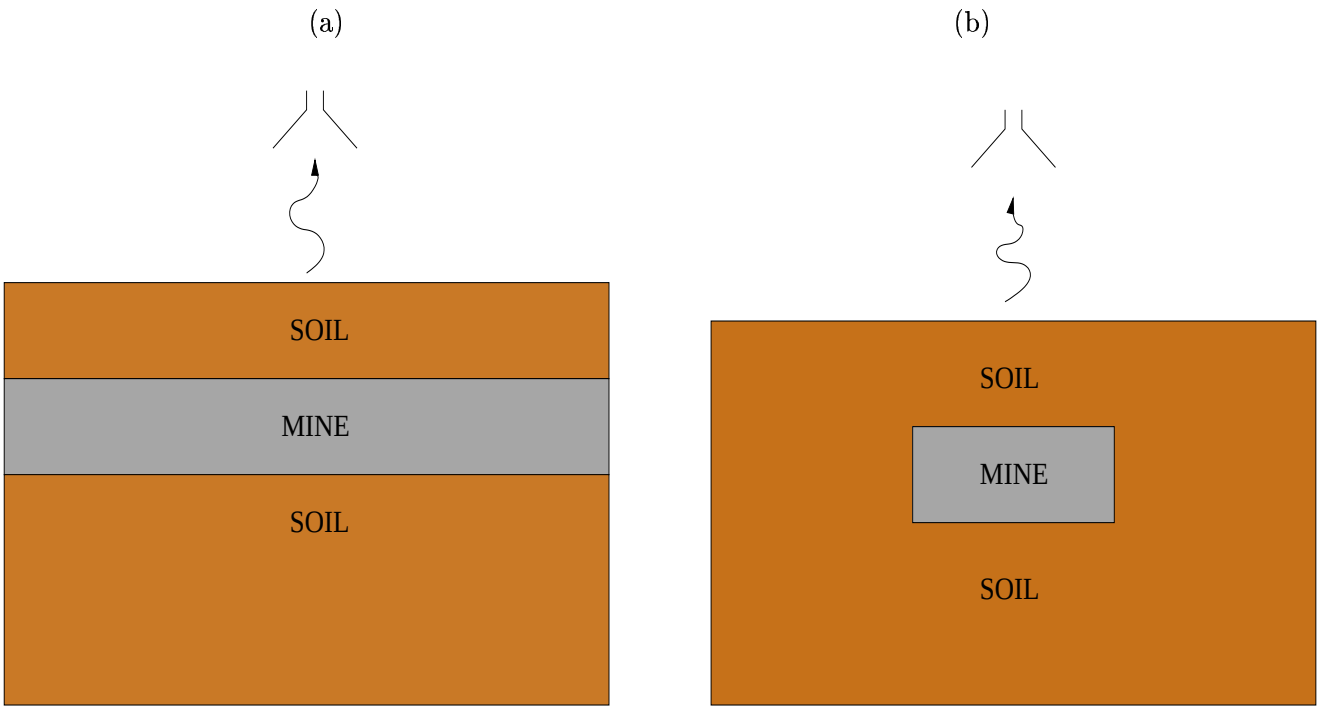


Figure 1: (a) Layered medium model for a sub-surface object (b) A finite size sub-surface object

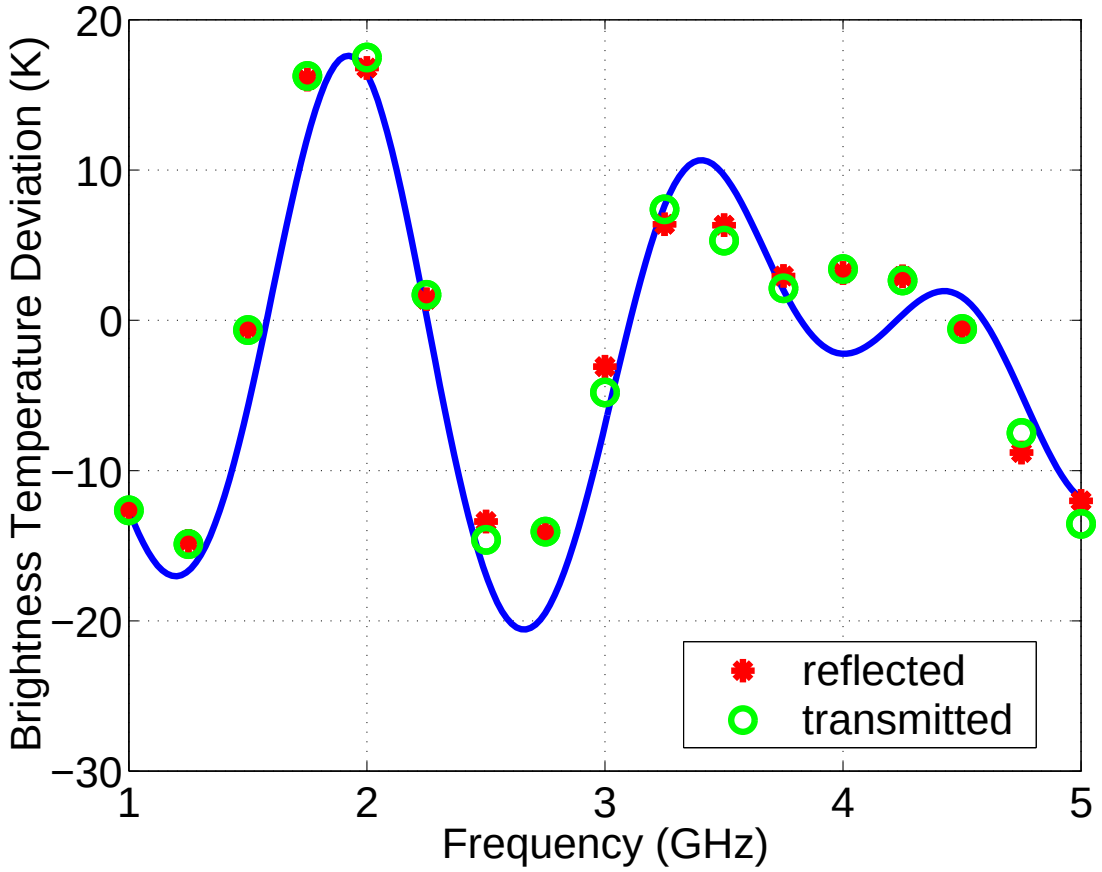


Figure 2: Brightness temperature deviation vs. frequency, numerical solution (discrete points) compared with layered model (smooth curve)

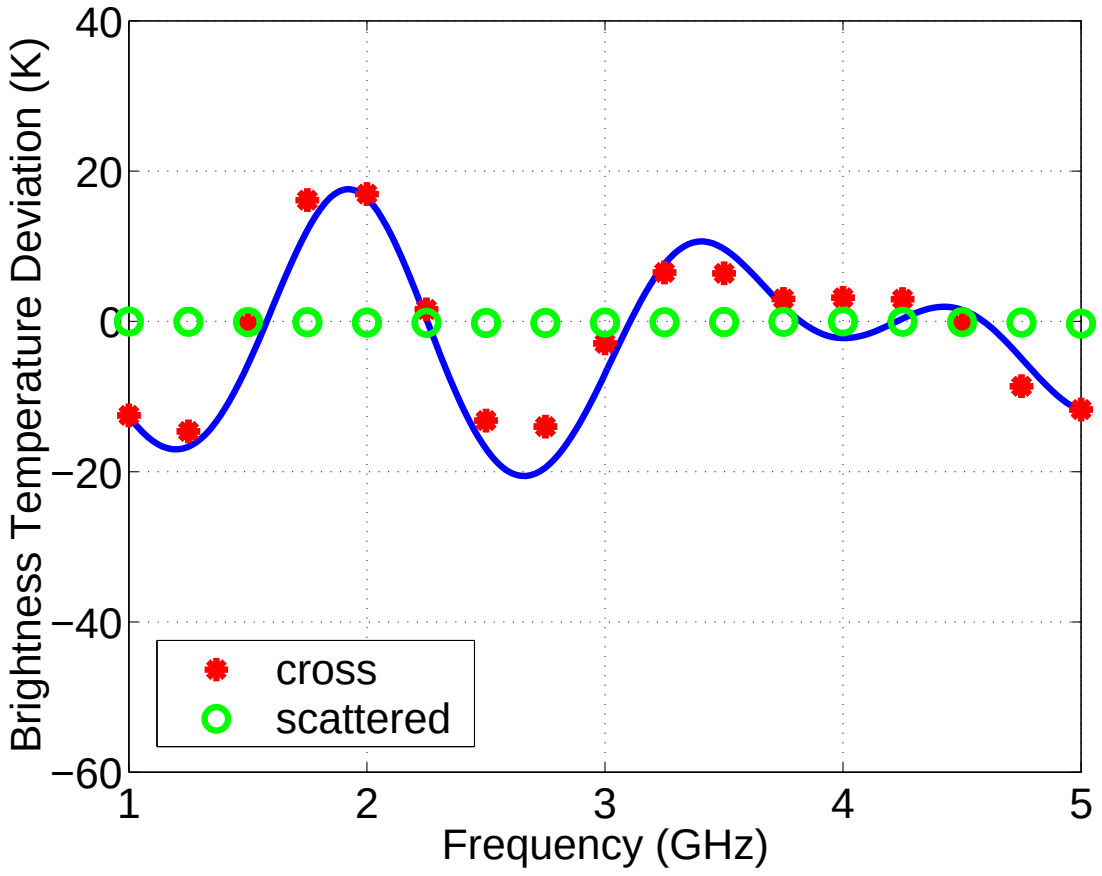


Figure 3: Contributions of scattered and cross- power terms to brightness temperature deviations using the fields above the ground. Numerical solution (discrete points) compared with total brightness deviation from layered model (smooth curve).

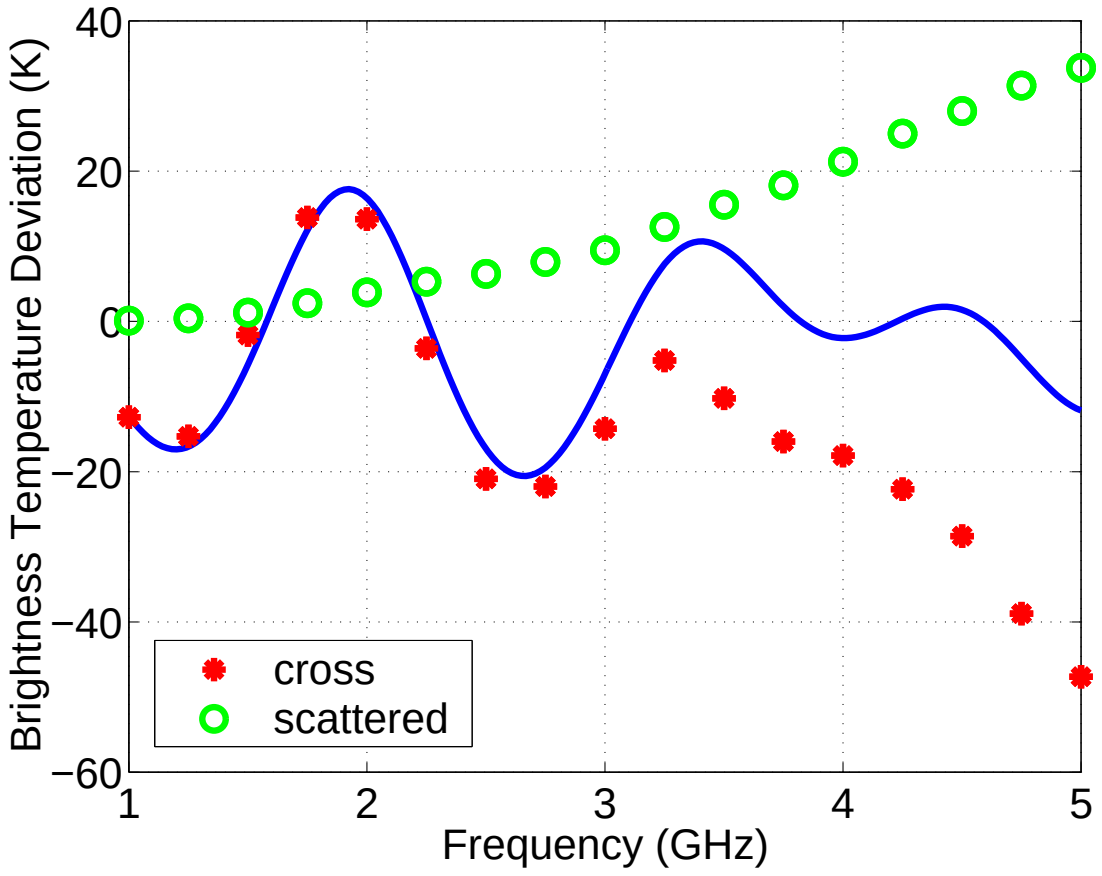


Figure 4: Contributions of scattered and cross- power terms to brightness temperature deviations using the fields below the ground. Numerical solution (discrete points) compared with total brightness deviation from layered model (smooth curve).

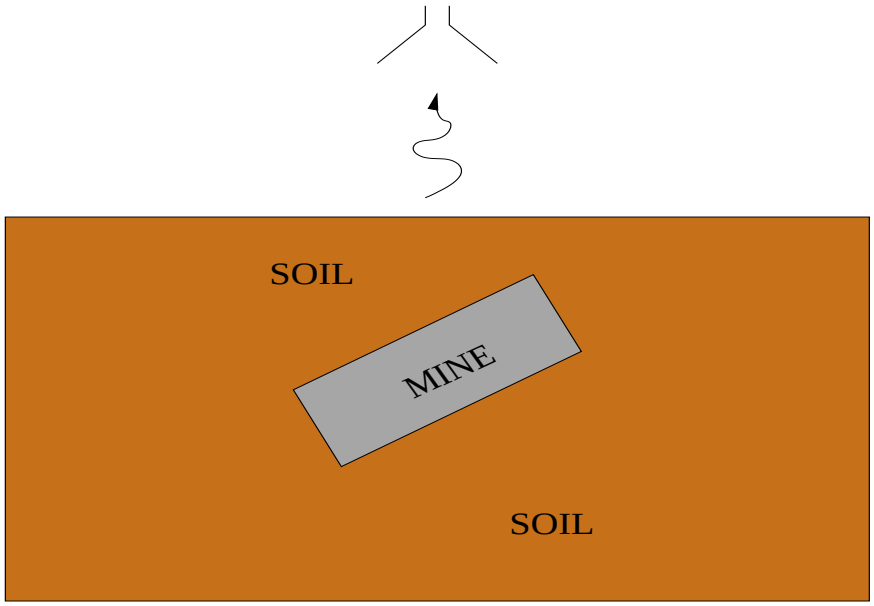


Figure 5: Geometry of tilted sub-surface object

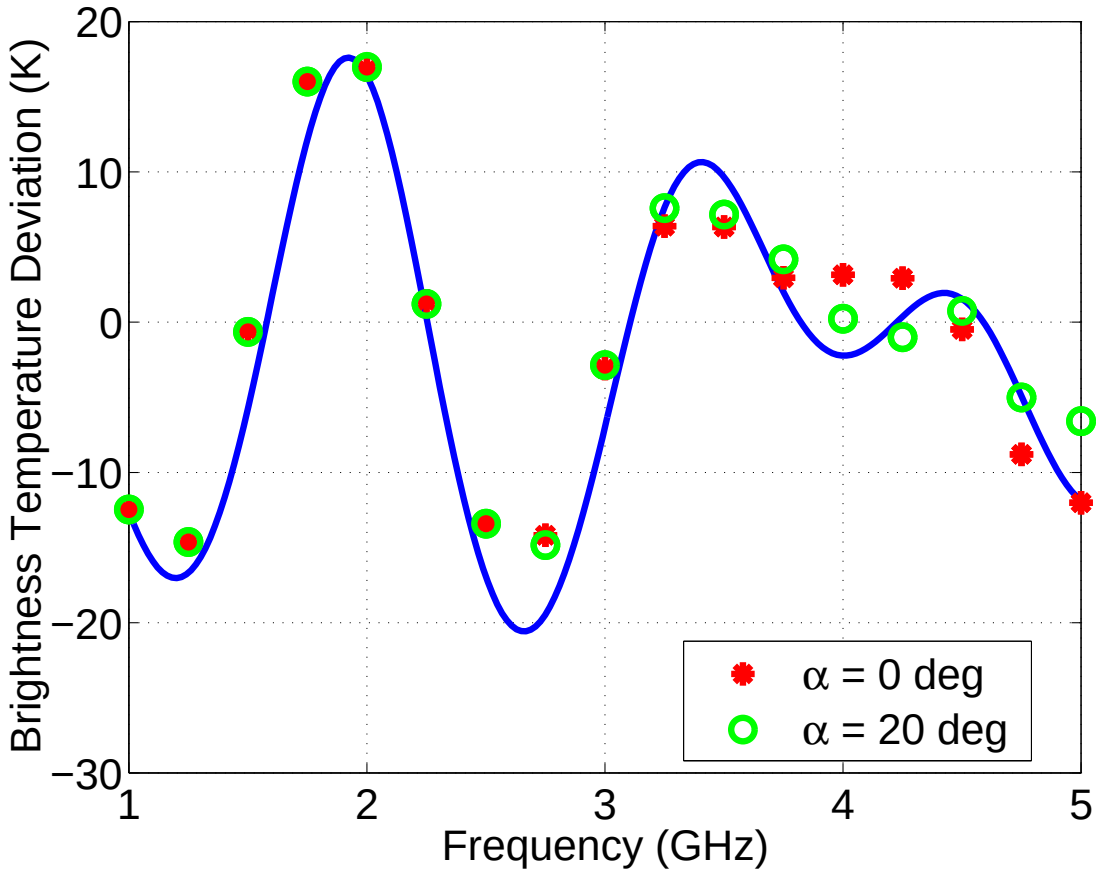


Figure 6: Comparison of brightness temperatures for tilted and untilted objects with layered medium model

Anal. Chem. **1998,** *70,* 1440−1443

Correspondence

Differential Detection of Enantiomeric Gaseous Analytes Using Carbon Black−Chiral Polymer Composite, Chemically Sensitive Resistors

Erik J. Severin, Robert D. Sanner,† Brett J. Doleman, and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Carbon black−chiral polymer composites were used to provide diagnostic differential resistance responses in the presence of enantiomers of chiral gaseous analytes. Vapors of (+)-2-butanol and (−)-2-butanol, (+)-α-pinene and (−)-α-pinene, (+)-epichlorohydrin and (−)-epichlorohydrin, and methyl (+)-2-chloropropionate and methyl (−)-2-chloropropionate were generated and passed over a chemically sensitive carbon black−poly((*R*)-3-hydroxybutyrate-*co*-(*R*)-3-hydroxyvalerate) (77% butyrate) composite resistor. Each enantiomer of a pair produced a distinct relative differential resistance change on the chiral detector, whereas both enantiomers of a set produced identical signals on achiral carbon black−poly(ethylene-*co*-vinyl acetate) (82% ethylene) detectors.

We have previously reported the use of carbon black−polymer composites for array-based vapor sensing applications.[1] In such an array, no individual detector responds solely to a specific molecule, but the collective response of the entire array of detectors yields a unique fingerprint for the vapor of interest. Such arrays are often referred to as "electronic noses" and are not designed in advance to perform a specific task but are, instead, developed to classify, identify, and quantify vapors on the basis of pattern recognition algorithms.[2−7] This approach to vapor sensing takes advantage of the collective output of an array of broadly responsive detectors. In the polymer composite array configuration, the signal transduction is extremely simple: swelling of the polymeric phase of the composite, in the presence of a vapor, leads to an increase in the electrical resistance of the composite, which is monitored using simple electronics.

An ideal detector array would produce a unique signature for every molecule to which it was exposed. To make progress toward such a system, it is necessary to include detectors that probe important, but possibly subtle, molecular parameters such as chirality. None of the polymer-based conducting composite detectors reported to date are chiral, so enantiomers would not be differentiable on arrays of such detectors. We demonstrate herein the use of chiral polymers in carbon black−polymer composites to achieve the differential detection of enantiomers. The materials described represent additional detector elements that would be part of a larger detector array, thus broadening the discrimination ability of such arrays toward enantiomeric pairs of analytes.[1]

EXPERIMENTAL SECTION

The carbon black used in the composites was Black Pearls 2000, a furnace black material donated by Cabot Co. (Billerica, MA). The polymer used in the chiral composites was poly((*R*)-3-hydroxybutyrate-*co*-(*R*)-3-hydroxyvalerate) (77% butyrate) and was obtained from the Goodfellow Corp. (Berwyn, PA). The achiral polymer used for control experiments was poly(ethylene-*co*-vinyl acetate) (82% ethylene) (Polysciences Inc., Warrington, PA). The enantiomeric pairs examined were (+)-2-butanol and (−)-2-butanol (Aldrich, Milwaukee, WI), (+)-α-pinene and (−)-α-pinene (Fluka, Ronkonkoma, NY), (+)-epichlorohydrin and (−)-epichlorohydrin (Aldrich), and methyl (+)-2-chloropropionate and methyl (−)-2-chloropropionate (Aldrich).

An apparatus that provided known partial pressures of the vapors was constructed of general laboratory glassware. This consisted of a bubbler made from small, 12-mL centrifuge tubes with conical bottoms that were filled to a depth of 3 cm (∼2 mL). Into the 1.5-cm-diameter mouth of the tube was affixed a two-hole rubber stopper. In each hole was a 5-mm-o.d. glass tube, one of which extended to the bottom of the bubbler and served as the gas inlet, and the other of which extended past the stopper by only a few millimeters and served as the gas exit. The carrier gas was nitrogen, obtained from a commercial gas supply tank. The measurements were performed at room temperature, which was 23 ± 1 °C.

The carrier gas was introduced through the glass tube which extended to the bottom of the bubbler apparatus and was bubbled through the solvent. The resulting vapor was carried out of the bubbler, diluted by blending with a controlled background flow of pure carrier gas, and then introduced into a sensing chamber.

† Current address: Mail Code L-325, Lawrence Livermore National Laboratory, Livermore, CA 94550.

(1) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298.
(2) Zaromb, S.; Stetter, J. R. *Sens. Actuators* **1984**, *6*, 225.
(3) Lundstrom, I.; Erlandsson, R.; Frykman, U.; Hedborg, E.; Spetz, A.; Sundgren, H.; Welin, S.; Winquist, F. *Nature* **1991**, *352*, 47.
(4) Shurmer, H. V.; Gardner, J. W. *Sens. Actuators B* **1992**, *8*, 1.
(5) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators B* **1994**, *18*, 211.
(6) Gardner, J. W.; Hines, E. L.; Tang, H. C. *Sens. Actuators B* **1992**, *9*, 9.
(7) Nakamoto, T.; Fukuda, A.; Moriizumi, T. *Sens. Actuators B* **1993**, *10*, 85.

S0003-2700(97)00757-9 CCC: $15.00 © 1998 American Chemical Society
Published on Web 02/20/1998

This chamber consisted of a glass tube (22 cm long with a 2.6-cm inner diameter), to which inlet and outlet sidearms had been attached. The detectors were introduced into the chamber through a 24/40 standard taper ground-glass opening attached at one end of the chamber. The chamber was then sealed with a ground-glass stopper through which electrical lead wires for the detectors had been sealed. The gas flow rates were controlled using needle valves and stopcocks.

To prepare the detector substrates, two parallel bands of gold, 50−100 nm thick and separated by 5 mm, were deposited onto conventional 7.5-cm × 2.5-cm glass slides. The slides were then cut into strips to produce 0.7-cm × 2.5-cm pieces of glass, with each strip of glass having one pair of Au leads.

The detectors were made from a solution of the polymer into which carbon black had been suspended. Here 125 mg of the polymer was dissolved in 10 mL of tetrahydrofuran, and carbon black (42 mg) was added to produce a composition of 75% polymer and 25% carbon black by weight of solids. A single solution that contained the polymer and the carbon black was used to prepare all the detectors of a given composition that were used in this work. Detectors used to analyze pinene vapors were fabricated slightly differently, having films made from a suspension with a carbon black loading of 30 wt % of solids. In both cases, an aliquot of the suspension was spin-coated, at 1000 rpm, onto a glass substrate using a Headway (Garland, TX) spin-coater, and the resulting film was allowed to dry in air. One coating of the suspension was applied to each substrate, yielding a film thickness of ∼1 μm, as determined by atomic force microscopy, except for pinene detectors, for which six coats of suspension were applied, producing films ∼6 μm in thickness. The detectors with extra coatings gave higher signal-to-noise ratios when used for pinene.

The dc resistance of each detector was determined as a function of time using a simple two-point resistance configuration. Contacts were made to the gold lines by pressure-contacting electrical leads using flat-jawed alligator clips. Resistance data were acquired using a Hydra 2620A data acquisition unit (John Fluke Mfg. Co., Everett, WA) which was interfaced to a personal computer. All of the films had resistance values below the 10 MΩ limit of the Hydra 2620A.

To initiate an experiment, five copies of a given detector type were placed into the glass chamber, and a background flow of nitrogen was introduced until the resistance of the detectors stabilized. Solvent vapor streams were then passed over the detectors. The background and analyte flow rates were monitored using two flow meters (Gilmont Instruments, Inc.), which had ranges of 0.2−15.0 and 0.0015−0.310 L min^{-1}, respectively. In a typical experiment, resistance data on the detectors were collected for 150 s, with the background gas flowing (typically about 1−2 L min^{-1}) to serve as a baseline. This was followed by a 150-s data collection while the detectors were exposed to the diluted analyte vapor stream (typically about 200−300 mL min^{-1}). The detectors were then given 200−300 s to recover, during which time pure background gas was passed through the chamber. The exposure times varied somewhat, but steady-state values of resistance change were always reached for any given exposure time. Resistances for all detectors in a given trial were monitored contemporaneously through the use of the multiplexing capabilities of the Hydra multimeter. Results were obtained by running

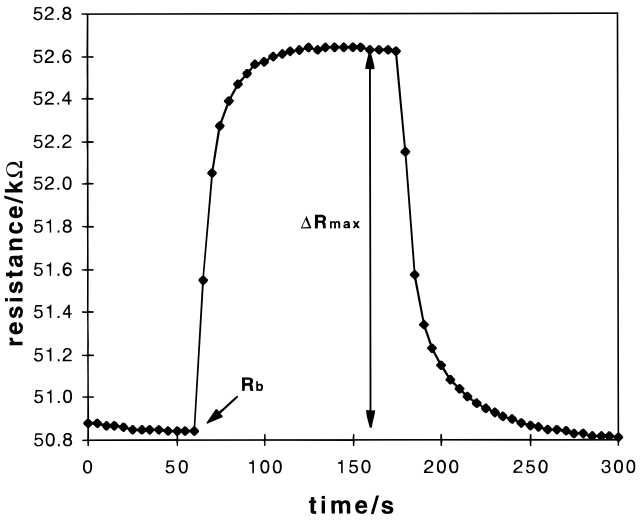


Figure 1. Typical chiral detector response upon exposure to 5 ppth of (+)-2-butanol.

two trials (except for epichlorohydrin for which three trials were run) of five exposures each, with the trials performed on different days. Each analyte was exposed to five copies of the detector simultaneously, and the results were averaged to obtain the reported data set. In the case of epichlorohydrin, a third trial was run using 300-s exposure and recovery times to investigate whether longer time cycles produced different response signals, but the data were essentially identical to those obtained using the shorter detector cycle times described above. The exposures were made at the following concentrations: epichlorohydrin, 3 parts per thousand (ppth); α-pinene, 1 ppth; methyl-2-chloropropionate, 2 ppth; and 2-butanol, 4 ppth. Concentrations were calculated by diluting gas streams to known volumes, and the analyte concentrations were determined from measurements of the flow rate of the gas and the rate of mass loss of the solvent.[1]

The achiral control detectors were made from benzene solutions of poly(ethylene-*co*-vinyl acetate) (82% ethylene) into which carbon black had been suspended. Glass slides with gold contacts were coated by dipping the slide into the suspension. Three coatings were applied to each slide. The polymer concentration was 10 mg mL^{-1}, and the carbon black loading was 30 wt % of solids. Results were obtained by running one trial of four (for epichlorohydrin and methyl 2-chloropropionate) or five (for 2-butanol and α-pinene) exposures. Five detectors were exposed simultaneously to each analyte, and the results were averaged to obtain the reported data set. The control exposures were made at the following concentrations: epichlorohydrin, 4 ppth; α-pinene, 1 ppth; methyl 2-chloropropionate, 3 ppth; and 2-butanol, 4 ppth.

RESULTS AND DISCUSSION

Figure 1 represents a typical response of a chiral detector to 2-butanol. All the detectors displayed an increase in resistance upon exposure to the vapor and returned to their baseline values after the vapor was removed. In all of the experiments performed, the change in resistance was quite rapid, taking less than 20 s to reach 75% of the final resistance value (taken after 150 s). The responses were analyzed by calculating the maximum differential response value, ΔR_{max}, observed during the exposure period, dividing it by the baseline value of the resistance, R_b, (taken as

Table 1. Representative Data[a] for the (+) and (−) Enantiomers of 2-Butanol

detector no.	$Q_{(-)}$[b]	$Q_{(+)}$[b]	$\Delta Q_{(+/-)}$	Q_{rel}, %	α	$\Delta_{1,2}(\Delta G)$, kJ/mol
			Trial 1			
1	3.41 (0.01)	3.60 (0.03)	0.19	5.5	1.06	−0.133
2	3.29 (0.02)	3.52 (0.03)	0.23	7.0	1.07	−0.167
3	2.89 (0.02)	3.10 (0.02)	0.21	7.3	1.07	−0.174
4	2.96 (0.02)	3.17 (0.03)	0.21	7.0	1.07	−0.166
5	2.75 (0.03)	2.97 (0.04)	0.22	8.0	1.08	−0.189
average			0.21 (0.02)	7.0 (0.9)	1.07 (0.01)	−0.17 (0.02)
			Trial 2			
1	2.71 (0.04)	2.97 (0.04)	0.26	9.5	1.10	−0.224
2	2.59 (0.02)	2.87 (0.04)	0.28	11	1.11	−0.253
3	2.55 (0.03)	2.79 (0.04)	0.24	9.5	1.10	−0.224
4	2.62 (0.03)	2.86 (0.04)	0.24	9.2	1.09	−0.216
5	2.49 (0.03)	2.73 (0.04)	0.24	9.6	1.10	−0.225
average			0.25 (0.02)	9.7 (0.6)	1.10 (0.01)	−0.23 (0.01)

control detector no.	$Q_{(-)}$[b]	$Q_{(+)}$[b]	$\Delta Q_{(+/-)}$
1	2.71 (0.02)	2.73 (0.06)	0.02
2	2.50 (0.01)	2.50 (0.01)	0.00
3	2.58 (0.03)	2.56 (0.02)	−0.02
4	3.02 (0.03)	3.04 (0.01)	0.02
average			0.00 (0.02)

[a] The numbers in parentheses are estimated standard deviations. [b] $Q_{(+)}$ and $Q_{(-)}$ correspond to the percent relative differential resistance change averaged over five exposures for the (+) and (−) enantiomers, respectively. α is defined as K_1/K_2, where K_1 is defined as the larger partition coefficient.

Table 2. Differences in Percent Response,[a] $\Delta Q_{(+/-)}$, Observed for Chiral Detectors during Exposure to the (+) and (−) Enantiomers of Four Test Vapors

analyte	trial	$\Delta Q_{(+/-)}$	Q_{rel}, %	avg α	avg $\Delta_{1,2}(\Delta G)$, kJ/mol
2-butanol	1	0.21 (0.02)	7.0 (0.9)	1.07 (0.01)	−0.17 (0.02)
	2	0.25 (0.02)	9.8 (0.6)	1.10 (0.01)	−0.23 (0.01)
	control	0.02 (0.01)			
α-pinene	1	0.15 (0.01)	7.6 (1.5)	1.08 (0.01)	−0.18 (0.03)
	2	0.15 (0.04)	8.0 (1.1)	1.08 (0.01)	−0.19 (0.02)
	control	0.02 (0.01)			
epichlorohydrin	1	0.22 (0.01)	6.9 (0.6)	1.07 (0.01)	−0.16 (0.02)
	2	0.19 (0.03)	5.6 (0.6)	1.06 (0.01)	−0.13 (0.01)
	3	0.24 (0.03)	7.1 (1.2)	1.07 (0.01)	−0.17 (0.03)
	control	0.03 (0.03)			
methyl 2-chloro-propionate	1	0.26 (0.03)	9.1 (1.2)	1.09 (0.01)	−0.21 (0.03)
	2	0.26 (0.01)	8.9 (0.6)	1.09 (0.01)	−0.21 (0.01)
	control	0.02 (0.01)			

[a] The values tabulated are the percent response values that were averaged over five nominally identical detectors, each of which had been exposed four (for epichlorohydrin and methyl 2-chloropropionate) or five (for 2-butanol and α-pinene) times to the analyte vapor. The other quantities were calculated using the formulas and definitions given in the text.

the resistance value just before the exposure began), and the result expressed as a percent change in resistance, Q:

$$Q = (\Delta R_{max}/R_b) \times 100 \quad (1)$$

$Q_{(+)}$ and $Q_{(-)}$ correspond to the percent relative differential resistance response for the (+) and (−) enantiomers, respectively. A representative data set, for the enantiomers of 2-butanol, is presented in Table 1, while Table 2 summarizes the data for all of the enantiomers studied in this work.

As reported in Tables 1 and 2, statistically significant differences in detector response were observed when enantiomers of a given analyte were exposed to the chiral carbon black-polymer composite detectors. In contrast, no statistically significant difference in response was observed when achiral detectors were exposed to these same pairs of enantiomers. The magnitude of the differentiation between enantiomers can be quantified by considering the relative difference in response, Q_{rel}, between the enantiomers:

$$Q_{rel} = [\Delta Q_{(+/-)}/Q_{(-)}] \quad (2)$$

where $\Delta Q_{(+/-)}$ is the difference in percent response of a detector when exposed to the (+) and (−) enantiomers of an analyte, respectively. As defined, $\Delta Q_{(+/-)}$ was always positive because, in our samples, $Q_{(+)}$ was always greater than $Q_{(-)}$.

In gas chromatography, the partition coefficient, K, is defined as $K = C_s/C_v$, where C_s is the concentration of solute in the sorbent phase and C_v is the concentration in the vapor phase, at equilibrium. In our experimental protocol, C_v is constant since the vapor stream is continuously being replenished by the vapor generation apparatus. Therefore, we can define a ratio, α, of the partition coefficients between the two enantiomers, as follows:

$$\alpha = K_1/K_2 = [C_s/C_v]_1/[C_s/C_v]_2 = [C_s]_1/[C_s]_2 \quad (3)$$

To remain consistent with the gas chromatography literature, K_1 is defined as the larger partition coefficient, ensuring that $\alpha > 1$.[8] This corresponds to the analyte which gave the larger response, which, in this work, was always the (+) enantiomer.

The differences in steady-state response for each of the enantiomeric pairs can be attributed to the differences in the free energy, ΔG, of sorption for each enantiomer into the chiral carbon black-polymer composite (i.e., $(\Delta G_1 - \Delta G_2)$, or $\Delta_{1,2}(\Delta G)$). The difference in ΔG of sorption is related to the ratio of the partition coefficients, as expressed below:[8]

$$\Delta_{1,2}(\Delta G) = -RT \ln(\alpha) \quad (4)$$

We have shown that the composite chemiresistor detectors respond linearly to gaseous analyte concentrations over at least a factor of 10^2 in concentration; therefore, C_s is proportional to Q.[1,9] And, since the vapor concentrations of each enantiomer of an analyte were identical, with only the amount sorbing into the polymer matrix being different, we can relate α to $Q_{(+)}$ and $Q_{(-)}$ as follows:

$$\alpha = [C_s]_1/[C_s]_2 = Q_{(+)}/Q_{(-)} \quad (5)$$

The values of $\Delta_{1,2}(\Delta G)$ calculated using the percent response data and the relationships of eqs 2−5 are given in Table 2. These

(8) Schurig, V. *J. Chromatogr. A* **1994**, *666*, 111.
(9) Severin, E. J.; Doleman, B. J.; Lewis, N. S., to be submitted.

values are similar to the minimum $-\Delta_{1,2}(\Delta G)$ values (~0.1 kJ mol^{-1}) observed for enantiomers in chiral gas chromatography.[8]

In summary, we have shown that the resistance response of carbon black-polymer composite detectors can be extended to differentiate between enantiomers in the vapor phase. This behavior increases the number of molecular characteristics of a vapor-based analyte that can be probed by a carbon black-polymer composite sensor array. The enhancement in classification ability arising from the use of these chiral detectors in an array configuration will be highly task-dependent, and quantification of the separation ability of enantiomers in specific application scenarios will be reported separately.

ACKNOWLEDGMENT

R.D.S. thanks Lawerence Livermore National Laboratory for a Professional Research and Teaching leave. B.J.D. thanks the Natural Science and Engineering Research Council of Canada for a 1967 Centennial Fellowship. We gratefully acknowledge support for this work by DARPA, NASA, and the Army Research Office.

Received for review July 15, 1997. Accepted January 13, 1998.

AC970757H

Anal. Chem. **2000,** *72,* 658−668

Articles

An Investigation of the Concentration Dependence and Response to Analyte Mixtures of Carbon Black/Insulating Organic Polymer Composite Vapor Detectors

Erik J. Severin, Brett J. Doleman, and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

The responses relative to an air background of carbon black/polymer composite vapor detectors have been determined as a function of the concentration of a homologous series of alcohols (n-$C_nH_{2n+1}OH$, $1 \leq n \leq 8$), a homologous series of alkanes (n-C_nH_{2n+2}, $5 \leq n \leq 10$ and $n = 12, 14$), and a set of diverse solvent vapors. In all cases, the steady-state relative differential resistance responses, $\Delta R/R_b$, of the carbon black/polymer composite vapor detectors were well-described by a linear relationship with respect to the analyte partial pressure, at least over the tested concentration range ($P/P^\circ = 0.005-0.03$, where P° is the vapor pressure of the analyte). When two vapors in air were simultaneously presented to the detectors, the $\Delta R/R_b$ response, relative to an air background, was the sum of the $\Delta R/R_b$ values obtained when each analyte was exposed separately to the carbon black/polymer composite detectors under study. Similarly, when an analyte was exposed to the detectors on top of a background level of another analyte, the $\Delta R/R_b$ values of the array of detectors were very close to those obtained when the test analyte was exposed to the detectors only in the presence of background air. The initial training requirements from the array response output data of such detectors are minimized because the $\Delta R/R_b$ response pattern produced by the analyte of concern can be associated uniquely with that odor, under the conditions explored in this work.

Arrays of several types of vapor detectors are actively being explored to produce an "electronic nose".[1−5] In this type of system architecture, no individual detector is highly selective toward an individual analyte, as would be the case in the traditional "lock and key" approach to chemical sensing. Instead, each detector responds to many analytes, and each analyte elicits a response from many detectors. The resulting odor signature from the array of broadly cross responsive detectors is used to classify, and in some cases quantify, the analyte of concern. Detector modalities that have been employed in this architecture include surface acoustic wave (SAW) devices,[6−9] tin oxide detectors,[10−12] electrically conductive organic polymers,[2,13,14] coated fiber-optic detectors,[15] polymer-coated micromirrors,[16,17] quartz crystal microbalances (QCMs),[18,19] and carbon black−polymer composite chemiresistors.[1]

These types of broadly responsive detector arrays can be useful in at least two generic categories of sensing tasks. In one mode of operation, the array is only required to sense changes in an odor relative to a known prior condition. The changes of interest may have many different physical and/or chemical origins, some of which may not be anticipated in advance, but all of which should optimally be probed by the vapor detector array. This mode of operation is useful for applications in quality control and quality assurance of foodstuffs, fragrances, consumer goods, and similar

(1) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298.
(2) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1995**, *92*, 2652−2656.
(3) Grate, J. W.; Klusty, M.; McGill, R. A.; Abraham, M. H.; Whiting, G.; Andonian-Haftvan, J. *Anal. Chem.* **1992**, *64*, 610−624.
(4) Gardner, J. W.; Bartlett, P. N. *Sens. Actuator, B* **1994**, *18*, 211−220.
(5) Walt, D. R.; Dickinson, T.; White, J.; Kauer, J.; Johnson, S.; Engelhardt, H.; Sutter, J.; Jurs, P. *Biosens. Bioelectron.* **1998**, *13*, 697.
(6) Nakamoto, T.; Fukuda, A.; Moriizumi, T. *Sens. Actuator, B* **1993**, *10*, 85−90.
(7) Ballantine, D. S.; Rose, S. L.; Grate, J. W.; Wohltjen, H. *Anal. Chem.* **1986**, *58*, 3058−3066.
(8) Grate, J. W.; Abraham, M. H. *Sens. Actuator, B* **1991**, *3*, 85−111.
(9) Grate, J. W.; Martin, S. J.; White, R. M. *Anal. Chem.* **1993**, *65*, 5, A987−A996.
(10) Gardner, J. W.; Barlett, P. N. *Sensors and Sensory Systems for an Electronic Nose*; Kluwer Academic Publishers: Dordrecht, 1992.
(11) Gardner, J. W.; Shurmer, H. V.; Corcoran, P. *Sens. Actuator, B* **1991**, *4*, 117−121.
(12) Corcoran, P.; Shurmer, H. V.; Gardner, J. W. *Sens. Actuator, B* **1993**, *15*, 32−37.
(13) Shurmer, H. V.; Corcoran, P.; Gardner, J. W. *Sens. Actuator, B* **1991**, *4*, 29−33.
(14) Pearce, T. C.; Gardner, J. W.; Friel, S.; Bartlett, P. N.; Blair, N. *Analyst* **1993**, *118*, 371−377.
(15) White, J.; Kauer, J. S.; Dickinson, T. A.; Walt, D. R. *Anal. Chem.* **1996**, *68*, 2191−2202.
(16) Butler, M. A.; Ricco, A. J.; Buss, R. *J. Electrochem. Soc.* **1990**, *137*, 1325−1326.
(17) Hughes, R. C.; Ricco, A. J.; Butler, M. A.; Pfeifer, K. B. *J. Biochem. Biotechnol.* **1993**, *41*, 77−85.
(18) Slater, J. M.; Paynter, J. *Analyst* **1994**, *119*, 191−195.
(19) Slater, J. M.; Watt, E. J. *Analyst* **1991**, *116*, 1125−1130.

10.1021/ac9910278 CCC: $19.00 © 2000 American Chemical Society
Published on Web 01/21/2000

applications.[14,20−22] For such purposes, the detector response need only be reproducible from trial to trial, and no constraints on the form of the detector response are necessarily required to perform the task at hand.

In another operational mode, a detector array could be used to identify a signature of an odor in the field on the basis of a comparison of the array response to the response signature that was recorded and stored for that analyte during a prior training/calibration run.[23−25] Such applications might include providing a warning when a particular odor becomes present above a certain concentration level in the vapor phase, tracking and/or localization of an odor in the environment, or determining the concentration of an analyte in a simple, but relatively time-independent, effluent mixture. In these types of applications, it is highly advantageous to utilize detectors that have a linear output signal in response to variations in the concentration of a particular odor, so that the pattern type allows identification of the odor while the pattern height can be straightforwardly related to the odor concentration. It is even more advantageous if the array response to the odor of concern is the same in the absence and presence of other odors. In this fashion, the initial training requirements from the array response output data are minimized because the pattern produced by the analyte of concern can be associated uniquely with that odor regardless of the changing environmental conditions under which the analysis is performed.

Prior work in our laboratory has demonstrated that insulating organic polymers interspersed with domains of electrical conductors can provide chemically sensitive detector materials that can be used to produce an "electronic nose" array.[1,2] The conducting polymer composites have been formed using either organic, inorganic, or carbonaceous materials as the conducting phase. Sorption of organic solvent vapors into these types of detectors produces a characteristic, reversible resistance change in the detector element.[1] Because every organic polymer will have a characteristic gas/polymer partition coefficient in response to the presence of a particular odor, a collection of insulating organic polymers provides a diversity in detector materials that produces the diagnostic response pattern of the detector array. Under certain circumstances, analysis of the pattern of signals produced by the detector array then allows information on odor classification and concentration to be extracted through signal processing methods.[26]

In this work, we describe the results of an extensive set of experiments designed to investigate the behavior of arrays of conductive polymer composite detectors when presented with a broadly construed, generic set of test organic vapors at varying analyte concentrations. In addition, we have probed the response when the detectors are exposed to various concentrations of members of homologous series of alkanes or alcohols. Additionally, the detector response properties have been investigated during exposure to various binary vapor mixtures to ascertain whether an array response pattern for a pure odor is transferable, weighted by the mole fraction of its vapor in an analyte mixture, to binary mixtures of analytes. Finally, we describe the results of experiments in which a small but rapidly changing odor concentration has been superimposed upon a relatively slowly varying baseline odor concentration.

EXPERIMENTAL SECTION

A. Materials. The carbon black used in the composites was Black Pearls 2000 (BP2000), a furnace black material that was generously donated by Cabot Co. (Billerica, MA). The following polymers were used in the composites (listed as detector number, polymer): **1**, poly(4-vinylphenol); **2**, poly(styrene-*co*-allyl alcohol), 5% hydroxy; **3**, poly(α-methylstyrene); **4**, poly(vinyl chloride-*co*-vinyl acetate), 10% vinyl acetate; **5**, poly(*N*-vinylpyrrolidone); **6**, poly(vinyl acetate); **7**, poly(methyl vinyl ether-*co*-maleic anhydride); **8**, poly(carbonate−bisphenol A); **9**, poly(styrene); **10**, poly(styrene-*co*-maleic anhydride), 50% styrene; **11**, poly(vinyl butyral); **12**, poly(sulfone); **13**, poly(methyl methacrylate); **14**, poly(vinylidene chloride-*co*-acrylonitrile), 80% vinylidene chloride; **15**, poly(caprolactone); **16**, poly(ethylene-*co*-vinyl acetate), 82% ethylene; **17**, poly(ethylene oxide); **18**, poly(butadiene), 36% *cis*-1,4-, 55% *trans*-1,4-, 9% vinyl-1,2-; **19**, poly(epichlorohydrin); **20**, poly(styrene-*co*-butadiene), 28% styrene; **21**, addition product of sodium menthoxide to poly(pentafluorostyrene); **22**, (+)-isopinocampheol-derivatized poly(*p*-chloromethylstyrene); **23**, poly(fluorostyrene); **24**, poly(styrene-*co*-isoprene) (Figure 1). All polymers were purchased from Polysciences Inc. or Aldrich Chemical Co. and were used as received, except polymers 20−23, which were kindly supplied by Prof. Robert H. Grubbs of Caltech. The solvents used in this study all were reagent grade and were used as received.

B. Fabrication of Detectors. Two substrates were used for the detectors. In one configuration, two parallel bands of gold, 50−100 nm thick and separated by either 1 or 5 mm, were deposited onto conventional 7.5 cm × 2.5 cm glass slides (Corning Inc.). The slides were then cut into strips to produce 0.7 cm × 2.5 cm pieces of glass, with each strip of glass having one pair of Au leads spaced 1 or 5 mm apart. In the second configuration, a commercial surface-mounting breadboard was slightly modified to be used as the substrate. The commercial product ("Surfboards") consisted of parallel leads of metal deposited onto the circuit board material. These leads were soldered to pins that were on 0.10 in. centers. The commercial product was cut into pairs of leads and was then coated with the composite films.

The detector films were made from a solution of the polymer into which carbon black had been suspended. A total of 160 mg of one of the insulating polymers (Figure 1) was dissolved in 20 mL of solvent, and carbon black (40 mg) was then suspended in this solution, to produce a composition of 80% polymer and 20% carbon black by weight of solids. The solvent was generally tetrahydrofuran, benzene, or methylene chloride, depending on the solubility of the polymer. The solutions were sonicated for 5 min to suspend the carbon black. Aromatics and chlorinated solvents yielded very good suspensions of the carbon black. A single solution that contained the polymer and the carbon black

(20) Schweizerberberich, P. M.; Vaihinger, S.; Gopel, W. *Sens. Actuator, B Chem* **1994**, *18*, 282−290.
(21) Gardner, J. W.; Shurmer, H. V.; Tan, T. T. *Sens. Actuator, B* **1992**, *6*, 71−75.
(22) Shurmer, H. V.; Gardner, J. W.; Chan, H. T. *Sens. Actuator* **1989**, *18*, 361−371.
(23) Charlesworth, J. M.; Riddell, S. Z.; Mathews, R. J. *J. Appl. Polym. Sci.* **1993**, *47*, 653.
(24) Nieuwenhuizen, M. S.; Harteveld, J. L. N. *Sens. Actuator, A* **1994**, *44*, 219.
(25) Grate, J. W.; Rosepehrsson, S. L.; Venezky, D. L.; Klusty, M.; Wohltjen, H. *Anal. Chem.* **1993**, *65*, 1868.
(26) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177−4190.

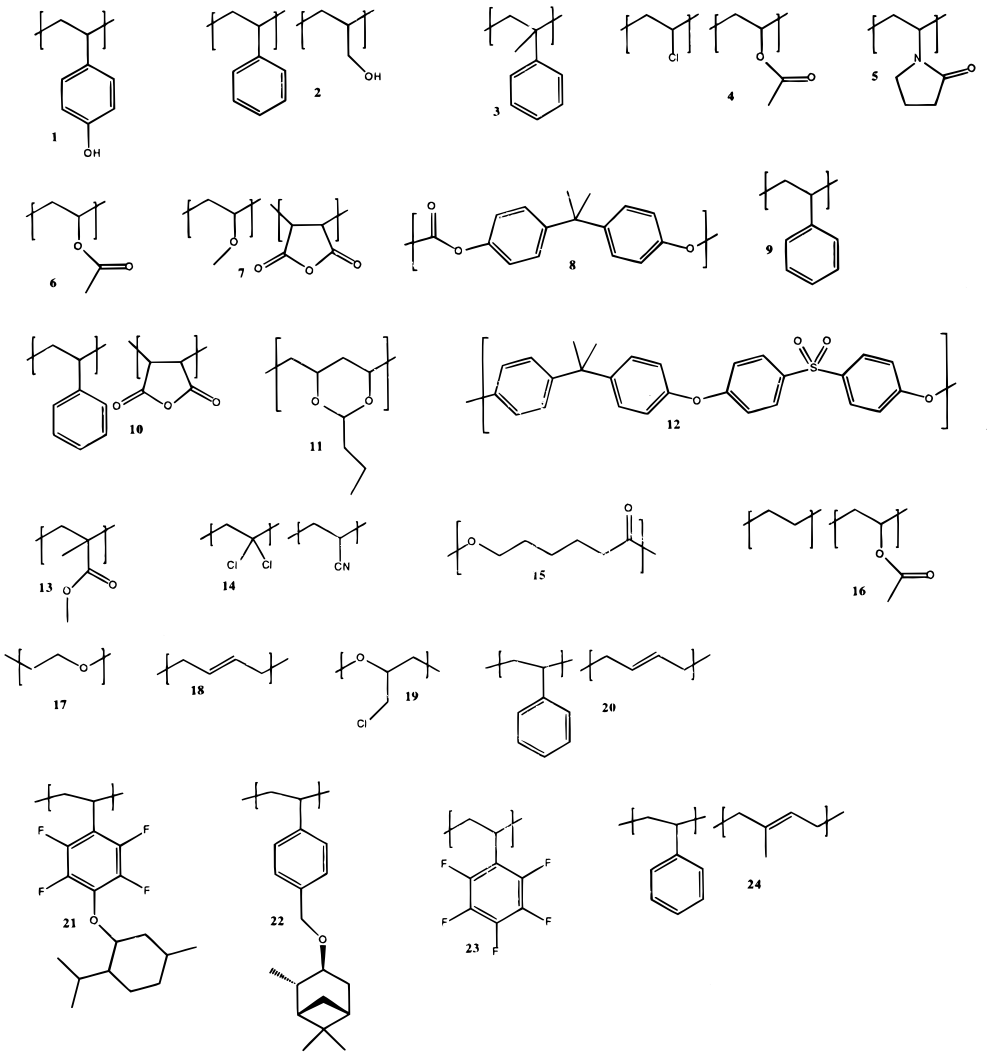


Figure 1. Structures of the polymers used in this work. Listed as detector number, polymer: **1**, poly(4-vinylphenol); **2**, poly(styrene-*co*-allyl alcohol), 5% hydroxy; **3**, poly(α-methylstyrene); **4**, poly(vinyl chloride-*co*-vinyl acetate), 10% vinyl acetate; **5**, poly(*N*-vinylpyrrolidone); **6**, poly(vinyl acetate); **7**, poly(methyl vinyl ether-*co*-maleic anhydride); **8**, poly(bisphenol A−carbonate); **9**, poly(styrene); **10**, poly(styrene-*co*-maleic anhydride), 50% styrene; **11**, poly(vinyl butyral); **12**, poly(sulfone); **13**, poly(methyl methacrylate); **14**, poly(vinylidene chloride-*co*-acrylonitrile), 80% vinylidene chloride; **15**, poly(caprolactone); **16**, poly(ethylene-*co*-vinyl acetate), 82% ethylene; **17**, poly(ethylene oxide); **18**, poly(butadiene), 36% *cis*-1,4-, 55% *trans*-1,4-, 9% vinyl-1,2-, **19**, poly(epichlorohydrin); **20**, poly(styrene-*co*-butadiene), 28% styrene; **21**, addition product of sodium menthoxide to poly(pentafluorostyrene); **22**, (+)-isopinocampheol-derivatized poly(*p*-chloromethylstyrene); **23**, poly(fluorostyrene); **24**, poly(styrene-*co*-isoprene)

was used to prepare all the detectors of a given composition that were used in this work. An aliquot of the suspension was spin coated, at 1000 rpm, onto a glass substrate using a Headway (Garland, TX) spin coater, and the resulting film was allowed to dry in air. Multiple coatings of the suspension were applied to each substrate to yield detectors having resistance values of approximately a few hundred kiloohms. For the fiberglass substrates, the film was applied by dip-coating the substrate two or three times until the desired resistance was achieved. Before use, the detectors were dried in open air and then were placed in air flowing at 20 L min^{-1} for 12−24 h.

C. Instrumentation and Apparatus. An automated flow system consisting of LabVIEW software, a Pentium computer, and electronically controlled solenoid valves and mass flow controllers was used to produce and deliver selected concentrations of solvent vapors to the detectors.[27] To obtain the desired analyte concentration, a stream of carrier gas was passed through a bubbler that had been filled with the solvent of choice. Saturation of the carrier gas with the solvent vapor was verified through measurement of the rate of mass loss of the solvent in the bubbler.[28] The vapor-saturated carrier gas was then diluted with pure carrier gas through the use of mass flow controllers (MKS Instruments, Inc). Calibrations of the flow system using a flame ionization detector (model 300 HFID, California Analytical Instruments, Inc.) verified that the analyte concentrations delivered to the sensors were those expected from the settings of the mass flow controllers.

The carrier gas for all experiments was oil-free air, obtained from the general compressed air laboratory source, containing

(27) Severin, E. J. Ph.D. Thesis, California Institute of Technology: Pasadena, CA, 1999.

(28) Atkins, P. W. *Physical Chemistry*; W. H. Freeman and Co.: New York, 1994.

1.10 ± 0.15 ppth (parts per thousand) of water vapor. The air was filtered to remove particulates, but deliberately was not dehumidified or otherwise purified. Fluctuations in laboratory temperature, 21.5 ± 1.5 °C, could cause a ~10% error in setting and controlling the vapor concentrations between nominally identical exposures over the course of the data collection analyzed in this work. No temperature control of the apparatus or of the carbon black−polymer composite detectors was performed. The flow rate of the vapor stream entering the exposure chamber (~1 L in total volume) was maintained at 15 L min^{-1}.

D. Measurements. The dc electrical resistance of each detector was monitored in response to the presence of various test vapors and mixtures of vapors. Resistance measurements were performed using a simple two-point configuration across the gold leads that bridged the sensing element. The detectors were multiplexed through a Keithley model 7001 channel switcher to a Keithley model 2002 multimeter that measured the dc resistance of each detector once every 3−5 s, with the exact time interval depending on the particular experiment.

To initiate an experiment, the detectors were placed into the flow chamber and a background flow of compressed air was introduced until the resistance of the detectors stabilized. Each exposure consisted of a three-step process that began with 60 s of air flow to achieve a smooth baseline resistance. After this period, the detectors were exposed to solvent vapor at a controlled concentration in flowing air. The solvent exposure was then followed by a flow of clean air for a time equal to the total exposure time, to restore the baseline resistance values. For the linearity studies, the 60 s baseline period was followed by 240 s of exposure to the test analyte. To probe the dependence of the detector response on the order of presentation, in some measurements of the mixture studies, the exposure phase consisted of two parts. In the sequential mixture measurements, the first analyte (denoted as s_1) was exposed for 120 s, at which time the second solvent, s_2, was introduced and exposed for an additional 120 s. During the exposure of the second analyte, the first analyte was continually flowing (this protocol is denoted as s_1, $s_1 + s_2$). In the measurements, when a mixture of two analytes was exposed simultaneously to the sensors (denoted $s_1 + s_2$), the two analytes of the mixture were presented to the detectors for a total of 240 s.

In studies of mixtures, the eight bubblers of the system were divided into two sets of four bubblers each. One mass flow controller was present for "set A" and one for "set B" (Table 1). One-way valves ensured that significant gas back flow did not occur during the experiments. Analytes in the same solvent set could not be exposed simultaneously to the detectors. Therefore, 16 pairs of solvents were available for use in the first set of mixture studies. Only six solvents were used in the second mixture study, three in each set, so nine solvent pairs were available. The detectors used for the eight-solvent experiment were formed from polymers **1**−**18**, **21**, and **23** (Figure 1). The detectors used for the six-solvent experiment, the alcohol linearity study, and the alkane linearity study were formed using polymers **8**, **12**, and **15**−**24** (Figure 1). In all experiments, one copy of each type of detector was used.

In both the eight-solvent and six-solvent mixture experiments, the detectors were exposed to individual solvents (s_1), to pairs of solvents presented simultaneously ($s_1 + s_2$), and to one solvent followed by addition of another solvent (s_1, $s_1 + s_2$). The individual solvents and the preselected pairs of s_1, s_2 solvents were exposed to the detectors at analyte concentrations that corresponded to 0.5, 1.0, and 1.5% of each solvent's vapor pressure, $P°$. In the six-solvent experiment, individual solvents were additionally presented at 2.0 and 2.5% of $P°$. Solvents forming every compositionally distinct binary mixture were permuted in their order of presentation to the detectors, so that for each solvent pair (one from set A and one from set B) at every distinct analyte concentration, the trials included the exposure protocol s_A, $s_A + s_B$ as well as the exposure protocol s_B, $s_B + s_A$. Each unique exposure protocol, for each type of mixture and pure analyte presentation, was repeated 5 times. The eight-solvent experiment thus contained 2280 total exposures (8 solvents, 3 concentrations, 5 repeats of each for the individual solvent exposures, $6 \times 8 \times 5$ simultaneous mixture exposures, and $16 \times 8 \times 2 \times 5$ sequential mixture exposures). The six-solvent experiment contained 1365 total exposures ($6 \times 5 \times 5$ individual solvent exposures, $3 \times 9 \times 5$ simultaneous mixture exposures, and $9 \times 6 \times 2 \times 5$ sequential mixture exposures). Within each experiment, every exposure was assigned a randomly generated index number using the Microsoft Excel random number generator. The exposures were then presented to the detector array in ascending order of the assigned index values.

Table 1. Two Groups of Solvents Used in the Eight-Solvent Binary Mixture Study and the Six-Solvent Binary Mixture Study[a]

set A	set B
Eight-Solvent Experiment	
benzene	chloroform
ethyl acetate	ethanol
heptane	hexane
methanol	toluene
Six-Solvent Experiment	
benzene	nitrobenzene
2-propanol	chloroform
cyclohexanone	heptane

[a] Binary mixtures were formed between solvents of set A and solvents of set B of each group. Solvents common to one set could not be paired.

In the studies designed to quantify the detector response as a function of analyte concentration, two homologous series of vapors, one consisting of straight-chain alcohols and the other of straight-chain alkanes, were exposed to the detectors. The following alcohols were used: methanol, ethanol, 1-propanol, 1-butanol, 1-pentanol, 1-hexanol, 1-heptanol, and 1-octanol. In a separate run, *n*-pentane, *n*-hexane, *n*-heptane, *n*-octane, *n*-nonane, *n*-decane, *n*-dodecane, and *n*-tetradecane were used. In another, related set of experiments, the broad test set of solvents used in the studies of mixtures (Table 1) was exposed to the detectors over a wider concentration range ($0.005P° \leq P \leq 0.03P°$) than was used in the runs to determine the detector's response to mixtures of these particular solvent vapors. Additionally, one run with the straight-chain alcohols was performed using vapor concentrations that were in the range $0.01P° \leq P \leq 0.06P°$. In each of these experiments, each unique presentation of an analyte was repeated 10 times, with the entire presentation order (within

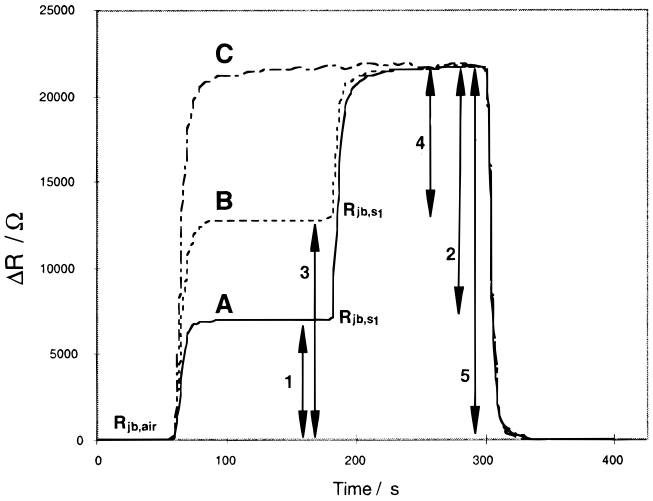


Figure 2. Representative differential resistance responses for three types of vapor presentations to a poly(ethylene-*co*-vinyl acetate)−carbon black composite vapor detector. (A) Exposure to benzene at $P/P^\circ = 0.02$ ($\Delta R_{js1,max}$ indicated by arrow 1) followed by exposure to benzene at $P/P^\circ = 0.02$ and chloroform at $P/P^\circ = 0.02$ ($\Delta R_{js2,max}$ indicated by arrow 2). The combined response, $\Delta R_{js1,max} + \Delta R_{js2,max}$, is indicated by arrow 5. (B) Exposure to chloroform at $P/P^\circ = 0.02$ ($\Delta R_{js1,max}$ indicated by arrow 3) followed by exposure to chloroform at $P/P^\circ = 0.02$ and benzene at $P/P^\circ = 0.02$ ($\Delta R_{js2,max}$ indicated by arrow 4). The combined response, $\Delta R_{js1,max} + \Delta R_{js2,max}$, is also indicated by arrow 5. Arrow 1 ≈ arrow 4; arrow 3 ≈ arrow 2. (C) Benzene at $P/P^\circ = 0.02$ and chloroform at $P/P^\circ = 0.02$ both presented simultaneously to the detector (response, $\Delta R_{js1+s2,max}$ is again indicated by arrow 5).

a run) randomized with respect to solvents, concentrations of solvents, and repeated exposures to a solvent.

E. Data Processing. Sample responses for a single exposure and for a sequential mixture exposure are shown in Figure 2. Although the resistance of each detector was sampled once every 3−5 s during each exposure, only the maximum relative differential resistance change, $\Delta R_{js,max}/R_{jb,air}$, where $\Delta R_{js,max}$ produced by exposure to an individual solvent is the maximum resistance change of the *j*th detector during exposure to solvent s and $R_{jb,air}$ is the baseline resistance of the *j*th detector exposed to the initial 60 s period of exposure to background air, was used in analysis of the data. In the mixture studies when solvents were exposed sequentially to the detectors, three separate $\Delta R/R$ values, $\Delta R_{j,s1,max}/R_{jb,air}$, $\Delta R_{j,s2,max}/R_{jb,s1}$, and $\Delta R_{j,s1,s1+s2}/R_{jb,air}$ were calculated from the data from each exposure protocol s_1, $s_1 + s_2$ (Figure 2).

For these solvents and detectors, the exposure time was sufficiently long that the maximum response value, $\Delta R_{js,max}/R_{jb}$, was a very good approximation to the change in the steady-state resistance value of the detectors in response to the specified analyte concentration relative to the baseline resistance of the detector in an air background flow alone. Examples of the temporal dependence of individual carbon black/insulating polymer composite detectors are shown in Figure 2. For some exposures in the eight-solvent system, the value $R_{jb,s1}$ had not completely reached steady state. Therefore, to calculate $\Delta R_{js2,max}/R_{jb,s1}$ in those cases, the slope of the resistance values 30 s prior to the start of the exposure was calculated and subtracted from the R_{js2} values. If this correction were not made, then the detector's response to s_2 would have been overestimated.

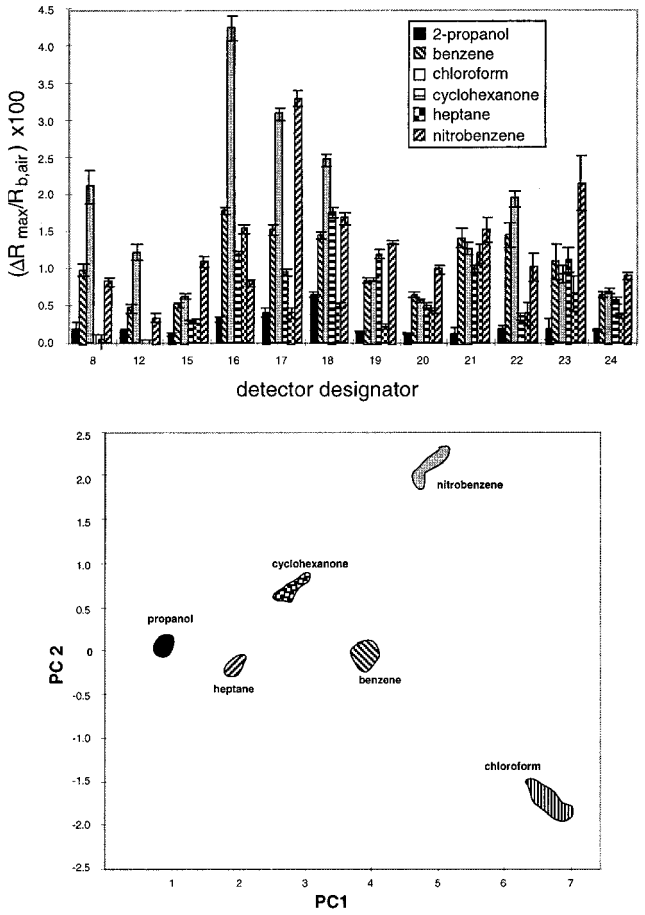


Figure 3. (a) histogram of the maximum relative differential resistance response of 12 carbon black/polymer composite detectors exposed to *n*-heptane, cyclohexanone, benzene, chloroform, nitrobenzene, and 2-propanol each presented at $P/P^\circ = 0.03$ in air. Each analyte was presented 10 times to the array, with the order of presentation randomized over all repetitions of all test solvents. (b) Results from the exposures described in (a) as represented by the first two dimensions of principal component space, which contain 96% of the total variance in the data. The ellipsoids contain 95% of the data for each analyte in principal component space.

RESULTS

A. Linearity of Detector Response for Pure Odors. Figure 3a displays the maximum relative differential resistance data, $\Delta R_{js,max}/R_{jb,air}$, for a 12-element conducting organic polymer composite detector array toward a series of test analytes when each analyte was maintained at a partial pressure, *P*, in air equal to 3% of its vapor pressure, P°, at 22 °C. Each analyte can be seen to produce a distinct $\Delta R_{js,max}/R_{jb,air}$ response pattern on the array of conducting polymer composite detectors. Principal component analysis was used in order to aid visualization of the differences between $\Delta R_{js,max}/R_{jb,air}$ patterns produced by the various analytes.[29] Figure 3b presents the $\Delta R_{js,max}/R_{jb,air}$ data in principal component space, with the axes representing the first and second principal components of the data set. All analytes were well-separated from each other on the basis of the differences between their characteristic $\Delta R_{js,max}/R_{jb,air}$ response patterns on the array of detectors.

The concentration of each analyte was then varied over six even steps in the range $0.005P^\circ \le P \le 0.03P^\circ$. Figure 4 depicts

(29) Hecht, H. G. *Mathematics in Chemistry: An Introduction to Modern Methods*; Prentice Hall: Englewood Cliffs, NJ, 1990.

Table 2. Statistics for $\Delta R/R_b$ Response vs $P/P°$ for Representative Polymers and Analytes[a]

	poly(butadiene)				poly(epichlorohydrin)		
	R^2	intercept	slope		R^2	intercept	slope
1-propanol	0.9998	−0.0136	0.2200	ethanol	0.9956	0.0024	0.0367
benzene	1.0000	−0.0429	0.5017	1-propanol	0.9990	−0.0048	0.0542
chloroform	0.9998	−0.0563	0.8509	1-butanol	0.9997	−0.0031	0.0629
cyclohexanone	1.0000	−0.0668	0.6128	1-pentanol	0.9992	−0.0060	0.0670
n-heptane	0.9993	−0.0003	0.1682	1-hexanol	0.9998	−0.0036	0.0703
nitrobenzene	0.9995	−0.0397	0.5656	1-heptanol	0.9991	−0.0033	0.0691
n-hexane	0.9994	−0.0344	0.1788				
n-heptane	0.9999	−0.0175	0.1659				
n-octane	0.9994	−0.0223	0.1688				
n-nonane	0.9995	−0.0294	0.1713				
n-decane	0.9988	−0.0162	0.1599				
n-dodecane	0.9997	−0.0136	0.1478				

[a] Correlation coefficients, intercepts, and slopes for three sets of analytes exposed at $P/P° = 0.005-0.03$.

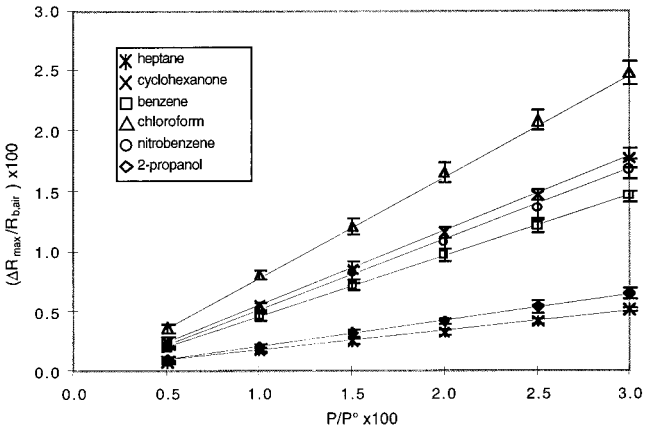


Figure 4. Average maximum relative differential resistance responses, $\Delta R_{js,max}/R_{jb,air}$, of composite detector films consisting of carbon black and poly(butadiene), when exposed to *n*-heptane, cyclohexanone, benzene, chloroform, nitrobenzene, and 2-propanol, each at $P/P° = 0.005-0.03$ in air in six even steps. Each analyte was presented 10 times to the array, with the order of presentation randomized over all repetitions of all test solvents. The error bars represent 1σ values computed from 10 exposures at each $P/P°$.

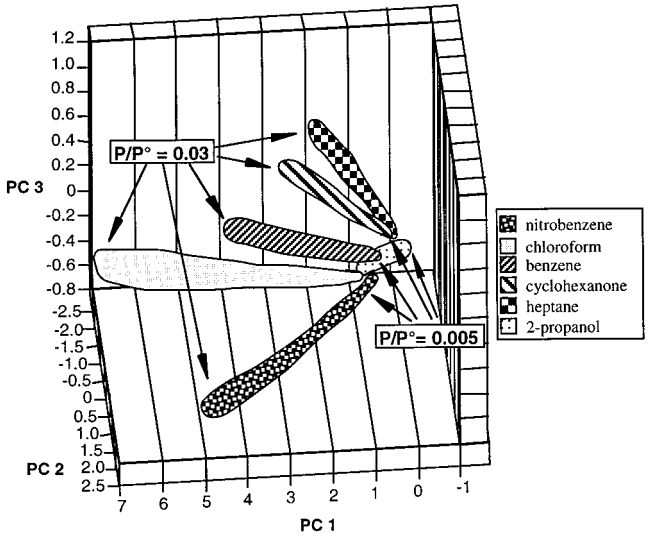


Figure 5. Data in principal component space from a 12-detector array exposed to *n*-heptane, cyclohexanone, benzene, chloroform, nitrobenzene, and 2-propanol each at $P/P° = 0.005-0.03$ in air in six even steps. The first three principal components depicted contained 98% of the total variance in the data. The ellipsoids contain 95% of the data for each analyte. Each analyte was presented 10 times to the array, with the order of presentation randomized over all repetitions of all test solvents.

the responses of a representative detector to all of the test solvent vapors. The data were well fit by a linear dependence of $\Delta R_{js,max}/R_{jb,air}$ on $P/P°$ over the $P/P°$ ranges probed in this experiment. A summary of the correlation coefficients calculated for these lines is presented in Table 2, while statistics for all of the sensor−analyte combinations are provided in the Supporting Information that accompanies this work. For some sensor−analyte combinations, the correlation coefficients were low because the sensor exhibited only a very small response to the analyte. For example, poly(sulfone) had a small response to nonpolar solvents and so the correlation coefficients for these presentations are low. Similarly, essentially no response was exhibited by poly(sulfone) to dodecane. The intercepts of such plots were statisticallly indistinguishable from zero for all sensor−analyte combinations investigated.

Figure 5 presents the concentration-dependent $\Delta R_{js,max}/R_{jb,air}$ response data for the entire detector array in principal component space. For each test vapor, the analytes produced a unique signal response pattern, with the pattern direction in principal component space diagnostic of the analyte and the pattern height proportional to the analyte concentration in the vapor phase. This behavior is further illustrated by normalization of the detector response patterns with respect to analyte concentration according to eq 1,

$$S_{js} = (\Delta R_{js,max}/R_{jb,air})(P°/P) \quad (1)$$

where S_{js} is the normalized signal for 12 detector films exposed to benzene, chloroform, and nitrobenzene, each presented at $P/P° = 0.005-0.03$ in six even steps. As can be seen from Figure 6, the characteristic S_{js} pattern of each test vapor was maintained, within experimental error, as the analyte concentration was varied.

Additional experiments were performed using a homologous series of alkanes, and then using a homologous series of alcohols, as test analytes. Figures 7 and 8 display the $\Delta R_{js,max}/R_{jb,air}$ values for selected detectors. The statistical information on these runs is summarized in Table 2. Again the data were well fit by a linear dependence of $\Delta R_{js,max}/R_{jb,air}$ on $P/P°$ over the $P/P°$ range probed in these experiments.

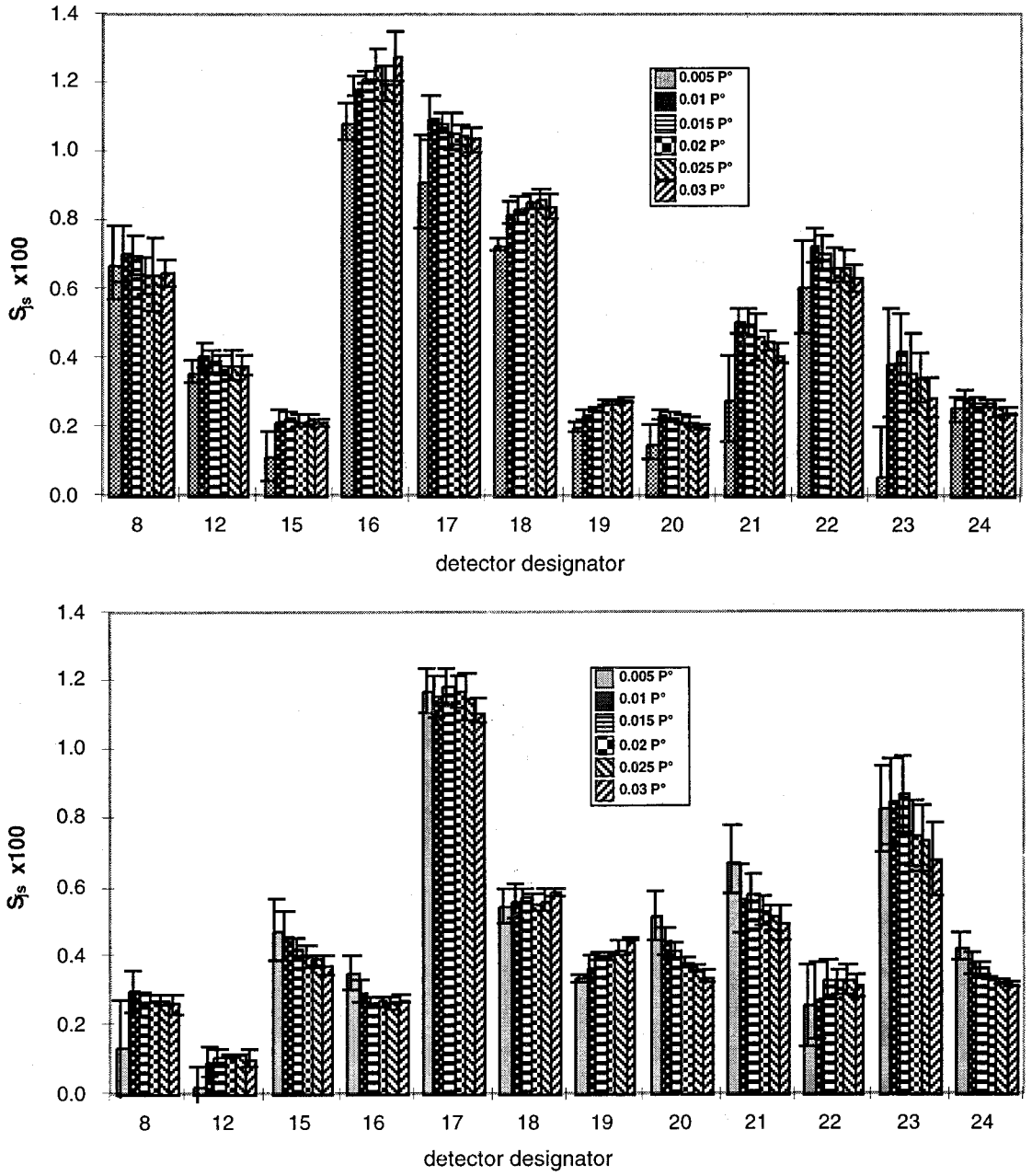


Figure 6. Histogram of the average normalized response of a 12-element array of carbon black/polymer detector films exposed to two analytes; (a) chloroform and (b) nitrobenzene, each presented 10 times at $P/P° = 0.005-0.03$ in air in six even steps. The data were normalized according to eq 1 in the text.

Figure 9a shows that all of the test alcohols could all be distinguished from one another visually in principal component space when the responses of all detectors in the array are considered. Additionally, like the analytes in the broad test set, the normalized patterns of $\Delta R_{js,max}/R_{jb,air}$ were essentially invariant as the analyte concentration was varied. Identical behavior was observed for the alkanes, as seen in Figure 9b. Thus, the $\Delta R_{js,max}/R_{jb,air}$ pattern type is diagnostic of the analyte and the pattern height indicates the concentration of each of these analytes, at least under the conditions of these test runs.

B. Detector Response to Analytes in the Presence of Background Odors. The response of the detectors to various test vapors was also investigated when the detectors were first exposed to, and then maintained in the presence of, a fixed concentration of another solvent vapor. Figure 10 exhibits the $\Delta R_{js,max}/R_{jb,air}$ values displayed by poly(ethylene-*co*-vinyl acetate) and poly(caprolactone) detectors in response to varying concentrations of heptane in the range $0.005P° \le P \le 0.025P°$, relative to an air background gas flow. The responses for heptane vapor at $0.005P° \le P \le 0.015P°$ in air were then recorded when the detector was exposed to the analyte gas stream in the presence of a constant background gas that consisted of air with 2-propanol, benzene, or cyclohexanone at $P/P° = 0.005$, 0.010, and 0.015 for each background gas. As displayed in Figure 10, $\Delta R_{jheptane,max}/R_{jb,s1}$ and $\Delta R_{jheptane,max}/R_{jb,air}$ were essentially constant for s_1 = benzene, cyclohexanone, and 2-propanol at the three values of $P/P°$. Figure 11 shows the same result in principal component space for the responses of the entire array of detectors, illustrating

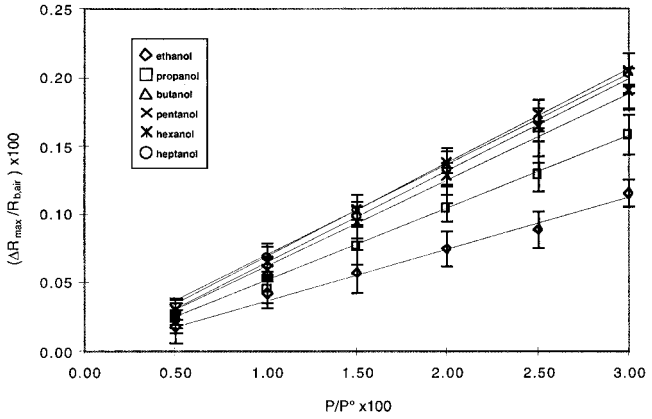


Figure 7. Maximum relative differential resistance responses, $\Delta R_{js,max}/R_{jb,air}$, of composite detector films consisting of carbon black and poly(epichlorohydrin), when exposed to ethanol, 1-propanol, 1-butanol, 1-pentanol, 1-hexanol, and 1-heptanol each at P/P° = 0.005−0.03 in six even steps in air. Each analyte was presented 10 times to the array, with the order of presentation randomized over all repetitions of all test solvents. The error bars represent 1σ values computed from 10 exposures at each P/P°.

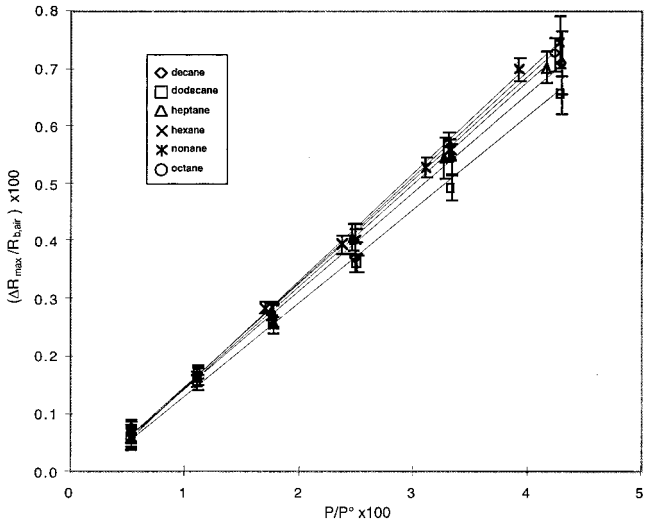
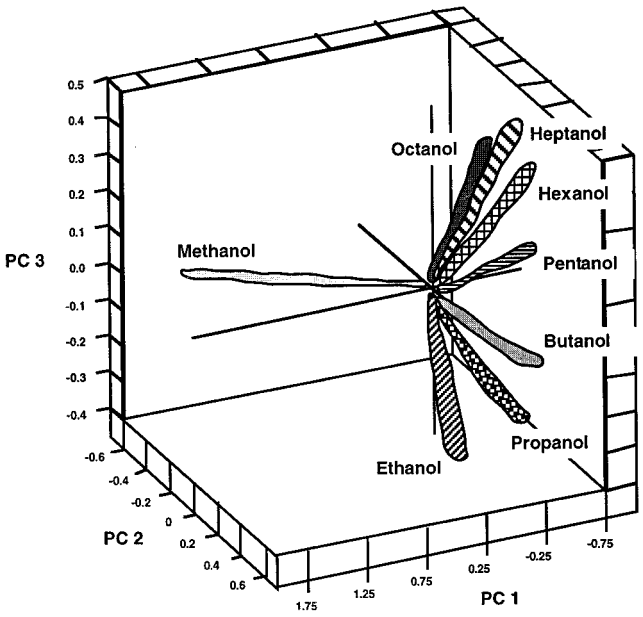


Figure 8. Maximum relative differential resistance responses, $\Delta R_{js,max}/R_{jb,air}$, of composite detector films consisting of carbon black and poly(butadiene), when exposed to *n*-dodecane, *n*-decane, *n*-nonane, *n*-octane, *n*-heptane, and *n*-hexane each at P/P° = 0.005−0.03 in six even steps in air. Each analyte was presented 10 times to the array, with the order of presentation randomized over all repetitions of all test solvents. The error bars represent 1σ values computed from 10 exposures at each P/P°.

that this behavior is characteristic of the response pattern in the detector array as well as of the individual detectors displayed in Figure 10.

C. Detector Response to Binary Analyte Mixtures. Figure 12 shows the $\Delta R_{s2,max}/R_{jb,s1}$ and $\Delta R_{s1,s1+s2,max}/R_{jb,air}$ values of a carbon black/poly(ethylene oxide) detector to mixtures of benzene and heptane. For this detector for both the $\Delta R_{s2,max}/R_{jb,s1}$ and $\Delta R_{s1,s1+s2,max}/R_{jb,air}$ values, s_1 and s_2 were each presented to the detectors at P/P° = 0.005, 0.010, and 0.015. The linear dependence of $\Delta R_{js,max}/R_{jb}$ on P/P° exhibited by an individual detector was maintained when the analyte was a constituent of a binary solvent mixture. The lines that have been drawn in Figure 12 to connect the data points also correspond to the change in



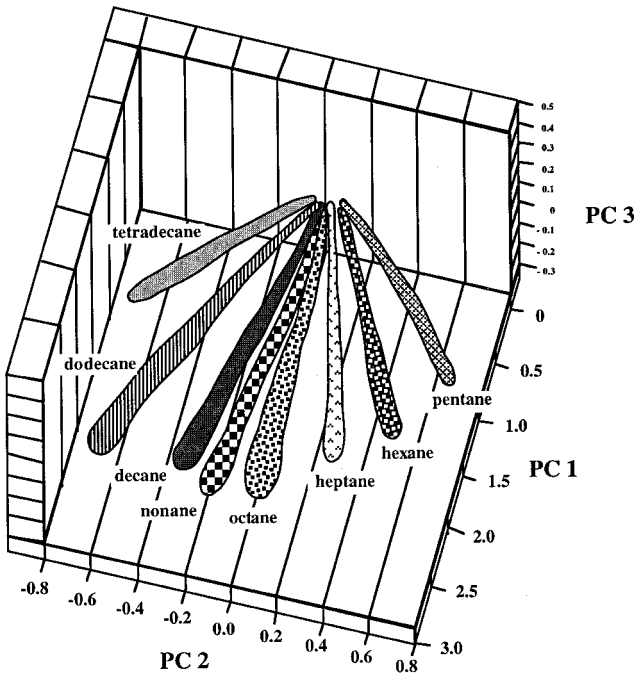


Figure 9. (a) Data in principal component space from a 20-detector array exposed 10 times each to methanol, ethanol, 1-propanol, 1-butanol, 1-pentanol, 1-hexanol, 1- heptanol, and 1-octanol each at P/P° = 0.005 to 0.03 in air in 27 even steps. The first three principal components contain 99% of the total variance in the data. The ellipsoids contain 99% of the data for each analyte. (b) Data in principal component space from a 20-detector array exposed 5 times each to *n*-tetradecane, *n*-dodecane, *n*-decane, *n*-nonane, *n*-octane, and *n*-heptane each at P/P° = 0.005−0.03 in air in 27 even steps. The first three principal components contain 99% of the total variance in the data. The ellipsoids contain 99% of the data for each analyte. All presentations in each set were randomized over all repetitions of all test solvents.

response that would be expected based on the $\Delta R_{js,max}/R_{jb,air}$ behavior of the detector when presented with corresponding changes in the concentration of the individual solvent vapor in an air background. Additionally, the total $\Delta R_{js1+s2,max}/R_{jb,air}$ response to two solvents relative to a background air baseline was

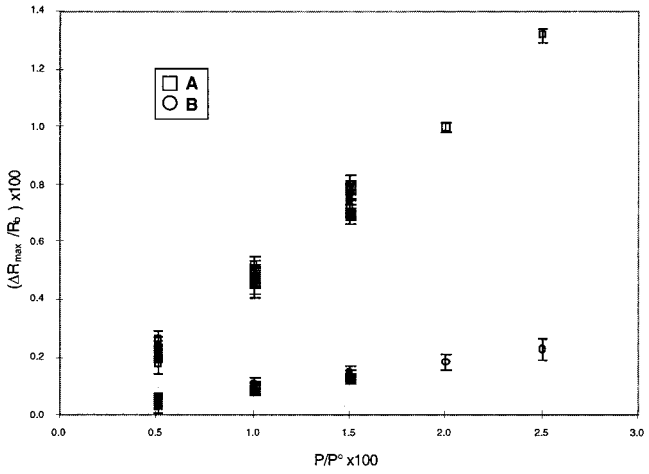


Figure 10. Maximum relative differential resistance responses, $\Delta R_{js,max}/R_{jb,air}$, of composite detector films consisting of carbon black and (A) poly(ethylene-*co*-vinyl acetate) and (B) poly(caprolactone), when exposed to *n*-heptane at $P/P^\circ = 0.005-0.025$ in air in five even steps (represented by the open symbols). Additional exposures (solid symbols) to *n*-heptane were performed at $P/P^\circ = 0.005$, 0.01, and 0.015 while the detector film was exposed to either benzene, cyclohexanone, or 2-propanol at $P/P^\circ = 0.005$, 0.01, or 0.015.

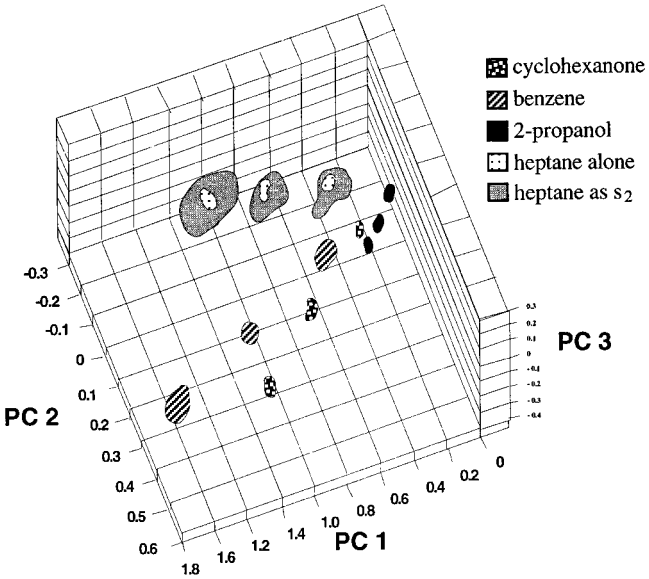


Figure 11. Data in principal component space from a 12-detector array exposed to *n*-heptane, benzene, cyclohexanone, or 2-propanol at $P/P^\circ = 0.005$, 0.01, and 0.015, and to exposures of *n*-heptane at $P/P^\circ = 0.005$, 0.01, and 0.015 while the detector film was exposed to either benzene, cyclohexanone, or 2-propanol each at $P/P^\circ = 0.005$, 0.01, or 0.015. The first three principal components contain 98% of the total variance in the data. The ellipsoids contain 95% of the data for each analyte. Each analyte was presented five times to the array, with the order of presentation randomized over all repetitions of all exposure types.

independent of whether the two solvents were exposed simultaneously or sequentially to the detector. Furthermore, in the case of sequential solvent vapor exposures, the maximum relative differential response values for a given solvent were independent of the order in which the solvents were presented to the detector. Figure 13 shows similar data, in principal component space, that were produced by an entire array of carbon black/polymer composite detectors during individual analyte exposure, and

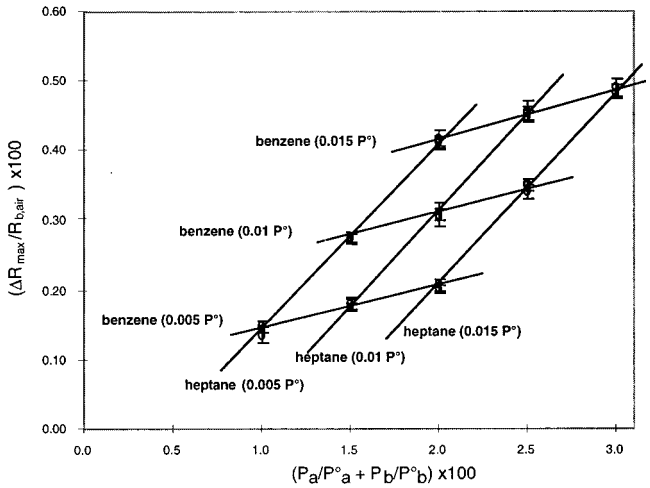


Figure 12. Maximum relative differential resistance responses of a poly(ethylene-*co*-vinyl acetate)−carbon black composite detector film when exposed to simultaneous and sequential binary mixtures of benzene at $P/P^\circ = 0.005$, 0.01, or 0.015, and *n*-heptane at $P/P^\circ = 0.005$, 0.01, or 0.015. Each of the nine binary mixture combinations was presented five times to the array, with the order of presentation randomized over all repetitions. The error bars represent 1σ values computed from five exposures at each P/P°.

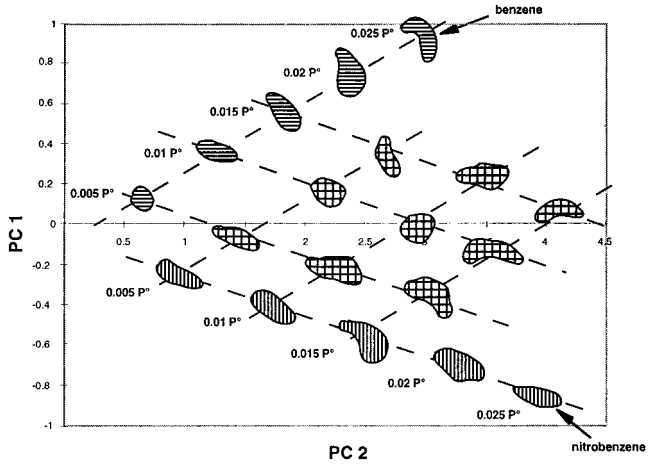


Figure 13. Data in principal component space from a 12-detector array exposed to benzene at $P/P^\circ = 0.005-0.025$ in air in five even steps, nitrobenzene at $P/P^\circ = 0.005-0.025$ in air in five even steps, and binary mixtures of benzene at $P/P^\circ = 0.005$, 0.01, or 0.015, and nitrobenzene at $P/P^\circ = 0.005$, 0.01, or 0.015. The first three principal components contain 99.6% of the total variance in the data. The ellipsoids contain 95% of the data for each analyte. Each analyte was presented five times to the array, with the order of presentation randomized over all repetitions of all exposure types. The error bars represent 1σ values computed from five exposures at each P/P°.

simultaneous and sequential exposures of binary mixtures of benzene and nitrobenzene. Similar behavior was observed for all nine binary mixtures explored in this work (see part D in the Experimental Section for a description of the binary mixtures explored).

DISCUSSION

A. Linearity of Detector Response vs Analyte Concentration. The linearity in $\Delta R_{js,max}/R_{jb}$ response of the conducting organic polymer composite detectors versus the concentration of a pure analyte is readily understood based on the signal trans-

duction mechanism of these types of vapor detectors. Sorption of the vapor into the detector leads to swelling of the polymer, which then produces an increase in the electrical resistance through the network of conducting regions in the composite film. Although the absolute $\Delta R_{js,max}$ of the composite is sensitive to the fractional loading of the conductive filler in the insulating polymer of the conductive material,[30−32] the relative swelling of the film in response to the presence of an analyte vapor should remain constant provided that the filler material does not significantly affect the properties of the insulating portion of the composite. Under such conditions, the ratiometric quantity $\Delta R_{js,max}/R_{jb}$ is expected to be the key parameter that characterizes the response of conducting polymer composite vapor sensors to various analytes of interest. The present work quantitatively confirms these expectations.

On the basis of the expectations discussed above, for small fractional film swellings, the observed $\Delta R_{js,max}/R_{jb}$ response should be a linear function of the concentration of the vapor that partitions into the film. This appears to be the case for the solvents studied during the course of this work. This type of behavior has been observed for poly(pyrrole) conducting polymer vapor sensors[33] and for vapor sensors that monitor the capacitance change of dielectric polymer films in response to the presence of vapor analytes, where again the response is a linear function of the analyte concentration.[34−36] Polymer films that are exposed to analytes that either bind very strongly to the polymer or that induce significant structural distortions in the chains of the polymeric material could certainly produce a saturation of the detector response at concentrations well below the vapor pressure of the analyte; however, such behavior was not observed for any of the solvents or detectors explored in this work.

For mixtures, as long as the concentration of analyte molecules is dilute in the polymer film, the linear swelling relationship as a function of the analyte concentration in the vapor phase is expected to be a good microscopic description of the signal transduction properties of the detectors when exposed to combinations of these same gaseous analytes. Thus, the swelling response of a polymer to binary analyte mixtures is expected to be a weighted linear combination of the response to the individual analytes in the vapor phase. Previous work in our laboratory has shown that the fraction of the partial pressure of the odorant, as opposed to the concentration of the odorant, is the key variable in determining the response of the carbon black/organic polymer composite vapor detectors.[37] Thus, to first order, the response of a polymer composite detector array to a mixture of solvents should be readily obtained by calculating the fractional composition of the constituents in the mixture relative to their individual vapor pressures under the experimental test conditions of concern. This additive behavior is, in fact, in excellent accord with experimental observations for the response of the conducting polymer composite arrays to the binary mixtures studied during the course of this work.

B. Implications for Algorithm Development/Pattern Recognition Requirements. All architectures that rely on array-based sensing require some type of training set and signal-processing algorithm in order to classify and/or identify an analyte upon presentation to the detector array. In this respect, the performance and range of applicability of such detector arrays is intimately coupled to the data reduction algorithms and computational capabilities that are required to achieve the sensing task of concern.

The minimum possible training set, and the minimum requirements on computational capabilities to analyze a mixture or to classify and/or identify a particular analyte, are clearly achieved when the detector response is a linear function of the analyte concentration and when the differential detector response to the analyte of concern is independent of whether other analytes are present in the environment. Both of these conditions were met for the carbon black/organic polymer composite chemiresistor response characteristics over the ranges of concentrations and for the ranges of analyte background concentrations that were explored during the course of this study. This behavior contrasts with the properties reported for tin oxide chemiresistors[38] or for dye-impregnated organic polymer coatings on fiber optics,[5,38] whose responses are nonlinear with analyte concentration and/or with variations in environmental background. Such nonlinearities imply that significantly more computational resources and algorithm development will be required to achieve similar system performance in varying background environments or when an analyte concentration is to be quantified either alone or in a mixture of vapors. The exact tradeoffs imposed by more complex data reduction and more involved computational requirements, relative to the opportunity to exploit possibly increased information content of a richly varying signal response pattern, will be array and task specific and will require a detailed analysis for the specific task of interest.

For odors that are more complex compositionally than simple binary or ternary mixtures of analytes, it could be envisioned that a single array-based detector response fingerprint would not be sufficient to produce a unique vector decomposition of the mixture into the signatures of each of the components of a training set of vapors. Thus, one response pattern might not be sufficient to provide a unique solution to the chemical composition of the odor mixture of concern. For example, if most of the variance among the data is contained in three to five principal components and if the cannonical variance tracks the total variance in the data, then mixtures of only three to five components can be decomposed uniquely from the use of the equilibrium response data alone. It is likely that, even for complex odors, useful information will be obtained, however, if some temporal or spatiotemporal variation in the composition of the odor is present. Under such conditions, changes in detector response can be identified with individual

(30) Anderson, J. E.; Adams, K. M.; Troyk, P. R. *J. Non-Cryst. Sol.* **1991**, *131*, 587−592.

(31) Godovski, D. Y.; Koltypin, E. A.; Volkov, A. V.; Moskvina, M. A. *Analyst* **1993**, *118*, 997−999.

(32) Kirkpatrick, S. *Rev. Mod. Phys.* **1973**, *45*, 574−588.

(33) Persaud, K. C.; Travers, P. J. *Arrays of Broad Specificity Films for Sensing Volatile Chemicals*; Kress-Rogers, E., Ed.; CRC Press, Inc.: New York, 1997; pp 563−592.

(34) Ralston, A. R.; Tobin, J. A.; Bajikar, S. A.; Denton, D. D. *Sens. Actuator, B* **1994**, *22*, 139.

(35) Boltshauser, T.; Leme, C. A.; Baltes, H. *Sens. Actuator, B* **1993**, *15*, 75.

(36) Denton, D. D.; Senturia, S. D.; Anolik, E. S.; Scheider, D. *Digest of Technical Papers, third Int. Conf. on Solid-State Sensors and Actuators (Transducers '85)* **1985**, 202−205.

(37) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5442−5447.

(38) Harsanyi, G. *Polymer Films in Sensor Applications*; Technomic Pub. Co., Inc.: Basel, Switzerland, 1995.

portions of the analyte based on their differential response patterns relative to the integrated baseline response of the odor on the detector array. The detector response characteristic that is least demanding on the signal processing and computational resources under such circumstances is when the pattern for an analyte remains linearly proportional to the analyte concentration regardless of the composition, or concentration, of the other components of the background ambient. This behavior was observed experimentally for the conducting polymer composite detectors for the various solvents and background ambient vapors evaluated in this work.

CONCLUSIONS

Under the conditions of this study, carbon black/organic polymer composite vapor detectors displayed a linear steady-state relative differential resistance signal in response to changes in the concentration of analyte vapor in the gas phase. This behavior was observed relative to either an air background or a background that contained an organic solvent vapor in air. Moreover, the steady-state relative differential resistance response patterns produced by an array of carbon black/polymer composite detectors upon exposure to a test series of binary mixtures of analytes were the arithmetic sums of the maximum relative differential resistance responses that were obtained upon independent exposure of the array to each individual component of the mixture. This behavior implies that, under our test conditions, a relatively simple algorithm and training set, based on identifying a solvent vapor through its pattern type and quantifying the vapor concentration through the pattern height, would be sufficient to identify and quantify the test vapors and test vapor mixtures studied in this work.

ACKNOWLEDGMENT

This work was supported by the National Aeronautics and Space Administration, the Army Research Office, the Defense Advanced Research Projects Agency, and the Department of Energy, B.J.D. acknowledges the Government of Canada for an NSERC 1967 Centennial Graduate Fellowship.

SUPPORTING INFORMATION AVAILABLE

Statistics for all of the sensor−analyte combinations. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review September 7, 1999. Accepted November 16, 1999.

AC9910278

Anal. Chem. **2000,** *72,* 2008−2015

Relationships among Resonant Frequency Changes on a Coated Quartz Crystal Microbalance, Thickness Changes, and Resistance Responses of Polymer−Carbon Black Composite Chemiresistors

Erik J. Severin and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

The relationships among frequency changes on a film-coated quartz crystal microbalance, thickness changes, and dc resistance changes have been investigated for carbon black−insulating polymer composite vapor detectors. Quartz crystal microbalance (QCM) measurements and ellipsometry measurements have been performed simultaneously on polymer films that do not contain carbon black filler to relate the QCM frequency change and the ellipsometrically determined thickness change to the analyte concentration in the vapor phase. In addition, quartz crystal microbalance measurements and dc resistance measurements on carbon black composites of these same polymers have been performed simultaneously to relate the QCM frequency change and dc electrical resistance response to the analyte concentration in the vapor phase. The data indicate that the dc resistance change is directly relatable to the thickness change of the polymers and that a variety of analytes that produce a given thickness change produce a constant resistance change for each member of the test set of polymers investigated in this work.

Carbon black−insulating organic polymer composite films have been employed previously as components of an array of vapor detectors for use in an "electronic nose".[1] In this approach, the response of an array of broadly cross-responsive vapor detectors is analyzed using standard chemometric methods to yield diagnostic patterns that allow classification and quantification of analytes in the vapor phase. Arrays of such detectors have been shown to be highly discriminating, even between very structurally similar analytes, and have also been shown for many test vapors to exhibit a linear steady-state dc resistance response to analyte concentration. Thus, under these conditions, the pattern type allows identification of the vapor and the steady-state pattern height allows quantification of the analyte of concern.[1−3]

The resistance response of such composites can, in general, be understood by percolation theory, which relates the resistance response of a composite consisting of an insulating polymer filled with regions of an electrical conductor to the change in volume fraction of the conducting (filler) phase of the composite.[4−7] The goal of the present work was to elucidate the factors that control the resistance change of such films in response to a change in the concentration of a vapor that is exposed to the detector. In polymer-coated quartz crystal microbalances, the frequency change of the detector is primarily determined by the change in the mass of analyte sorbed into the polymer film for relatively small frequency shifts and/or small changes in the viscoelastic properties of the film.[8a,b] Polymer-coated surface acoustic wave devices, utilize changes in sorbed mass and modulus of the polymer film to produce the detected signal.[8c] The hypothesis that was challenged in this work is that the volume change, and thus the fractional swelling, of the polymer film upon exposure to a test vapor is the key variable that determines the change in dc electrical resistance of the carbon black−polymer composite detectors.

To test this hypothesis, we performed measurements to determine the resonant frequency changes on a film-coated quartz crystal microbalance (QCM), the thickness changes, and the resistance changes of various composite and noncomposite polymer films exposed to a variety of test organic vapors. The resonant frequency changes and the dc electrical resistance changes of a set of carbon black−organic polymer composite films were determined on a QCM. QCM measurements and thickness measurements using fixed-wavelength ellipsometry methods were then performed on clear (non-carbon-black-filled) films formed from the same polymers. Relationships between the two sets of measurements were facilitated because at a given analyte concentration in the vapor phase, the measured QCM resonant frequency changes were very similar for polymers that did, and did not, contain the carbon black filler material.

(1) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298.
(2) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2652−2656.
(3) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177−4190.
(4) Anderson, J. E.; Adams, K. M.; Troyk, P. R. *J. Non-Cryst. Sol.* **1991**, *131*, 587−592.
(5) Ast, D. G. *Phys. Rev. Lett.* **1974**, *33*, 1042−1045.
(6) Godovski, D. Y.; Koltypin, E. A.; Volkov, A. V.; Moskvina, M. A. *Analyst* **1993**, *118*, 997−999.
(7) Kirkpatrick, S. *Rev. Mod. Phys.* **1973**, *45*, 574−588.
(8) (a) Lu, C. In *Applications of Piezoelectric Quartz Crystal Microbalances*; Lu, C. C., Ed.; Elsevier: New York, 1984; Vol. 7, pp 19−61. (b) Buttry, D. A. In *Electroanalytical Chemistry: A Series of Advances*; Bard, A. J., Ed.; Marcel Dekker: New York, 1991; Vol. 17, pp 1−85. (c) Grate, J. W.; Klusty, M.; McGill, R. A.; Abraham, M. H.; Whiting, G.; Andonian-Haftvan, J. *Anal. Chem.* **1992**, *64*, 610−624.

10.1021/ac991026f CCC: $19.00 © 2000 American Chemical Society
Published on Web 04/01/2000

Table 1: Correlation Coefficients, Slopes, Intercepts, Intercept Errors, and Slope Errors for the Eight Solvents and Two Polymer Systems Used in This Work[a]

PCL composite film	Δf^*_{max} vs P/P°					$\Delta R_{max}/R_b$ vs P/P°					$\Delta R_{max}/R_b$ vs Δf^*_{max}				
	R^2	intcpt	slp	intcpt err	slp err	R^2	intcpt	slp	intcpt err	slp err	R^2	intcpt	slp	intcpt err	slp err
hexane	0.9997	0.00	3.64	0.006	0.034	0.9989	0.00	17.99	0.066	0.344	0.9991	0.01	11.61	0.059	0.197
2-propanol	0.9934	0.02	3.44	0.028	0.161	0.9933	0.04	15.93	0.130	0.754	0.9999	−0.05	10.88	0.018	0.070
benzene	0.9976	0.38	15.18	0.078	0.429	0.9968	0.40	76.98	0.456	2.520	0.9987	−1.52	11.89	0.330	0.252
dichloromethane	0.9998	−0.06	20.91	0.046	0.159	0.9993	−0.60	71.15	0.317	1.099	0.9987	−0.38	7.98	0.420	0.165
chloroform	0.9998	0.19	46.08	0.078	0.384	0.9988	0.21	145.40	0.604	2.962	0.9984	−0.38	7.40	0.694	0.170
hexafluorobenzene	0.9937	0.54	13.65	0.113	0.628	0.9969	0.78	35.03	0.203	1.124	0.9987	−0.58	5.99	0.158	0.127
dibromomethane	0.9974	−0.04	55.04	0.280	1.620	0.9960	−1.31	105.29	0.668	3.861	0.9991	−1.23	4.49	0.319	0.079
bromoform	0.9989	0.41	76.41	0.245	1.481	0.9982	−1.16	137.74	0.558	3.375	0.9983	−1.88	4.22	0.558	0.101

PCL clear film	Δf^*_{max} vs P/P°					$\Delta h_{max}/h_b$ vs P/P°					$\Delta h_{max}/h_b$ vs Δf^*_{max}				
	R^2	intcpt	slp	intcpt err	slp err	R^2	intcpt	slp	intcpt err	slp err	R^2	intcpt	slp	intcpt err	slp err
hexane	0.9981	0.01	1.84	0.012	0.031	0.9909	0.02	0.55	0.008	0.021	0.9926	0.01	0.70	0.007	0.024
2-propanol	0.9971	−0.02	3.67	0.040	0.081	0.9932	−0.03	1.10	0.019	0.037	0.9987	−0.02	0.70	0.008	0.010
benzene	0.9956	0.18	10.40	0.039	0.273	0.9928	0.05	3.02	0.015	0.101	0.9967	0.00	0.68	0.011	0.015
dichloromethane	0.9992	−0.13	17.70	0.041	0.192	0.9963	−0.03	3.65	0.018	0.084	0.9950	0.00	0.48	0.020	0.013
chloroform	0.9984	0.00	42.96	0.121	0.685	0.9967	−0.03	8.43	0.033	0.189	0.9984	−0.03	0.46	0.023	0.007
hexafluorobenzene	0.9973	0.15	9.20	0.097	0.175	0.9984	−0.02	1.43	0.012	0.021	0.9985	−0.04	0.36	0.011	0.005
dibromomethane	0.9980	0.56	44.16	0.154	0.667	0.9989	−0.01	5.88	0.015	0.064	0.9993	−0.08	0.31	0.012	0.003
bromoform	0.9977	−0.05	66.71	0.187	1.256	0.9977	−0.05	7.94	0.022	0.151	0.9981	−0.04	0.28	0.020	0.005

PEO composite film	Δf^*_{max} vs P/P°					$\Delta R_{max}/R_b$ vs P/P°					$\Delta R_{max}/R_b$ vs Δf^*_{max}				
	R^2	intcpt	slp	intcpt err	slp err	R^2	intcpt	slp	intcpt err	slp err	R^2	intcpt	slp	intcpt err	slp err
hexane	0.9985	0.01	0.97	0.003	0.014	0.9990	−0.02	9.89	0.025	0.117	0.9970	−0.06	23.83	0.044	0.479
2-propanol	0.9972	0.01	2.19	0.009	0.054	0.9992	0.11	22.57	0.048	0.305	0.9979	−0.02	24.13	0.079	0.527
benzene	0.9996	0.02	7.92	0.013	0.072	0.9963	0.13	69.49	0.358	2.006	0.9967	−0.07	20.59	0.342	0.561
dichloromethane	0.9993	0.13	16.60	0.073	0.255	0.9991	0.50	91.89	0.464	1.615	0.9986	−0.20	12.97	0.591	0.284
chloroform	0.9985	0.22	39.47	0.182	0.879	0.9975	1.34	210.58	1.260	6.096	0.9991	0.16	12.52	0.758	0.213
hexafluorobenzene	0.9979	0.11	6.91	0.034	0.185	0.9997	0.25	24.76	0.047	0.259	0.9968	−0.12	8.37	0.157	0.272
dibromomethane	0.9985	0.38	45.18	0.179	1.028	0.9975	0.47	143.09	0.726	4.176	0.9985	−0.73	7.43	0.586	0.168
bromoform	0.9989	0.23	55.20	0.179	1.069	0.9992	0.40	155.11	0.420	2.510	0.9998	−0.24	6.59	0.235	0.058

PEO clear film	Δf^*_{max} vs P/P°					$\Delta h_{max}/h_b$ vs P/P°					$\Delta h_{max}/h_b$ vs Δf^*_{max}				
	R^2	intcpt	slp	intcpt err	slp err	R^2	intcpt	slp	intcpt err	slp err	R^2	intcpt	slp	intcpt err	slp err
hexane	0.9975	0.00	0.60	0.006	0.014	0.9922	0.00	0.26	0.004	0.011	0.9945	0.00	1.01	0.004	0.036
2-propanol	0.9986	0.01	2.12	0.012	0.040	0.9980	0.02	0.92	0.006	0.020	0.9989	0.01	1.01	0.004	0.017
benzene	0.9966	0.05	5.72	0.021	0.167	0.9929	0.03	2.41	0.013	0.102	0.9963	0.01	0.99	0.010	0.030
dichloromethane	0.9916	0.05	14.55	0.094	0.600	0.9892	0.04	3.88	0.028	0.182	0.9970	0.02	0.63	0.015	0.015
chloroform	0.9975	0.04	38.52	0.052	0.789	0.9957	0.02	9.94	0.018	0.267	0.9983	0.01	0.61	0.011	0.010
hexafluorobenzene	0.9986	−0.01	2.05	0.016	0.031	0.9981	0.00	0.39	0.003	0.007	0.9989	0.01	0.44	0.003	0.006
dibromomethane	0.9978	0.13	40.44	0.062	0.742	0.9974	0.03	7.48	0.013	0.150	0.9985	0.00	0.43	0.010	0.006
bromoform	0.9905	0.06	51.26	0.139	1.972	0.9905	0.01	8.61	0.023	0.331	0.9986	0.00	0.39	0.009	0.006

[a] Mass-normalized maximum resonant frequency change, Δf^*_{max}, vs the fraction of the analyte's vapor pressure (Δf^*_{max} vs P/P°), relative differential resistance increase vs the fraction of the analyte's vapor pressure ($\Delta R_{max}/R_b$ vs P/P°), relative differential thickness increase vs the fraction of the analyte's vapor pressure ($\Delta h_{max}/h_b$ vs P/P°), relative differential resistance increase vs mass-normalized maximum resonant frequency change ($\Delta R_{max}/R_b$ vs Δf^*_{max}), and relative differential thickness increase vs mass-normalized maximum resonant frequency change ($\Delta h_{max}/h_b$ vs Δf^*_{max}) are tabulated.

EXPERIMENTAL SECTION

QCM crystals (10 MHz, blank diameter = 13.7 mm) with a custom electrode pattern were obtained from International Crystal Manufacturing (ICM), Oklahoma City, OK. The standard oscillation electrodes were configured at 90° angles to make room for two other tabs that would serve as electrodes for resistance measurements of the carbon black−polymer composite films (Figure 1). The crystals were polished to a surface roughness of less than 5 μm, which produced a mirrorlike finish on the gold electrodes. To facilitate reflection of the ellipsometer's laser beam when the crystals were used with transparent films during the thickness measurements, one oscillator electrode was larger than the other (larger electrode diameter = 7.8 mm, smaller electrode diameter = 5.1 mm). The resistance tabs were not used during the thickness vs QCM frequency measurements on films that were not filled with carbon black. Similarly, the ellipsometer was not used during the resistance vs QCM frequency measurements in which optically opaque, carbon-black-filled, composite films were used.

The vapor stream was produced by passing general laboratory compressed air through analyte solvents contained in custom bubblers. The solvents were of HPLC quality (Aldrich Chemical Co.) and were used as received. Saturation of the vapor with solvent was confirmed by mass loss experiments.[9] The solvent-saturated air was then diluted to the desired concentration with house air. The air flows through the bubbler and in the back-

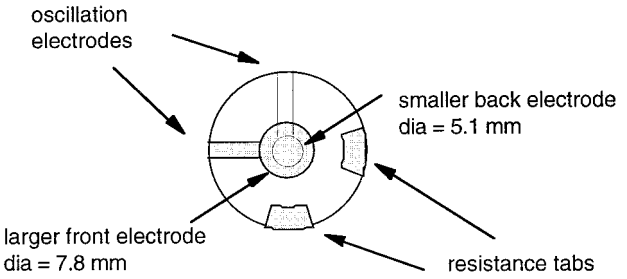


Figure 1. Custom 10 MHz quartz crystal microbalance with oscillation electrodes and tabs for reading the resistance of the composite film. Shaded areas indicate regions coated with Au. The larger electrode was used to facilitate ellipsometry measurements.

ground gas were regulated by needle valves, and the flows in both streams were monitored with Gilmont rotamers (VWR Scientific). The concentration of analyte in the vapor stream was independently verified using a calibrated flame ionization detector (California Analytical, Santa Ana, CA).

Two polymers were used in this work, poly(caprolactone) (PCL) and poly(ethylene oxide) (PEO). Films of these polymers that contained carbon black were used for the resistance measurements, while transparent, pure polymer films were used for the thickness measurements. All films were cast from standard solutions that consisted of 160 mg of polymer dissolved in 20 mL of benzene to which 40 mg of carbon black was added to make composite films (resulting in solutions that were 20 wt % carbon black). All solutions were sonicated for at least 5 min immediately prior to casting the films. The polymer films were spun-cast on a spin coater (Headway Research, Garland, TX) at 2000 rpm, and the average film thickness was obtained by profilometry (Dektak 3030, Sloan Technology Corp., Santa Barbara, CA).

The QCM crystals were weighed before and after film application using a Cahn microbalance (resolution 0.001 mg; Cahn C-35, Orion Research, Beverly, MA) to obtain the masses of the films that were deposited over the active electrode (5.1 mm diameter area) of the QCM. The PCL clear film mass was 40 μg cm^{-2} and 375 nm thick, while the PCL−carbon black composite film mass was 185 μg cm^{-2} with a baseline resistance of ≈12 kΩ. The PEO clear film had a mass of 120 μg cm^{-2} and a thickness of 1090 nm, while the PEO−carbon black composite film was 19 μg cm^{-2} with a baseline resistance of ≈16 kΩ. Using the clear polymer film areas and the mass and thickness values above, densities for the clear films of PEO and PCL were calculated and agreed with literature values for these polymers.

Resistances were measured using a 2002 digital multimeter (Keithley, Cleveland, OH), and the resonant frequency of the QCM was obtained using a 5384A frequency counter (Hewlett-Packard, Palo Alto, CA). Ellipsometry measurements were taken on an L116C ellipsometer (Gaertner Scientific, Chicago, IL). Optical constants were obtained for each surface before the films were applied. The index of refraction of each polymer film was taken from the literature. The absorption coefficient for the film was obtained using the two-angle technique,[10,11] which also provided an independent measurement of the index of refraction and thickness of the film. The film thicknesses obtained by ellipsometry agreed to within 10% with the values obtained by profilometry.

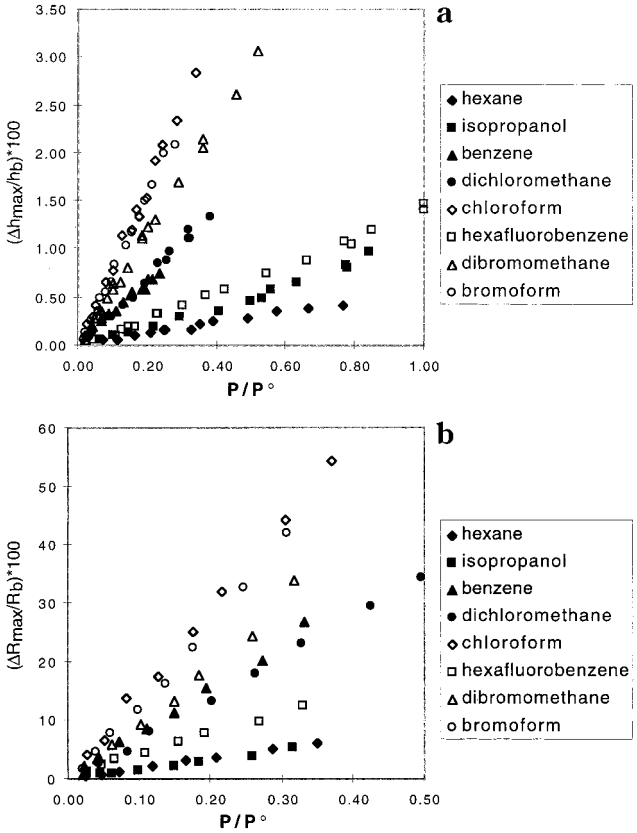


Figure 2. (a) Relative differential thickness increase for a pure PCL film vs fraction of analyte vapor pressure exposed to the film. (b) Differential relative resistance increase in a PCL−carbon black composite vs fraction of analyte vapor pressure exposed to the film.

To initiate an experiment, a baseline value was recorded for the QCM resonant frequency, resistance, and/or thickness of the film. The film was then exposed to analyte vapor until steady-state values were reached as determined by constant output readings from the instruments. The data were recorded manually for convenience. Thickness measurements were taken three to five times after steady state had been reached for a given vapor, and the average result was recorded for both the baseline and the steady-state, solvent-exposed values.

RESULTS

Figure 2a shows the relative thickness change, $\Delta h_{max}/h_b$, where Δh_{max} is the thickness change of the film during exposure to the analyte vapor and h_b is the baseline thickness of the film in air prior to analyte exposure, of poly(caprolactone) films as a function of the fraction of the analyte's vapor pressure, $P/P°$. The series of test vapors used in these experiments is representative of a broad test set of analytes that have been used previously to investigate the discrimination ability of arrays of conducting polymer composite vapor detectors.[1−3] The data of Figure 2a are well-fit to a linear dependence of $\Delta h_{max}/h_b$ on $P/P°$ (Table 1). Figure 3a shows similar data for poly(ethylene oxide) films.

Figures 2b and 3b depict the steady-state relative differential resistance responses, $\Delta R_{max}/R_b$, where ΔR_{max} is the resistance

(9) Atkins, P. W. *Physical Chemistry*; W. H. Freeman and Co.: New York, 1994.
(10) Comfort, J. C.; Urban, F. K.; Barton, D. *Thin Solid Films* **1996**, *291*, 51.
(11) Urban, F. K. *Appl. Surf. Sci.* **1988**, *33*, 934.

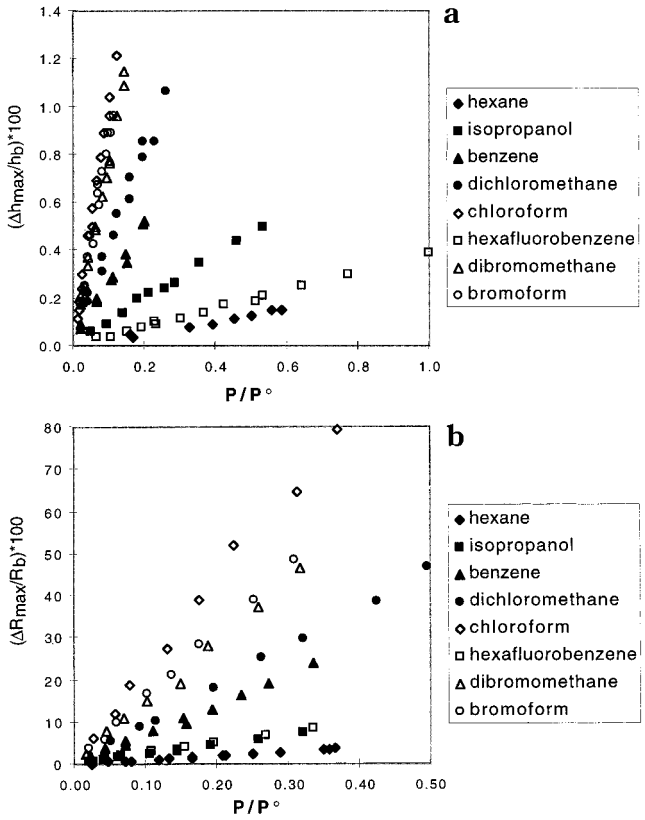


Figure 3. (a) Relative differential thickness increase for a pure PEO film vs fraction of analyte vapor pressure exposed to the film. (b) Differential relative resistance increase in a PEO−carbon black composite vs fraction of analyte vapor pressure exposed to the film.

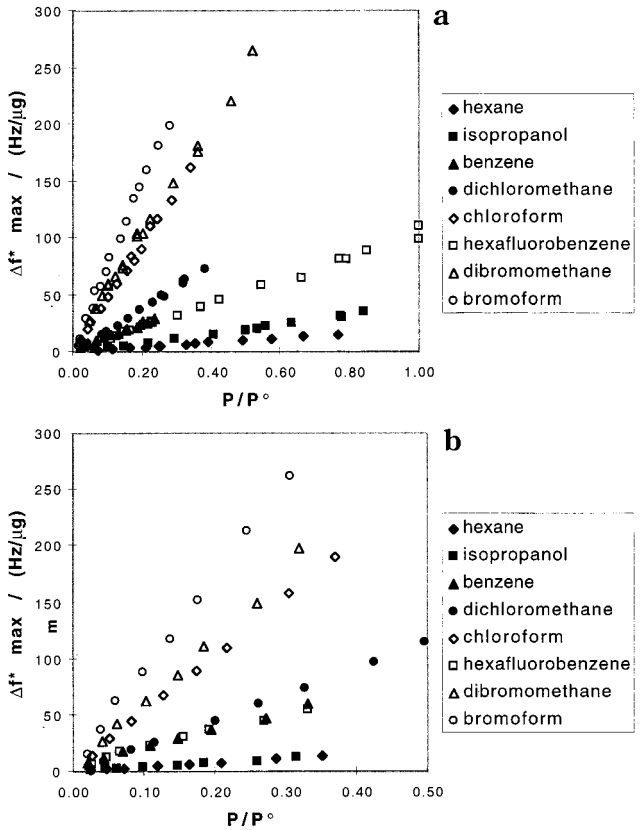


Figure 4. Mass-normalized maximum resonant frequency change vs fraction of analyte vapor pressure exposed to the film for (a) a PCL film without carbon black and (b) a PCL−carbon black composite.

change of the film during exposure to the analyte vapor and R_b is the baseline resistance of the film in air prior to analyte exposure, of carbon-black filled poly(caprolactone) and poly(ethylene oxide) films, respectively, as a function of the analyte concentration, for the same set of test analytes. Over the concentration ranges probed in the experiment, the data are well-fit by straight lines passing though the origin (Table 1).

Figure 4 depicts the mass-normalized maximum resonant frequency change, Δf^*_{max}, of the poly(caprolactone) films on a QCM crystal during exposure to the analyte vapor. The frequency shifts were all negative upon sorption of analyte, and only absolute values of Δf are reported herein. The observed resonant frequency change was normalized by the mass of the films (as determined by the Cahn microbalance mesurements) in the active QCM area to remove any variability due to the use of different film thicknesses and/or film masses between experiments on a given type of polymer. Figure 5 depicts the same data for poly(ethylene oxide) films. Data are depicted for films of polymer that were and were not, respectively, filled with carbon black. Again the data are well-fit by straight lines over the analyte concentration range of experimental interest (Table 1). For all the solvents, the Δf^*_{max} value of a pure polymer film was the same as the Δf^*_{max} value of the analogous carbon-black-filled composite to within the error in the measurements. For example, Figure 6 depicts the Δf^*_{max} value as a function of P/P° for $CHCl_3$ sorbed in poly(caprolactone)− and poly(ethylene oxide)−carbon black composites and in pure polymer films without carbon black added.

DISCUSSION

Figure 7a depicts a plot of the relative differential dc resistance change of the poly(caprolactone) film, from electrical measurements, as a function of the fractional swelling of the polymer, as determined by optical ellipsometry measurements. The same analysis for a second poly(caprolactone) film is shown in Figure 7b to illustrate the variance in the data. For both polymer systems, the slopes and intercepts of the $\Delta R_{max}/R_b$ vs P/P° data for the composite films were used to predict what values of $\Delta R_{max}/R_b$ would be expected for the P/P° values used in the measurements for the nonfilled polymer films. Likewise, the slopes and intercepts of the $\Delta h_{max}/h_b$ vs P/P° data for the nonfilled polymer films were used to predict what values of $\Delta h_{max}/h_b$ would be expected for the P/P° values used in the measurements for the composite films. The predicted values of $\Delta R_{max}/R_b$ were then plotted vs the predicted $\Delta h_{max}/h_b$ values at the corresponding P/P° values of the analytes. As displayed in Figure 7, the data are linear and roughly fall on the same line for all of the test vapors investigated in this work. For each film, some solvents do not lie on the common line, but this is presumed to be due to experimental error in the delivery of the vapor.[12] A robust interpretation of the relatively small deviations of the behaviors of the various analytes on a given polymer film vs the common line would require implementation of methods that could determine the resistance and thickness changes simultaneously on one detector, and such methods were not available in this study. The data of Figure 7 clearly indicate that, regardless of the analyte used, a given

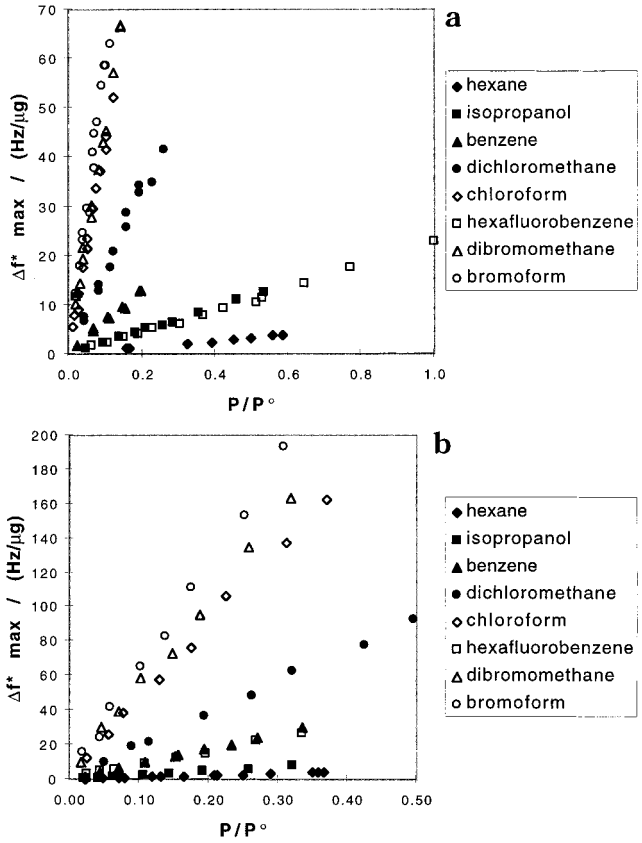


Figure 5. Mass-normalized maximum resonant frequency change vs fraction of analyte vapor pressure exposed to the film for (a) a PEO film without carbon black and (b) a PEO−carbon black composite.

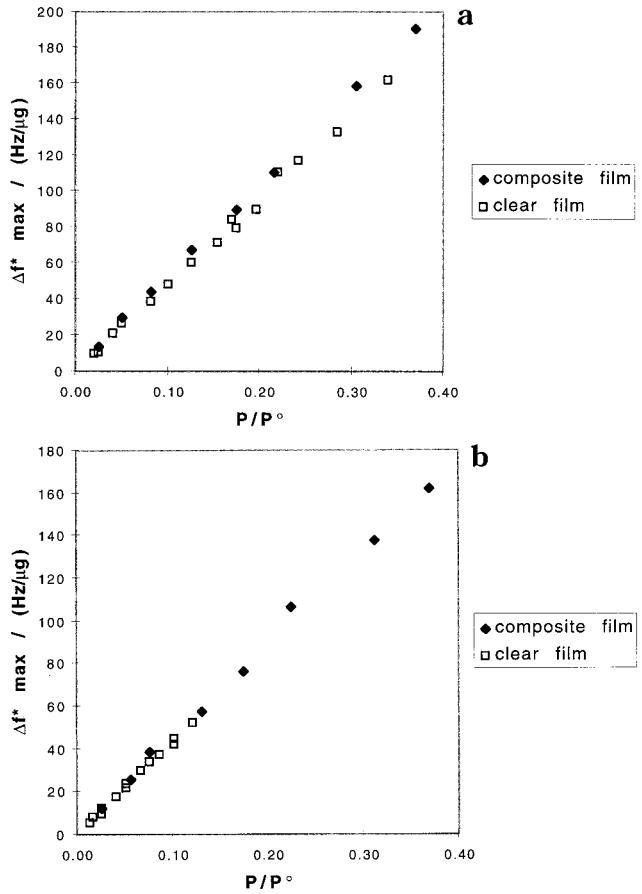


Figure 6. Mass-normalized maximum resonant frequency change vs fraction of analyte vapor pressure exposed to the film for (a) PCL−carbon black composites and PCL films without carbon black and (b) PEO−carbon black composites and PEO films without carbon black.

fractional thickness change of the polymer produces a given steady-state relative differential resistance response of the corresponding carbon-filled composite, at least for the polymer−analyte combinations explored in this work. Thus, the hypothesis of concern—that volumetric film swelling is the key variable determining $\Delta R_{max}/R_b$ in the composite carbon black−insulating polymer detectors—seems to be confirmed from the data obtained in this work, at least for the analytes and polymers investigated to date. Also, these data indicate that the relationship between relative thickness change and steady-state relative differential resistance change is linear, at least over the range of analyte concentrations investigated in this work.

One complicating factor is that the thickness measurements obtained in this work were performed on pure polymeric materials, while the $\Delta R_{max}/R_b$ measurements were performed on carbon-black-filled polymer composites. The assumption made above in interpreting the data of Figure 7 is that the volumetric swellings of the polymers are similar whether or not the material is filled with carbon black. Given the linear relationship deduced between $\Delta h_{max}/h_b$ and $\Delta R_{max}/R_b$ and the low likelihood that, over a range of analytes and concentrations, two separate functional dependencies of swelling on analyte concentration would precisely counteract each other to yield the data of Figure 7, this assumption seems quite reasonable. Given the linear dependence of $\Delta R_{max}/R_b$ on $P/P°$ that has been observed for other test analytes,[13] it seems reasonable to assume that the relationship between relative volumetric swelling and relative differential resistance measurements is extendible, at least to first order, for those composite−analyte combinations as well.

An independent check on the validity of the relationship between swelling in the carbon-black-filled composites and the pure polymer films is available from the QCM resonant frequency measurements. The relationship between $\Delta R_{max}/R_b$ vs Δf^*_{max} and $\Delta h_{max}/h_b$ vs Δf^*_{max} is linear as seen in Figure 8 for PCL and in Figure 9 for PEO (see also Table 1). The slopes and intercepts of the $\Delta R_{max}/R_b$ vs Δf^*_{max} data for the composite films were used to predict what values of $\Delta R_{max}/R_b$ would be expected for the Δf^*_{max} values measured for the nonfilled polymer films at the various analyte concentrations used in the measurements. Likewise, the slopes and intercepts of the $\Delta h_{max}/h_b$ vs Δf^*_{max} data for the nonfilled polymer films were used to predict what values of

(12) The relationship between the resistance change and the volume swellling depends on the carbon-black loading and other parameters involved in making the films. Under controlled conditions where several films are made in a single batch process from a single carbon black−polymer suspension, the variability in response between films to a given analyte concentration is typically less than 10%. The higher variability between the two films of Figure 7 results from the fact that they were made on two separate occasions with no attempt to control fully the deposition process or the suspension properties for consistency between batches.

(13) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 658−668.

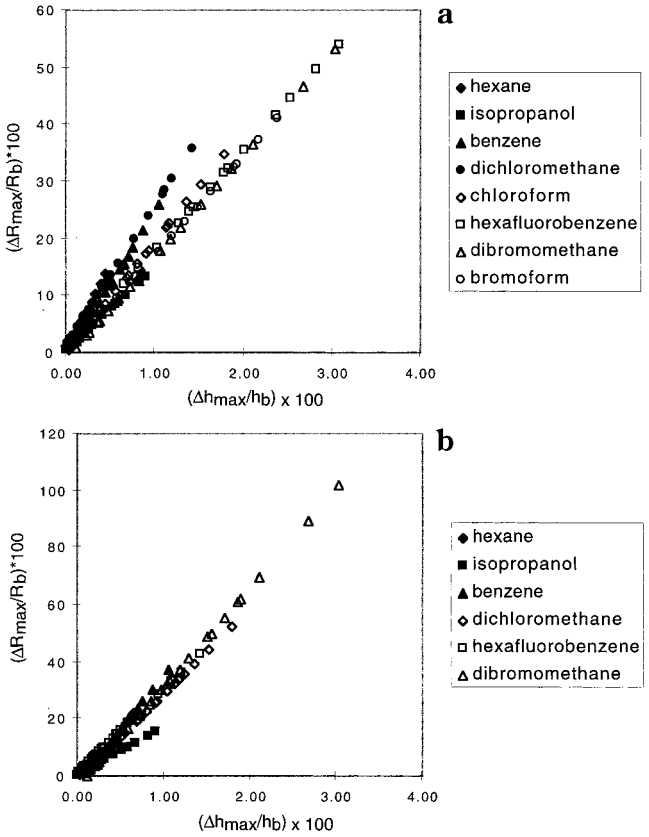


Figure 7. Relative differential resistance increase for a PCL−carbon black composite film vs relative differential thickness increase for a PCL clear film when both films were exposed to various analytes at various fractional vapor pressures, correlated by the analyte fractional vapor pressure for two separate (a, b) PCL films.

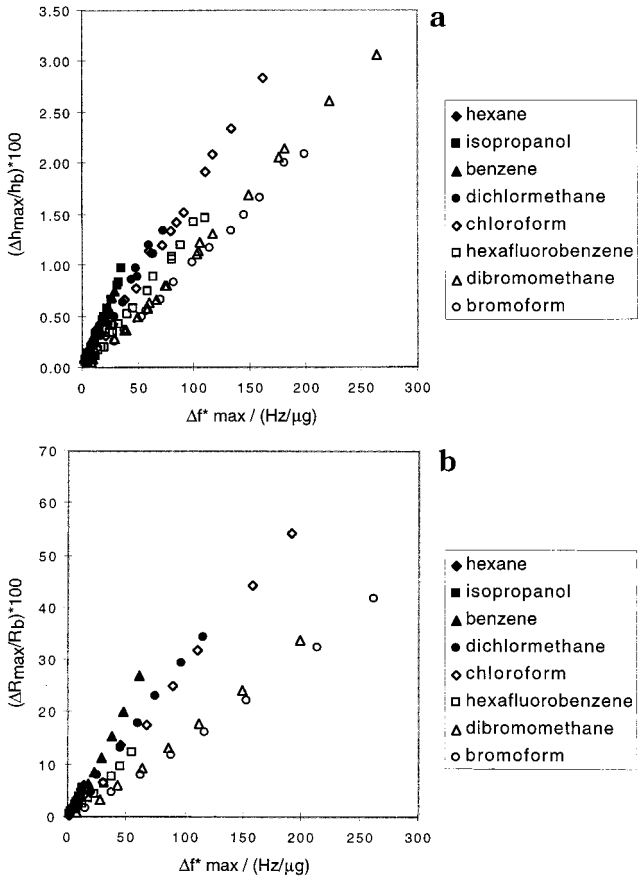


Figure 8. (a) Relative differential thickness increase vs mass-normalized maximum resonant frequency change for a PCL film exposed to various analyte pressures. (b) Relative differential resistance increase vs mass-normalized maximum resonant frequency change for a PCL film exposed to various analyte pressures.

$\Delta h_{max}/h_b$ would be expected for the Δf^*_{max} values measured for the composite films at the various analyte concentrations used in the measurements. The predicted values of $\Delta R_{max}/R_b$ were then plotted vs the predicted $\Delta h_{max}/h_b$ values at the corresponding P/P° values of the analytes. As displayed in Figure 10, the data for each solvent are linear and roughly fall on the same line for all of the test vapors investigated in this work. This strongly implies the presence of a correlation between volume change and resistance change in these composite films. This is a stronger indicator than the correlation using P/P° because the Δf^*_{max} value for each presentation for each film was taken simultaneously with the $\Delta R_{max}/R_b$ and $\Delta h_{max}/h_b$ measurements. The correlation calculated from P/P° presented above was less precise because of variance in the flow system, whereas any changes in the concentration of the exposed analyte would be reflected in the Δf^*_{max} value as well.

In our work, the frequency shift of the polymer-coated QCM crystals arising from sorption of the analyte vapor was <2% of the resonant frequency of the polymer-coated crystal. Under such conditions, prior work has concluded that mechanical losses are minimal and that the frequency shifts are predominantly due to changes in mass uptake.[8a] Although under such conditions the frequency shifts observed in the QCM data can be related, through the proportionality between Δf_{max} and Δm_{max} implied by the Sauerbrey equation,[8a,b] to the fractional mass uptake of these films, the validity of this relationship is not necessary to support any of the key conclusions of our study. Regardless of the physical phenomena that produce a shift in the resonant frequency of the polymer-coated QCM crystals upon vapor sorption, it is clear that the key variable correlating with the $\Delta R_{max}/R_b$ responses of various analytes for a given type of polymer is not Δf_{max}, Δf^*_{max}, or $\Delta m_{max}/m_b$ (with Δm_{max} deduced from Δf_{max} through the Sauerbrey equation) but instead that the experimentally observed correlation is with $\Delta h_{max}/h_b$.

Further support for the swelling-induced resistance change hypothesis can be obtained by investigating the relationship between $\Delta R_{max}/R_b$ and $\Delta h_{max}/h_b$ as a function of analyte density. As seen in Figures 11 and 12, the slopes of the $\Delta h_{max}/h_b$ vs Δf^*_{max} lines and the $\Delta R_{max}/R_b$ vs Δf^*_{max} lines depend linearly on the density (as measured in the pure liquid phase) of the sorbing species. These data are in agreement with recently reported results that were obtained in parallel with our study, in which the relative differential resistance response of carbon-black-filled poly(ethylene oxide) composites was shown to correlate with the density of the gaseous analyte (as measured in its pure liquid phase).[14]

These data support the hypothesis that the resistance response is primarily induced by a change in the volume of the film as reflected in the thickness change. A straight line of any slope for $[\Delta f^*_{max}/(\Delta R_{max}/R_b)]$ vs density that goes through the origin

(14) Swann, M. J.; Glidle, A.; Cui, L.; Barker, J. R.; Cooper, J. M. *Chem. Commun.* **1998**, 2753−2754.

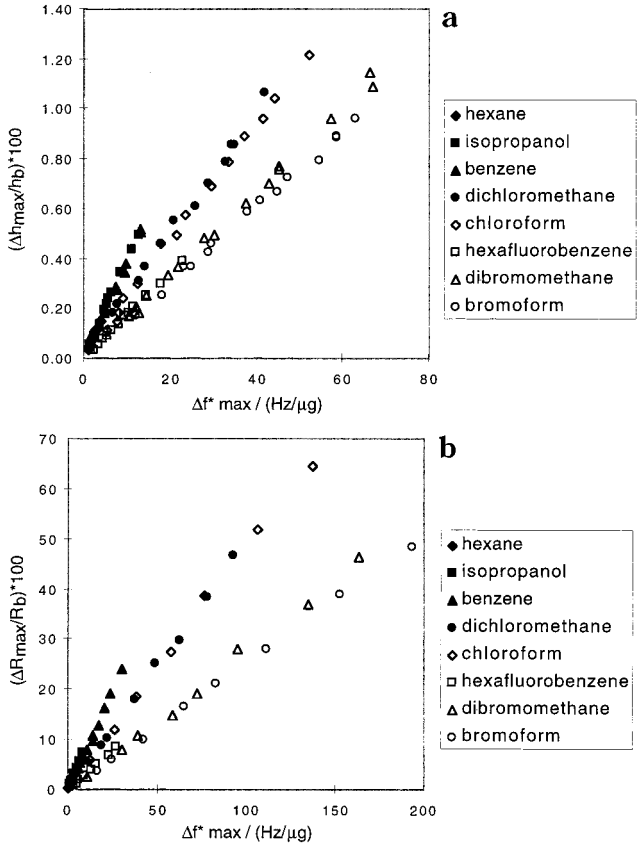


Figure 9. (a) Relative differential thickness increase vs mass-normalized maximum resonant frequency change for a PEO film exposed to various analyte fractional vapor pressures. (b) Differential relative resistance increase vs mass-normalized maximum resonant frequency change for a PEO film exposed to various analyte fractional vapor pressures.

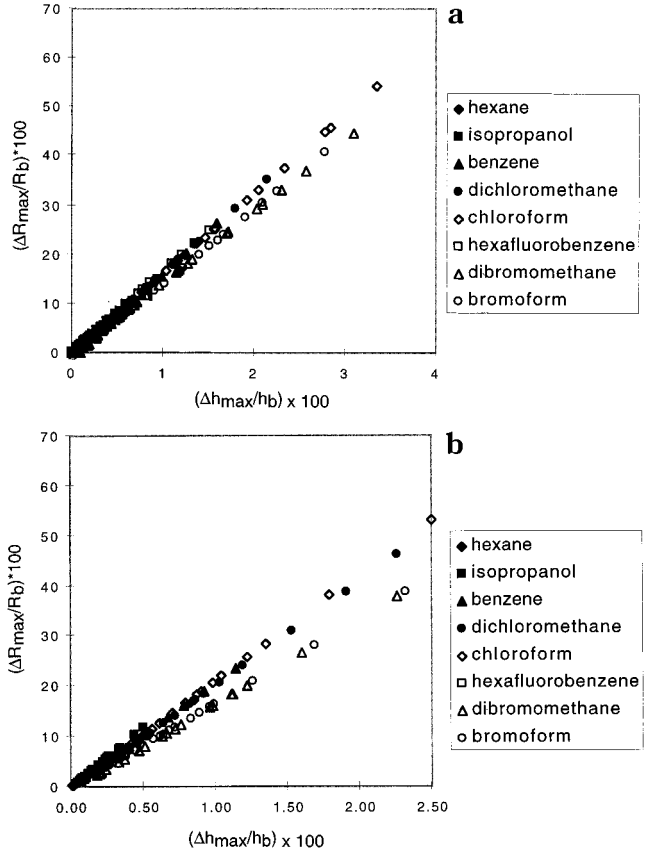


Figure 10. Relative differential resistance increase for a polymer−carbon black composite film vs relative differential thickness increase for a polymer clear film when both films were exposed to various analytes at various pressures, correlated by the mass-normalized maximum QCM resonant frequency change in each film recorded during those analyte exposures for (a) PCL and (b) PEO films.

would imply a precise correlation between the density and the detector response. The $[\Delta f^*_{max}/(\Delta R_{max}/R_b)]$ ratio for hexafluorobenzene is larger in all cases, most likely because the molecules do not chemisorb into the polymer matrix in proportion to the amount that physisorbs because molecular interactions between the perfluoroinated analyte and the polymer chains are not likely to be sufficiently favorable energetically to disrupt the polymer interchain interactions. This would cause an increase in QCM resonant frequency response for hexafluorobenzene (due to adsorption) without a concomitant increase in resistance or thickness response (which requires absorption), leading to larger $[\Delta f^*_{max}/(\Delta R_{max}/R_b)]$ and $[\Delta f^*_{max}/(\Delta h_{max}/h_b)]$ ratios for that solvent.

Generally, the slope of the line for the thickness response vs the density is about an order of magnitude larger than the slope of the line for the related resistance response measurements. In both the thickness and resistance measurements the Δf^*_{max} responses are similar; therefore, the difference in slopes is due to differences in relative response between the thickness and resistance measurements. In all cases, the relative differential resistance response is greater than the relative thickness change for a given Δf^*_{max} change. This finding is consistent with percolation theory, which relates the fractional volume change of a conductor in a composite to a fractional resistivity change of that composite for a given initial conductor volume fraction. We are unable to make direct comparisons with percolation theory because we do not have a complete understanding of the morphology of the carbon black in the composites; however, these data are consistent with reasonable values for the variables in the percolation theory equation for high-conductivity carbon black.[15−17]

An implication of these findings is that low-density analytes will cause a larger resistance response in our detectors for a given Δf^*_{max} value. We have shown in prior work that the amount of analyte that sorbs into these detector films is a function of the fraction of vapor pressure of the analyte.[18] This P/P° dependence accounts for most of the response by a detector to an analyte, but the differences in response by a detector to a set of analytes are due to differences in chemical affinity between the polymer film and the analytes as well as the molecular properties of the analytes such as their molecular volume. Therefore, at the same level of sorption (mass uptake), a lower density analyte will be easier to detect than a higher density analyte.

In conclusion, we have shown that the composite detectors respond according to the volume change of the composite film

(15) Ali, M. H.; AboHashem, A. *J. Mater. Process. Technol.* **1997**, *68*, 163.
(16) Ali, M. H.; AboHashem, A. *J. Mater. Process. Technol.* **1997**, *68*, 168.
(17) Ali, M. H.; AboHashem, A. *Plast. Rubber Compos. Process. Appl.* **1995**, *24*, 47.
(18) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5442−5447.

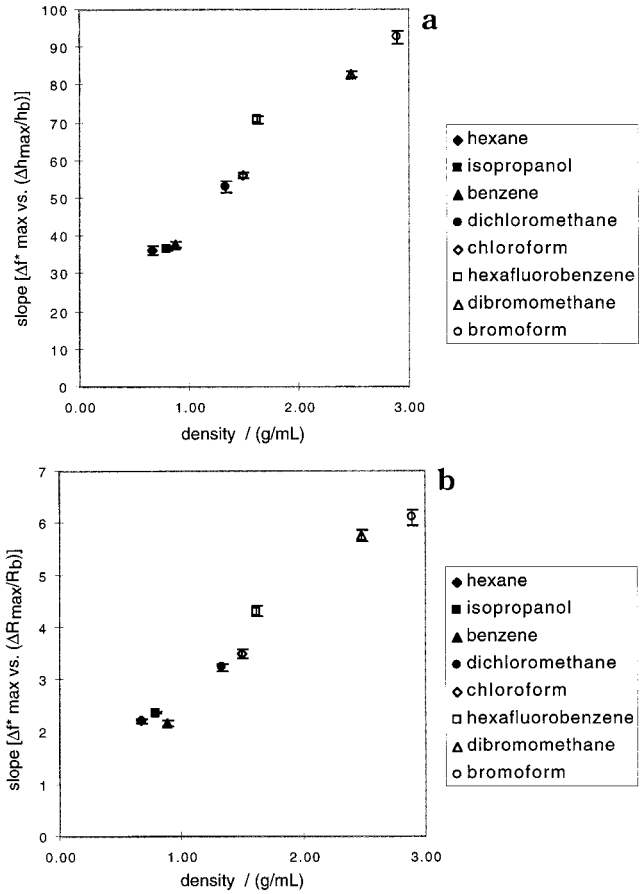


Figure 11. (a) Value of the slope of the line corresponding to $[\Delta f^*_{max}/(\Delta h_{max}/h_b)]$ for a clear PCL film for various analyte presentations at various analyte fractional vapor pressures vs the analyte liquid-phase density for the exposed analyte. (b) Value of the slope of the line corresponding to $[\Delta f^*_{max}/(\Delta R_{max}/R_b)]$ for a PCL−carbon black composite film for various analyte presentations at various analyte fractional vapor pressures vs the analyte liquid-phase density for the exposed analyte.

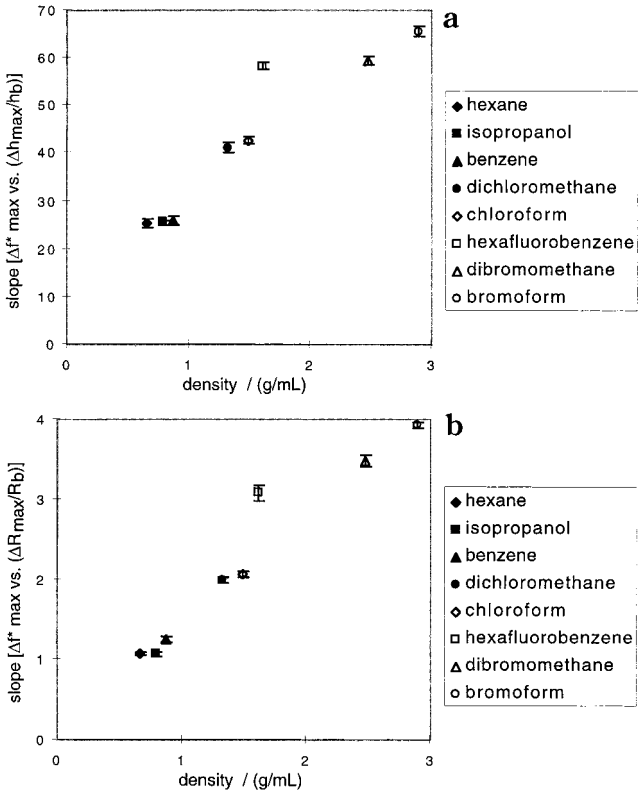


Figure 12. (a) Value of the slope of the line corresponding to $[\Delta f^*_{max}/(\Delta h_{max}/h_b)]$ for a clear PEO film for various analyte presentations at various analyte pressures vs the analyte liquid-phase density for the exposed analyte. (b) Value of the slope of the line corresponding to $[\Delta f^*_{max}/(\Delta R_{max}/R_b)]$ for a PEO−carbon black composite film for various analyte presentations at various analyte pressures vs the analyte liquid-phase density of the exposed analyte.

as evidenced by a linear dependence on the analyte densities of the slopes of the lines for the thickness and resistance responses vs film-coated QCM resonant frequency change and by a linear relationship between percent resistance change and percent thickness change when these two are correlated by the film-coated QCM resonant frequency changes. Additionally, we have developed a single-element densitometer that can be used to characterize in a convenient manner a molecular property of many different types of analytes presented to these types of detectors.

ACKNOWLEDGMENT

This work was supported by the National Aeronautics and Space Administration, the Army Research Office, and the Defense Advanced Research Projects Agency.

Received for review September 7, 1999. Accepted January 11, 2000.

AC991026F

Anal. Chem. **2000,** *72,* 3181−3190

Preparation and Properties of Vapor Detector Arrays Formed from Poly(3,4-ethylenedioxy)thiophene−Poly(styrene sulfonate)/Insulating Polymer Composites

Gregory A. Sotzing, Shawn M. Briglin, Robert H. Grubbs,* and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Poly(3,4-ethylenedioxy)thiophene−poly(styrene sulfonate) (PEDOT−PSS) was used as the conductive component in a matrix of chemically different insulating polymers to form an array of vapor detectors. Such composites produced larger relative differential resistance responses when exposed to polar analytes than did the corresponding carbon black filled polymer composite detectors. However, the PEDOT−PSS composites produced smaller responses than carbon black composites when exposed to nonpolar analytes. The resolving power of a PEDOT−PSS detector array was compared to that of a carbon black composite array for a broadly construed set of organic vapors. The PEDOT−PSS array exhibited better, on average, discrimination between pairs of polar analytes and polar/nonpolar analytes than did the carbon black composite array. The carbon black composite array outperformed the PEDOT−PSS array in discriminating between nonpolar compounds. The addition of PEDOT−PSS composites to an array of carbon black composite detectors therefore can produce improved overall discrimination in a vapor sensor system when used in tasks to differentiate between of a broad set of analyte vapors.

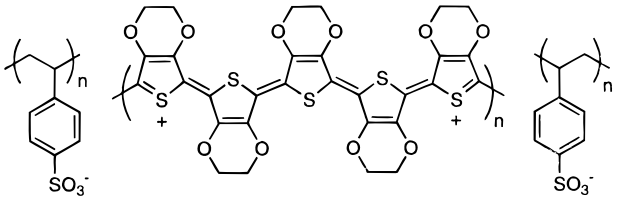
Sensor arrays have recently attracted significant interest for classification and quantification of analytes in liquid and gaseous environments.[1−16] Such "artificial noses" generally consist of an array of detectors, in which any individual detector responds to many analytes and any individual analyte elicits a response from many detectors.[17−21] In this architecture, the array response output provides a characteristic fingerprint for each analyte, and detection of this analyte involves use of pattern recognition methods on the array data stream.

Intrinsically conductive polymers are an attractive class of materials for use in such sensor arrays.[22−26] However, obtaining chemical diversity in intrinsically conductive polymers involves significant synthetic effort, such as modification of the polymer backbone or modification of substituient groups. In addition, polymer backbones commonly used in intrinsically conductive polymers, such as polypyrrole and polythiophene, have electrical properties that are quite sensitive to humidity and to other environmental variations. Another approach is to use composites formed from a conducting material dispersed into a variety of insulating polymeric phases. Carbon black, Ag, Au, or conducting organic polymers have all been used as the conductive phases in such composites.[27] A wide variety of insulating polymers can be used as components of these types of detector elements, resulting in a chemically diverse set of resistive detectors having good vapor classification properties.[27] One shortcoming of the existing carbon black composite preparations is that it is not straightforward to prepare composites using highly polar polymers because the hydrophobic carbon particles do not disperse well into polar

(1) Gardner, J. W.; Pearce, T. C.; Friel, S.; Bartlett, P. N.; Blair, N. *Sens. Actuators, B* **1994**, *18*, 240−243.
(2) Pearce, T.; Gardner, J. *Analyst* **1998**, *123*, 2047−2055.
(3) Gardner, J. W.; Gardner, W. E. *Insight* **1997**, *39*, 865−869.
(4) Pearce, T. C.; Gardner, J. W.; Friel, S.; Bartlett, P. N.; Blair, N. *Analyst* **1993**, *118*, 371−377.
(5) Baldacci, S.; Matsuno, T.; Toko, K.; Stella, R.; De Rossi, D. *Sens. Mater.* **1998**, *10*, 185−200.
(6) Hodgins, D.; Simmonds, D. *J. Autom. Chem.* **1995**, *17*, 179−185.
(7) Hirschfelder, M.; Ulrich, D.; Hoberg, E.; Hanrieder, D. *Gartenbauwissenschaft* **1998**, *63*, 185−190.
(8) Simon, J. E.; Hetzroni, A.; Bordelon, B.; Miles, G. E.; Charles, D. J. *J. Food Sci.* **1996**, *61*, 967.
(9) Broda, S.; Schnitzler, W. *Dtsch. Lebensm.-Rundsch.* ***1998***, *94*, 13−16.
(10) Stuetz, R. M.; White, M.; Fenner, R. A. *J. Water Serv. Res. Technol.-Aqua* **1998**, *47*, 223−228.
(11) Stuetz, R. M.; Engin, G.; Fenner, R. A. *Water Sci. Technol.* **1998**, *38*, 331−335.
(12) Stuetz, R. M.; Fenner, R. A.; Engin, G. *Water Res.* **1999**, *33*, 453−461.
(13) Stuetz, R. M.; Fenner, R. A.; Engin, G. *Water Res.* **1999**, *33*, 442−452.
(14) Chandiok, S.; Crawley, B.; Oppenheim, B.; Chadwick, P.; Higgins, S.; Persaud, K. *J. Clin. Pathol.* **1997**, *50*, 790−791.
(15) Gardner, J. W.; Craven, M.; Dow, C.; Hines, E. L. *Meas. Sci. Technol.* **1998**, *9*, 120−127.
(16) Persaud, K.; Pisanelli, A.; Szyszko, S.; Reichl, M.; Horner, G.; Rakow, W.; Keding, H.; Wessels, H. *Sens. Actuators, B* **1999**, *55*, 118−126.
(17) Gardner, J. W.; Bartlett, P. N. *Sens. Actuators, B* **1994**, *18*, 211−220.
(18) Hodgins, D. *Sens. Actuators, B* **1995**, *27*, 255−258.
(19) Gopel, W. *Mikrochim. Acta* **1997**, *125*, 179−196.
(20) Groves, W. A.; Zellers, E. T. *Am. Ind. Hyg. Assoc.* **1996**, *57*, 1103−1108.
(21) Persaud, K.; Khaffaf, S.; Payne, J.; Pisanelli, A.; Lee, D.; Byun, H. *Sens. Actuators, B* **1996**, *36*, 267−273.
(22) Reddinger, J.; Reynolds, J. *Adv. Polym. Sci.* **1999**, *145*, 57−122.
(23) Roncali, J. *Chem. Rev.* **1992**, *92*, 711−738.
(24) *Handbook of Conducting Polymers*, 2nd ed.; Skotheim, T. A., Elsenbaumer, R. L., Reynolds, J. R., Eds.; Marcel Dekker: New York, 1998; Vol. 1.
(25) Bartlett, P. N.; Gardner, J. W.; Whitaker, R. G. *Sens. Actuators, A* **1990**, *23*, 911−914.
(26) Freund, M. S.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2652−2656.
(27) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298−2312.

10.1021/ac991079x CCC: $19.00 © 2000 American Chemical Society
Published on Web 06/09/2000

Chart 1. Structure of Poly(3,4-ethylenedioxy)-Thiophene−Poly(styrene sulfonate) (PEDOT−PSS)



media. In addition, the mean diameter of unagglomerated carbon black particles is approximately 20−50 nm, so that films of carbon black composites having thicknesses below 200 nm are difficult to prepare reproducibly and have not displayed good electrical resistance properties to date. Finally, it would be desirable to have detectors that show increased sensitivities toward polar compounds to improve the magnitude of the differences in response properties between polar and nonpolar detectors in an array. This would allow more robust analyte classification through the use of pattern recognition methods on the resulting data stream.

In this work, we report the fabrication of a stable and diverse array of chemiresistor-type vapor detectors from soluble, processable, polar conducting organic polymers. The detectors in the sensor array consist of either a blend or composite of commercially available poly(3,4-ethylenedioxy)thiophene−poly(styrene sulfonate) (PEDOT−PSS) (Chart 1) with a number of different insulating polymers. PEDOT−PSS was chosen as the conductive polymer used in this study for several reasons. PEDOT is relatively inexpensive and is sold commercially (Baytron P) as a colloidal suspension in water.[28,29] The poly(styrene sulfonate) (PSS) that is intimately associated through electrostatic attractions with the p-doped poly(3,4-ethylenedioxy)thiophene contributes to the stability of the colloid. PEDOT−PSS films cast from solution have been shown to exhibit superior environmental stability[30−32] due to both the electron richness concentrated in the polymer backbone and the inertness of the thiophene heterocycle (compared to that of pyrrole) as well as the high conductivity of the material itself.[33−38] Furthermore, because this material is tractable, fabrication of sensors is greatly simplified compared to alternative techniques for conductive polymer detector fabrication. For example, in previous work, the fabrication of polypyrrole detectors has entailed the polymerization of the monomer directly onto the detector substrate, making it difficult to ensure the same properties of the conductive polymer phase from trial to trail and within the composites in every detector in the array.

The perfomance of an array of such detectors was quantified using resolution factors and these values were compared to the performance of an array of carbon black composite detectors that utilized the same insulating matrixes. The ultimate purpose of the PEDOT−PSS array is not to replace carbon black composite detectors but to incorporate PEDOT−PSS composites into a system that provides a wider and more diverse electronic nose array for improved overall vapor detection performance.

EXPERIMENTAL SECTION

A. Materials. The PEDOT−PSS (0.5% PEDOT and 0.8% PSS by weight) was supplied as a colloidal suspension in water and was used as received (AG Bayer Inc.). The carbon black used in the composites was Black Pearls 2000 (BP2000), which was generously donated by Cabot Co. (Billerica, MA). The insulating polymers used in this study are shown in Figure 1. Poly(vinyl acetate) (M_n 260 000), poly(epichlorohydrin) (MW 700 000), poly(ethylene oxide) (MW 100 000), poly(vinyl butyral) (M_n 110 000), poly(*n*-butyl methacrylate) (M_n 180 000), poly(caprolactone) (M_n 30 000), poly(diallyl phthalate), poly(2-hydroxyethyl methacrylate) (M_n 300 000), vinyl chloride (91%)/vinyl acetate (6%)/vinyl alcohol (3%) copolymer (M_n 70 000), hydroxypropyl cellulose (MW 60 000), poly(2-butadiene) (MW 100 000), cellulose propionate (M_n 200 000), poly(vinyl alcohol) (16 000), poly(methyl vinyl ether) (50% solids in water, M_n 90 000), and poly(styrene) (M_n 45 000) were used as received from Scientific Polymer Products, Inc. (SP^2). Poly(4-vinylphenol) (MW 8000) was used as received from Aldrich Chemical Co. The solvents used in this study were methanol, ethanol, ethyl acetate, nitromethane, acetonitrile, acetone, tetrahydrofuran, chloroform, hexane, benzene, methoxybenzene, toluene, chlorobenzene, trifluoromethylbenzene (TFMbenzene), benzaldehyde, and nitrobenzene. These solvents were purchased from Aldrich and were used without further purification.

B. Apparatus. An automated flow system was used to deliver a diluted stream of solvent vapor to the detectors. The flow system consisted of LabVIEW 5.0.1, a pentium computer, electronically controlled solenoid valves, three mass flow controllers, and eight bubblers with coarse frits.[39] The background carrier gas was oil-free air obtained from the general compressed air laboratory source (1.10 ± 0.15 parts per thousand (ppth) water vapor) controlled via a 50 L min^{-1} mass flow controller (MKC Inc.). The air was filtered before entry into the mass flow controllers but was not dehumidified. To obtain the desired concentration of analyte in the gas phase, a stream of carrier gas was passed though a bubbler that contained the desired solvent. The eight bubblers of the system were divided into two banks of four, with each bank being controlled by a separate 500 mL min^{-1} mass flow controller (MKC Inc.). Saturation of the gas flow through each bubbler was validated both by measurements of the rate of mass loss of the solvent in the bubbler and through calibration of the flow reaching the detector chamber using a flame ionization detector (model 300 HFID, California Analytical Instruments, Inc.). The flame ionization detector was calibrated with toluene standards (50, 100, 200, and 1000 ppm) that were purchased from Matheson Inc.. The calibrations indicated that the flow system had an ∼5% random error in the delivery of a preset concentration of toluene between various trials, and no attempts were made to correct for this equipment-derived error in the reported detector response data. The temperature during data collection was 21.5

(28) Lerch, K.; Jonas, F.; Linke, M. *J. Chim. Phys. PCB* **1998**, *95*, 1506−1509.
(29) Jonas, F.; Morrison, J. *Synth. Met.* **1997**, *85*, 1397−1398.
(30) Jolly, R.; Pairis, S.; Petrescu, C. *J. Chim. Phys. PCB* **1998**, *95*, 1400−1405.
(31) Rannou, P.; Nechtschein, M. *Synth. Met.* **1999**, *101*, 474−474.
(32) Sapp, S.; Sotzing, G.; Reynolds, J. *Chem. Mater.* **1998**, *10*, 2101−2108.
(33) Ghosh, S.; Inganas, O. *Synth. Met.* **1999**, *101*, 413−416.
(34) Ghosh, S.; Rasmusson, J.; Inganas, O. *Adv. Mater.* **1998**, *10*, 1097−1099.
(35) Aleshin, A.; Kiebooms, R.; Heeger, A. *Synth. Met.* **1999**, *101*, 369−370.
(36) Lefebvre, M.; Qi, Z.; Rana, D.; Pickup, P. *Chem. Mater.* **1999**, *11*, 262−268.
(37) Morvant, M.; Reynolds, J. *Synth. Met.* **1998**, *92*, 57−61.
(38) Deleeuw, D.; Kraakman, P.; Bongaerts, P.; Mutsaers, C.; Klaassen, D. *Synth. Met.* **1994**, *66*, 263−273.
(39) Severin, E. J. Ph.D. Thesis, California Institute of Technology, 1999.

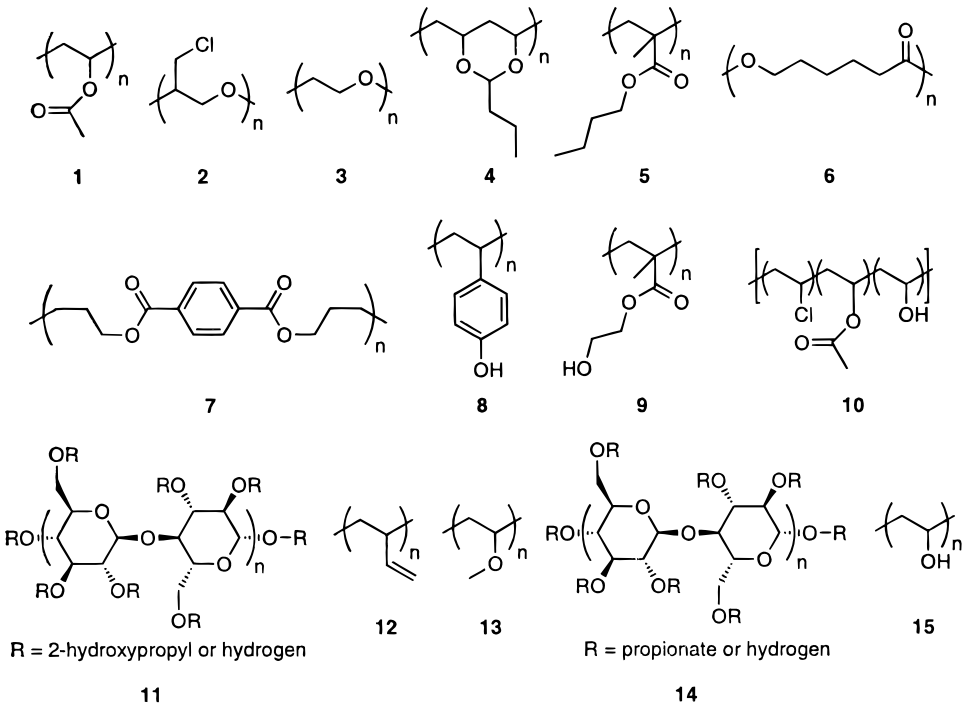


Figure 1. Insulating polymers that were were mixed with PEDOT−PSS or carbon black to yield the composite detector films evaluated in this work.

± 0.5 °C, but no active temperature control was maintained over the detectors.

C. Detector Fabrication. Sensor substrates were fabricated by evaporating 200 nm of chrome and then 800 nm of gold onto glass microscope slides using 2.5-mm-wide drafting tape as a mask. After evaporation, the mask was removed and the slides were baked at 300 °C for 12 h. The glass slides were then cut to yield the detector substrates.

The concentrations of the insulating polymer in the respective solvent ranged from 1 to 10 mg mL^{-1}. Solvents were either tetrahydrofuran (polymers **1**, **6**, **8**, **10**, and **12**), acetone (polymers **2**, **4**, **5**, **7**, and **14**), methanol (polymer **9**), or water (polymers **3**, **11**, **13**, and **15**). Five grams of the as-received PEDOT−PSS solution was added to each of the polymer solutions (typically 10 mL in volume; exact details are contained in Table 1 of the Supporting Information). The stock solution of PEDOT−PSS was diluted to obtain the blends for polymer **9** having <40% PEDOT−PSS. In these experiments, the PEDOT−PSS stock solution was diluted to obtain the desired weight percentage of PEDOT relative to the insulating polymers, but the volume of PEDOT solution was maintained at 5 mL. The PEDOT−PSS solution was added dropwise to the solution containing the insulating polymer. All aqueous solutions were stable, whereas aqueous solutions of polymer and organic solvents resulted in precipitation of polymer at ∼2 weeks of storage for polymers **1**, **2**, **4**, **6**, **10**, and **12**.

Thin films (between 90 and 900 nm thick, as measured by profilometry) were obtained by spin coating (Headway Research Inc.) the solution of PEDOT−PSS and insulating polymer onto the substrate. The solution was applied dropwise via a Pasteur pipet while the substrate was spinning at 1000 rpm. The film thickness was increased by continuing to add drops of the polymer solution until the dc resistance of the resulting film was 100−200 kΩ. After fabrication, all of the detectors were placed in a stream of dry air for 40 h to allow for the off-gassing of solvent vapor. Upon inspection under a Leica StereoZoom 6 optical microscope (30×), PEDOT−PSS loaded films obtained using polymers **2**, **3**, **5**, **7**, **9**. **11**, and **13**−**15** appeared to be homogeneous, whereas with polymers **1**, **4**, **6**, **8**, **10**, and **12** the films were heterogeneous in that PEDOT−PSS clumps and strands were noticeable. Not all polymers that are soluble in water-miscible organic solvents

Table 1.

PEDOT−PSS film	R^a (kΩ)	R^b (kΩ)	σ (S cm^{-1})	σ^b (S cm^{-1})
1	4.75 (9)	63 (3)	121	9.1
1	5.41 (8)	53 (1)	360	36
2	38.8 (5)	924 (5)	18.9	0.791
2	44.4 (6	650 (3)	29.4	2.0
3	232 (4)	4300	19.5	1.0
3	157 (1)	2800	10.9	0.63
4	14.0 (2)	87 (6)	66.3	11
4	8.6 (1)	60. (5)	64	9.2
5	13.0 (1)	270 (6)	142	6.9
5	10.9 (1)	80. (9)	198	27
6	6.8 (1)	37 (5)	260	49
6	8.7 (1)	46 (4)	160	30.
7	22.1 (4)	210 (1)	53.0	5.7
7	31.3 (2)	290 (6)	92.9	10.
8	7.12 (9)	31.5 (6)	152	12.8
8	7.6 (1)	27 (3)	38	11
9	213 (2)	1210 (6)	11.5	2.09
9	298 (5)	1160 (5)	11.5	2.95
15	4700 (2)	3300 (7)	0.32	0.48
15	5300 (2)	3300	0.22	0.34
15[c]	3520 (3)	2600	0.321	0.44
15[c]	6100 (5)	2900	0.43	0.90
P[d]	42.5 (4)	420 (2)	35.1	3.6
P[d]	38.4 (7)	380 (2)	23.9	2.4

[a] Average resistance values obtained from a minimum of four current settings. Standard deviations are given in parentheses as the last digits of the average. [b] Average resistance values and conductivities measured after 16 months of storage in the dark. [c] Fraction of poly(vinyl alcohol) to PEDOT−PSS is 80 to 20 (w/w). [d] Pure PEDOT−PSS.

made good detector films. For instance, attempts to prepare PEDOT−PSS composite films using poly(sulfone), poly(vinyl chloride), poly(carbonate), and poly(styrene) were not fruitful. Despite the fact that all of these polymers were soluble in tetrahydrofuran, a precipitate formed upon addition of the PEDOT−PSS colloidal water solution. The polymers listed in Figure 1 all made detectors that exhibited good electrical responses to the various analytes of interest in this work.

Carbon black composite detectors were prepared by first dissolving ∼200 mg of insulating polymer in the appropriate solvent (the same solvents were used as those described above for PEDOT−PSS composite preparation) and then adding 50 mg of carbon black such that the overall composition of the solution was 80% insulating polymer−20% carbon black by weight. These solutions were sonicated for ∼20 min, and detector films were then spin cast by covering the detector substrate with solution. As for the PEDOT films, solution was deposited until the films had resistance values of 100−200 kΩ. The resulting film thicknesses were in typically in the range of 100−500 nm. Detectors were then placed in a stream of dry air for 40 h.

D. Conductivity and Mass Uptake Measurements. Conductivity measurements were performed using the four-point collinear array technique. The current was supplied to the outer leads using a galvanostat (Pine Instrument Co.) and was monitored using a multimeter (Fluke 87). A multimeter was used to measure the voltage drop across the inner leads. To ensure that the materials followed ohmic behavior, the voltage drops were measured for at least at four different applied currents for each film. The substrate used for the measurement consisted of four evaporated gold leads deposited onto a glass backing. The distance between each set of gold leads was 0.212 cm. All films were prepared by the spin-coating technique described above.

For determining the resistance response to the presence of solvent vapors, the resistive film detectors were housed in an aluminum chamber, and two electrical leads were connected via alligator clips to each detector. The leads were multiplexed through a Keithley model 7001 channel switcher to a Keithley model 2002 multimeter that measured the resistance of each detector approximately once every 3 s.

QCM crystals (10 MHz, blank diameter 13.7 mm) with a custom electrode pattern were obtained from International Crystal Manufacturing (ICM) in Oklahoma City, OK. The resonant frequency of the QCM was obtained using a HP 53181A frequency counter (Palo Alto, CA). Shielded cables were used between the crystal, the oscillator circuit, and the frequency counter. The standard oscillation electrodes were configured at 90° angles to make room for two other tabs that served as electrodes for resistance measurements of the carbon black−polymer composite and PEDOT−PSS−polymer composite films. The crystals were polished to a surface roughness of less than 5 μm, which produced a mirrorlike finish on the gold electrodes.

E. Flow System Measurements. Flow system experiments were carried out using the apparatus described above. All sensors were purged with carrier gas for 2 h before each separate flow system experiment. Prior to data collection, all of the detectors were subjected to 10 exposures to each of the 16 different solvents used in this study, with each analyte at 5% of its vapor pressure, $p°$, at room temperature.

The data presented in this paper were obtained during seven separate flow system experiments. Five of the flow system experiments consisted of 80 exposures to 8 separate analytes (10 exposures to each analyte) at a constant activity (5% of the analyte's vapor pressure). The concentrations of analyte in this run were (in units of ppth) as follows: methanol, 6.63; ethanol, 3.21; ethyl acetate, 4.87; nitromethane, 1.90; acetonitrile, 4.90; acetone, 12.9; tetrahydrofuran, 9.02; chloroform, 10.7; hexane, 9.74; benzene, 5.21; methoxybenzene, 0.181; toluene, 1.55; chlorobenzene, 0.641; trifluoromethylbenzene, 0.217; benzaldehyde, 0.059; and nitrobenzene, 0.011. Each exposure consisted of a three-step process that began with 240 s of air flow over the detector followed by a 300-s period in which the analyte was present in the vapor stream. Finally, the detectors were exposed to background air for another 240 s. All exposures for each separate flow system experiment were fully randomized with respect to analyte identity. Two copies of each detector were used, and the responses for both detectors of a given composition were combined with the response to multiple exposures to a given analyte to produce a reported response quantity for that analyte/detector combination. Only one PEDOT−PSS detector using polymers **3** and **5** and only one carbon black detector using polymer **2** were functional throughout the first five flow system runs, so the responses for these compositions are only reported for these single detectors.

The flow system experiment for the study of response versus analyte concentration was similar to that described above except that two solvents (methanol and nitromethane) were used at 5, 3, 1, and 0.5% of each analyte's vapor pressure. This is equivalent to concentrations of 6.63, 3.98, 1.32, and 0.66 ppth for methanol and 1.90, 1.14, 0.380, and 0.190 ppth for nitromethane. Exposures for this experiment were not randomized and followed the order of being exposed first to methanol, from lowest to highest concentration (a total of 40 exposures, 10 exposures at each of the 4 different concentrations), followed by the exposures to nitromethane, again from the lowest to highest concentration (a total of 40 exposures, 10 exposures at each of the 4 different concentrations). The calculated responses reported for each concentration were an average of 20 data points, 10 exposures for each of the pair of nominally identical detectors of a given composition.

The experiment using different loading levels of PEDOT−PSS with polymer **9** was performed using four copies of each detector type. Five exposures were performed for each of the 2 analytes (methanol and acetonitrile) so that the reported responses are an average of 20 data points. The exposures to the analytes were not randomized but instead five exposures to methanol at 3% (3.98 ppth) of its vapor pressure were performed, followed by five exposures to acetonitrile at 3% (2.98 ppth) of its vapor pressure.

F. Data Processing. Data were processed using specially written programs in Microsoft Excel. All data were first corrected to remove any drift in the baseline resistance. This drift correction was also applied to the resistance data obtained for the detector during exposure to that particular analyte. Then the maximum relative differential resistance change, $\Delta R_{ij,\max}/R_b$, was calculated, where $\Delta R_{ij,\max}$ is the maximum differential resistance change, taken as the average of five data points about the maximum measured differential resistance of the jth detector during the ith exposure, and R_b is the drift-corrected baseline resistance of the

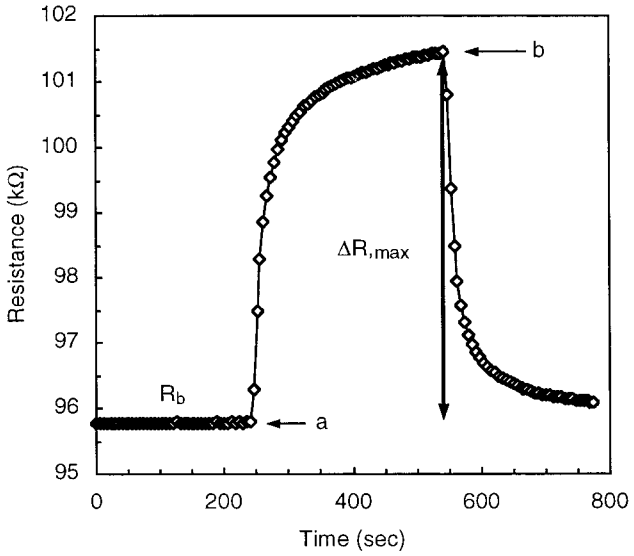


Figure 2. Exposure of a PEDOT−PSS/**7** detector to methanol (at $P/P^o = 0.05$). At time a, the methanol vapor was introduced into the carrier gas that was directed over the detectors. At time b, the methanol vapor was removed from the carrier gas. The relative differential resistance change is defined as $\Delta R_{max}/R_b$, where R_b is the baseline resistance and ΔR_{max} is the maximum resistance change upon exposure of the detector to solvent vapor.

detector prior to the exposure. A sample response for a PEDOT−PSS/**7** composite detector is shown in Figure 2.

G. Quantification of Detector Array Performance. Statistical methods based on cluster analysis using the Fisher linear discriminant methodology were used to analyze the detector array data.[40,41] A resolution factor for any solvent pair can be obtained along any vector, $\vec{w}$, from the vector projection onto $\vec{w}$ of the distance between the cluster centroids, $d_{\vec{w}}$, divided by the square root of the sum of the squared projected standard deviations, $\sigma_{a,\vec{w}}$ and $\sigma_{b,\vec{w}}$, for data arising from repeated exposures for two different analytes, *a* and *b*. The resulting numerical resolution factor along $\vec{w}$ is defined as

$$\mathrm{rf} = d_{\vec{w}}/\sqrt{\sigma^2_{a,\vec{w}} + \sigma^2_{b,\vec{w}}}$$

The Fisher linear discriminant algorithm searches for the vector, $\vec{w}$, such that the rf value is maximized along this optimal discriminant vector.[41] Assuming a Gaussian distribution relative to the mean value of the data points in a given cluster, the probabilities of correctly identifying an analyte as *a* or *b* are approximately 72, 92, and 98% from a single presentation when analytes *a* and *b* are separated with resolution factors of 1.0, 2.0, or 3.0, respectively. Data extracted from multiple exposures of an analyte only estimate the statistical distributions of the clustered data, which may lead to an overestimation of rf. Overestimations are typically less than 30% for an array consisting of 14 detectors and the given number of solvent exposures, with the overestimations dropping to less than 3% in cases involving a single detector.

(40) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177−4190.

(41) Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; John Wiley & Sons: New York, 1973.

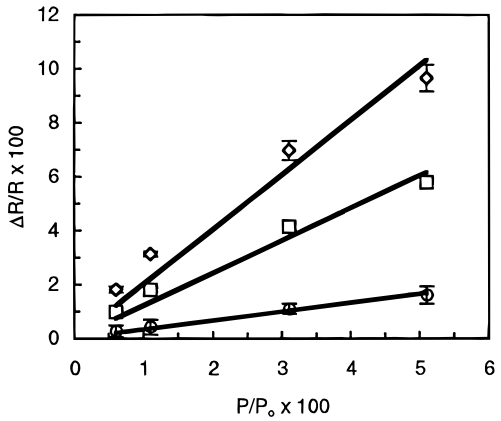


Figure 3. Relative differential resistance response, $\Delta R/R_b$, for three different PEDOT−PSS composite detectors as a function of the concentration of methanol in the gas phase. The methanol (having a vapor pressure P^o) was maintained at a partial pressure *P* in a stream of air flowing over the detectors. Key: squares, PEDOT−PSS/**1**; diamonds, PEDOT−PSS/**2**; circles, PEDOT−PSS/**3**.

Especially large rf values should be treated with caution as they could be overestimated by larger amounts.

For the resolution factors calculated for both the PEDOT−PSS detector array and the carbon black composite array, the best detector of the two copies that was prepared with each of the first nine insulating polymers, as determined by the lowest standard deviation in the response across the 10 exposures to a given analyte, was used. Comparisons of resolution factors were made between a nine-detector array of PEDOT−PSS composites composed of insulating polymers **1**−**9** and a nine-detector array of carbon black composite detectors. The same nine insulating polymers were used in both array types.

RESULTS

A. Conductivity, Short-Term Drift, Long-Term Drift, and dc Noise Levels of the PEDOT−PSS Composite Films. Table 1 lists the dc conductivities measured for several PEDOT−PSS films loaded with 40 wt % PEDOT−PSS. Films exhibiting the highest conductivities were those prepared using poly(vinyl acetate) (**1**), poly(*n*-butyl methacrylate) (**5**) and poly(caprolactone) (**6**). These films exhibited conductivities ranging from approximately 100 to 360 S cm^{-1}, whereas films of PEDOT−PSS itself had a conductivity of 30 S cm^{-1}. The lowest conductivities were obtained for PEDOT−PSS/**15**, PEDOT−PSS/**9**, and PEDOT−PSS/**3**.

Table 1 also lists the conductivities for the same PEDOT−PSS films after 16 months of storage under atmospheric conditions in the dark. Over this time period, all samples exhibited a decline of ∼1 order of magnitude in their conductivities, with the exception of the PEDOT−PSS/**15** samples.

In flow system experiments, the mean short-term drift of the baseline resistances for the entire set of detectors in a PEDOT−PSS detector array was less than 38 ± 30 Ω min^{-1}, while the detector with the largest drift in the array showed a short-term baseline resistance change of 180 Ω min^{-1}. For comparison, the average baseline drift for the array of carbon black/insulating polymer composite detectors was 105 Ω min^{-1}. The baseline drift was negative for all of the PEDOT−PSS and carbon black composite detectors. The average dc noise level (in an ∼1-Hz bandwidth) for all detectors across the PEDOT−PSS array was

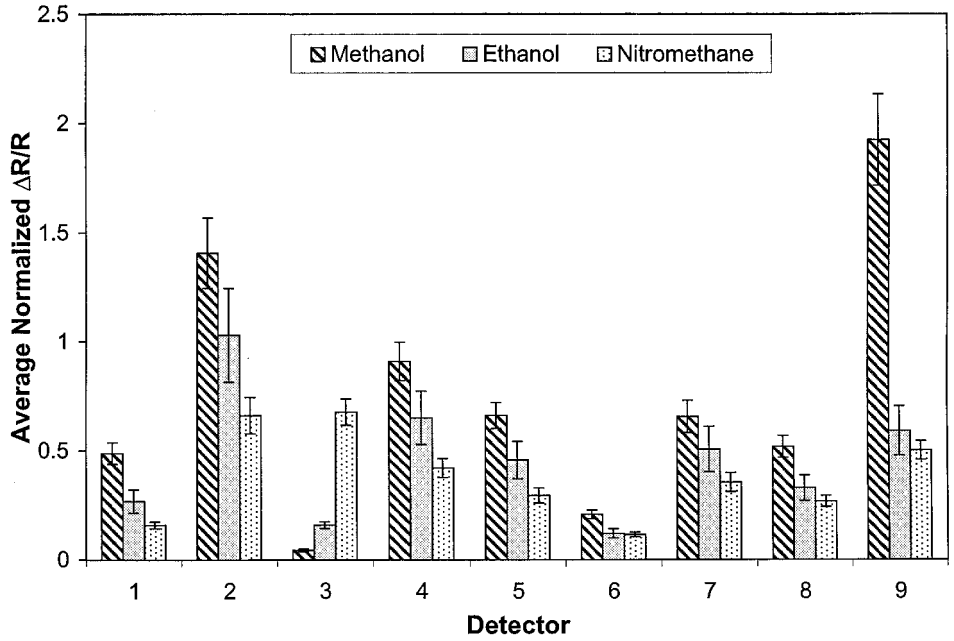


Figure 4. Histograms of $\Delta R_{max}/R_b$ obtained for a PEDOT−PSS detector array upon multiple exposures to ethanol, methanol, and nitromethane each at $P/P^o = 0.05$. The detector number indicates which polymer was used in forming the PEDOT−PSS/polymer composite sensor, with the detector number indicating which polymer in Figure 1 was used to form the composite detector film. Error bars are reported as one standard deviation unit of the response averaged over 10 presentations of each analyte to the detector array.

Table 2. PEDOT−PSS Composite Detector Sensitivities to Methanol and Nitromethane

detector	slope (methanol)	R^2	slope (nitromethane)	R^2
PEDOT−PSS	2.4	0.99	0.28	0.99
1	1.1	0.99	0.21	0.99
2	1.7	0.99	0.26	0.99
3	0.28	0.99	0.35	0.99
4	1.5	0.99	0.30	0.99
5	1.2	0.99	0.26	0.99
6	0.80	0.99	0.18	0.99
7	1.3	0.99	0.31	0.99
8	1.0	0.98	0.26	1.00
9	1.9	0.93	0.26	1.00

~1 part in 10 000, which is comparable to the noise levels obtained for the detectors in the carbon black composite array under similar conditions.

B. Array-Based Vapor Sensing Using PEDOT−PSS Composite Sensors. Figure 3 displays a plot of the relative differential resistance response versus analyte concentration for three different PEDOT−PSS detectors exposed to methanol vapor. The fitting statistics of these data are summarized in Table 2. The pristine PEDOT−PSS and PEDOT−PSS/**9** detectors were the most sensitive to methanol vapor, displaying a sensitivity value of $(\Delta R_{max}/R_b)/(P/P^o)$ (where P is the partial pressure of the analyte in the carrier gas and P^o is the vapor pressure of the analyte) of 2.4 and 1.9, respectively. The most sensitive detectors to nitromethane were PEDOT−PSS/**3** and PEDOT−PSS/**7**, which had $(\Delta R_{max}/R_b)/(P/P^o)$ values of 0.35 and 0.31, respectively.

Figure 4 presents the $\Delta R/R$ fingerprints obtained for three different analyte vapors exposed to an array composed of nine compositionally different PEDOT−PSS composite detectors. The data displayed in the figure are the $\Delta R_{max}/R_b$ response values that have been normalized with respect to the maximum $\Delta R_{max}/R_b$ response value obtained across all detectors for a given solvent vapor exposure. The purpose of normalization was to remove the concentration dependence of the data and to isolate the differences in the patterns of array response observed for each of the three solvents. Each separate analyte clearly produced a unique response pattern on this array of detectors.

Table 3 lists the relative differential resistance responses for all PEDOT−PSS detectors during exposure to each of 16 different analytes. The detectors were especially sensitive to the more polar analytes such as methanol, ethanol, nitromethane, acetonitrile, and acetone. The relative response of the different PEDOT−PSS composite detectors to the various analytes was primarily dictated by the response properties of the PEDOT−PSS itself. The composites did however affect this basic response behavior somewhat, such that some PEDOT−PSS composite detectors showed lower sensitivities to a given analyte than the pristine PEDOT detectors, while others exhibited higher sensitivites. For example, PEDOT−PSS/**11** showed a significantly enhanced response to all of the analytes in comparison to the pristine PEDOT−PSS detector.

The differences in response properties between different PEDOT−PSS composites were sufficient to provide robust analyte vapor classification information from the array response data. Table 4 lists the resolution factors obtained for an array of nine PEDOT−PSS composites consisting of the first nine insulating polymers listed in Figure 1. In this analysis, the data were not normalized but all analytes were evaluated at a fixed fraction of their vapor pressure. The best resolution factor obtained was 360, for the pairwise discrimination between methanol and hexane. In general, the best resolution factors were observed for the discrimination of polar analyte pairs and for polar/nonpolar analyte pairs, while low-resolution factors were encountered for the discrimination of pairs of nonpolar of compounds. The polar group

Table 3. Average Perent $\Delta R_{max}/R_b$ Values for PEDOT−PSS Detectors upon Exposure to 16 Different Analytes at 5% of Their Vapor Pressure[a]

detector	hexane	benzene	toluene	methoxybenzene	chloroform	chlorobenzene	ethyl acetate	THF
1	−0.012 (6)	0.07 (4)	0.05 (7)	0.20 (8)	0.12 (9)	0.2 (1)	0.1 (1)	−0.012 (6)
2	−0.03 (1)	0.14 (8)	0.07 (8)	0.3 (1)	0.2 (1)	0.4 (2)	0.2 (2)	−0.02 (1)
3	−0.03 (1)	0.1 (1)	0.04 (3)	0.3 (1)	0.07 (4)	0.4 (2)	0.14 (6)	−0.02 (1)
4	−0.012 (7)	0.1 (1)	0.06 (7)	0.3 (1)	0.2 (1)	0.3 (2)	0.2 (2)	−0.012 (6)
5	−0.010 (5)	0.1 (1)	0.05 (6)	0.2 (1)	0.17 (9)	0.3 (1)	0.2 (2)	−0.011 (6)
6	−0.006 (4)	0.06 (8)	0.05 (7)	0.16 (7)	0.10 (9)	0.2 (1)	0.1 (1)	−0.004 (2)
7	−0.010 (7)	0.1 (2)	0.07 (8)	0.3 (2)	0.2 (1)	0.3 (2)	0.3 (1)	−0.010 (7)
8	−0.02 (1)	0	0.09 (9)	0.3 (1)	0.2 (1)	0.3 (1)	0.2 (2)	−0.009 (5)
9	−0.03 (1)	0.2 (1)	0.1 (3)	0.29 (9)	0.2 (3)	0.5 (2)	0.2 (2)	−0.03 (2)
10	−0.019 (9)	na[b]	0.1 (1)	na	0	na	0.4 (3)	−0.03 (2)
11	1.8 (9)	na	1.9 (6)	na	3.3 (6)	na	6 (2)	3.5 (8)
12	−0.02 (1)	na	−0.01 (1)	na	0.3 (3)	na	0.7 (5)	−0.03 (1)
13	0.8 (2)	na	0.7 (4)	na	1.4 (3)	na	4 (1)	1.8 (5)
14	−0.007 (3)	na	0.2 (1)	na	0.4 (2)	na	0.8 (8)	0.3 (3)
15	−0.049 (9)	0.14 (8)	0.03 (3)	0.4 (1)	0.01 (1)	0.6 (3)	0.08 (3)	−0.05 (2)
P[c]	−0.03 (2)	0.01 (1)	0.1 (3)	0.20 (6)	0.04 (4)	0.09 (5)	0.11 (9)	−0.03 (2)

detector	TFMbenzene	benzaldehyde	acetone	ethanol	methanol	nitrobenzene	acetonitrile	nitromethane
1	0.1 (1)	0.17 (8)	0.5 (2)	2.1 (4)	5.6 (3)	0.08 (5)	1.4 (1)	1.2 (1)
2	0.2 (3)	0.3 (2)	0.8 (4)	3.6 (5)	9.6 (6	0.2 (1)	2.3 (4)	2.1 (5)
3	0.3 (2)	0.3 (2)	0.76 (9)	1.5 (2)	1.6 (3)	0.2 (1)	2.0 (1)	2.3 (3)
4	0.2 (2)	0.3 (2)	0.8 (3)	3.3 (6)	8.1 (8)	0.1 (1)	2.5 (2)	2.0 (2)
5	0.2 (1)	0.3 (1)	0.7 (2)	2.5 (6)	6.2 (8)	0.2 (1)	1.9 (3)	1.5 (3)
6	0.1 (1)	0.2 (1)	0.5 (2)	1.6 (3)	3.9 (4)	0.09 (7)	1.5 (2)	1.2 (2)
7	0.2 (2)	0.3 (2)	0.9 (4)	3.3 (4)	7.4 (5)	0.1 (1)	2.5 (3)	2.0 (3)
8	0.2 (2)	0.2 (1)	0.8 (2)	2.5 (4)	6.0 (6)	0.11 (8)	2.0 (2)	1.7 (2)
9	0.2 (2)	0.3 (2)	0.7 (3)	3.1 (3)	12 (1)	0.18 (9)	2.6 (2)	2.1 (2)
10	na	na	0.8 (4)	3.1 (9)	11 (2)	na	na	na
11	na	na	9.4 (9)	14 (2)	17 (2)	na	na	na
12	na	na	1.4 (8)	5.7 (1)	14 (2)	na	na	na
13	na	na	6.3 (6)	9 (1)	12 (1)	na	na	na
14	na	na	1.7 (8)	5.5 (8)	11 (2)	na	na	na
15	0.2 (2)	0.23 (6)	0.38 (5)	1.49 (9)	−0.105 (7)	0.15 (5)	−0.06 (2)	1.50 (9)
P[c]	0.03 (3)	0.14 (4)	0.5 (1)	2.7 (7)	11 (1)	0.09 (3)	1.1 (4)	1.8 (3)

[a] Standard deviations for the last digit of the values are given in parentheses. [b] na, not available. [c] Pure PEDOT−PSS.

Table 4. Resolution Factors Calculated Using an Array of Nine Different PEDOT−PSS Composite Detectors

	benzene	toluene	methoxy-benzene	chloro-form	chloro-benzene	EtOAc	THF	TFM-benzene	benzal-dehyde	acetone	ethanol	methanol	nitro-benzene	aceto-nitrile	nitro-methane
hexane	4.3	4.7	10.0	6.6	3.6	16.0	2.0	4.6	9.3	42	58	360	7.2	130	51
benzene		3.2	3.1	5.2	2.6	3.7	5.0	1.6	1.7	11	27	147	2.3	88	32
toluene			4.8	4.0	3.1	4.4	4.9	2.6	3.3	27	44	59	4.5	33	35
methoxybenzene				11	3.5	6.0	14	4.9	2.3	6.6	23	83	2.6	62	24
chloroform					4.3	10	5.7	3.0	5.5	22	37	78	7.7	36	32
chlorobenzene						4.0	3.8	2.3	3.7	6.8	16	144	2.6	50	12
EtOAc							14	2.5	4.3	32	47	75	4.6	24	29
THF								7.9	7.4	37	62	240	6.4	280	52
TFMbenzene									1.9	42	58	360	2.7	130	51
benzaldehyde										6.0	27	86	3.4	20	23
acetone	-										24	71	6.3	26	24
ethanol	-											49	25	14	14
methanol													120	66	77
nitrobenzene														38	37
acetonitrile															20

of analytes, taken as those having a dielectric constant, ϵ, of >17, consisted of the following (dielectric constants in parentheses): nitromethane (37.3), acetonitrile (36.6), nitrobenzene (35.6), methanol (33.0), ethanol (25.3), acetone (21.0), and benzaldehyde (17.9). The nonpolar analytes, taken as those having $\epsilon < 10$, consisted of the following: trifluoromethylbenzene (9.2), tetrahydrofuran (7.5), ethyl acetate (6.1), chlorobenzene (5.7), chloroform (4.8), methoxybenzene (4.3), toluene (2.4), benzene (2.3), and hexane (1.9). The average resolution factor obtained for the discrimination between pairs of polar analytes was 34, with a minimum of 3.4 obtained for the pairwise discrimination of benzaldehyde/nitrobenzene. The average resolution factor for distinguishing polar from nonpolar analytes was calculated to be 47, with a minimum rf of 1.7 obtained for the pairwise discrimination of benzaldehyde/benzene. Discrimination of nonpolar/nonpolar compounds displayed an average resolution factor of 5.0, with a minimum of 1.6 for the discrimination between benzene and trifluoromethylbenzene.

Table 5. Average Percent $\Delta R_{max}/R_b$ Values for Carbon Black Composite Detectors upon Exposure to 16 Different Analytes at 5% of Their Vapor Pressure[a]

detector	hexane	benzene	toluene	methoxybenzene	chloroform	chlorobenzene	ethyl acetate	THF
1	0.2 (2)	0.4 (1)	0.3 (2)	0.3 (1)	1.0 (3)	0.3 (5)	0.54 (8)	0.49 (7)
2	0.49 (7)	0.79 (4)	0.8 (2)	1.43 (6)	0.6 (1)	0.96 (5)	0.63 (8)	0.8 (1)
3	1.4 (3)	2.4 (8)	2.7 (6)	4 (2)	6 (1)	4 (1)	1.6 (2)	2.1 (4)
4	2.0 (6)	2.7 (3)	2.7 (7)	2.1 (3)	6.0 (7)	2.5 (3)	3.4 (4)	3.9 (3)
5	7 (2)	11 (2)	11 (3)	12 (2)	21 (8)	14 (2)	6 (1)	9 (2)
6	0.9 (2)	1.5 (1)	1.5 (2)	2.3 (2)	1.8 (2)	2.2 (1)	0.80 (4)	1.02 (8)
7	0.4 (3)	0.8 (4)	0.7 (3)	0.7 (4)	1.4 (8)	0.8 (3)	1.0 (5)	1.1 (6)
8	0.3 (3)	0.1 (1)	0.4 (3)	0.4 (6)	0.8 (3)	0.2 (1)	0.9 (3)	1.1 (4)
9	0.1 (1)	0.08 (6)	0.2 (2)	1 (3)	0.3 (2)	0.08 (5)	0.4 (2)	0.21 (7)

detector	TFMbenzene	benzaldehyde	acetone	ethanol	methanol	nitrobenzene	acetonitrile	nitromethane
1	0.11 (4)	0.22 (5)	0.76 (5)	0.46 (6)	0.62 (4)	0.22 (5)	1.03 (3)	1.4 (3)
2	0.63 (7)	1.07 (5)	0.54 (3)	0.14 (5)	0.16 (5)	1.16 (7)	0.25 (5)	0.23 (6)
3	1.9 (7)	3 (1)	1.5 (3)	1.3 (2)	1.5 (3)	4 (2)	1.6 (5)	2.3 (8)
4	1.4 (3)	1.6 (4)	3.0 (3)	2.4 (3)	2.5 (3)	1.4 (3)	1.9 (3)	1.8 (2)
5	11 (2)	7.6 (5)	4.2 (8)	1.7 (4)	1.3 (3)	9.1 (4)	2.3 (4)	3.0 (5)
6	1.3 (1)	2.1 (1)	0.69 (3)	0.37 (3)	0.38 (3)	2.5 (1)	0.55 (9)	0.7 (1)
7	0.4 (1)	0.7 (3)	1.3 (7)	0.6 (3)	0.9 (2)	0.7 (3)	1.6 (6)	1.6 (6)
8	0.09 (9)	0.2 (1)	2.6 (5)	2.5 (5)	3.5 (5)	0.1 (1)	2.7 (7)	1.7 (2)
9	0.05 (3)	0.12 (7)	0.8 (2)	1.7 (3)	5.0 (9)	0.11 (6)	2.0 (3)	1.5 (3)

[a] Standard deviations for the last digit of the values are given in parentheses.

Table 6. Resolution Factors Calculated Using an Array of Nine Different Carbon Black Composite Detectors

	benzene	toluene	methoxy-benzene	chloro-form	chloro-benzene	ethyl acetate	THF	TFM-benzene	benzal-dehyde	acetone	ethanol	methanol	nitro-benzene	aceto-nitrile	nitro-methane
hexane	7.0	4.0	12	11	10	6.2	6.7	8.5	11	11	18	19	13	100	30
benzene		7.9	15	12	30	35	11	20	28	37	89	70	26	71	49
toluene			7.9	8.1	14	7.9	5.2	11	11	12	16	26	12	27	25
methoxybenzene				15	20	28	17	19	25	39	56	54	17	71	73
chloroform					21	21	9.0	23	17	37	24	38	18	27	18
chlorobenzene						51	20	21	31	60	120	110	30	92	49
EtOAc							5.2	24	42	10	30	35	28	41	33
THF								20	18	14	24	31	19	22	21
TFMbenzene									18	41	49	58	21	47	31
benzaldehyde										46	94	57	5.2	46	38
acetone											18	33	34	19	24
ethanol												14	50	17	33
methanol													34	21	24
nitrobenzene														35	30
acetonitrile															21

When normalized responses were used, the average rf across all pairwise discriminations was 11, with a minimum of 1.6 (compared to 33 and 1.6 for the unnormalized responses used for the rf calculations in Table 4). The same trends in the pairwise discriminations for the rf calculations using the unnormalized data were observed in the rf values obtained using the normalized responses.

C. Array-Based Vapor Sensing Using Carbon Black Composite Detectors. Carbon black composite detectors consisting of the first nine polymers listed in Figure 1 were studied to compare their sensitivities and pairwise analyte-resolving properties to those of a PEDOT−PSS detector array formed from the same nine insulating polymers. Table 5 lists the relative differential resistance responses for these carbon black composite detectors upon exposure to each of 16 different analytes. As for the PEDOT−PSS array, these experiments were performed at a constant activity ($P/P^o = 0.05$) of all analytes. The response trends are in accord with data reported earlier for carbon black composite detectors in that the composites most responsive to polar analytes contained the most polar polymers, whereas the detectors most sensitive to the nonpolar analytes contained the most nonpolar polymers in the array. For example, the largest response to methanol was obtained for the carbon black composite detectors formed from polymers **8** and **9**, both of which are good hydrogen bond donors and acceptors. The most sensitive detector to benzene was CB/**5**. The unsaturated backbone of **5**, in addition to a butyl side chain, contributes to this detector's effectiveness in responding to this class of analytes.

The resolution factors for a nine-detector carbon black composite array consisting of the first nine polymers in Figure 1 were calculated using the relative differential resistance values of Table 5. The average of all of the resolution factors listed in the table is 30, with a minimum rf of 4.0 obtained for the pairwise discrimination of toluene and hexane (Table 6). The average rf for discrimination between pairs of polar solvents is 30, with the worst polar solvent pair being nitrobenzene and benzaldehyde, whereas, the average rf for discrimination of polar/nonpolar analyte pairs was 37. The average rf for the pairs of nonpolar analytes was

Table 7. Average Percent $\Delta R/R$ and Percent $\Delta f/f'$ for 10 Exposures of Carbon Black and PEDOT−PSS Composite Detectors to Four Analytes at 5% of Each Analyte's Vapor Pressure[a]

	poly(vinyl butyral)				poly(2-hydroxyethyl methacrylate)			
	carbon black		PE−DOT PSS		carbon black		PE−DOT PSS	
	$\Delta R/R$	$\Delta f/f'$	$\Delta R/R$	$\Delta f/f'$	$\Delta R/R$	$\Delta f/f'$	$\Delta R/R$	$\Delta f/f'$
acetone	3.2 (5)	−1.2 (1)	1 (2)	−0.4 (1)	1.8 (4)	−2.4 (7)	0.7 (4)	−0.3 (1)
methanol	1.7 (2)	−0.62 (7)	14 (2)	−0.95 (9)	5.6 (7)	−1.3 (2)	17 (2)	−1.35 (6)
THF	4.7 (4)	−1.59 (8)	−0.2 (2)	−0.31 (4)	0.5 (1)	−0.8 (2)	−0.4 (4)	−0.08 (2)
toluene	3.0 (4)	−1.20 (6)	0.3 (2)	−0.22 (4)	0.05 (7)	−0.2 (4)	−0.2 (2)	−0.06 (1)

[a] Standard deviations for the last digit of the values are given in parentheses.

calculated to be 16. The resolution factors for the carbon black detector array were also calculated using normalized responses. Again the same trends in rf factors were observed. The average resolution factor obtained across all pairs using the normalized data was 35, slightly higher than that obtained with the unnormalized responses.

D. Resistance and Quartz Crystal Microbalance Measurements. To assess whether the enhanced sensitivity to polar analytes was predominantly due to an increased mass uptake of the PEDOT-containing films or to an increased $\Delta R/R$ sensitivity of the PEDOT-containing materials, mass uptake measurements were performed simultaneously with resistance change measurements. As displayed in Table 7, data were collected for carbon black composites and PEDOT composites using two representative polymers, poly(vinyl butyral) and poly(2-hydroxyethyl methacrylate), during exposures to acetone, methanol, THF, and toluene.

Significant differences were observed between the analyte sorption properties and the relative differential resistance responses of the PEDOT-containing detectors relative to the properties of the carbon black composites. For example, methanol partitioned slightly more favorably into a PEDOT/poly(vinyl butyral) detector than into a carbon black/poly(vinyl butyral) detector, but toluene partitioned more strongly into the carbon black composite than the PEDOT-containing film. Similar behavior was observed for poly(2-hydroxyethyl methacrylate) composites. These data suggest an increased polarity of the PEDOT films relative to the carbon black composites, so that PEDOT films produce increased partition coefficients for polar analytes but decreased partition coefficients for nonpolar analytes.

This increased analyte sorption process was not, however, sufficient to explain the improved $\Delta R/R$ response of the PEDOT-containing films to the polar analytes. In fact, as shown in Table 7, the relative differential resistance response with respect to $\Delta f/f'$, which is proportional to the relative mass change of the detector film, is much enhanced for the polar compounds, such as methanol, on the PEDOT films relative to the analogous carbon black composites. This indicates that, for polar analytes, the PEDOT-containing films possess an inherent amplification of the sorption-induced signal transduction relative to carbon black composites.

DISCUSSION

In general, the PEDOT−PSS detectors were more sensitive than the respective carbon black detectors to polar compounds.

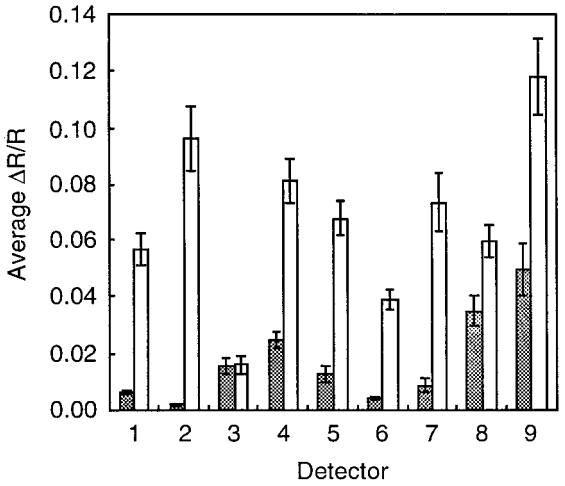


Figure 5. Comparison between the $\Delta R_{max}/R_b$ response of carbon black (shaded) and the PEDOT−PSS (not shaded) detector arrays to methanol at $P/P^o = 0.05$. The detector number indicates which polymer was used in forming the composite sensor, with the detector number indicating which polymer in Figure 1 was used to form the composite detector film. Error bars represent 1σ values obtained for >10 presentations of each analyte to the detector array.

Figure 5 shows the relative differential resistance responses obtained upon exposure to methanol at 5% of its vapor pressure to both a nine-detector PEDOT−PSS array and a nine-detector carbon black composite array. All responses for the PEDOT−PSS array were larger than those for the carbon black detectors made using the same insulating polymer. The least responsive of the conductive polymer detectors, PEDOT−PSS/**6**, was only slightly less responsive than the most responsive carbon black detector toward this particular analyte vapor.

The differential resistance versus analyte sorption measurements show that, in fact, the enhanced signal displayed by PEDOT films when exposed to polar analytes arises predominantly from an increased sensitivity in the conduction mechanism of the PEDOT to the presence of polar analytes relative to that displayed by carbon black composites. Accordingly, better signal/noise performance for polar analytes can be obtained using PEDOT-containing detectors than using the test set of carbon black composites alone for vapor detection tasks. Using the resolution factors calculated using the Fisher linear discriminant, the PEDOT−PSS detector array was shown to outperform, on average, the carbon black sensor array in the pairwise discrimination of polar analytes and the pairwise discrimination of polar/nonpolar analyte pairs, whereas in the carbon black detector array outperformed the PEDOT−PSS array in distinguishing between non-

Table 8. Summary of Pairwise Resolution Factors for Unnormalized, Normalized, and Randomly Combined Arrays[a]

detectors (1-9) Carbon Black					(10-18) PEDOT-PSS				(across all 16 analyte pairs)				(polar vs. polar)			(polar vs. nonpolar)			(nonpolar vs. nonpolar)		
									MAX	MIN	AVG	STDEV	MAX	MIN	AVG	MAX	MIN	AVG	MAX	MIN	AVG
unnormalized																					
1	2	3	4	5	6	7	8	9	120	4.0	30	22	94	14	34	120	10	37	51	4.0	16
10	11	12	13	14	15	16	17	18	360	1.6	33	53	120	6.0	39	360	1.7	47	16	1.6	5.5
normalized																					
1	2	3	4	5	6	7	8	9	82	2.4	25	16	52	11	29	82	8.9	31	43	2.4	14
10	11	12	13	14	15	16	17	18	45	1.3	7.0	7.1	32	2.9	13	45	1.6	7.6	6.1	1.3	2.6
random combinations, normalized																					
15	8	4	11	3	6	10	16	12	170	1.6	25	26	170	8.2	39	140	2.8	29	24	1.6	9.4
9	2	15	1	13	7	4	10	5	72	1.9	20	14	60	10	28	72	7.6	23	31	1.9	11
16	4	3	18	11	5	12	9	8	81	1.9	19	14	56	10	23	81	6.7	23	30	1.9	10
12	17	4	10	13	18	3	14	15	1100	1.7	38	110	120	8.0	34	1100	3.0	57	15	1.7	6.8
7	17	4	2	10	1	16	9	13	84	1.7	17	13	51	9.5	21	84	6.5	22	23	1.7	6.8
14	18	12	11	6	10	17	1	13	230	1.4	22	30	68	12	30	230	1.4	29	13	1.5	4.9
4	8	16	14	12	13	2	11	6	90	1.7	25	19	90	8.6	37	83	2.9	29	39	1.7	12
5	1	7	17	18	13	3	14	9	160	2.0	32	36	110	10	37	160	6.9	46	14	2.0	6.8
15	1	16	4	3	14	17	2	7	56	2.1	18	13	56	14	28	54	5.1	21	21	2.1	8.2
3	7	18	10	6	15	2	9	13	61	1.8	17	12	55	12	28	61	2.3	19	29	1.8	8.4
3	17	10	13	7	5	16	18	12	98	2.1	25	23	77	12	30	98	8.5	34	18	2.1	7.2
9	6	17	3	13	2	15	7	5	80	1.9	23	15	45	10	25	80	8.4	29	36	1.9	12
14	2	13	9	17	16	4	18	7	110	1.4	19	18	39	10	21	110	5.7	26	23	1.4	7.0
16	8	17	7	10	11	13	6	4	69	2.0	19	14	64	10	27	69	5.3	23	30	2.0	9.1
5	9	12	8	6	1	3	17	4	82	2.0	26	19	81	15	31	82	8.6	33	51	2.0	12
11	5	16	13	9	10	8	3	14	150	1.6	28	37	130	6.6	36	150	5.0	39	11	1.6	4.4
6	10	7	13	3	11	9	12	8	94	1.6	20	17	94	11	33	79	2.0	23	17	1.6	7.0
9	5	16	2	11	10	12	7	18	88	1.4	21	15	45	11	25	88	6.7	27	41	1.4	10
9	2	12	16	4	1	17	7	18	77	1.2	17	11	41	10	21	77	5.9	22	22	1.2	7.1
9	3	15	12	11	18	14	4	10	120	1.4	19	17	84	8.0	26	120	2.8	24	17	1.4	7.0
Random Array AVG									150	1.7	22		77	10	29	150	5.2	29	25	1.7	8.3
Random Array MEDIAN									89	1.7	20		66	10	28	83	5.5	26	23	1.7	7.7

[a] Detectors 1−9 are carbon black composites, and detectors 10−18 are the respective PEDOT−PSS composites.

polar analyte pairs. This behavior is a direct consequence of the increased sensitivity to polar compounds that is obtained by use of the PEDOT materials as opposed to the carbon black-filled composites.

Combinations of PEDOT−PSS detectors and carbon black detectors were also evaluated for their performance in vapor detection tasks. To avoid bias in the selection of an array of detectors, nine PEDOT-containing detectors and nine carbon black composite detectors were selected, each containing one member of polymers **1**−**9** in Figure 1. Nine-member arrays were then randomly selected from these 18 detectors. The rf values for pairwise resolution of the test solvents were then tabulated and the performance of these arrays in resolving solvents pairwise were compared to the performance of nine-member arrays containing only the PEDOT composite or the carbon black composite detectors.

As displayed in Table 8, the median of the average pairwise resolution factors for a randomly chosen array generally fell between the average pairwise resolution factors for the original nine-member carbon black and PEDOT arrays. These results show that, in most cases, the addition of a small number of carbon black detectors to the PEDOT array can increase its performance in discrimination of a broadly construed set of test solvents to approximately the same level as the array containing only carbon black detectors. The addition of carbon black detectors to an array of stable and well-behaved PEDOT PSS detectors also increases the performance significantly in the average pairwise resolution of nonpolar analytes from other nonpolar analytes.

CONCLUSIONS

Poly(3,4-ethylenedioxy)thiophene-based composite chemiresistor detector arrays have been shown to be able to classify a test set of 16 analyte vapors from the relative differential resistance responses produced upon exposure to the analytes. The stability of the conductivity, low noise levels, and chemical diversity that can be achieved by using a multitude of insulating polymers in the composites make PEDOT−PSS an attractive material for use as the conducting component of composite chemiresistor arrays. The PEDOT−PSS composite detector array outperformed a carbon black composite array consisting of the same insulating polymers in the mean discrimination of pairs of polar compounds, whereas the carbon black array outperformed the PEDOT−PSS composite detector array in discriminating between pairs of nonpolar analytes. The increased sensitivity of PEDOT-containing films to polar analytes resulted predominantly from increased sensitivity of the electrically conductive PEDOT regions to the sorption of analyte, as opposed to increased mass uptake of analyte by the composite film.

ACKNOWLEDGMENT

The authors thank NASA, DARPA, and an Army MURI for financial support.

SUPPORTING INFORMATION AVAILABLE

Assorted $\Delta R_{max}/R_b$ values for various detector/analyte combinations. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review September 20, 1999. Accepted March 2, 2000.

AC991079X

Highly Sensitive Detection and Discrimination of Biogenic Amines Utilizing Arrays of Polyaniline/Carbon Black Composite Vapor Detectors

Gregory A. Sotzing, Jennifer N. Phend, Robert H. Grubbs,* and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received November 1, 1999

We have previously reported the construction of an "electronic nose" consisting of an array of chemically sensitive resistors in which each resistor is composed of a conductor (typically carbon black, CB) dispersed into an insulating organic polymer.[1] In this architecture, sorption of an odorant into the insulating polymeric phase of the composite swells the material and produces a characteristic increase in the direct current (dc) electrical resistance response of the detector. These types of conducting polymer composite vapor detectors have been shown to exhibit trends in odor detection thresholds that generally parallel the human nose.[2] This thermodynamically based relationship holds for vapors of alkanes, alcohols, esters, carboxylic acids, and ketones, which all exhibit detection threshold behavior in human olfactory tests similar to that observed for the conducting polymer composite-based electronic noses.[2,3]

A striking exception to this trend is the behavior of biogenic amine odorants such as putrescine, cadaverine, and spermine. Although the average equilibrium sorption behavior, and thus the detection thresholds, of these odorants in the carbon black/insulating organic polymer composite detectors are close to those of analogous alkanes and alcohols, humans can detect biogenic amines at concentrations that are 10^3-10^4 lower than the corresponding chain length alkanes, alcohols, or ketones.[4] Without a mechanism to obtain increased sensitivity to these classes of compounds, it will not be possible for the conducting polymer composite-based electronic nose to reproduce faithfully the general odor classification or intensity perception characteristics of the human olfactory system. In addition, biogenic amines have been related to the freshness of foodstuffs such as meats,[5−9] cheeses,[10] alcoholic beverages,[11,12] and other fermented foods.[13−15] Additionally, aniline and *o*-toluidine have been reported to be biomarkers for patients having lung cancer,[16] whereas dimethylamine and trimethylamine have been reported to be the cause of the "fishy" uremic breath odor experienced by patients with renal failure.[17] We report herein sensor film chemistry that has allowed us to obtain a *million-fold* enhancement in sensitivity toward these biogenic amines, such that chemiresistor-based electronic noses can now significantly outperform the detection capabilities of human olfaction and other ambient-pressure detection methods for this important class of analytes.

Instead of using an insulating organic polymer in the composite chemiresistor films, we sought to exploit the possibility of the amines interacting with, and manipulating chemically, the electrical properties of a conducting organic polymer used in a composite material. In this fashion, a small degree of sorption could induce a significant change in the conductivity of the organic conducting regions of the composite. To explore this possibility, we used polyaniline, a proton-dopable conducting polymer, as the polymeric phase of a carbon black/polyaniline composite. Polyaniline was chosen because its half-oxidized form, the emeraldine base ($y = 0.5$), is rendered electrically conductive upon reaction with a strong acid. The conductive form of polyaniline, commonly referred to as the emeraldine salt (ES), has been reported to deprotonate to the emeraldine base and become insulating in alkaline environments.[18−20] Because of this unique doping/undoping behavior, polyaniline is particularly sensitive to both acids and bases and has been used as a gas sensor for ammonia and HCl vapors as well as for a solution pH detector.[21−24] We report the preparation and unique properties of five types of polyaniline/carbon black chemiresistive detectors in this work.

(1) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298.
(2) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci., U.S.A.* **1998**, *95*, 5442.
(3) Doleman, B. J. Ph.D. Thesis, California Institute of Technology, 1999.
(4) Devos, M.; Patte, F.; Rouault, J.; Laffort, P.; Van Gemert, L. J. *Standardized Human Olfactory Thresholds*; Oxford University Press: New York, 1990.
(5) Hernandezjover, T.; Izquierdopulido, M.; Vecianangues, M. T.; Vidalcarou, M. C. *J. Agric. Food Chem.* **1996**, *44*, 2710.
(6) Hernandezjover, T.; Izquierdopulido, M.; Vecianangues, M. T.; Vidalcarou, M. C. *J. Agric. Food Chem.* **1996**, *44*, 3097.
(7) Hernandezjover, T.; Izquierdopulido, M.; Vecianangues, M. T.; Marinefont, A.; Vidalcarou, M. C. *J. Agric. Food Chem.* **1997**, *45*, 2098.
(8) Vecianangues, M. T.; Marinefont, A.; Vidalcarou, M. C. *J. Agric. Food Chem.* **1997**, *45*, 4324.
(9) Vecianangues, M. T.; Marinefont, A.; Vidalcarou, M. C. *J. Agric. Food Chem.* **1997**, *45*, 2036.
(10) Muir, D. D.; Hunter, E. A.; Banks, J. M. *Milchwissenschaft* **1997**, *52*, 85.
(11) Izquierdopulido, M.; Hernandezjover, T.; Marinefont, A.; Vidalcarou, M. C. *J. Agric. Food Chem.* **1996**, *44*, 3159.
(12) Nouadje, G.; Simeon, N.; Dedieu, F.; Nertz, M.; Puig, P.; Couderc, F. *J. Chromatogr. A* **1997**, *765*, 337.
(13) Draisci, R.; Giannetti, L.; Boria, P.; Lucentini, L.; Palleschi, L.; Cavalli, S. *J. Chromatogr. A* **1998**, *798*, 109.
(14) Horneromendez, D.; Garridofernandez, A. *Analyst* **1994**, *119*, 2037.
(15) Kirschbaum, J.; Busch, I.; Bruckner, H. *Chromatographia* **1997**, *45*, 263.
(16) Preti, G.; Labows, J. N.; Kostelc, J. G.; Aldinger, S.; Daniele, R. *J. Chromatogr. Biomed. Appl.* **1988**, *432*, 1.
(17) Simenhoff, M. L.; Burke, J. F.; Saukkonen, J. J.; Ordinario, A. T.; Doty, R. *N. Engl. J. Med.* **1977**, *297*, 132.
(18) Epstein, A. J.; Macdiarmid, A. G. *Makromol. Chem., Macromol. Symp.* **1991**, *51*, 217.
(19) Ray, A.; Richter, A. F.; Macdiarmid, A. G.; Epstein, A. J. *Synth. Met.* **1989**, *29*, E151.
(20) Yoshino, K.; Gu, H. B. *Jpn. J. Appl. Phys. 1* **1986**, *25*, 1064.
(21) Dogan, S.; Akbulut, U.; Yalcin, T.; Suzer, S.; Toppare, L. *Synth. Met.* **1993**, *60*, 27.
(22) Dhawan, S. K.; Kumar, D.; Ram, M. K.; Chandra, S.; Trivedi, D. C. *Sens. Actuators B* **1997**, *40*, 99.
(23) Kukla, A. L.; Shirshov, Y. M.; Piletsky, S. A. *Sens. Actuators B* **1996**, *37*, 135.
(24) (a) Laranjeira, J. M. G.; Deazevedo, W. M.; Dearaujo, M. C. U. *Anal. Lett.* **1997**, *30*, 2189. (b) Sotomayor, M. D. T.; Depaoli, M. A.; Deoliveira, W. A. *Anal. Chim. Acta* **1997**, *353*, 275−280.

10.1021/cm990694e CCC: $19.00 © 2000 American Chemical Society
Published on Web 02/29/2000

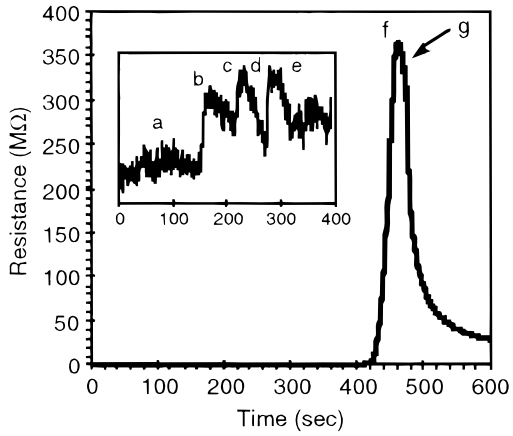


Figure 1. Resistance response of an EM−DBSA(1:0.5)/CB (80:20) detector exposed to water (a), acetone (b), methanol (c), ethyl acetate (d), butanol (e), and butylamine (f), all at 0.05% of their saturated vapor pressure. The ordinate on the inset was multiplied by a factor of $\approx 1 \times 10^6$ from the ordinate on the main plot of the figure. The value at the low end of the ordinate of the insert is 0.1682 MΩ, and the value at the high end of the ordinate is 0.1687 MΩ. The baseline value of R in the inset is 0.1683 MΩ. At point g, the sensor was removed from the chamber.

Emeraldine salt/carbon black suspensions were prepared by first dissolving the emeraldine base (Polysciences, Inc.) into hexafluoro-2-propanol. Sulfonic acid (reported as the mole ratio of polyaniline repeat to sulfonic acid) was then added,[25,26] and sufficient carbon black was introduced to produce a suspension that was 20 wt % carbon black and 80 wt % polymer, excluding the weight of the solvent. The solution was then sonicated for 20 min to disperse the carbon black. Substrate properties and other measurement methods have been described previously.[27]

Figure 1 depicts a series of headspace exposures (performed by injecting 1 mL of analyte-saturated air into a 2000 mL sealed glass container that contained a magnetically controlled fan to mix the headspace air in the vessel) of an ES−dodecylbenzenesulfonic acid (DBSA) (1:0.5)/CB detector to water (14.7 ppm), acetone (145 ppm), methanol (79 ppm), ethyl acetate (59 ppm), butanol (4 ppm) and butylamine (58 ppm). Each analyte was present at the same activity, corresponding to 0.05% of its vapor pressure. The steady-state relative differential resistance responses, defined as $\Delta R/R_b$, where R_b is the baseline resistance before exposure to the solvent vapor and ΔR is the change in resistance upon exposure to the solvent, were calculated to be 0, 0.0018, 0.0012, 0.0015, 0, and 2.25×10^3, respectively. Thus, the response of the detector is approximately 6 orders of magnitude greater to butylamine than to the five nonamine analytes. Composite detectors made from carbon black and insulating organic polymers, which can only respond by analyte-induced swelling effects on the insulating organic polymeric phase of the composite film, showed responses for all six analytes that were similar in magnitude to the responses displayed in the inset of Figure 1 of the polyaniline/carbon black composite detectors to nonamine vapors.

Detection thresholds, defined as the response required to obtain a signal-to-noise ratio of 3, were experimentally determined via headspace exposures to be 10 parts per trillion (ppt) for butylamine and 1 ppt for cadaverine on an ES−DBSA (1:0.5)/CB composite film. For comparative purposes, the human detection threshold for butylamine has been reported to be between 0.1 and 1 ppm,[28−30] and pristine emeraldine salt chemiresistive detectors have been reported to have detection thresholds of 1 ppm to ammonia.[23,31] Despite the pseudoreversibility of the ES−DBSA (1:0.5)/CB responses (see Figure 1) at high (between 50 parts per million and 1 part per thousand) concentrations of butylamine, the responses were reversible to exposures of butylamine at vapor phase concentrations between 10 ppt and 700 ppb. Exposure times under 30 s were generally sufficient to produce 90% of the limiting value of the electrical resistivity response. The variation in magnitude of the signals observed for a set of nominally identical detectors during exposures to a given vapor concentration of amine was less than 10% of the mean response of the detectors in such trials, with measurements typically performed on >5 detectors and for >10 exposures to each analyte.

The reversibility of the amine detectors can be explained in terms of the equilibrium established between the polyaniline emeraldine salt, and the amine analyte. At low concentrations of amine vapor, i.e., below 700 ppb for ES−DBSA (1:0.5)/CB, the polymer is not expected to undergo large changes in its secondary structure as a result of amine sorption. However, at concentrations above this limit, the secondary structure could undergo significant changes (i.e., breakup of crystallinity) that may not be restored on a rapid time scale and that may account for the quasireversibility of the signals at the higher amine concentrations.

To quantify the sorption behavior of these composites, a quartz crystal microbalance (QCM) with a 10 MHz resonant frequency was used to assess the mass uptake characteristics of butylamine, butanol, and acetone onto an ES−DBSA (1:0.5)/CB composite film. The response to both butanol and acetone was rapid (2 s to achieve 100% of the response) and reversible, whereas the response to butylamine was slower (10 s to achieve 80% of the response) and was only pseudoreversible. By using the Saurbrey equation and by assuming that modulus changes in the film are minimal, the mass uptakes for acetone, butanol, and butylamine were calculated to be 17.1, 23.2, and 272 ng cm^{-2}, respectively. Thus, although the sorption equilbria favor the amine relative to acetone or butanol, the sorption effect can only account for a factor of $\approx 10^1$ out of the $\approx 10^6$ enhancement in dc resistance sensitivity of these composites to amine vapors.

Evidence supporting an amine-induced change in the conductivity of the conducting organic polymer compo-

(25) Hopkins, A. R.; Rasmussen, P. G.; Basheer, R. A. *Macromolecules* **1996**, *29*, 7838.
(26) Hopkins, A. R.; Rasmussen, P. G.; Basheer, R. A.; Annis, B. K.; Wignall, G. D.; Hamilton, W. A. *Synth. Met.* **1998**, *97*, 47.
(27) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177.
(28) Hellman, T. M.; Small, F. H. *Chem. Eng. Prog.* **1973**, *69*, 75.
(29) Hellman, T. M.; Small, F. H. *J. Air Pollut. Control Assoc.* **1974**, *24*, 979.
(30) Laing, D. G.; Panhuber, H.; Baxter, R. I. *Chem. Senses Flavour* **1978**, *3*, 149.
(31) Lubentsov, B. Z.; Timofeeva, O. N.; Khidekel, M. L. *Synth. Met.* **1991**, *45*, 235.

nent of the chemiresistors was obtained through the use of UV−visible spectrophotometry. The spectrum for an ES−DBSA (1:0.5)/CB detector in air was identical to that previously reported for the emeraldine salt.[25] In contrast, after the injection of 1 μL of saturated butylamine vapor into a sealed 4.5 cm^3 quartz cuvette, the ES/CB film displayed absorption maxima at 340 and 660 nm. These two maxima have been observed previously for the emeraldine base and have been attributed to the $\pi-\pi^*$ transition (340 nm) and the polaron/bipolaron transition, respectively, in such materials.[25,32]

Classification and differentiation between amines was achieved by using an array of ES/CB detectors, each of which contained a chemically different polyaniline material to produce a diversity in response properties over the detector array. The $\Delta R/R_b$ values in response to aniline and butylamine, respectively, were as follows: ES−methanesulfonic acid (1:1)/CB, 0.389 ± 0.037 and 17.27± 2.6; ES−ethanesulfonic acid (1:1)/CB, 0.441± 0.0048 and 8.6 ± 0.9; ES−*p*-toluenesulfonic acid (1:0.5)/CB, 0.892 ± 0.067 and 8.6 ± 0.9; ES−*p*-dodecylbenzenesulfonic acid (1:0.5)/CB, 1.61± 0.081 and (6.7± 0.7) × 10^4; and ES−*p*-dodecylbenzenesulfonic acid (1:1)/CB, 16.31 ± 1.33 and 215 ± 52. Normalization of the $\Delta R/R_b$ values was performed to remove the concentration dependence of the data so that only changes in pattern type, as opposed to pattern height, contribute to the differences between the signatures of the different amines on the detector array. For these two amines, the Fisher linear discriminant method[27,33] yielded an resolution factor of 5300, where a resolution factor of 3 corresponds to a 98% probability of correctly identifying one analyte from the other. Thus, in this test case, the analytes were essentially perfectly separated from each other based on their distinctive patterns on the detector array.

The enhanced sensitivity obtained to this class of compounds from the ES/CB detectors indicates the flexibility in detector design that is made possible by using organic conductors with inorganic conductors to form composite chemiresistors as elements of vapor sensor arrays. Obtaining analogous sensitivity improvements for these classes of compounds using polymer-coated surface acoustic wave resonators[34,35] would require the introduction of amine-selective binding sites into the polymer film or the establishment of very large amplification effects due to extreme changes in the acoustic modulus of the polymer film upon analyte sorption. Mass-based detectors, such as QCMs[36] or micromachined cantilevers,[37] coated with these same polymer films could not display the enhanced sensitivity to amines reported here, given the small increase in sorption properties exhibited by these types of films for amine vapors relative to other classes of compounds. The superior dc electrical resistance sensitivity of these materials to biogenic amines now means that it is possible to exceed the sensitivities of organoleptic panels to these chemically important components of a variety of odors. Applications of these detectors in a variety of analytical scenarios, as well as exploitation of the chemiresistor-based approach described herein to obtain enhanced sensitivity toward other classes of biomedically important odorants, are currently in progress in our laboratories.

Acknowledgment. We acknowledge support from the Army Research Office (under a MURI grant program) and DARPA and thank Dr. C. R. Lewis for insightful discussions and suggestions during the conception and execution of this work.

Supporting Information Available: Frequency shift of an ES-DBSA-coated crystal, UV-vis spectrum of EM-DBSA, and pattern of relative differential resistance signals. This material is available free of charge via the Internet at http://pubs.acs.org.

CM990694E

(32) Rannou, P.; Gawlicka, A.; Berner, D.; Pron, A.; Nechtschein, M.; Djurado, D. *Macromolecules* **1998**, *31*, 3007.
(33) Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; John Wiley & Sons: New York, 1973.
(34) Penza, M.; Milella, E.; Anisimkin, V. *Sens. Actuators B* **1998**, *47*, 218.
(35) Zellers, E.; Han, M. *Anal. Chem.* **1996**, *68*, 2409.
(36) Chance, J.; Purdy, W. *Thin Solid Films* **1998**, *335*, 237.
(37) Ho, J.; Fang, Y.; Wu, K.; Hsieh, W.; Chen, C.; Chen, G.; Ju, M.; Lin, J.; Hwang, S. *Sens. Actuators B* **1998**, *50*, 227.

A 2-Dimensional Generalized Likelihood Ratio Test for Land Mine and Small Unexploded Ordnance Detection

Ping Gao and Leslie M. Collins

Abstract

The fundamental goals of land-mine and small unexploded ordnance (UXO) detection are to achieve a high probability of detection (P_d) and a low probability of false alarm (P_{fa}). Conventional methods usually fulfill the first goal at the cost of a high P_{fa}. In our previous work [1], [2], [3], we have shown that a Bayesian decision theoretic approach can be applied to improve the detectibility of land mines and small UXO targets using a single spatial sample of the electromagnetic induction (EMI) sensor data. In this paper, we present an alternative approach which significantly improves P_d at a fixed P_{fa} by utilizing features that capture the physical nature of EMI data within a statistical signal processing framework. The method we develop is a two-dimensional generalized likelihood ratio test (2-D GLRT) which utilizes spatial information from the sensor output. To illustrate the performance improvement, results obtained with the 2-D GLRT detector are compared to those for the standard threshold test for single channel time-domain sensor data, as well as the energy detector, the integral detector, and the single location generalized likelihood ratio test (1-D GLRT) detector for multi-channel time-domain EMI sensor data.

Keywords

Detection theory, Electromagnetic induction, Land mine detection, UXO detection.

P. Gao and L. Collins are members of the Electrical and Computer Engineering Department, Duke University, Durham, NC 27708-0291, U.S.A. Telephone: 919-660-5260, Fax: 919-660-5293, E-mail: pgao@ee.duke.edu, lcollins@ee.duke.edu.

Special symbols used in this paper

ω_n

$\alpha, \alpha_1, \alpha_0, \beta$

σ^2, δ

$\underline{r}$, $\underline{S_1}$, $\underline{S_0}$, $\underline{\hat{S}_1}$, $\underline{\hat{S}_0}$

λ

$\prod_i$

$\underline{\mu}_{d_1}, \underline{\mu}_{d_0}, \underline{\mu}_{y_1}, \underline{\mu}_{y_0}$

$\mathbf{\Sigma}_{d_1}, \mathbf{\Sigma}_{d_0}, \mathbf{\Sigma}_{y_1}, \mathbf{\Sigma}_{y_0}$

$\underline{\hat{\mu}}_{d_1}$ and $\underline{\hat{\mu}}_{d_0}$

$\hat{\mathbf{\Sigma}}_{d_1}, \hat{\mathbf{\Sigma}}_{d_0}, \hat{\mathbf{\Sigma}}_{y_1}$, and $\hat{\mathbf{\Sigma}}_{y_0}$

The paper contains total of 44 pages, 4 tables, 13 figures.

I. Introduction

The goal of any detection system is to achieve a high probability of detection (P_d) while at the same time maintaining a low probability of false alarm (P_{fa}). This is particularly important in land mine detection where P_d is required by the United Nations to be 99.6%, and P_{fa} is directly proportional to the time and cost to remediate a site. At the current rate of clearance, 1,100 years will be required to remove all land mines that are already emplaced (This statistic further assumes that no additional land mines are emplaced.) Thus, reducing the false alarm rate is of immediate importance. However, it is most often the case that land mine detectors which achieve high P_d do so at the cost of high P_{fa}. This is because conventional mine detection technologies simply seek anomalies caused either by land mines or clutter, but do not exploit the nature of the physical signature or the statistics of the sensor response due to mines and clutter. In this paper, we present an approach which significantly reduces P_{fa} at a fixed P_d by incorporating the underlying physics of electromagnetic induction (EMI) sensors into a statistical signal processing framework. This novel method is a two-dimensional generalized likelihood ratio test (2-D GLRT) algorithm, which utilizes the spatial measurements taken with an EMI sensor across a local area.

There are several devices which either have been used or have been proposed for use in land mine detection: magnetometer, infrared imager, electromagnetic induction (EMI) sensor, and ground penetrating radar (GPR). Among these sensors, the most well established is the EMI sensor. An EMI system is essentially a metal sensor which records the electromagnetic induction field that is the response from underground objects, clutter, etc. due to an incident electromagnetic field. An EMI system can detect mines which contain metal, as well as unexploded ordnance (UXO) or metallic anthropic clutter. In order to detect such targets, the EMI system normally operates at low frequencies ($<$ 1MHz), at which the conductivity- and permeability-dependent skin depth of the materials varies significantly [4], [12]. Furthermore, at these frequencies the displacement current is weak enough to be neglected [6]. Hence, the response of the pulsed EMI system, $\mathbf{r}$, at each location surveyed with the sensor can be modeled as a superposition of weighted resonant

responses:

$$\mathtt{r} = \sum_{n=1}^{N} A_n e^{-j\omega_n t} \tag{1}$$

where ω_n is the n-th natural resonant frequency of the object and A_n is the initial magnitude of the response corresponding to that natural resonant frequency. In practice, the real part of ω_n is very small, and thus can be ignored [6]. Also, the late time field, which is the field recorded by EMI sensors, is dominated by the lowest mode, which is approximately an exponential damping. Therefore, to a first order approximation the response can be modeled as:

$$\mathtt{r} = Ae^{-\alpha t} \tag{2}$$

where A is the initial magnitude of the response and α is the dominant natural resonant frequency. A is strongly dependent on the excitation level, the depth, and the orientation of the underground objects [6], [12]. The resonant frequency, α, can be used to identify land mines because it is a function of conductivity and permeability, which are unique to each metal type [4], [12], [6]. Generally, the response from a relatively high metal content mine has a lower natural resonant frequency than that of anthropic clutter, *i.e.* the decay rate of a target signature is slower.

In this paper, several detection approaches are explored for two types of EMI sensors: (1) multi-channel time-domain EMI systems and (2) single channel integrated time-domain EMI systems. The detection approaches are validated using field data collected in conjunction with the *DARPA Backgrounds Clutter Data Collection Experiment* [7]. One example of a prototype multi-channel time-domain pulsed EMI sensor is the Geonics EM61-3D sensor, which was used to collect data in the *DARPA* experiment. This sensor samples the induced response at 20 geometrically spaced time gates from $320\mu s$ to $30ms$ following the incident pulse [7]. Thus, the received signal from the Geonics EM61-3D sensor can be expressed as:

$$\underline{\mathtt{r}} = Ae^{-\alpha \underline{t}} \tag{3}$$

where $\{t_i\}$, an element of $\underline{t}$, is the sampling time, $i = 1, 2, ..., 20$. An example of the second type of sensors, a single channel time-domain EMI system, is the Geonics EM61 sensor, also used in the *DARPA* study. This sensor integrates the time-domain induced response

within a pre-determined range of time to obtain a scalar value at each survey location. Therefore, its response can be expressed approximately as:

$$\mathtt{r} = \sum_{i=0}^{N} A e^{-\alpha(t_0 + i\Delta t)} \tag{4}$$

where t_0 is the initial time for the integration, and the integration ends at time $t_0 + N\Delta t$.

The traditional approach to mine detection using data from a single channel EMI sensor is to perform a threshold test on the data obtained at each individual survey location. As stated before, the phenomenon exploited to distinguish mines from clutter is that the decay rate of the target response is generally slower than that of clutter. Discrimination based on the initial magnitude A is not investigated in this paper; however, experimental data indicates that, on average, A for mines is greater than that of clutter. Thus, after integrating the response, the output from a single channel integrated time-domain EMI sensor due to a target is usually greater than that of clutter. An extension of this approach to multi-channel EMI sensor data is to perform a threshold test on either the energy present in the received signal, or on the integral (sum) of the sampled values at each survey location. These two approaches essentially convert the multi-channel responses to single channel data in order to make a decision.

In addition to the conventional methods, there have been several studies on mine detection using statistical methods. In [1], [2], [3], [8], [9], [10], the statistical characterization and modeling of mine fields have been addressed. Reference [11] introduced an area-based "δ-technique", which incorporates spatial information through the energy levels of an EMI sensor response at the center point under test and its immediate neighbors. This technique utilizes the number of neighbors whose energy level is lower than that of the center location under test as the criterion of whether a mine is present or not. This technique is similar to CFAR detection [18].

In our previous work [1], [2], [3], we have applied signal detection theory to generate both the one-dimensional likelihood ratio test (LRT) and the GLRT for the Geonics EM61 and EM61-3D sensor data. The probability density functions describing the sensor response

to target and clutter were used to formulate the likelihood ratio at each surveyed location. We have shown that the performance of the two detection approaches were the same experimentally and theoretically for multi-channel sensor data [3], [17]. In addition, the 1-D GLRT has been shown to improve performance dramatically over the performance achieved with the standard tests on multi-channel sensor data. We have also proven that a threshold test on the raw data is the optimal processor for single channel EMI sensor data from a single location, even when the decay rates for target and clutter are not deterministic parameters [1].

In this paper, a statistical approach incorporating both the underlying physics of EMI sensors [4], [5], [12], [13], [14] and statistical signal processing theory [15], in which the statistics of local data surrounding the location under test are utilized, is presented. Since the minefields are sampled spatially, *i.e.* the sensor head is moved throughout the candidate area, we have hypothesized that the accuracy of the mine detectors would be improved when spatial information is incorporated into processor. The results demonstrate that this hypothesis is valid.

The remainder of this paper is organized as follows. In order to illustrate the improvement of the 2-D GLRT, it is compared to several other detection schemes including the threshold, energy, and integral detectors, the δ-technique, and the 1-D GLRT. Each of these detectors is described in Sec. II. In Sec. III we introduce a 2-dimensional GLRT detector for land mine and small UXO detection. The field data utilized to analyze performance is described in Sec. IV. Next, the results obtained by applying each of these detectors to the experimental data are presented in Sec. V. Finally, the results are discussed in Sec. VI.

II. 1-Dimensional Generalized Likelihood Ratio Test and Other Detectors

First, several simple detection approaches are described. It is necessary to investigate these detection techniques in order to evaluate the significance of the performance im-

provement achieved by the 2-D GLRT. The distributions of the outputs of multi-channel time-domain EMI sensors can be modeled with multivariate Gaussian density functions [1]. $\underline{r}|H_1 \sim \mathcal{N}(\underline{S_1}(A_1, \alpha_1), \sigma^2 I)$ and $\underline{r}|H_0 \sim \mathcal{N}(\underline{S_0}(A_0, \alpha_0), \sigma^2 I)$ where $\underline{r}$ is the output of the sensor as a function of time, A_1 and A_0 are the initial magnitude and α_1 and α_0 are the natural resonant frequencies under H_1 (target-present hypothesis) and H_0 (no-target) hypothesis, $\underline{S_1}(A_1, \alpha_1) = A_1 e^{-\alpha_1 \underline{t}}$, $\underline{S_0}(A_0, \alpha_0) = A_0 e^{-\alpha_0 \underline{t}}$, which correspond to theoretical values of the sensor output of the target and clutter under the parameters of A_1, α_1, A_0, and α_0 at sample time $\underline{t}$, I is N by N identity matrix where N is 20 for data from the Geonics EM61-3D sensor, and σ^2 is variance of each element of $\underline{r}$. Since the sensor is assumed to be subject to independent identical Gaussian white noise, the variance of each element of $\underline{r}$ is the same. This hypothesis was verified by examining the field data [1]. Thus, the generalized likelihood ratio test for a single spatial measurement can be formulated [2] as:

$$\lambda(\underline{r}) = \frac{\prod_i e^{-\frac{[r_i - \hat{S}_{1i}(\hat{A}_1, \hat{\alpha}_1)]^2}{2\hat{\sigma}^2}}}{\prod_i e^{-\frac{[r_i - \hat{S}_{0i}(\hat{A}_0, \hat{\alpha}_0)]^2}{2\hat{\sigma}^2}}} \tag{5}$$

where r_i is the output of the sensors at time indices $i = 1, 2, ..., 20$, and $\hat{\sigma}^2$, $\hat{S}_{1i}(\hat{A}_1, \hat{\alpha}_1)$ and $\hat{S}_{0i}(\hat{A}_0, \hat{\alpha}_0)$ are the maximum likelihood estimates (MLE) of σ^2, $S_{1i}(A_1, \alpha_1)$ and $S_{0i}(A_0, \alpha_0)$.

Since A_1, α_1, A_0 and α_0 are random variables (rv's), a Least-Squares method was used to estimate A_1, α_1 for each H_1 item and A_0, α_0 for each H_0 item from data taken in the calibration area, an area which provided training data (details can be found in Sec. IV.A). The sample means of these estimates are denoted $\hat{A}_1, \hat{\alpha}_1, \hat{A}_0$, and $\hat{\alpha}_0$, which are the estimates of A_1, α_1, A_0 and α_0 used for Eqn. (5). Thus, $\hat{S}_{1i}(\hat{A}_1, \hat{\alpha}_1) = \hat{A}_1 e^{-\hat{\alpha}_1 t_i}$ and $\hat{S}_{0i}(\hat{A}_0, \hat{\alpha}_0) = \hat{A}_0 e^{-\hat{\alpha}_0 t_i}$. A GLRT was used, as the prior knowledge available from the calibration area regarding the distribution of the random parameters was not considered statistically reliable due to the paucity of emplaced targets in the test sites. (Only 14 metal targets were emplaced in the calibration area and could be used for the estimation of A_1 and α_1. However, the calibration area contained 886 locations which could be used to estimate A_0 and α_0.) It should be noted that 1-D GLRT detectors provide processors which in some instances achieve performance that is identical to that of the likelihood ratio test performed at single location [17], although in general they are not optimum [15], [16]. The optimal detector requires integration over all uncertain parameters, and

the computational expense associated with evaluation of the multidimensional integral usually precludes implementing an optimal detector.

The standard threshold test for single channel integrated time-domain EMI data was also implemented for the EM61 sensor, since in this case the single location optimal detector reduces to a threshold detector as shown in [1], [3]. Other standard detectors, such as the energy and integral detectors, were also implemented for the multi-channel time-domain EMI data. The energy detector is essentially a threshold test performed on the energy presented at each surveyed location. The integral detector is a processor which integrates, or sums the time samples at each location to obtain a scalar output and then thresholds the output.

The δ-technique is an approach which includes spatial information into the decision-making process in an ad-hoc fashion [11]. This method simply considers the number of immediate neighbors (up to 8 neighbors in the case of *DARPA* data) whose energy values are strictly less than that of the center location under test as the statistic used to make a mine or no-mine decision. When this number is greater than a pre-determined integer value (usually 7 or 8), the decision made is that a mine is present; otherwise a no-mine decision is made.

III. Detector Design - 2-Dimensional Likelihood Ratio Test

Generally, it can be assumed that data can be obtained at sample points which are laid out as a rectangular mesh as illustrated in Fig. 1. Our goal is to make a decision on whether there is a mine present at the center location, $\mathbf{X}$. The data that is available consists of the outputs from the sensor both at the center location and the surrounding locations. Three possible situations can occur:

(**i**) there is a mine buried under the center location; or

(**ii**) there is nothing buried in this area; or

(**iii**) there is clutter located in this area.

The fourth possibility is both a mine and clutter (potentially at different depths) are

present at the location under test. This scenario is treated as situation (**i**) in this work. Since an EMI sensor is essentially a metal detector, when an object which contains metal is present the sensor will record an induced response due to the object, in addition to background noise. In situation (**i**), because of the presence of a mine, the sensor response at the surrounding locations may also include a response due to the mine, *i.e.* part of the response may induced by the mine. As the sensor head is moved further and further away from the center point, the effect will vanish. Fig. 2 is a plot of the theoretical prediction of the EM61 sensor response from a metal object with length to diameter aspect ratio of 0.5 at a depth of 0.5 m. The response is plotted over a 2 m by 2 m area. The prediction is obtained using a EMI model based on a dipole approximation of the object (in free space) with the dimension of the Geonics EM61 sensor and with the object parallel to the ground. The overall level of the response shown in Fig. 2 varies depending on the strength of the sensor excitation. For situation (**ii**), the EMI sensor will only record background noise, and for situation (**iii**) the response from the position under test might be strong and the response at the surrounding locations might also be disturbed by the clutter response depending on the size and shape of the clutter.

To approach this detection problem statistically, two hypotheses are made. H_1 is the hypothesis that a target (a mine or UXO) is present at the location under test; and H_0 is the hypothesis that a target is not present, *i.e.*, either clutter or background alone is present. In our previous work, a mine/no-mine decision was made by only using data at the single position under test and tested by the (one-dimensional) likelihood ratio detector. The likelihood ratio test for a single sensor response, or location, is defined [15] as:

$$\lambda(r) = \frac{p(r|H_1)}{p(r|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \beta \tag{6}$$

where r is the received signal from the sensor at the location under test, $p(r|H_1)$ and $p(r|H_0)$ are the probability density functions under the target hypothesis, H_1, and under the clutter or no-mine hypothesis, H_0, respectively. When $\lambda(r)$ is greater than the threshold, β, the decision is H_1. When $\lambda(r)$ is less than β, the decision is H_0, or no target present. (In the case where $\lambda(r)$ is equal to β, either H_1 or H_0 can be assigned.)

In this work, we have hypothesized that for the land mine and small UXO detection problem the statistics of the sensor response obtained from the locations surrounding a test position may be a function of whether a target is present or not. This dependence will be a function of the target and sensor size, the data sampling grid, *etc.* Based on measured land mine, small UXO and clutter field data from EMI sensors, this hypothesis appears to be valid. Therefore, performance should improve when the information in the neighborhood of the test position is included in the decision process. Hence, we developed a 2-D GLRT detector, which incorporates not only the sensor response at the location under test, but also those obtained from the vicinity of the location under test. A GLRT, as opposed to a LRT, was used since parameters of the probability density functions describing the data are not known exactly and no reliable prior knowledge on these parameters are available. Thus, we chose to estimate the parameters based on training data in order to implement a GLRT.

In [1], [3], we have shown that for a single channel EMI sensor (for instance, the Geonics EM61) the optimum decision statistic at the point by point level is the received signal from the sensor. This decision statistic is equivalent to the one-dimensional LRT, and it can also be shown [17] that for multi-channel EMI data the one-dimensional LRT is equivalent to the 1-D GLRT under some practical assumptions. Hence, fusing data from the 1-D LRT/GLRT at various locations as opposed to using the sensor output or energy level, should result in a further improvement in performance.

Based on the above argument, the input, x, of the 2-D GLRT was chosen to be the output from the one-dimensional LRT at each of the spatial positions. For the single channel time-domain EMI sensor, this is the raw data from the sensor since the single location optimal detector reduces to a threshold test on the raw data as shown in [1], [3]. Since we have assumed that the sensor is subject to Gaussian noise, x is a Gaussian random variable. For the multi-channel EMI sensor, the data considered for the 2-D GLRT is the output of the 1-D GLRT. The 1-D GLRT is a simple version of the 1-D LRT which

avoids extensive computation and provides the same performance to that of the 1-D LRT under a set of assumptions [17]. As defined in [17], the 1-D GLRT is a matched filter-like processor. The output of the 1-D GLRT is the summation of N non-central Chi-square random variables [19] (N=20), and thus, can be approximated as a Gaussian random variable based on the Central Limit Theorem.

In the formulation of the 2-D GLRT, two physical properties of the input, x, are explored sequentially. The first property is the spatial change in x across a local area; the second property is the relative size of the center point compared to its neighbors. In fact, this cascade of two generalized likelihood ratio tests is first testing the mine *vs.* background hypothesis, and then testing the mine *vs.* clutter hypothesis.

Because of the small size of the target (between 10 cm and 20 cm) and the sensor dimensions of 1 m (EM61-3D) or 0.5 m (EM61), it was hypothesized that the target would only influence the response at nearby neighbors (within 1 m of the center of a target) [8]. Furthermore, since the input of the 2-D GLRT, x, is the likelihood ratio, its value at a "target present" location will theoretically be larger than those at the neighboring locations which are assumed not to contain targets. Hence, when the center location contains a target, which is often symmetric, the spatial pattern of x will resemble a "peaked" or hill shape, which has maximum value at the center location, and the value decreases from the center in all directions at an approximately equal rate (also see Fig. 2). If there is no metal located at the test area (background only case), the spatial pattern of x will be a spatial Gaussian field (assuming the sensor is subject to white Gaussian noise). For clutter, this spatial pattern will have no predictable pattern; it will depend on the clutter type, shape, size, amount of metal, *etc.* It might appear as a symmetric hill pattern, but can follow other shapes as well (*e.g.* asymmetric). Examples of these patterns from experimental data for target and no-target are shown in Figs. 3 and 4.

The second physical characteristic of a targets is that the absolute value of x for the target is substantially higher than those of its neighbors, and also that of the clutter across

the whole area. For example, a 50 gallon drum will not be considered as a target by the 2-D GLRT detector because even though the spatial pattern somewhat follows a symmetric hill shape, the rate of fall-off of x is not as fast as that for land mines or small UXO across the 2 m by 2 m area.

Therefore, to incorporate the two phenomenologies described above, the 2-D GLRT is formulated as a cascade of two GLRTs, each of which exploits one the two previously described properties. The first GLRT primarily tests situation (**i**) that is a target buried at the center of the testing area *vs.* situation (**ii**) that is nothing buried in the testing area, which effectively rejects noise and some of the clutter classes (very small items). The second GLRT tests situation (**i**) *vs.* situation (**iii**), which is clutter present case, to eliminate additional clutter whose spatial pattern appears to be a symmetric hill shape. Thus, the 2-D GLRT detector uses a cascade of two GLRTs formulated as follows.

A. Processing Stage 1

The first processing stage exploits the fact that the one-dimensional GLR data within a local area satisfies the symmetric hill pattern when a target is present. In order to achieve this goal, a new variable, $\underline{d}$, is formed as a vector of the differences between the input x at the center location and those values at the surrounding locations. Based on the sample speed and sensor size during experimental data acquisition, the grid size of data was chosen to be 1m by 0.4m, thus 10 surrounding locations are within 1 m (also the size of bigger sensor, EM61-3D) of the center location, as illustrated in Fig. 5. Hence, the vector of normalized differences from the center location, $\underline{d}$, is a 10 by 1 vector. Each element of $\underline{d}$ is defined by:

$$d_i = \frac{x_c - x_i}{x_c} = 1 - \frac{x_i}{x_c} \tag{7}$$

where $i = 1, 2, ..., 10$, x_c is the one-dimensional GLR value at the center location, x_i is the GLR value at i-th location as defined in Fig. 5.

Based on the statistics derived in [1], [3] and the Central Limit Theorem, we have assumed that x is a Gaussian random variable. In [20] and [21], the mathematical form of

the distribution of the ratio of two Gaussian random variables is derived. If $x_1 \sim \mathcal{N}(\mu_1, \sigma_1)$ and $x_2 \sim \mathcal{N}(\mu_2, \sigma_2)$, the correlation coefficient between x_1 and x_2 is ρ, and $v = \frac{x_1}{x_2}$, then the probability density function (pdf) of v is

$$\begin{aligned} f(v) &= \frac{1}{\pi}\frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{\sigma_1^2 - 2\rho v\sigma_1\sigma_2 + v^2\sigma_2^2}\exp\left[-\frac{1}{2}\frac{1}{1-\rho^2}\left(\frac{\mu_2^2}{\sigma_2^2} - 2\rho\frac{\mu_1\mu_2}{\sigma_1\sigma_2} + \frac{\mu_1^2}{\sigma_1^2}\right)\right] \\ &\quad + \frac{\sigma_1(\rho\mu_1\sigma_2 - \mu_2\sigma_1) + v\sigma_2(\rho\mu_2\sigma_1 - \mu_1\sigma_2)}{\pi(\sigma_1^2 - 2\rho v\sigma_1\sigma_2 + v^2\sigma_2^2)^{3/2}}\exp\left[-\frac{1}{2}\frac{(\mu_1 - v\mu_2)^2}{\sigma_1^2 - 2\rho v\sigma_1\sigma_2 + v^2\sigma_2^2}\right] \\ &\quad \times \int_0^{\frac{\sigma_1(\rho\mu_1\sigma_2 - \mu_2\sigma_1) + v\sigma_2(\rho\mu_2\sigma_1 - \mu_1\sigma_2)}{\sigma_1\sigma_2\{(1-\rho^2)(\sigma_1^2 - 2\rho v\sigma_1\sigma_2 + v^2\sigma_2^2)\}^{1/2}}} \exp\left[-\frac{1}{2}u^2\right]du \end{aligned} \tag{8}$$

In this problem, the two random variables (x_c and x_i) are assumed to be independent since noise is independent and identically distributed, thus, $\rho = 0$. Also, the variances of x_c and x_i are assumed to be equal, *i.e.* σ^2. Therefore, the pdf simplifies to

$$\begin{aligned} f(v) &= \frac{1}{\pi(1+v^2)}\exp\left(-\frac{\mu_1^2 + \mu_2^2}{2\sigma^2}\right) + \frac{\mu_2 + \mu_1 v}{\pi\sigma(1+v^2)^{3/2}} \\ &\quad \times \exp\left[-\frac{(\mu_1 - \mu_2 v)^2}{2\sigma^2(1+v^2)}\right] \times \int_0^{\frac{\mu_2 + \mu_1 v}{\sigma(1+v^2)^{1/2}}} \exp\left[-\frac{1}{2}u^2\right]du \end{aligned} \tag{9}$$

In [21], the detailed analytical error analysis is provided. Here, we show examples of a comparison of the accurate pdf and Gaussian approximation using estimates of μ_1, μ_2 and σ that were obtained from the calibration area data. Figs. 6 and 7 provide such comparisons for H_1 and H_0 case, respectively. Clearly, for this problem, a Gaussian approximation to the pdf is acceptable. Thus, we assumed the distribution of $\underline{d}$ is a multivariate Gaussian. Under H_1:

$$\underline{d}|H_1 \sim \mathcal{N}_{10}(\underline{\mu}_{d_1}, \mathbf{\Sigma}_{d_1}) \tag{10}$$

Under H_0:

$$\underline{d}|H_0 \sim \mathcal{N}_{10}(\underline{\mu}_{d_0}, \mathbf{\Sigma}_{d_0}) \tag{11}$$

where $\mathcal{N}_{10}$ represents 10-dimensional Gaussian distribution, $\underline{\mu}_{d_1}$ and $\underline{\mu}_{d_0}$ are the mean vectors, and $\mathbf{\Sigma}_{d_1}$ and $\mathbf{\Sigma}_{d_0}$ are the covariance matrices of $\underline{d}$ under H_1 and H_0, respectively. The sample means, $\hat{\underline{\mu}}_{d_1}$ and $\hat{\underline{\mu}}_{d_0}$, and the sample covariance matrices, $\hat{\mathbf{\Sigma}}_{d_1}$ and $\hat{\mathbf{\Sigma}}_{d_0}$, were calculated using the calibration area data (see Sec. IV), and then used as the estimated values of the true means and covariance matrices when calculating GLRs. The elements of the $\underline{d}$'s are not independent, especially under the H_1 hypothesis. The correlations between the elements are represented by the covariance matrix. This correlation provides an

added incentive to utilize the Gaussian formulation for the GLRT as opposed to assuming independence of the elements of $\underline{d}$ and using Eqn. (9).

Therefore, under these assumptions the decision statistic of the first processing stage has the form:

$$\Omega_1(\underline{d}) = \frac{p(\underline{d}|H_1, \hat{\underline{\mu}}_{d_1}, \hat{\mathbf{\Sigma}}_{d_1})}{p(\underline{d}|H_0, \hat{\underline{\mu}}_{d_0}, \hat{\mathbf{\Sigma}}_{d_0})} \tag{12}$$

$$= \frac{\frac{1}{(2\pi)^5|\hat{\mathbf{\Sigma}}_{d_1}|^{1/2}} e^{-\frac{1}{2}(\underline{d}-\hat{\underline{\mu}}_{d_1})^T \hat{\mathbf{\Sigma}}_{d_1}^{-1}(\underline{d}-\hat{\underline{\mu}}_{d_1})}}{\frac{1}{(2\pi)^5|\hat{\mathbf{\Sigma}}_{d_0}|^{1/2}} e^{-\frac{1}{2}(\underline{d}-\hat{\underline{\mu}}_{d_0})^T \hat{\mathbf{\Sigma}}_{d_0}^{-1}(\underline{d}-\hat{\underline{\mu}}_{d_0})}} \tag{13}$$

By taking the logarithm of Eqn. (12) and incorporating the constant terms into the threshold [15], the log generalized likelihood ratio simplifies to:

$$\Omega_1(\underline{d}) = -(\underline{d} - \hat{\underline{\mu}}_{d_1})^T \hat{\mathbf{\Sigma}}_{d_1}^{-1}(\underline{d} - \hat{\underline{\mu}}_{d_1}) + (\underline{d} - \hat{\underline{\mu}}_{d_0})^T \hat{\mathbf{\Sigma}}_{d_0}^{-1}(\underline{d} - \hat{\underline{\mu}}_{d_0}) \tag{14}$$

Since a logarithm is a monotonic increasing function, transformation from the original values to logarithmic values does not change the order of the corresponding values, and the resulting performance remains unchanged.

Fig. 8 illustrates the performance of the first processing stage evaluated on synthetic spatial data using the exact pdf of $\underline{d}$ and the Gaussian approximation. Similar performance is achieved when evaluating Ω_1 using Eqn. (12) by substituting the corresponding parameters, or using Eqn. (14), which assumes Gaussian distributed rv's.

B. Processing Stage 2

The generalized likelihood ratio used in the second processing stage is formulated to exploit the second property of the physical nature of the data mentioned above. Another vector, $\underline{y}$, is formed as:

$$\underline{y} = [x_c \quad \frac{x_c - m(\underline{x}_n)}{x_c}]^T \tag{15}$$

where x_c is the value of the one-dimensional GLR at the center location as before, $\underline{x}_n$ is a 10 by 1 vector which consists of the values of the one-dimensional GLR at the neighborhood locations around the center location as illustrated in Fig. 5, and $m(\underline{x}_n)$ is the

mean value of $\underline{x}_n$. The first element of $\underline{y}$ is the original x value at the center location. The statistics of x_c partially represent the likelihood of target present at that location since the x's are obtained from the one-dimensional LR, which is the best decision statistic at the single point by point level. The second element of $\underline{y}$ models the relative size of the center location compared to the locations surrounding it. For targets, this value should be greater than for clutter.

Under the bivariate Gaussian distribution assumption, which was chosen for similar reasons as in the formulation of the generalized likelihood ratio at the first stage, the log generalized likelihood ratio was formulated as:

$$\Omega_2 = -(\underline{y} - \underline{\hat{\mu}}_{y_1})^T \hat{\mathbf{\Sigma}}_{y_1}^{-1} (\underline{y} - \underline{\hat{\mu}}_{y_1}) + (\underline{y} - \underline{\hat{\mu}}_{y_0})^T \hat{\mathbf{\Sigma}}_{y_0}^{-1} (\underline{y} - \underline{\hat{\mu}}_{y_0}) \tag{16}$$

where $\underline{\mu}_{y_1}$ and $\underline{\mu}_{y_0}$ are the means, and $\mathbf{\Sigma}_{y_1}$ and $\mathbf{\Sigma}_{y_0}$ are the covariance matrices of $\underline{y}$ under H_1 and H_0, respectively. Again, the estimated values, $\underline{\hat{\mu}}_{y_1}$, $\underline{\hat{\mu}}_{y_0}$, $\hat{\mathbf{\Sigma}}_{y_1}$ and $\hat{\mathbf{\Sigma}}_{y_0}$, were sample means and sample covariance matrices obtained from the data in the calibration area (see Sec. IV).

C. Cascade of the Two Generalized Likelihood Ratios to Make a Decision

The two generalized likelihood ratios described above exploit the two physical properties of the target and clutter respectively. The overall 2-D GLRT was implemented by cascading the two individual GLRTs by:

- First eliminating the locations at which generalized likelihood ratio 1 is less than its threshold, β_1.
- Then calculating generalized likelihood ratio 2 for the remaining locations, the final decision is made by comparing generalized likelihood ratio 2 to its threshold, β_2.

In the first stage, those locations which do not have a symmetric hill pattern indicative of targets are classified as clutter or noise. Then, the second GLRT is applied to further confirm the decision, and the locations which do not satisfy the second statistical model are discarded from H_1 hypothesis. The choice of β_1 and β_2 determines the performance of the detector. If the threshold, β_1, of GLR 1 is chosen to detect all the targets, *i.e.* set equal

to the minimum value of mine data in calibration area, β_2 will dominate the probability of detection, on the other hand, if β_2 is chosen to detect all mines, P_d will depend on the value of β_1. The probability of false alarm for the sequential application of both GLRTs is smaller than that of either one of these two GLRTs performed alone. In this analysis, we use the minimum value of GLR 1 corresponding to mine-present hypothesis of each site as the β_1 value, then vary β_2 from its minimum to maximum value to generate an ROC curve. The values of β_1 and β_2 required to achieve the same performance (same P_d and P_{fa}) are different from site to site because the signature for targets at different sites is not the same. This procedure is similar to sensor fusion, the only difference is that here we fuse GLR 1 with GLR 2 only for locations where GLR 1 gives H_1.

IV. Experimental Details

A. Data

The objective of the *DARPA Background Clutter Data Collection Experiment* [7] was to collect data to aid in the understanding of the effects of clutter on system performance. During the course of the experiment, data was collected using four types of sensors: ground penetrating radar (GPR), electromagnetic induction (EMI), magnetometer, and infrared (IR). Data was collected at four sites distributed over two locations (Fort Carson, Colorado and Fort A. P. Hill, Virginia). The locations represented three different soil types, and exhibited substantial variability in man-made contamination (anthropic clutter). Three of the four sites had a small amount of anthropic clutter, while at one of the sites (Fort A. P. Hill) the amount of such man-made/anthropic clutter was substantial. In addition to the indigenous anthropic clutter, a few synthetic calibration targets, UXO items, and land mines were emplaced at each site.

A typical layout for a test site is illustrated in Fig. 9. The site measured 125 m by 100 m and consisted of three areas: a center square, side bars, and a calibration area. The center square, measuring 100 m by 100 m, was intended primarily for clutter characterization, although 15 calibration targets were placed there to aid equipment calibration and

location registration. Replicas of the calibration targets, along with synthetic unexploded ordnance (UXO) items and land mines were emplaced in the red, blue, yellow, and orange side bars on the sites. A complete description of the targets and their deployment can be found in [7]. In general, the red lane consisted of copies of the registration targets, other targets for calibration, and system-stressing targets. The blue side bar contained UXO and demagnetized spheres, the yellow side bar contained mines, infrared calibration sources, and dielectric targets. The orange side bar consisted of a combination of these, and other targets. The total area of the blue, yellow, and orange side bars was 1300 square meters. The calibration area consisted of a 30 m by 15 m portion of the red, blue, and yellow side bars. Table I lists the metal targets used in the *DARPA* experiments. From these targets and the size of the test sites, one can clearly see the motivation for these experiments was to examine the ability to detect high metal targets and to substantially decrease false alarms through system and algorithm improvements. Data from the calibration area has been extremely valuable in this work since it was used to obtain estimated means and covariance matrices, and to choose thresholds for the 2-D GLRT. The locations and descriptions of the targets in the calibration area (14 of which could be seen by EMI systems) and the center square (10 of which could be seen by EMI systems) were released to the public, the remaining locations (blue, yellow and orange side bars) have not been disclosed. In the analyses that are described below, only data from the calibration areas and the center square are processed to evaluate performance of the various detectors. Clearly, since only 10 targets were emplaced in the 105m x 100m center square, the main goal of these analyses was false alarm rate reduction.

B. Sensors

In this paper, we consider data from two classes of EMI sensors: sensors that integrate time-domain information to provide a single data point (standard metal detector), and those that provide a sampled portion of the time-domain waveform.

The first system, the Geonics EM61, consists of a single-channel pulsed induction system with a 0.5-m transmitter coil positioned approximately 0.3 m above the ground. Data is

received in both the transmitter coil, and a secondary receiver coil that is located 0.4 m above the transmitter. The system operates at a center frequency of 75 Hz; the received signal is integrated from 0.18 to 0.87 ms after each transmit pulse, and the result, a single data point, is stored for later processing. Data were collected along survey lines spaced 0.5 m in the easterly direction, and 0.2 m on the average in the northerly direction.

The second sensor, the Geonics EM61-3D, is a prototype, three-component, time-domain induction system having a 1 m square transmitter coil and three orthogonal 0.5 m receiver coils positioned approximately 0.3 m above the ground. The system operates at a center frequency of 7.5 Hz. Sensor output is measured and recorded at 20 geometrically spaced time gates, spanning a time range from 320 μs to 32 ms. Data were collected along survey lines spaced 1 m apart in the easterly direction and at a rate of 3 samples per second, or approximately 0.4 m, in the northerly direction.

V. Results

The data used to analyze the detection strategies were collected during the *DARPA Backgrounds Clutter Data Collection Experiment* [7] at four sites at two U.S. locations, Fort Carson, Colorado and Fort A.P. Hill, Virginia. The raw data were first divided into an appropriate raster based on the spatial separation of survey lines and the sampling rate of the sensors, and then all the responses collected in each grid were averaged.

The standard quantitative tool to evaluate performance of a detection algorithm is an receiver operating characteristic (ROC), which is a plot of probability of detection *vs.* probability of false alarm. The performance of each of the detection strategies is shown in terms of an ROC. Figs. 10 and 11 illustrate that the 10-location 2-D GLRT detector operating on the Geonics EM61 sensor data provides a substantial improvement in performance, compared to the standard threshold test at all four sites. In [1], [3], it is shown that the threshold test is the single location optimal detector for the integrated time-domain EMI sensors, thus no separate single location LRT performance is shown.

The improvements of the 2-D GLRT on Geonics EM61-3D sensor data, compared to the standard energy detector, the integral detector, and the one-dimensional GLRT, are demonstrated in Figs. 12 and 13.

Table II lists the actual threshold values used for β_1 for the first processing stage of the 2-D GLRT for all four test sites and two types of sensors. In Tables III and IV, the values of P_{fa} corresponding to P_d equal to 0.95 for various detection approaches are shown. Clearly, the improvement of 2-D GLRT is dramatic. The performance of the δ-technique is also shown in the tables. The decision rule was: if $m \geq 7$, H_1 is true; otherwise, H_0 is true, where m is the number of neighbors which have a smaller energy than that of the center location. The parameter m can be any integer from 1 to 8. Among these values, 8 can provide lowest P_{fa}; however, the trade-off is that at this value P_d does not usually achieve 100% since there are always some mines which have 1 or 2 neighbors which have higher energy than the center location due to noise. Therefore, 7 was used to obtain the results shown in the tables. It is clear that δ-technique can provide better detection performance than the standard detection methods, but the 2-D GLRT outperforms these alternative detection techniques. The 1-D GLRT performs better than a standard energy or integral test, and an energy detector is generally better than an integral detector. These results are consistent with the theoretical analysis of the performance of these detectors provided in [17].

VI. Conclusions

The above results indicate that the 2-D GLRT analysis can significantly reduce false alarm rates in land mine and small UXO detection scenarios using EMI sensor data. The performance improvement obtained for both multi-channel time-domain and single channel integrated time-domain EMI data were evaluated at four test sites, and were consistently high across all sites, even at the most highly cluttered Fort A.P. Hill FP20 site. This improvement occurs because the processor may correctly model some of the physical, statistical, and spatial properties of the output of EMI sensors.

Comparing the performance of different time-domain EMI sensors, single channel and multi-channel sensors, the multi-channel sensors provide slightly better performance. Even though the improvement is not significant at this point (because the multi-channel (EM61-3D) EMI is still a prototype sensor under development and the noise level at the receiving coil is high compared to the standard metal detector), it is believed that utilizing a time-domain waveform allows greater exploitation of the phenomenology embedded in the signatures.

For the multi-channel data, an energy detector provides better performance than an integral detector. This can be explained theoretically. At each time gate the noise that receiver is subject to is independent and identically distributed. Therefore, at lower signal levels, the Signal-to-Noise-Ratio (SNR) is less than that at higher signal levels. When applying the same weight to each channel and summing up all channels, (*i.e.* an integral detector), the influence of noise on the low-level signal is highlighted. The energy detector, however, assigns greater weight to the higher-level portions of the signal and less weight to the lower-level portions of the signal (squaring is not a linear, but a quadratic function). Thus, the noise present on the low-level signal will not interfere with the decision as much as it does for the integral detector.

Tables II and III list the performance of different detectors quantitatively. The reduction of P_{fa} for a fixed $P_d = 0.95$ realized by using the 2-D GLRT is at least a factor of 2. For highly cluttered sites, or noisy data, this factor can be as high as 20 over the standard test; a dramatic improvement. This demonstrates that the 2-D GLRT is robust in a noisy environment. The performance of another spatially-based detection strategy, called δ-technique, was also evaluated. This approach is easy to implement and does improve performance over non-spatial algorithms, but the 2-D GLRT provides better performance.

The mean and covariance matrices of the vectors $\underline{d}$ and $\underline{y}$ defined in Sec. II vary across sites. Because close sites provide similar results for the estimated mean and variance, these variations might be caused by the geological differences between the test sites. The

concentration of clutter also affects the variance of the data. To obtain such *a priori* knowledge before remediating an area at a particular site, one solution might be to take a sub-area close to the site as a calibration area to obtain the statistics of this site, then use these statistics to implement a 2-D GLRT. Another solution might be to find the parameters that affect those statistics and apply Bayesian analysis by assuming proper priors to obtain best estimates of those parameters or integrate over all the uncertain parameters.

Acknowledgments

This research has been supported by the Army Research Office under grant DAAH04-96-1-0448 (Demining MURI). The authors would like to acknowledge the valuable discussions with Dr. Lawrence Carin, Dr. Thomas Altshuler, Ms. Vivian George, Dr. Regina Dugan, Dr. Stacy Tantum, Dr. Erol Gelenbe, and Mr. Taskin Kocak regarding this work. Also, thanks to Dr. Yan Zhang for providing the EMI model code.

References

[1] L. M. Collins, P. Gao, and L. Carin, "An Improved Bayesian Decision Theoretic Approach for Land Mine Detection", IEEE Trans. Geoscience and Remote Sensing, vol. 37 No. 2, March 1998, pp.811-819.

[2] P. Gao and L. M. Collins, "Improved Signal Processing Approaches for Landmine Detection", Proceedings of SPIE, April 1998, Orlando, FL.

[3] P. Gao, Improved Approaches to Land Mine Remediation Using Signal Detection and Estimation Theory, Master's Thesis, Duke University, December, 1997.

[4] G. D. Sower, S. P. Cave, "Detection and Identification of Mines from Natural Magnetic and Electromagnetic Resonances", Proceedings of SPIE, Vol. 2496, Orlando, FL, April 1995.

[5] N. Geng, P. Garber, L. Collins, L. Carin, D. Hansen, D. Keiswetter, I.J. Won, Wideband Electromagnetic Induction for Metal-Target Identification: Theory, Measurement and Signal Processing, Duke University Technical Report, 1998.

[6] N. Geng, C. E. Baum, and L. Carin, "On the Low-Frequency Natural Response of Conducting and Permeable Targets", IEEE Trans. Geoscience and Remote Sensing, vol. 37 No. 1, January 1999, pp.347-359.

[7] V. George, T. Altshuler, A. Andrew, J. Nicoll, E. Cespedes, D. Butler, T. Broach, and R. Mehta, Background Data Collection Plan, DARPA/Defense Science Office, December 1996.

[8] T. E. Olson and C. E. Priebe, "Detection and Classification of mines via discriminant features and borrowed strength", Proceedings of SPIE, Vol. 3079, Orlando, FL, April 1997.

[9] C. E. Priebe, "Exploiting stochastic partitions for minefield detection", Proceedings of SPIE, Vol. 3079, Orlando, FL, April 1997.

[10] I. V. Basawa, "Unilateral Markov random fields for minefield modeling and detection", Proceedings of SPIE, Vol. 3079, Orlando, FL, April 1997.

[11] E. Gelenbe and T. Kocak, "Area-based results for mine detection", IEEE Trans. Geoscience and Remote Sensing, in press.

[12] C. E. Baum, Low-Frequency Near-Field Magnetic Scattering from Highly, but Not Perfectly, Conducting Bodies, Phillips Laboratory Interaction Note 499, November 1993.

[13] Y. Das, J. E. McFee and R. H. Cherry, "Time-domain response of a sphere in the field of a coil: Theory and experiment", IEEE Trans. Geoscience and Remote Sensing, vol. 22, pp.360-367, July 1984.

[14] Y. Das and J. E. McFee Analysis of an electromagnetic induction detector for real-time location of buried objects, IEEE Trans. Geoscience and Remote Sensing, vol. 28, pp.278-288, May 1990.

[15] H. L. Van Trees, Detection, Estimation, and Modulation Theory, New York: John Wiley and Sons, 1968. ch. 2, ch. 4. pp.19-116, pp.246-373.

[16] A. D. Whalen, Detection of Signals in Noise, New York: Academic Press, Inc. 1971. P.352

[17] P. Gao, and L. M. Collins, "A Theoretical Performance Analysis and Simulation Time-Domain EMI Data for Landmine Detection", submitted to IEEE Trans. Geoscience and Remote Sensing, 1997.

[18] L. L. Scharf, Statistical Signal Processing: Detection, Estimation, and Time Series Analysis, Massachusetts: Addison-Wesley, 1991. Ch.4, pp.140-149.

[19] N. L. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, vol.2, 2nd Ed. New York: John Wiley & Sons, Inc, ch.29, pp.433.

[20] E. C. Fieller, "The Distribution of the Index in a Normal Bivariate Population", Biometrika, **24**, pp.428-440, 1932.

[21] D. V. Hinkley, "On the Ratio of Two Correlated Normal Random Variables", Biometrika, **56**, pp.635-639, 1969.

List of Tables

I Targets emplaced in the center square and the calibration area, where 'Al' means aluminum, 'Fe' means iron and 'AP' mean anti-personnel. 26

II Threshold values used for the first step process, β_1, for all four test sites and two types of sensors, EM61 and EM61-3D sensor 27

III P_{fa} value when $P_d = 0.95$ for EM61 sensor 28

IV P_{fa} value when $P_d = 0.95$ for EM61-3D sensor 29

Target	Description	Depth to Top (cm)
Fe Sphere	4.875" diameter	5
Al Plate	4"x4"x0.5"	surface
Al Plate	8"x8"x1"	5
Fe Sphere	4.875" diameter	5
Al Plate	4"x4"x0.5"	surface
Al Plate	8"x8"x1"	5
Fe Sphere	4.875" diameter	5
Al Plate	4"x4"x0.5"	surface
Al Plate	8"x8"x1"	5
Fe Sphere	4.875" diameter	5
Al Plate	4"x4"x0.5"	surface
Al Plate	8"x8"x1"	5
Fe Sphere	4.875" diameter	5
Al Plate	4"x4"x0.5"	surface
Al Plate	8"x8"x1"	5
Fe Sphere	4.875" diameter	5
Al Plate	8"x8"x1"	5
Al Plate	4"x4"x0.5"	surface
Al Plate	8"x8"x1"	25
Al Plate	15"x15"x1"	100
Al Spherical Shell	10" diameter	60
Fe Sphere	4.875" diameter	55
Al Plate	white	surface
Al Plate	black	surface
OZM-3	Metal AP landmine	4

Fe Sphere	4.875" diameter	35
81 mm Mortar	UXO	30
152 mm Projectile	UXO	70
155 mm Projectile	UXO	55

TABLE I

TARGETS EMPLACED IN THE CENTER SQUARE AND THE CALIBRATION AREA, WHERE 'AL' MEANS ALUMINUM, 'FE' MEANS IRON AND 'AP' MEAN ANTI-PERSONNEL.

Site	EM61	EM61-3D
FP20	-13.51	-2.23
FP22	-30.64	-3.00
Seabee	-1.08	-17.23
Turkey Creek	-1.70	-3.47

TABLE II

THRESHOLD VALUES USED FOR THE FIRST STEP PROCESS, β_1, FOR ALL FOUR TEST SITES AND TWO TYPES OF SENSORS, EM61 AND EM61-3D SENSOR

Site	1-D LRT/GLRT	δ-technique ($m \geq 7$)	2-D GLRT
FP20	0.22	0.098	0.016
FP22	0.28	(cannot achieve P_d=0.95)	0.059
Seabee	0.066	0.038	0.038
Turkey Creek	0.12	0.048	0.0066

TABLE III

P_{fa} VALUE WHEN $P_d = 0.95$ FOR EM61 SENSOR

Site	Integral Detector	Energy Detector	1-D LRT/ 1-D GLRT	δ-technique ($m \geq 7$)	2-D GLRT
FP20	0.313	0.16	0.125	0.1	0.013
FP22	0.55	0.18	0.27	0.108	0.014
Seabee	0.287	0.099	0.057	0.047	0.016
Turkey Creek	0.158	0.12	0.086	0.059	0.008

TABLE IV

P_{fa} VALUE WHEN $P_d = 0.95$ FOR EM61-3D SENSOR

List of Figures

1 Schematic of data placement and data available for a testing location 32

2 Theoretical prediction of EM61 response from a metal object with an aspect ratio of 0.5 at a depth of 0.5m 33

3 Typical spatial response patterns of both mine/UXO and clutter/noise from single channel EMI sensor data 34

4 Typical spatial patterns of both mine/UXO and clutter/noise based on the 1-D GLR of multi-channel EMI sensor data 35

5 The neighborhood locations used when calculating generalized likelihood ratio 1 and 2 36

6 (Top) Probability density functions for the ratio of 2 Gaussian rv's. and a Gaussian approximation for a target present (H_1) case. (Bottom) Error between two pdfs shown in the top figure. 37

7 (Top) Probability density functions for the ratio of 2 Gaussian rv's. and a Gaussian approximation for a target not present (H_0) case. (Bottom) Error between two pdfs shown in the top figure. 38

8 A comparison of performance of the first stage processing using synthetic data based on the exact pdf of d and the Gaussian approximation 39

9 Typical site layout from the *DARPA* Background Clutter Data Experiment . 40

10 ROC of different detectors based on EM61 data at Fort A.P. Hill FP20 and FP22 sites. The left one, Fig. 10.a, is corresponding to Ft. A.P. Hill FP20 site, the right one, Fig. 10.b, is corresponding to FP22 site. Solid line = 2-D GLRT detector performance, dashed line = performance of standard threshold detector on the received data from EM61 sensor. 41

11 ROC of different detectors based on EM61 data at Fort Carson Seabee and Turkey Creek sites. The left one, Fig. 11.a, is corresponding to Ft. Carson Seabee site, the right one, Fig. 11.b, is corresponding to Turkey Creek site. Solid line = 2-D GLRT detector performance, dashed line = performance of standard threshold detector on the received data from EM61 sensor. 42

12 ROC of different detectors based on EM61-3D data at Fort A.P. Hill FP20 and FP22 sites. The left one, Fig. 12.a, is corresponding to Ft. A.P. Hill FP20 site, the right one, Fig. 12.b, is corresponding to FP22 site. Solid line = 2-D GLRT detector performance, dashed line = 1-D GLRT performance, dashed-dotted line = integral detector performance, dash-dot-dot line = energy detector performance. 43

13 ROC of different detectors based on EM61-3D data at Fort Carson Turkey Creek and Seabee sites. The left one, Fig. 13.a, is corresponding to Ft. Carson Seabee site, the right one, Fig. 13.b, is corresponding to Turkey Creek site. Solid line = 2-D GLRT detector performance, dashed line = 1-D GLRT performance, dashed-dotted line = integral detector performance, dash-dot-dot line = energy detector performance. 44

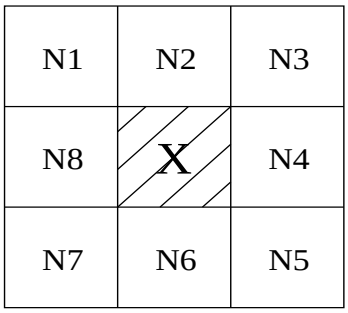


Fig. 1. Schematic of data placement and data available for a testing location

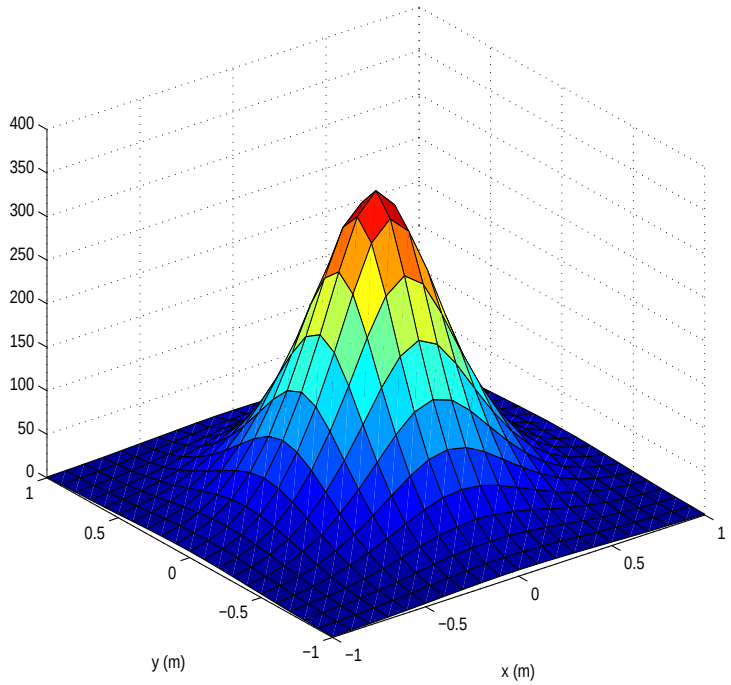


Fig. 2. Theoretical prediction of EM61 response from a metal object with an aspect ratio of 0.5 at a depth of 0.5m

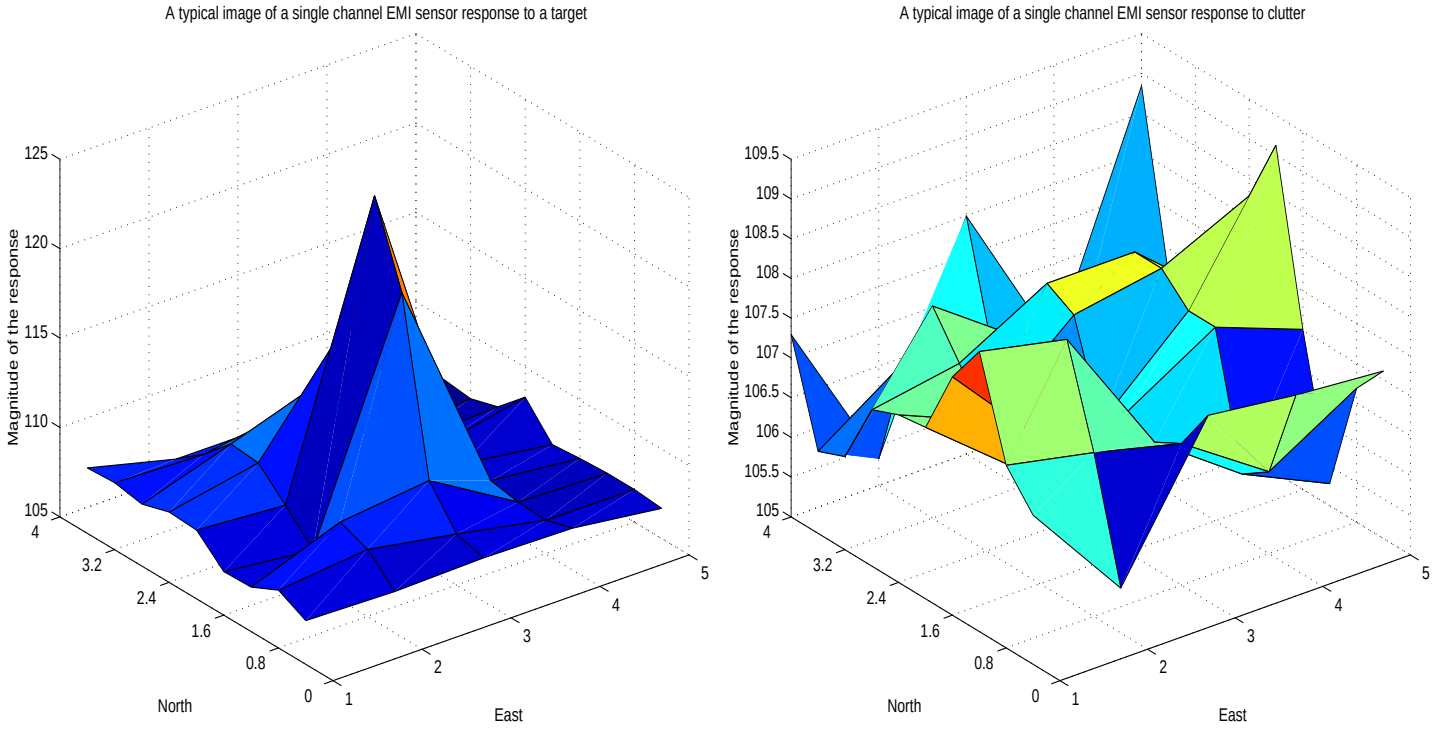


Fig. 3. Typical spatial response patterns of both mine/UXO and clutter/noise from single channel EMI sensor data

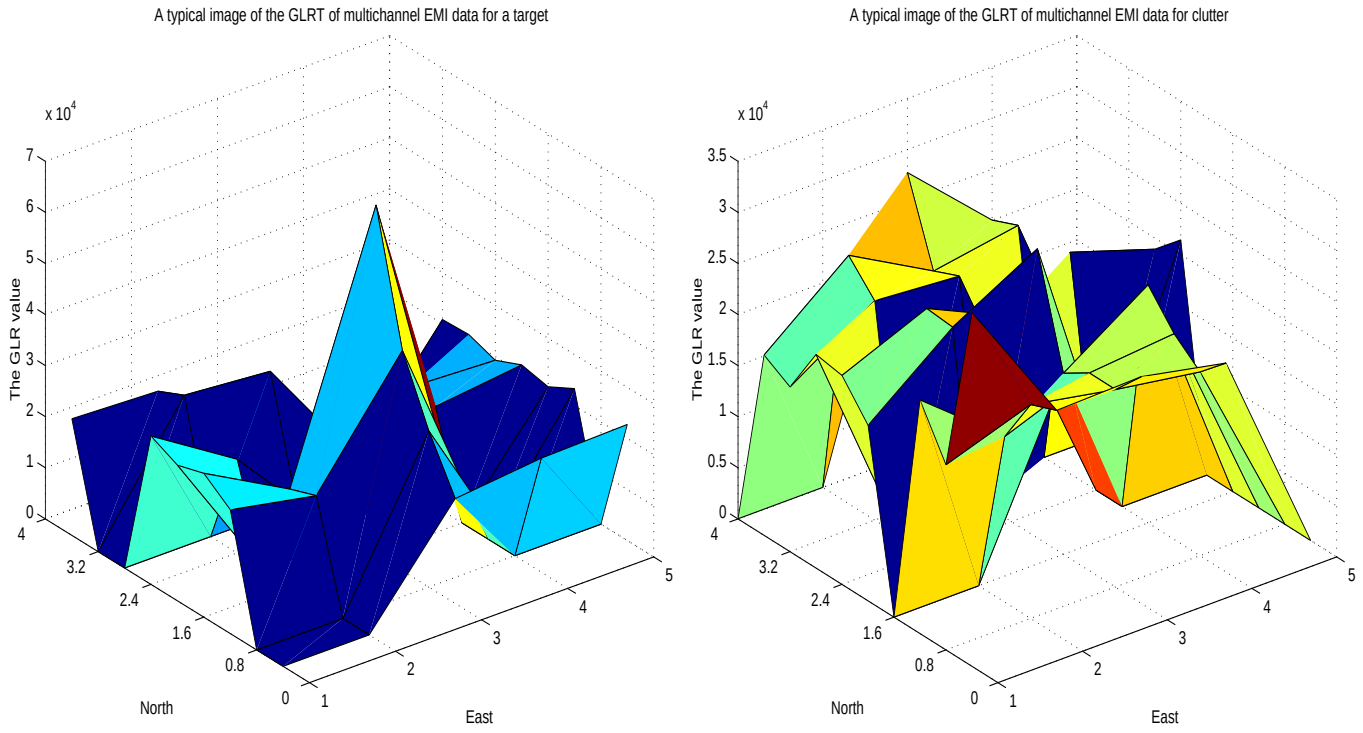


Fig. 4. Typical spatial patterns of both mine/UXO and clutter/noise based on the 1-D GLR of multi-channel EMI sensor data

	X7	
X1	X6	X8
X2	Center	X9
X3	X5	X10
	X4	

Fig. 5. The neighborhood locations used when calculating generalized likelihood ratio 1 and 2

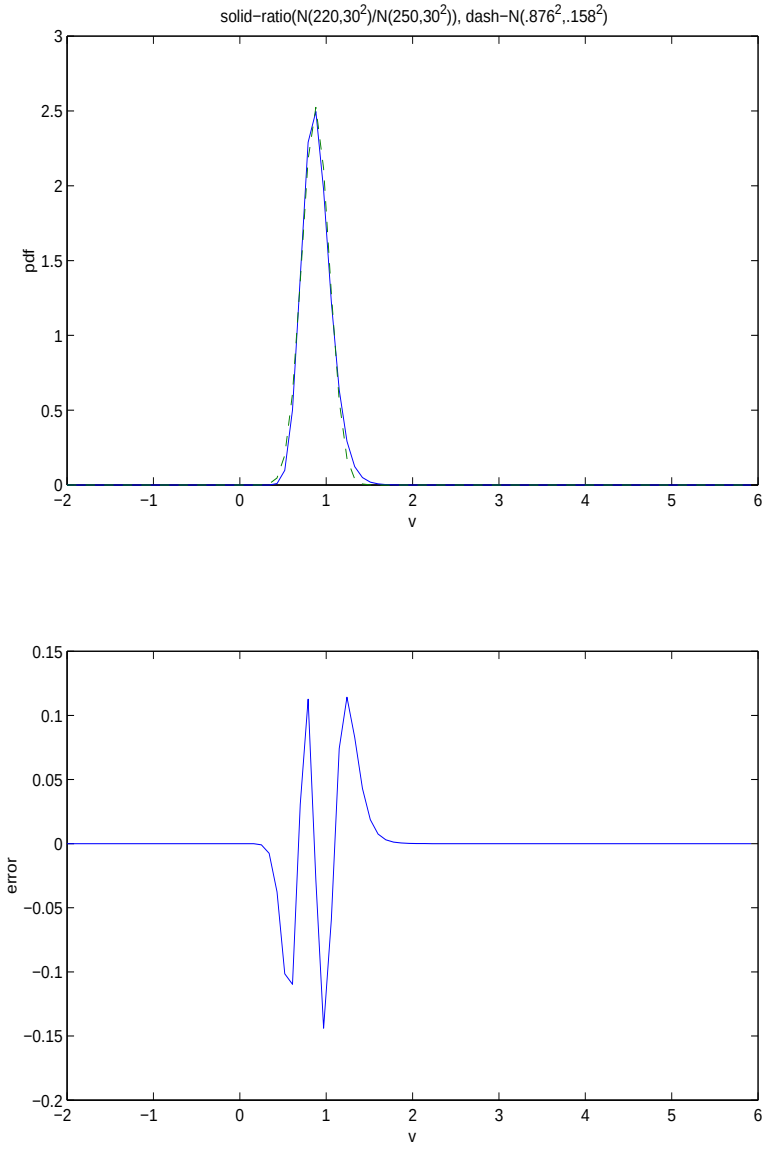


Fig. 6. (Top) Probability density functions for the ratio of 2 Gaussian rv's. and a Gaussian approximation for a target present (H_1) case. (Bottom) Error between two pdfs shown in the top figure.

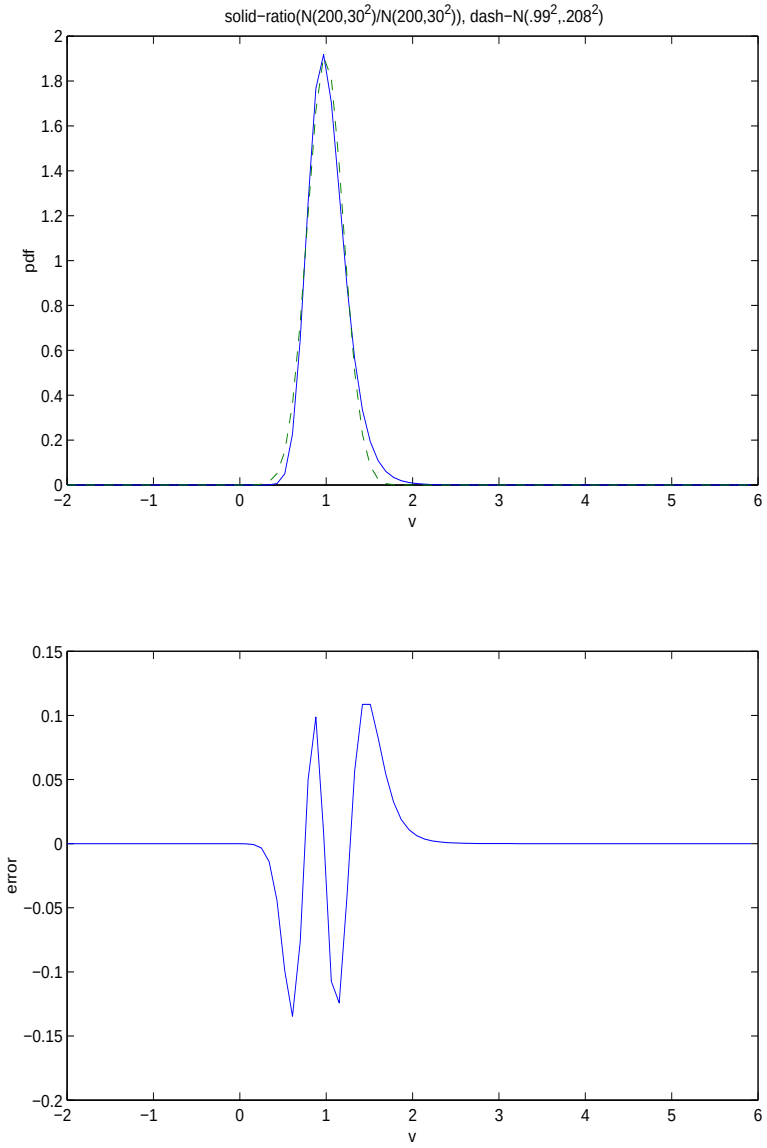


Fig. 7. (Top) Probability density functions for the ratio of 2 Gaussian rv's. and a Gaussian approximation for a target not present (H_0) case. (Bottom) Error between two pdfs shown in the top figure.

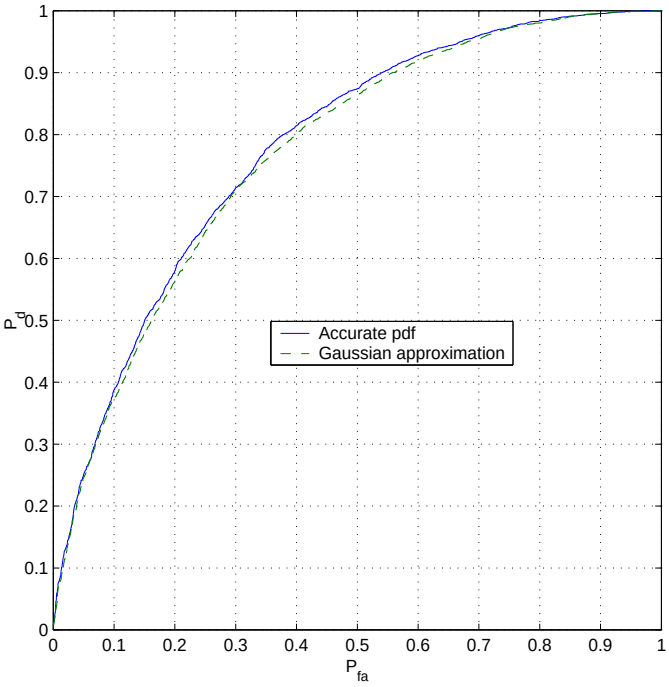


Fig. 8. A comparison of performance of the first stage processing using synthetic data based on the exact pdf of d and the Gaussian approximation

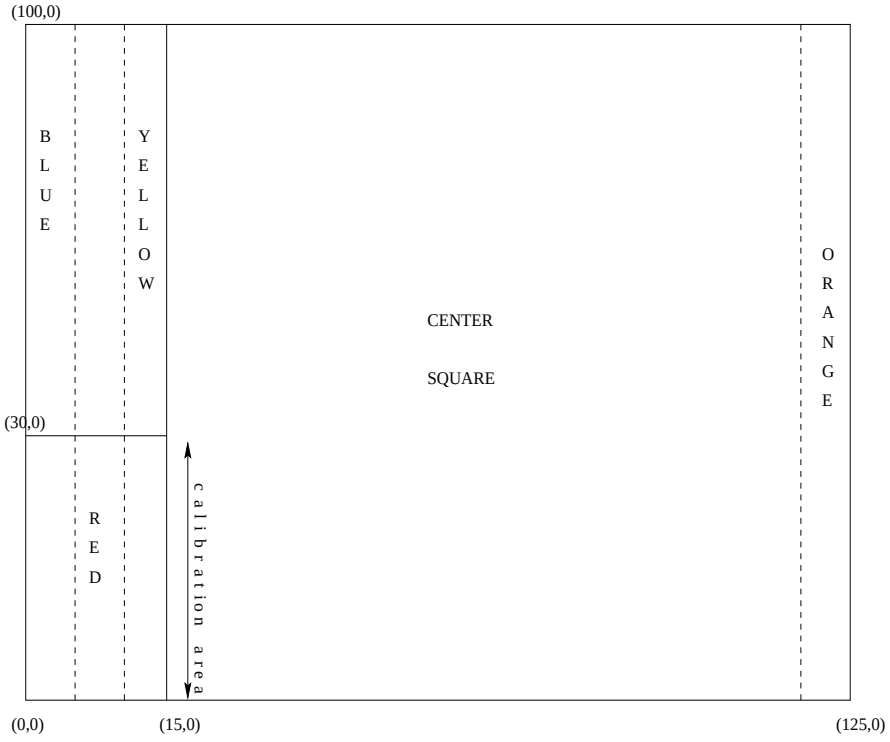


Fig. 9. Typical site layout from the *DARPA* Background Clutter Data Experiment

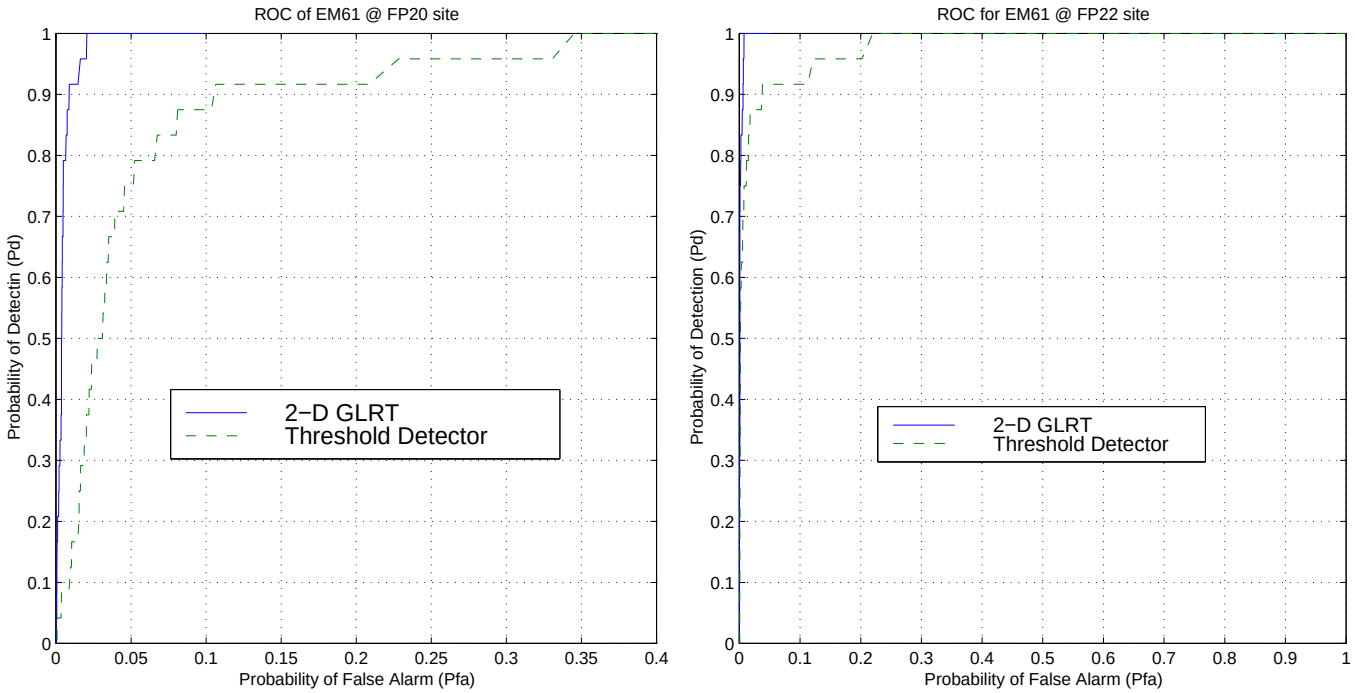


Fig. 10. ROC of different detectors based on EM61 data at Fort A.P. Hill FP20 and FP22 sites. The left one, Fig. 10.a, is corresponding to Ft. A.P. Hill FP20 site, the right one, Fig. 10.b, is corresponding to FP22 site. Solid line = 2-D GLRT detector performance, dashed line = performance of standard threshold detector on the received data from EM61 sensor.

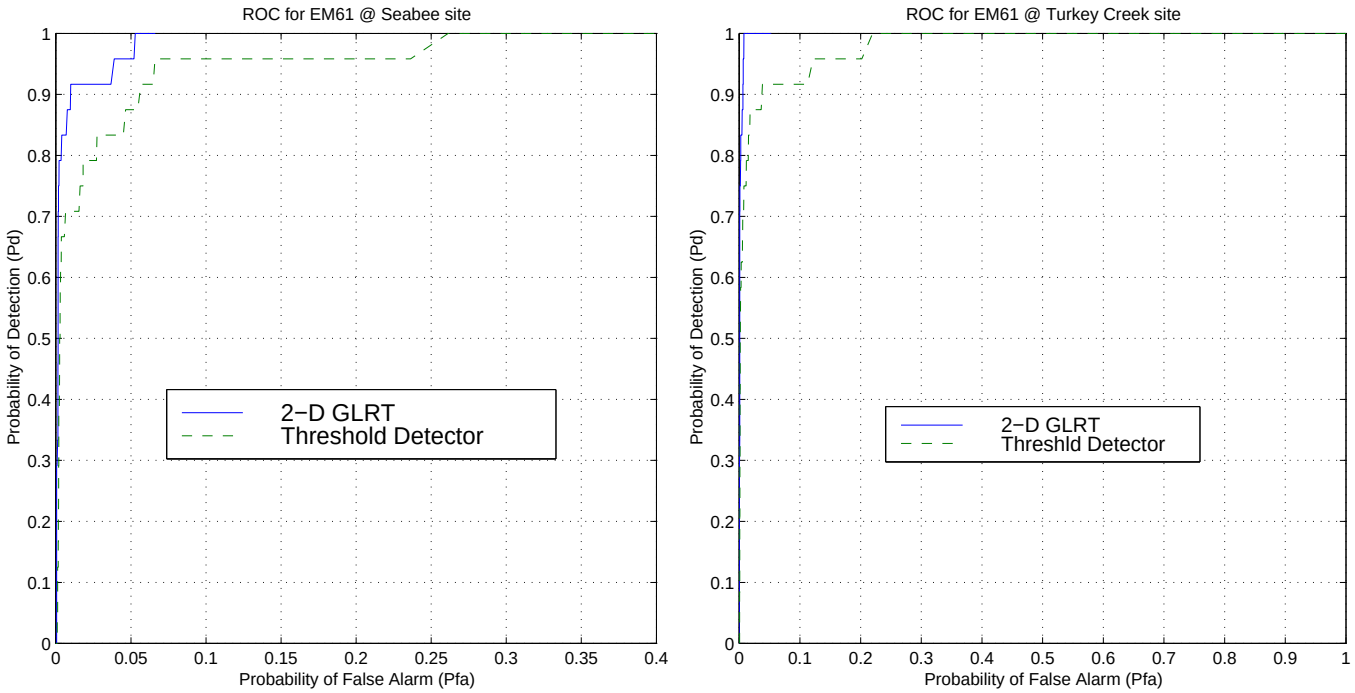


Fig. 11. ROC of different detectors based on EM61 data at Fort Carson Seabee and Turkey Creek sites. The left one, Fig. 11.a, is corresponding to Ft. Carson Seabee site, the right one, Fig. 11.b, is corresponding to Turkey Creek site. Solid line = 2-D GLRT detector performance, dashed line = performance of standard threshold detector on the received data from EM61 sensor.

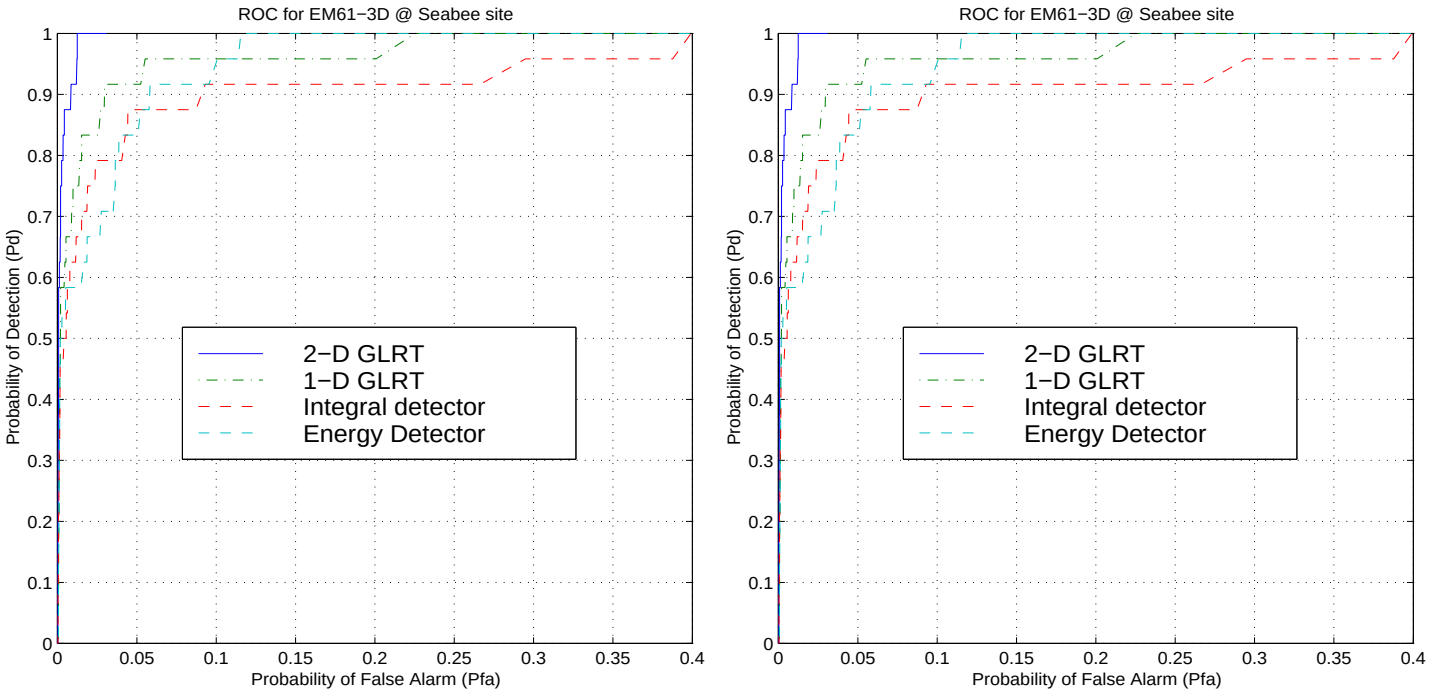


Fig. 12. ROC of different detectors based on EM61-3D data at Fort A.P. Hill FP20 and FP22 sites. The left one, Fig. 12.a, is corresponding to Ft. A.P. Hill FP20 site, the right one, Fig. 12.b, is corresponding to FP22 site. Solid line = 2-D GLRT detector performance, dashed line = 1-D GLRT performance, dashed-dotted line = integral detector performance, dash-dot-dot line = energy detector performance.

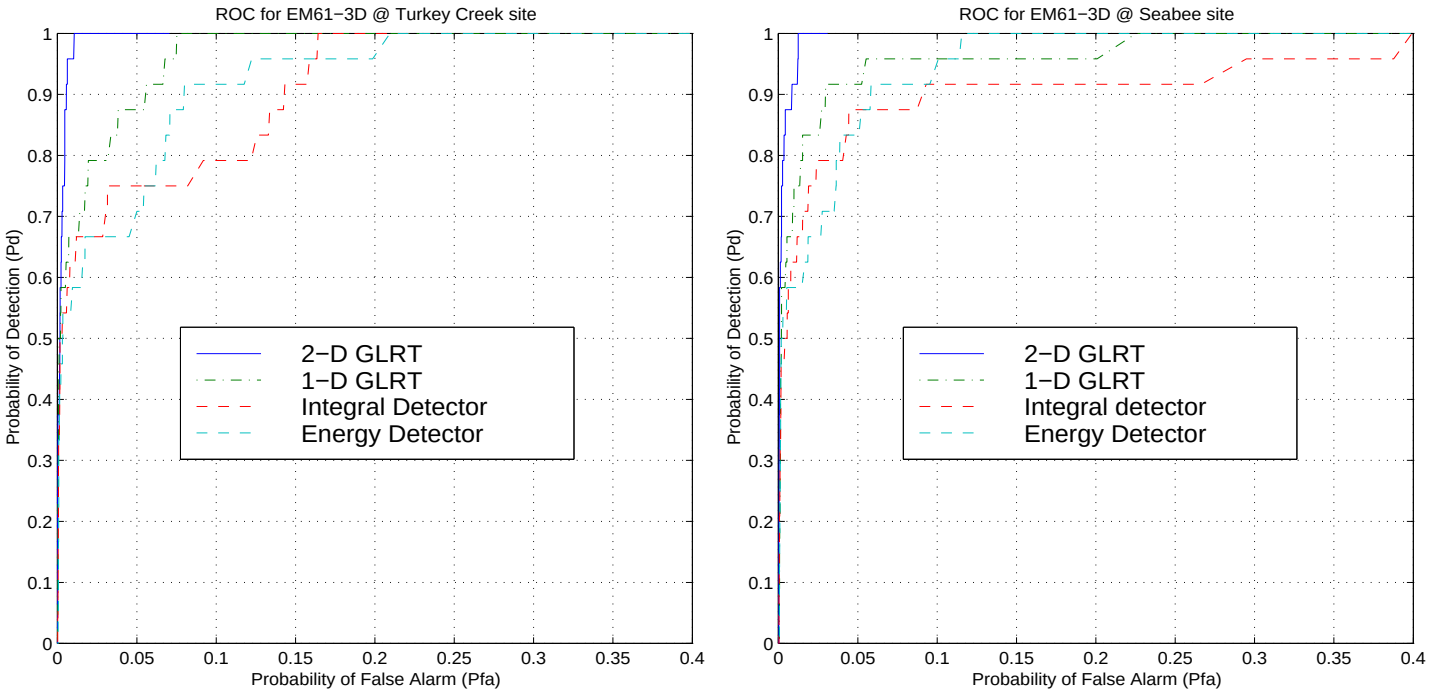


Fig. 13. ROC of different detectors based on EM61-3D data at Fort Carson Turkey Creek and Seabee sites. The left one, Fig. 13.a, is corresponding to Ft. Carson Seabee site, the right one, Fig. 13.b, is corresponding to Turkey Creek site. Solid line = 2-D GLRT detector performance, dashed line = 1-D GLRT performance, dashed-dotted line = integral detector performance, dash-dot-dot line = energy detector performance.

An acousto-electromagnetic sensor for locating land mines

Waymond R. Scott, Jr.[a], Chistoph Schroeder[a] and James S. Martin[b]

[a]School of Electrical and Computer Engineering
[b]School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

A hybrid technique is presented that simultaneously uses both electromagnetic and acoustic waves in a synergistic manner to detect buried land mines. The system consists ofan electromagnetic radar and an acoustic source. The acoustic source causes both the mine and the surface of the earth to be displaced. The electromagnetic radar is used to detect these displacements and, thus, the mine. To demonstrate the viability of this technique, an experimental system has been constructed. The system uses anelectrodynamic transducer to induce an acoustic surface wave, a tank filled with damp sand to simulate the earth, asimulated mine, and a radar to measure the vibrations. The technique looks promising; we have been able to detect both simulated antipersonnel mines and antitank mines buried in damp sand from the experimental results obtained with the system.

Keywords: land mine, mine detection, acoustic, ground penetrating radar, GPR

1. INTRODUCTION

A technique to detect land mines that simultaneously uses both electromagnetic and acoustic waves in a synergistic manner is currently being investigated. The synergism has the potential to significantly enhance the signature of the mine with respect to the clutter and make it possible to detect a mine that would be impossible to detect by purely electromagnetic or purely acoustic means.

The configuration of the system currently being studied is shown in figure 1. The system consists ofan electromagnetic radar and an acoustic source. The acoustic source induces an acoustic (seismic) wave into the earth. The acoustic wave causes both the mine and the surface of the earth to be displaced. The displacement of the mine is different than the earth, because the acoustic properties of the mine are quite different than those of the earth. The displacement of the surface of the earth when a mine is present is different than when it is not present because of the waves scattered from the mine. The electromagnetic radar is used to detect these displacements and, thus, the mine. This idea has been discussed previously [1, 2], but it has not been seriously investigated.

In addition, the radar can be simultaneously used in a mode in which only the electromagnetic waves are used to identify (image) the mine. In this electromagnetic only mode, the radar is a conventional ground penetrating radar (GPR). Thus, both the acoustic and electromagnetic properties of the mine can be used to differentiate it from the earth. It may be possible to locate or identify mines with the GPR that are invisible to the acoustic system or visa versa. Or it may me possible to combine the output of both the GPR and the acoustic system to lower the false alarm rate of the system. For example, a wet packed soil is generally very lossy for electromagnetic waves; thus, it is difficult to find the mines using the electromagnetic only mode. However, these soils tend to be lower loss for acoustic waves; thus, the acousto-electromagnetic mode may perform well in these soils. The opposite is also true. A loose dry soil is generally verylossy for acoustic waves; thus, it will be difficult to find the mine in this soil using theacousto-electromagnetic mode. However these soils tend to be low loss for electromagnetic waves; thus, the electromagnetic only mode may perform well in these soils.

An experimental model for the system has been constructed and automated. The experimental model is being used to demonstrate the viability of the technique and to study the interactions of the acoustic and electromagnetic waves with buried mines. The technique looks promising; we have been able to detect both simulated antipersonnel mines and antitank mines buried in damp sand. However, additional investigation of the technique is needed to determine the capabilities of the technique in more varied conditions. A two-dimensional finite-difference time-domain (FDTD) model for the acoustic waves has been also developed and is being used to help understand the interactions of the acoustic waves and the mines.

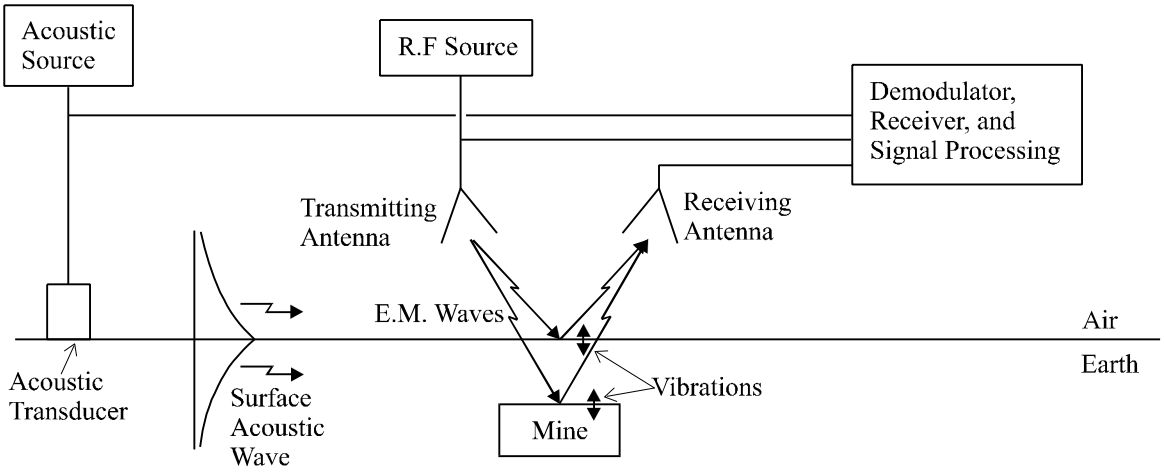


Figure 1. Acousto-electromagnetic mine detection system with the acoustic transducer placed on the surface of the earth.

2. EXPERIMENTAL MODEL

A radar has been designed and built to measure the displacements of the surface of the earth and the mine. Figure 2 is a schematic diagram of the radar. The radar radiates electromagnetic waves that are reflected off of a vibrating interface. The reflected waves are received by the radar, and a Homodyne system is used to demodulate the signals. The vibrations are determined from these demodulated signals. The two biggest challenges to make this radar perform adequately for the mine detection system are 1) to make it sufficiently sensitive to be able to detect the small vibrations, and 2) to make the spot size (the area on the surface illuminated by the electromagnetic waves) sufficiently small. The radar can measure vibrations as small as 1nm (10^{-9} m) as currently configured. To obtain this sensitivity, the radar was designed to minimize the effects of noise, such as the phase noise of the source and the electromagnetic interference from low-frequency magnetic fields. The spot size of the radar must be smaller than approximately one half of a wavelength of the acoustic waves. Currently, a small spot size is obtained by using an open-ended waveguide as the antenna for the radar. This antenna produces a sufficiently small spot size when the open end of the antenna is placed within a few centimeters of the surface. This antenna is adequate for the experimental model, but probably will not be adequate for a fieldable system. We have investigated antennas that may be appropriate for a fieldable system, but we have not tried to incorporate one into the current radar. The radar can be configured in either a monostatic mode as in figure 2 or in a bistatic mode as in figure 1. Currently the radar is being used in the monostatic mode to get a smaller spot size, but the radar has been used and has performed well in the bistatic mode. The radar can be operated at frequencies between 2 GHz and 8 GHz; however, all of the results presented in this paper are obtained with the radar operating at 8 GHz. The vibration of the surface of the sand has been measured with both the radar and an accelerometer; these measurements were compared and found to be in good agreement.

A drawing of the top view of the experimental model is shown in figure 3. The model consists of a tank that is filled with damp sand that has been packed to a relatively uniform density. The tank is approximately 120 cm

wide, 120 cm deep, and 240 cm long. A transducer is placed on the surface of the sand and is used to launch the acoustic waves into the sand. The transducer is an electrodynamic shaker that is driven with a signal generator. The transducer has been coupled to the surface of the sand so that it preferentially launches acoustic surface waves. These surface waves travel across the surface of the tank and interact with a mine that is buried in the sand. The electromagnetic radar is used to measure the displacements caused by the acoustic waves. A x-y positioner is used to scan the radar over the surface of the sand. This system is under computer control, so it can scan the radar over the surface and record the data automatically.

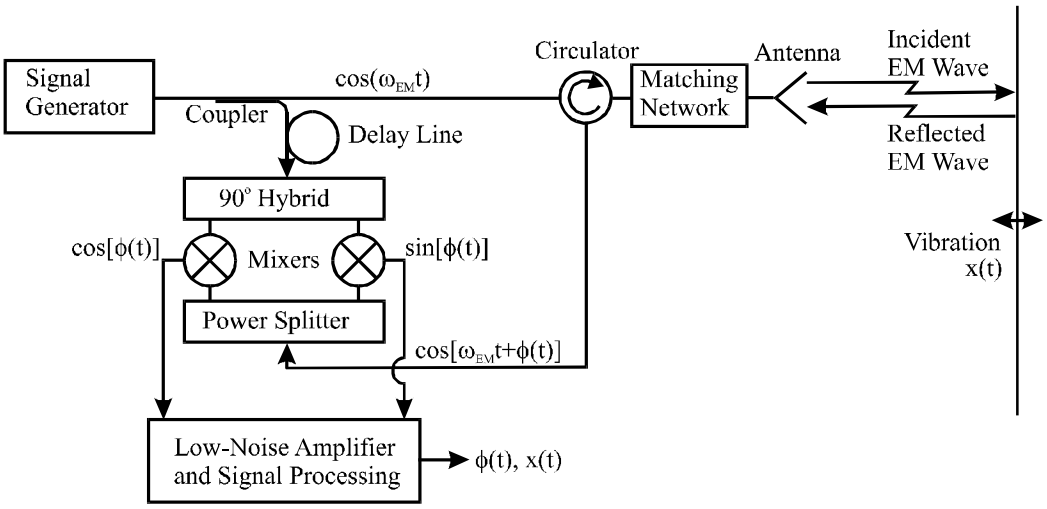


Figure 2 Schematic diagram of the radar

The region scanned with the radar is indicated on figure 3. The displacements are measured in this region as a function of time and position. The measurements are made on a uniform rectangular grid of discrete positions in the scanned region. The grid consists of 41 points in the y-direction that are spaced 2 cm apart and 121 points in the x-direction that are spaced 1 cm apart, for a total of 4961 points. The displacements were measured and recorded as a function of time at each of these points. Each of these time traces is averaged 256 times to reduce the noise. This noise is primarily due to background vibrations in the building that houses the model. The pulse repetition rate of the acoustic source must be significantly lower in the experimental setup than in the field, because of reverberation of the acoustic waves in the tank. It currently takes 24 to 48 hours to perform a complete scan; this is because the measurement is setup to obtain the maximum data integrity without concern for the scan time. The scan time can be greatly reduced by reducing the number of averages, increasing the pulse repetition rate, increasing the distance between sample positions on the surface, or arraying the radar so that multiple

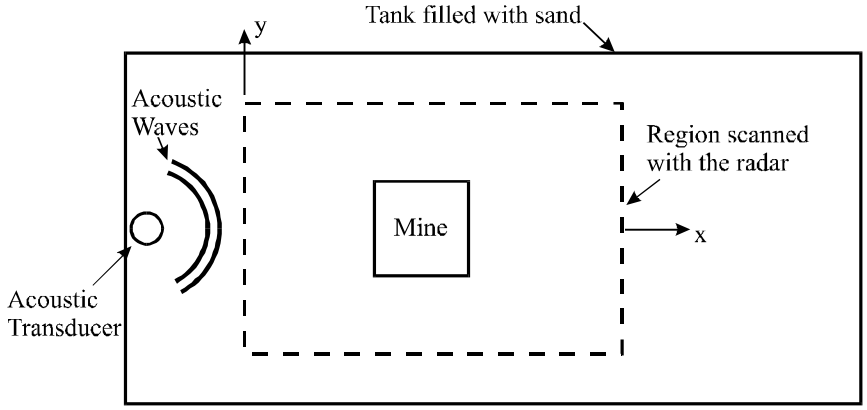


Figure 3 Top view of experimental model.

positions can be measured simultaneously. For example, if the number of averages is reduced by a factor of 8, the pulse repetition rate is increased by a factor of 5, the spacing between the points is increased by a factor of 2 in each direction, and the radar was arrayed to take 10 measurement points simultaneously; it would take approximately 1 to 2 minutes to perform a complete scan instead of 24 to 48 hours.

3. EXPERIMENTAL RESULTS

To date, the model has been used to study the interaction of the acoustic waves with two different mines: a simulated anti-tank mine made out of Acrylic plastic that is 30 cm wide, 30 cm long, and 7.5 cm high, and a simulated anti-personnel mine that is 7.6 cm in diameter and 2.5 cm high (the EM-3 inert mine). The anti-tank mine has been studied with it buried at three depths, and the anti-personnel mine has been studied at one depth. In these studies, the acoustic transducer is excited with a differentiated gaussian pulse. The pulse has a center frequency of 400 Hz for the anti-tank mine and a center frequency of 800 Hz for the anti-personnel mine.

Pseudo color graphs of the amplitude of the displacement of the surface are presented in figures 4 and 5 for four times. These results are for a simulated antitank mine buried in the sand. The top of the mine is flush with the surface of the sand for the graphs in figure 4, and the top of the mine is 5 cm below the surface for the graphs in figure 5. The position of the mine is indicated by the dotted white line. In these graphs, the color scale goes from black to blue to green to yellow to red to white. The smallest displacements are in black and the largest displacements are in white. At time #1, the acoustic wave is seen traveling toward the mine. At times #2 and #3, the wave has reached the mine, and a portion of the wave is going across the mine while the rest is going around the mine. Notice that the portion of the wave that is going across the mine is ahead of the portion that is around the mine. This is because the mine is much stiffer than is the sand; thus, the wave appears to travel faster across the mine. The mine actually moves as if it is a rigid body. The motion of the mine is essentially a rocking motion excited by the wave motion in the sand. At time #4, the wave is seen to be still going around the mine; however, a significant portion has passed across the mine. Notice that the wave that went across the mine is still ahead of the wave that went around the mine, and notice that the wave that went across the mine is smaller in amplitude than the one that went around the mine. Also notice that the displacements are smaller above the mine. The waves that are reflected from the mine can also be seen. The location and the shape of the mine are clearly evident in these graphs. We have also generated movies from the data recorded with the radar. In the movies, pseudo color graphs like those in figures 4 and 5 are stored and played back sequentially. The interaction of the waves with the mine is clearer in the movies than it is in figures 4 or 5; thus, it is easier to see the mine.

A waterfall graph of the displacement of the surface is presented in figure 6 for the anti-tank mine when the top of the mine is flush with the surface. In this graph, the displacement is plotted as a function of time for the 121 points spaced along the x-axis with y=0 (along the center of the scanned region). The plot for each of these 121 points is shifted vertically from the previous one. The plot for x=0 is at the bottom of the graph and the plot for x=120 cm is at the top of the graph. The region in which the mine is located is indicated in gray. The incident surface wave is seen to propagate toward the mine and to be reflected from the mine. The incident surface wave is also seen to travel across the mine. Notice that the wave travels faster across the mine; this is expected because the mine is stiffer than the sand. After the wave passes across the mine it is transmitted into the sand on the other side of the mine. Notice that the wave slows back down when it enters the sand. The incident pressure wave is also seen on the graph; however it is much smaller in amplitude than the surface wave. It is seen to propagate faster then the surface wave. The ripples on the top right of the graph are due to reflections of the incident wave off of the bottom of the tank. Note that the incident surface wave is spread out in time and more complex than the differentiated gaussian pulse that is the excitation for the transducer. This is partially due to the manner in which the wave is injected into the sand. We are investigating methods for injecting a more time-limited pulse.

When the top of the anti-tank mine is 10 cm below the surface, its effects can be seen in the recorded data; however, they are smaller and less obvious than when the mine is closer to the surface. This is because of the exponential decay in amplitude of the surface waves with depth. The rate of the exponential decay increases with increasing frequency; thus, only the lower frequency waves penetrate all the way down to the mine. The recorded

vibrations are high-pass filtered at 100 Hz; therefore, the frequencies that are the most sensitive to the mine have been removed. This is done because the tank that holds the sand has several strong resonances below 100 Hz that corrupt the measurements. These resonances will not occur when the measurement is made in the field; so it will be possible to use lower frequencies. Then the response of mine will be more obvious.

In an effort to enhance the signature of the mine, a simple signal-processing scheme was tried in which the forward traveling waves are filtered out leaving only the reflected waves. The reflected waves occur due to the interactions of the waves with a discontinuity in the sand such as a mine. Pseudo color graphs of the amplitude of the displacement of the surface of the sand are presented in figure 7 with the forward traveling waves filtered out. The top of the anti-tank mine is 10 cm below the surface. For time #1, the incident pulse has just reached the mine and the reflected pulse is forming. For time #2, the incident pulse has reached the mine and the reflected pulse is clearly seen.[1] Thus, by filtering out the forward traveling waves, the response of the mine is clearly visible even when it is buried 10 cm deep. Measurements have not been made with the mine buried deeper to date, but we are planing to make such measurements.

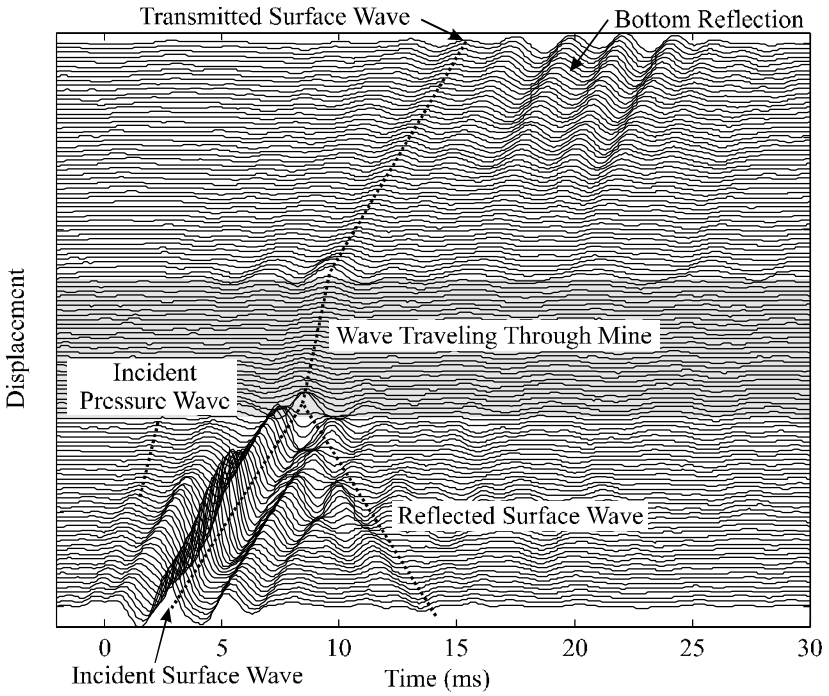


Figure 6 Waterfall graph of the displacement of the surface of the sand when the anti-tank mine is buried in th sand with y=0. The top of the mine is flush with the surface.

Pseudo color graphs of the amplitude of the displacement of the surface are presented in figure 7 for the anti-personnel mine. The forward traveling waves are filtered out, and the top of the mine is approximately 3 mm below the surface. For time #1, the incident pulse has just reached the mine and a reflected pulse is seen being formed. For time #2, the incident pulse has reached the mine and the reflected pulses are clearly seen. The effects of the anti-personnel mine can be clearly seen with the forward waves filtered out. The effects of the mine can be seen without the forward waves being filtered out, but the effects are smaller and less obvious. The effects of the anti-personnel mine would be stronger if a higher frequency incident pulse were used. So far we have been unable to get higher frequency pulses to propagate in the sand, because the sand is verylossy at these frequencies.

[1] Some of the smaller displacements seen in figure 7 are not due to reflections from the mine. They are an artifact of the method used to record the data. They are due to reflections off of the walls of the tank from previous incident pulses that have not died out. The measurement procedure has been modified to remove these artifacts.

We believe that the higher frequencies will propagate better in a soil that contains clay, because the clay will bind the particles together. We plan to investigate materials other than sand to model the soil.

4. NUMERICAL MODEL

A two-dimensional FDTD model for the acoustic waves has been developed. Figure 8 is a diagram of the FDTD model. The waves are injected with a point source, a perfectly matched layer (PML) is used to absorb the waves at the edge of the mesh, and a free surface boundary condition is used on the boundary between the soil and the air. The results from the numerical model are in fairly good agreement with those from the experimental model. The numerical model has been very useful in helping us to understand the interaction of the acoustic wave with the mines. With the numerical model, the waves can be observed below the surface, whereas, with the experimental model, the waves can only be observed on the surface. For example, when the anti-tank mine is buried 10 cm deep; the lower-frequency surface waves can be seen to interact with and be reflected by the mine, while the higher frequency surface waves can be seen to propagate relatively unaffected over the mine. A surface wave can be converted into a pressure wave and visa versa, when the waves interact with the mine. The conversions are more evident in the numerical model than in the experimental model. This is because of the greater signal to noise ratio and the much lower reverberations (reflections from the walls) of the numerical model.

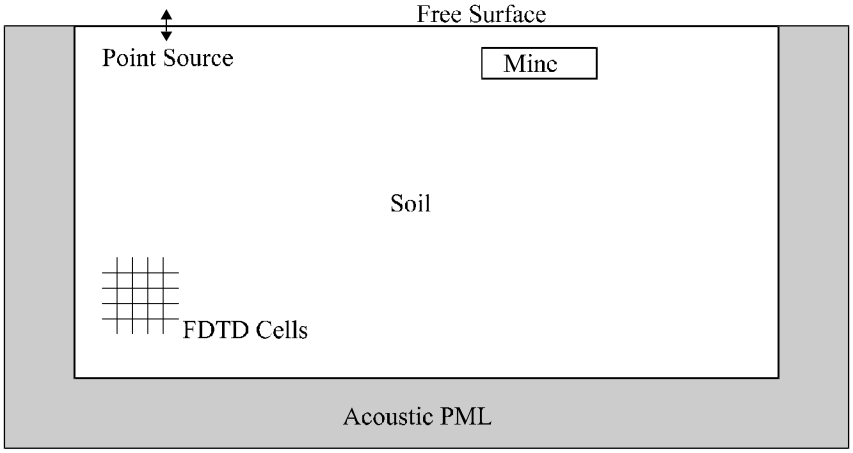


Figure 8 Diagram of the acoustic FDTD model.

Pseudo color graphs of the amplitude of the velocity of the particles in the soil are presented in figure 9 for the anti-tank mine buried 5 cm deep. The incident pulse is a differentiated gaussian pulse with a center frequency of 400 Hz. The soil models the sand that is used in the experimental model. At time #1, pressure (P), shear (S), and surface (R) waves are seen to be launched. The shear and surface waves overlap, because they propagate at approximately the same velocity (the surface wave propagates slightly slower than the shear wave). The surface wave is the more intense wave near the surface. The pressure wave is ahead of the surface/shear wave because it propagates faster. At time #2, the surface/shear wave is seen to have reached the mine. The portion of the wave that passed across the mine is seen to be ahead of the portion that is passing around the mine. Again this is because the wave travels faster in the mine than in the sand. Scattered pressure waves that are due to the interaction of the incident surface/shear wave with the mine are also seen. At time #3, the incident wave is seen to have almost passed by the mine. Scattered pressure, shear, and surface waves are seen propagating away from the mine. At time #4, the wave has passed by the mine. The scattered pressure, shear, and surface waves are still seen propagating away from the mine. Notice that a surface wave appears to be trapped above the mine, we have observed similar behavior in some of the experimental results.

The numerical model predicts a stronger response for the anti-personnel mine than what is obtained with the experimental model. We believe that this difference is due to the numerical model being two dimensional, whereas the experimental model is three dimensional. We are planning to extend the numerical model to three dimensions.

5. CONCLUSTIONS

The acousto-electromagnetic sensor has been investigated using both numerical and experimental modeling. The technique looks promising; we have been able to detect both simulated antipersonnel mines and antitank mines buried in damp sand from the experimental results. However, much more work is needed to understand the capabilities of the sensor in more varied conditions.

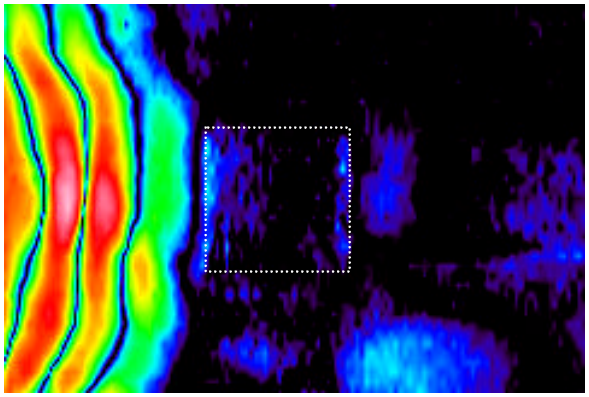
6. ACKNOWLEDGEMENTS

This work is supported in part by the US Army Research Office under contract DAAH04-96-1-0048.

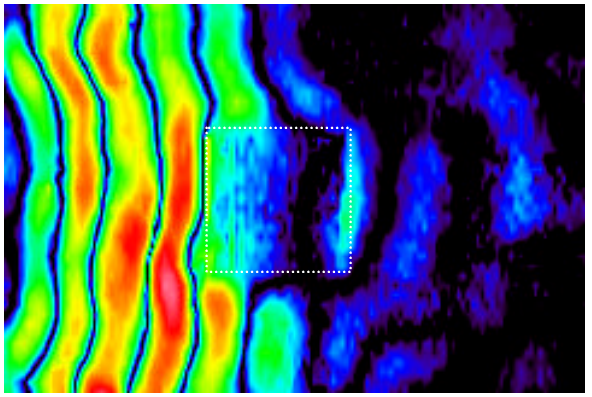
7. REFERENCES

[1] G.S. Smith, "Summary Report: Workshop on New Directions for Electromagnetic Detection of Non-Metallic Mines," Report for U.S. Army BRDEC and ARO, June 1992

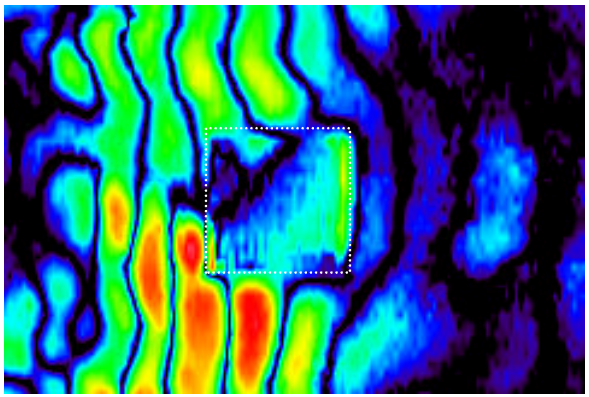
[2] C. Stewart, *Summary of Mine Detection Research,* Vol. I, pp. 172-179, Tech. Report 1636-TR, May 1960, U.S. Army Engineering Res. and Devel. Labs, Corps. of Eng., Belvoir, VA.



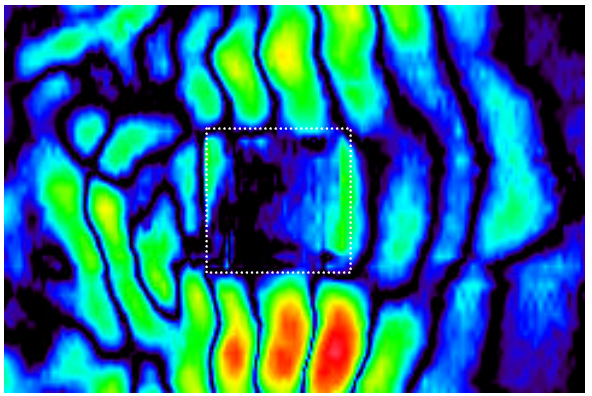
Time #1 – Acoustic pulse traveling toward the mine



Time #2 – Pulse interacting with the mine. The pulse is seen to travel faster across the mine.

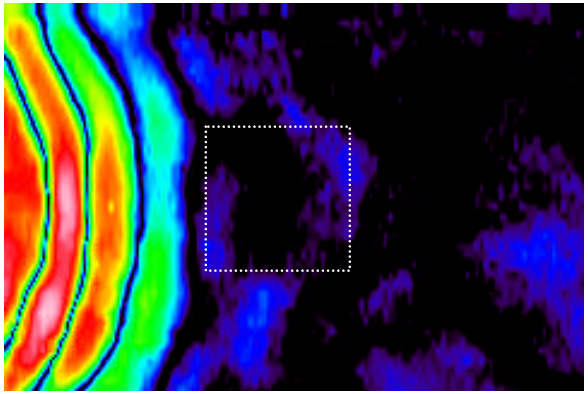


Time #3 – Pulse interacting with the mine. The pulse is seen to travel faster across the mine.

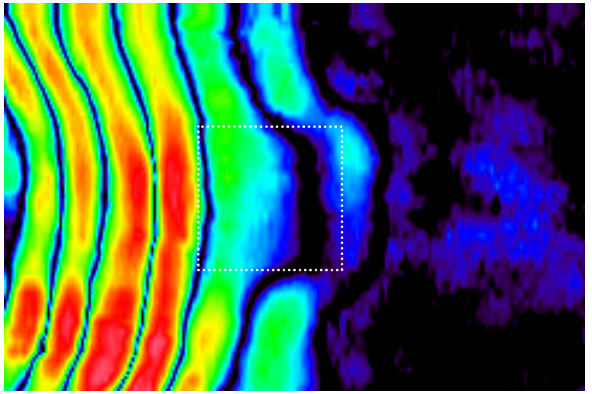


Time #4 – Pulse is reflected from, transmitted through, and travels around the mine

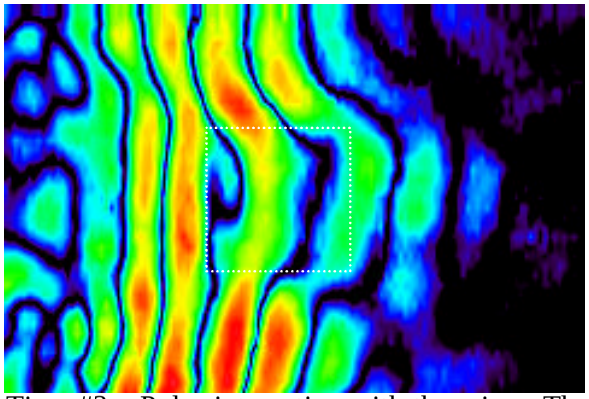
Figure 4 Pseudo color graphs of the amplitude of the displacement of the surface of the sand for four times when a simulated antitank mine is buried in the sand. The top of the mine is flush with the surface, and the mine is outlined by a white dotted line.



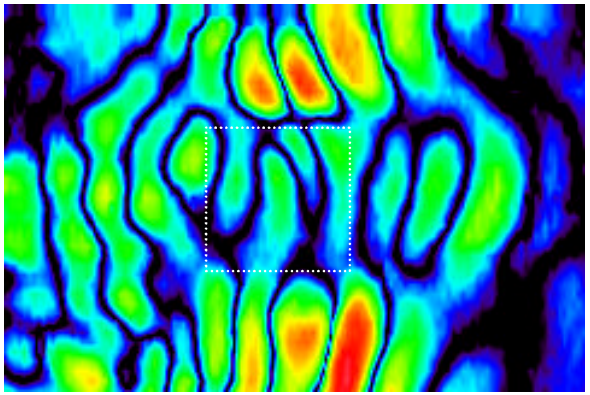
Time #1 – Acoustic pulse traveling toward the mine



Time #2 – Pulse interacting with the mine. The pulse is seen to travel faster across the mine.

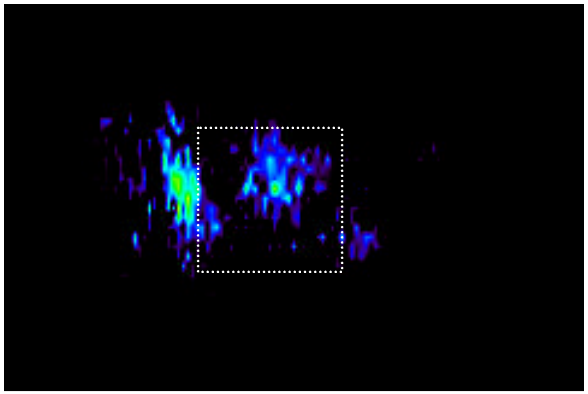


Time #3 – Pulse interacting with the mine. The pulse is seen to travel faster across the mine.

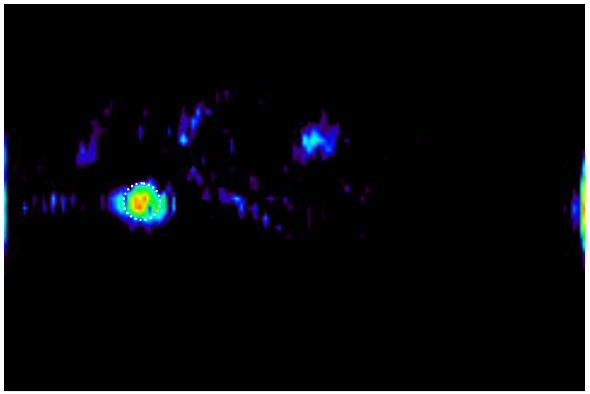


Time #4 – Pulse is reflected from, transmitted through, and travels around the mine

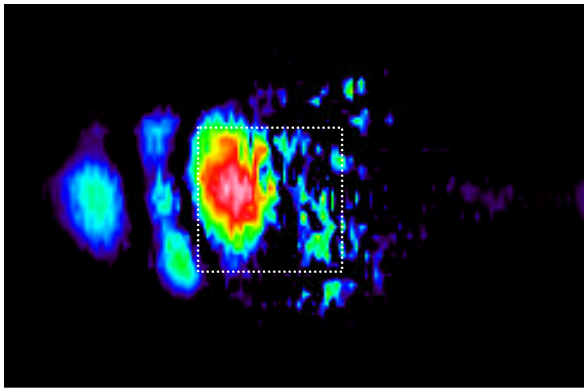
Figure 5 Pseudo color graphs of the amplitude of the displacement of the surface of the sand for four times when a simulated antitank mine is buried in the sand. The top of the mine is 5 cm below the surface, and the mine is outlined by a white dotted line.



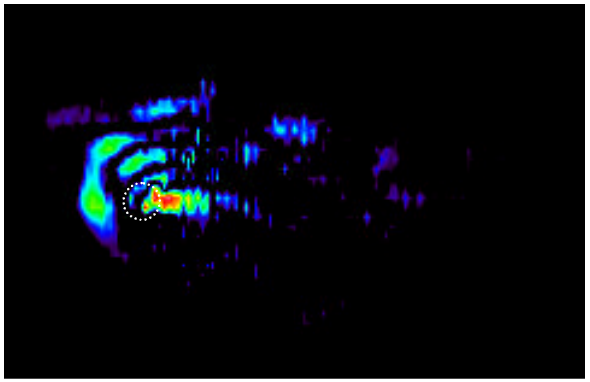
Time #1 – Acoustic pulse has just reached the mine. Anti-tank mine, buried 10 cm deep.



Time #1 – Acoustic pulse has just reached the mine. Anti-personnel mine, buried 3 mm deep.

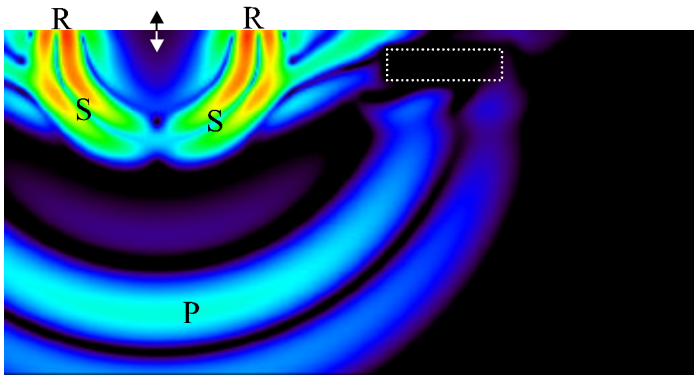


Time #2 – Acoustic pulse reflected from the mine. Anti-tank mine, buried 10 cm deep.

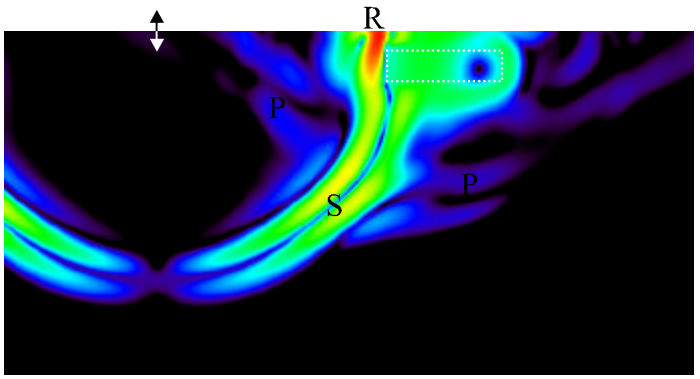


Time #2 – Acoustic pulse reflected from the mine. Anti-personnel mine, buried 3 mm deep.

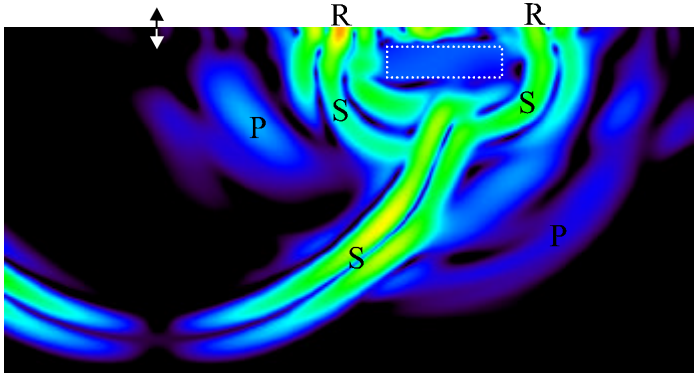
Figure 7 Pseudo color graph of the amplitude of the displacement of the surface with the forward traveling waves filtered outThe dotted lines indicate the position of the mines.



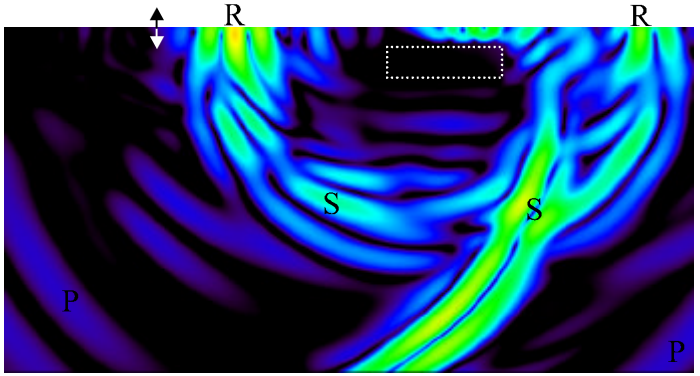
Time #1 – Surface/shear wave has not reached the mine.



Time #2 – Surface/shear wave has just reached the mine



Time #3 – Surface/shear wave is interacting with the mine.



Time #4 – Surface/shear wave has passed by and beer reflected from the mine.

Figure 9 Pseudo color graphs of the amplitude of the velocity of the particles in the soil for four times when a simulated antitank mine is buried in the sand. The top of the mine is 5 cm below the surface, and the mine is outlined by a white dotted line.

Beamforming Array for Detecting Buried Land Mines

Seung-Ho Lee and Waymond R. Scott, Jr.

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

ABSTRACT

A beamforming array is investigated for use in a radar system that is part of a hybrid acoustic/electromagnetic technique for detecting land mines. The radar is used to measure the surface displacement of the earth due to acoustic waves in the earth. The beamforming array is used to obtain small spatial resolution for the measurement of the displacement while allowing an adequate standoff distance for the radar. The tradeoffs between the resolution and the sidelobes of the array are investigated. Finite-difference time-domain and experimental models have been implemented to examine the feasibility of the beamforming array.

Keywords: Beamforming Array, Mine Detection, Displacement Measurement, FDTD

1. INTRODUCTION

This work is part of a project in which a land mine detection system is being studied that simultaneously uses both acoustic and electromagnetic waves.[1] The system consists of an electromagnetic radar and an acoustic source. An acoustic wave is launched into the earth, and the acoustic wave causes the surface of the earth and the mine to be displaced. The radar is used to measure the displacements and, thus, detect the mine. Currently the radar uses an open-ended waveguide for the antenna. However, this antenna must be placed within a few centimeters of the surface of the earth in order to get sufficiently small spatial resolution. In this work, we investigate using a beamforming array so that the antenna can be placed farther from the surface while maintaining sufficiently small spatial resolution. The array can be implemented either by constructing an array of N-elements and recording the signal from all of the elements simultaneously or by using a single element that is scanned to construct a synthetic array. The most practical solution will be to probably use a combination of these techniques and scan a subset of the N-element array to more quickly construct the synthetic array.

A simple analytical model is developed for the beamforming array. This model is used to investigate the tradeoffs between the resolution and the sidelobes of the array as a function of the height of the array and the method of focusing used in the beamforming. A finite-difference time-domain model (FDTD) and an experimental model have been implemented to examine the feasibility of the beamforming array. The results from these models are promising. The array is shown to have sufficient resolution for the mine detection technique mentioned above.

2. BEAMFORMING ARRAY

Figure 1 is a schematic diagram of an N-element beamforming array. The elements are placed on a plane parallel to the x-y plane, assumed to be identical, and located at the positions $\vec{R}_i'$. Assuming an isotropic source is located at the position $\vec{R}$, the voltage received by the i^{th} element can be expressed as

$$V_i(\vec{R}) = C\frac{\cos\psi_i}{|\vec{R}-\vec{R}_i'|}e^{-jk|\vec{R}-\vec{R}_i'|}, \tag{1}$$

where C is a constant. Here, the electric field is treated as a scalar, and the elements are assumed to have the simple angular dependence $\cos\psi_i$, where ψ_i is the angle between the propagation direction and the z-direction.

The terminal voltages of the elements of the array are summed with the appropriate weighting function to focus the array at the position $\vec{R}_f$. The phase of the weighting function is chosen so that the voltages will add constructively

E-mail and Tel.: S.H. Lee, gt2332b@prism.gatech.edu, 404-894-3123, W. R. Scott, Jr. : wrscott@ee.gatech.edu, 404-894-3084

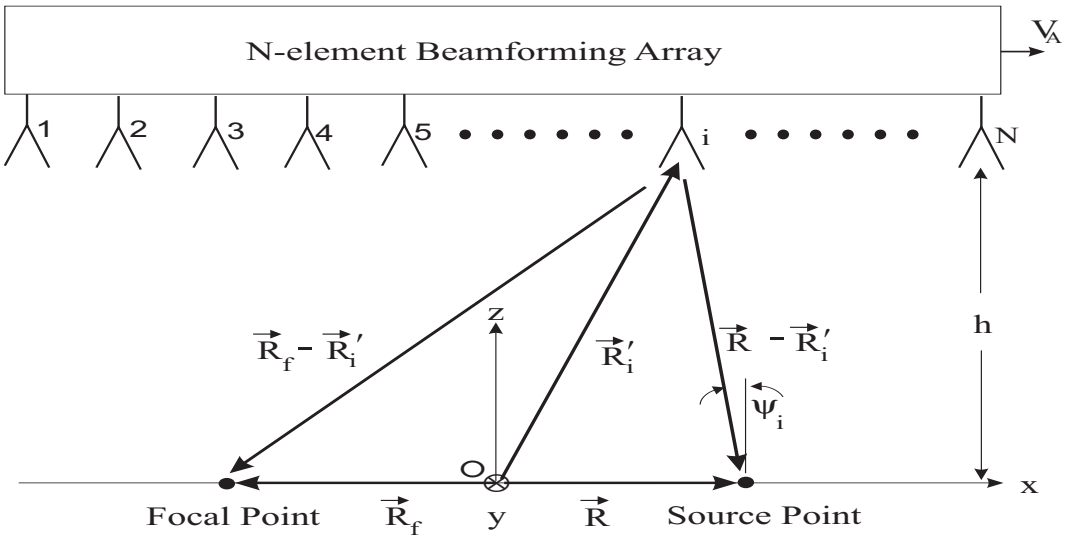


Figure 1. Beamforming array configuration

when the source is at the position $\vec{R_f}$, and the amplitude of the weighting function is chosen to control the sidelobe levels of the array. The output of the array is then

$$V_A(\vec{R}, \vec{R_f}) = \sum_{i=1}^{N} V_i(\vec{R}) W(\vec{R_i'}, \vec{R_f}) e^{jk|\vec{R_f} - \vec{R_i'}|}, \tag{2}$$

where W is a window function used to adjust the amplitude of the weighting function. We have determined empirically that a Gaussian window function performs well:

$$W(\vec{R_i'}, \vec{R_f}) = \exp\left(-\frac{1}{2\tau^2} |\hat{z} \times (\vec{R_f} - \vec{R_i'})|^2\right), \tag{3}$$

where τ is a parameter used to adjust the width of the window function. The window function is maximum directly above $\vec{R_f}$, $|\hat{z} \times (\vec{R_i'} - \vec{R_f})| = 0$, and it decreases with increasing distances from this point.

A graph of the output of the array as a function of the source position for several values of τ is shown in Fig.2. The array consists of 59 by 59 identical elements equally spaced on a square grid; they are spaced 0.47λ apart. The array is centered on the point $(x = 0, y = 0, z = h = 5.33\lambda)$. The isotropic source is scanned from $x = -10\lambda$ to $x = 10\lambda$ with $y = 0$ and $z = 0$. The array is focused at $x = y = z = 0$. Since the beamforming array is focused at the center of the scan, the maximum output is obtained when the isotropic source is located at the center of the scan as expected. The sidelobe levels are seen to be quite high, and the beamwidth is seen to be very narrow when the window is very broad ($\tau = 12.0\lambda$). As τ decreases, the sidelobe levels drop at the expense of wider beamwidths.

In Fig.3(b), the 3dB beamwidth and the first sidelobe level are plotted as a function of τ for three different values of h. The beamwidth is seen to decrease when τ is increased and when h is decreased. However, the first sidelobe level behaves in an opposite manner; the first sidelobe level is seen to increase when τ is increased and when h is decreased. Figure 3(a) is a similar graph for a 2-D case. Here, the elements and the source are assumed to have an infinite length in the y-direction. In this configuration, the beamforming array consists of 59 identical elements along the x-direction only. The results for the 2-D case are very close to those for the 3-D case.

Figure 4 shows the output of the array with $\tau = 2.67\lambda$ as a function of the x- and z-coordinates of the isotropic source. The configuration of the beamforming array is the same as the above 3-D case with $h = 5.33\lambda$. The array is seen to be focused at the origin; the beamwidth is narrowest along the line $z = 0$. However, note that the beamwidth does not change significantly when z is varied in a range of approximately, $-\lambda < z < \lambda$.

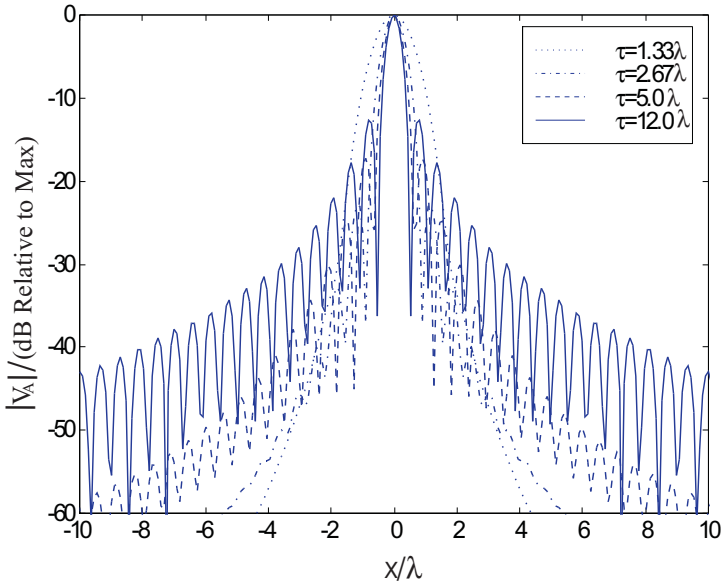


Figure 2. The output of the array as a function of the source position for different values of τ when h=5.33λ

3. FDTD MODEL

A two-dimensional FDTD model has been implemented to study the beamforming array. A diagram of the model is shown in Fig.5. A total field/scattered field formulation was used to inject an electromagnetic plane wave, an uniaxial perfectly matched layer[2] was used to truncate the boundaries, and the earth was modeled as a lossy half space. The surface of the earth is displaced due to an acoustic wave traveling in the earth. These displacements are included in the model. The electromagnetic plane wave is reflected from the boundary between the air and the earth, and the total field (both incident and reflected waves) is recorded at the observation points. The reflected wave is obtained by subtracting the incident wave from the total field; the incident wave is calculated in the implementation of the total field/scattered field interface.[3] The reflected electromagnetic waves are recorded both with and without the surface displacements.

The acoustic wave travels much slower than the electromagnetic wave. Thus, it is impractical to run the FDTD simulation for the time it would take the acoustic wave to travel across the model, because of the number of time steps that would be required. Fortunately, this is not necessary. In a given period of time, the acoustic wave travels 10^{-6} times as far as does the electromagnetic wave. The acoustic wave will also travel a very short distance with respect to its width during the total travel time of the electromagnetic wave. Thus, it is possible to decouple the time scales for the acoustic and the electromagnetic waves. The electromagnetic simulation is run with the acoustic wave at discrete positions on the boundary.

The surface displacements would seem to be very difficult to model in the FDTD method because they are very small compared to the size of a FDTD cell. However, it has been shown that displacements as small as 10^{-14} times smaller than a FDTD cell can be accurately modeled using the standard averaging scheme.[4] The actual displacements due to the acoustic wave are many orders of magnitude larger than this smallest displacement that can be modeled. So it is relatively easy to model the displacements.

The surface displacements are calculated from the waves recorded at the observation points to demonstrate the effectiveness of the beamforming array. Figure 6(a) is a graph of the actual displacement and the displacement calculated without beamforming for several values of h. Here the displacements are obtained by comparing the phase between the data obtained with and without the displacement in the model. The actual displacement of the surface is a differentiated Gaussian pulse with the maximum amplitude of $1\mu m$. When the observation points are located at $h = 1cm$ from the surface, the calculated displacement agrees well with the actual displacement. It is seen that when the observation points are relatively far ($h > 5cm$) from the boundary, the displacement obtained from the

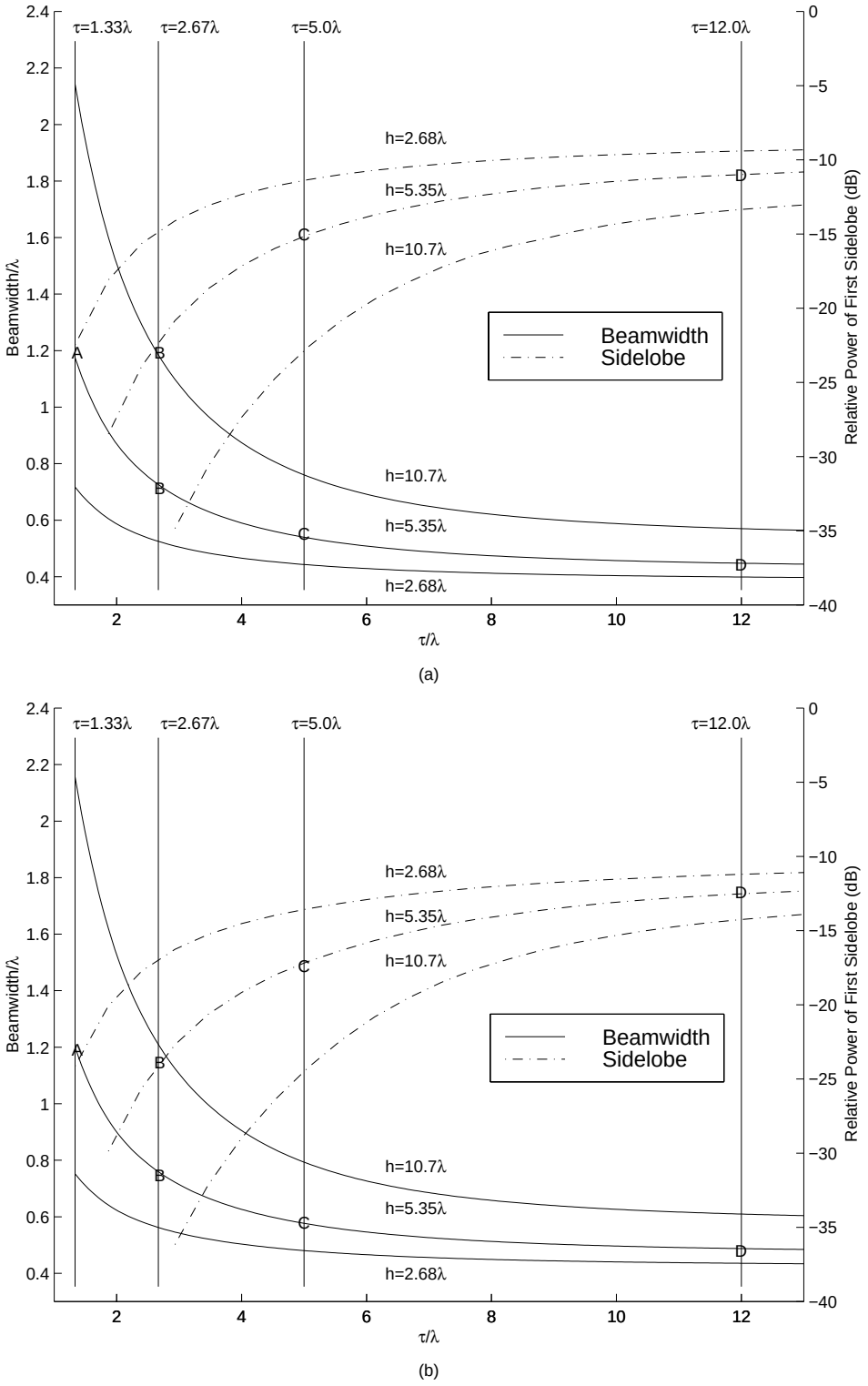


Figure 3. The beamwidth and the first sidelobe level versus τ : (a) 2-D Case (b) 3-D Case

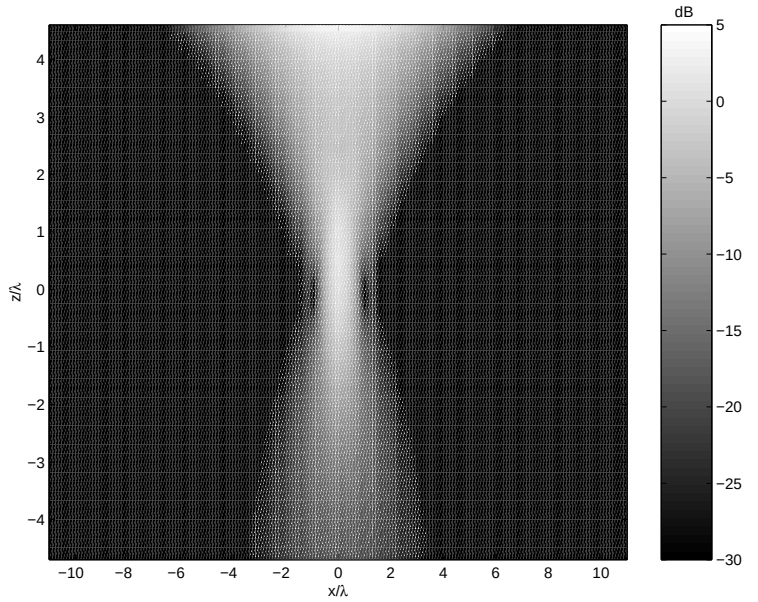


Figure 4. The output of the array when $h = 5.33\lambda$ and $\tau = 2.67\lambda$

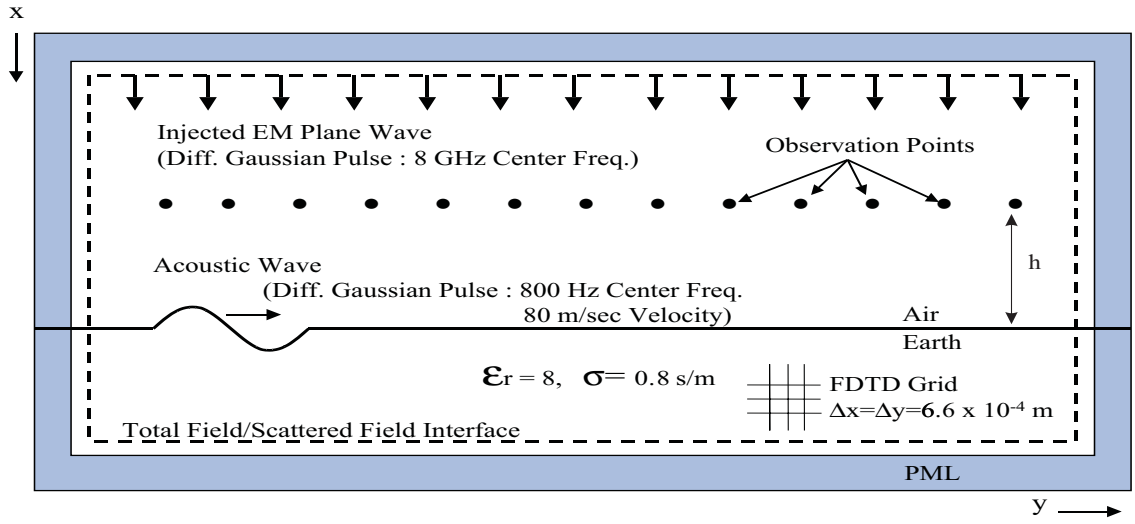


Figure 5. Geometry for FDTD model being investigated

phase at the observation points is not a good replica of the actual displacement of the surface. In order to solve this problem, the beamforming array technique was used to obtain the good image of the displacements.

Figure 6(b) is a graph of the actual displacement and the displacement calculated with beamforming at $h = 20cm$ for several values of τ. The pulses in the reconstructed displacements are seen to be somewhat wider than those in the actual displacement. The pulse is seen to be only slightly wider when $\tau = 12.0\lambda$ and is seen to be much wider when $\tau = 1.33\lambda$. This can be explained using Fig.3(a), and the points A,B,C,D on the graph that correspond to the values of τ and h used in this example. The beamwidth for $\tau = 12.0\lambda$ is seen to be relatively small and is seen to increase with decreasing values of τ. Thus, the increased width of the pulse is seen to be due to the increased beamwidth of the array. A ripple is seen on the reconstructed displacements. The amplitude of the ripple is seen to increase with increasing τ. This is due to the sidelobe levels increasing with increasing τ as shown in Fig.3(a).

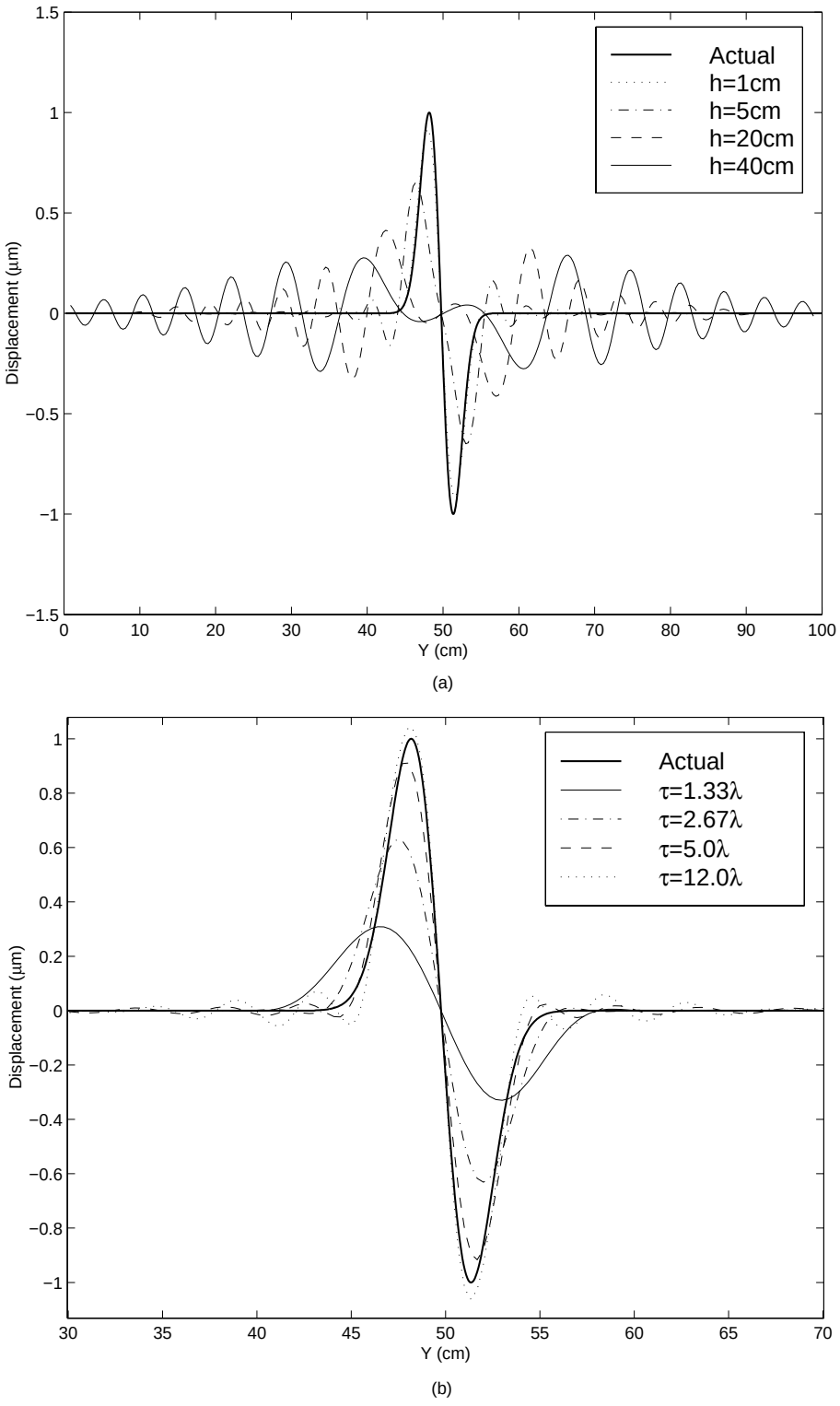


Figure 6. Actual and computed boundary displacements : (a) Without beamforming, (b)With beamforming when $h = 20cm$

4. EXPERIMENTAL RESULTS

A schematic diagram of the experimental model is shown in Fig.7. It consists of a sandbox and a radar which is scanned over the surface of the sand to implement a synthetic beamforming array. To beamform this array, the phase term in Eq.(2) is doubled to account for the two way travel time of the wave.

The radar is monostatic and uses an open-ended waveguide as the antenna. A circulator is used to help isolate the transmitted and received signals. A HP8720D network analyzer is used to measure the magnitude and phase of the ratio of the received to transmitted signal. The antenna was placed 5.33λ above the surface of the sand and the radar was operated at a frequency of 8GHz. The radar is scanned using a x-y positioner. The response of the radar is measured on a grid of 59×59 equally spaced points that are spaced 0.47λ apart. The scanned region is approximately 1m square. In the sand below the scanned region, one large and two smaller symbols are dug into the sand. A photograph of the symbols is shown in Fig.8(a). Both the large symbol and the smaller symbols are 0.16λ high.

The reconstructed height of the surface of the sand calculated from the raw measured data without beamforming is shown in Fig.8(b). The symbols are seen to be very blurred. The reconstructed height of the surface of the sand calculated using the beamforming is shown in Fig.8(c)-(f) for four values of τ. The results with $\tau = 1.33\lambda$, 2.67λ and 5.0λ are seen to be clearly better than the result without beamforming. Note that the resolution of the images increases with increasing value of τ because of the decreasing beamwidth of the array. However, note that the noise in the image increases with increasing values of τ because of the increasing sidelobes of the array. The image with $\tau = 12.0\lambda$ is seen to be very noisy. This is predicted by the results in Fig.3. The sidelobes are clearly more important for the 3-D experimental results than for the 2-D numerical results.

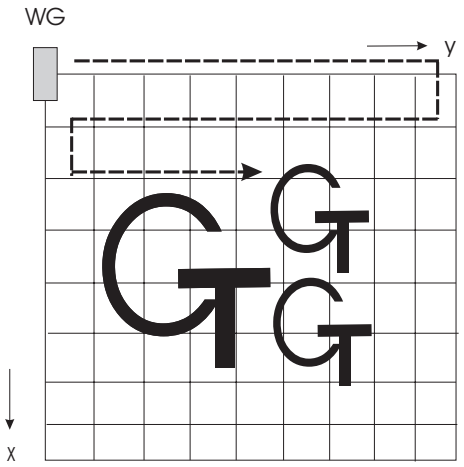


Figure 7. Grid of the region scanned

5. CONCLUSION

The feasibility of the beamforming array has been investigated using the FDTD method and the experimental model. By using the beamforming array, very small displacements could be detected in the numerical model, and the surface of the ground was reconstructed in the experimental model. Future plans include extending the experimental model to include time-varying displacements so it can be incorporated into the mine detection system.

ACKNOWLEDGMENTS

This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0048.

REFERENCES

1. W. R. Scott, Jr., C. Schroeder, and J. Martin, "An acousto-electromagnetic sensor for locating land mines," *SPIE, AeroSense, Detection and Remediation Technologies for Mines and Minelike Targets III* , pp. 176–186, April 1998.
2. S. Gedney, "An anisotropic perfectly matched layer-absorbing medium for the truncation of the fdtd lattices," *IEEE Trans. on Antennas and Propagation* **44**, pp. 1630–1639, 1996.
3. A. Taflove, *Computational Electrodynamics:The Finite-Difference Time-Domain Method*, Artech House, Boston.London, 1995.
4. S. H. Lee and W. R. Scott, Jr., "Modeling vibrating boundaries with the finite-difference time-domain method," *Proc. of the 1998 URSI Radio Science Meeting* , p. 70, June 1998.

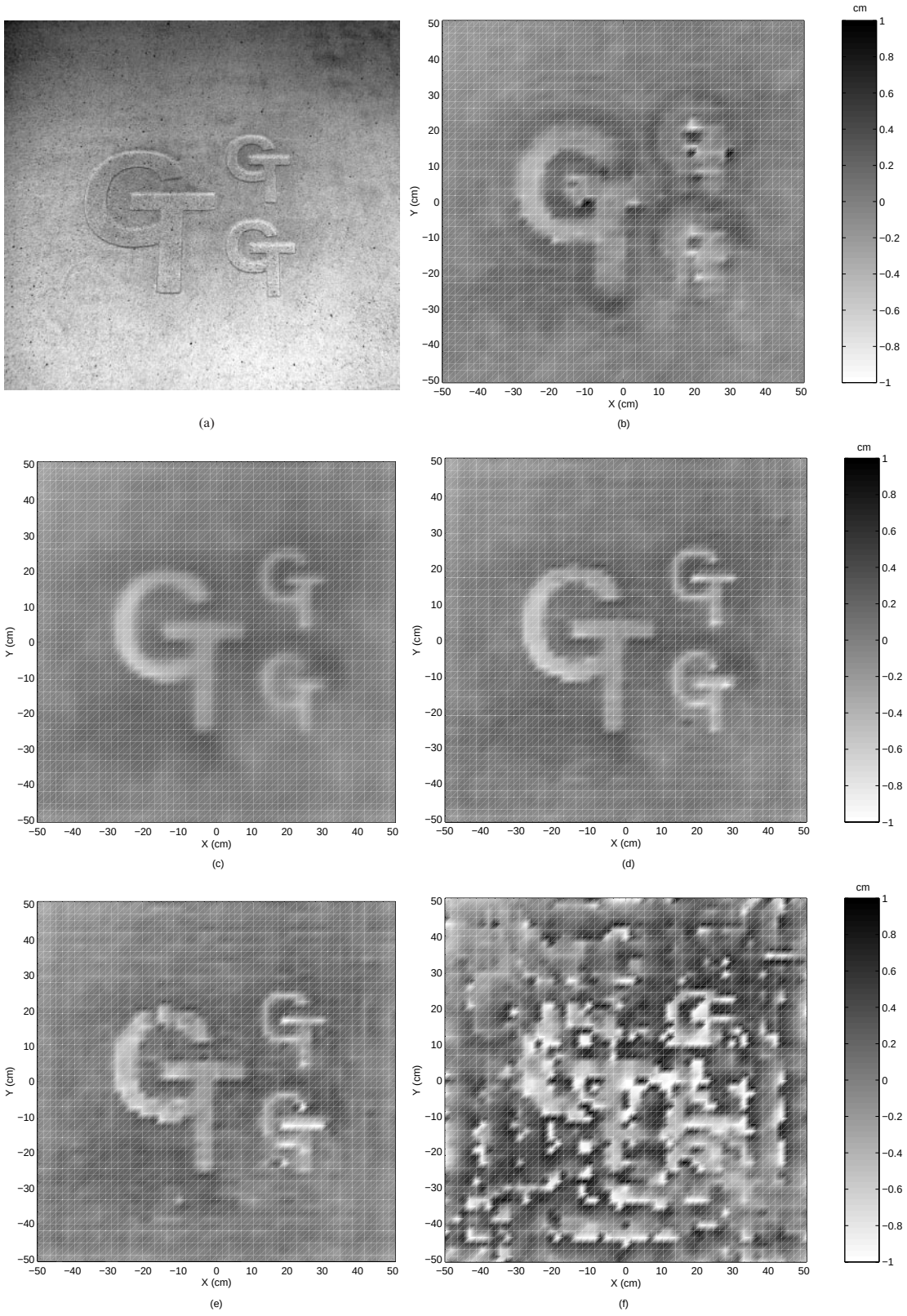


Figure 8. Reconstructed surface of the sand when h=20cm : (a) Photograph of the surface of the sand, (b) Raw, (c) $\tau = 1.33\lambda$, (d) $\tau = 2.67\lambda$, (e) $\tau = 5.0\lambda$, (f) $\tau = 12.0\lambda$

Proceedings of SPIE, Vol. 3710, April 1999

Experimental investigation of the acousto-electromagnetic sensor for locating land mines

Waymond R. Scott, Jr.[a] and James S. Martin[b]

[a]School of Electrical and Computer Engineering
[b]School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

A hybrid technique is presented that simultaneously uses both electromagnetic and acoustic waves in a synergistic manner to detect buried land mines. The system consists of an electromagnetic radar and an acoustic source. The acoustic source causes both the mine and the surface of the earth to be displaced. The electromagnetic radar is used to detect these displacements and, thus, the mine. To demonstrate the viability of this technique, experimental models have been constructed. The models use an electrodynamic transducer to generate an acoustic surface wave, a tank filled with damp sand to simulate the earth, simulated mines, and a radar to measure the vibrations. The technique looks promising; we have been able to measure the interactions of the acoustic waves with both simulated antipersonnel mines and antitank mines buried in damp sand. We have measured strong resonances in some of the mines; these resonances are shown to help differentiate the mine from clutter.

Keywords: land mine, mine detection, acoustic, ground penetrating radar, GPR

1. INTRODUCTION

Seismic/elastic techniques show considerable promise for the reliable detection of all types of buried mines, even low-metal anti-personnel mines. The reason for this is that mines have mechanical properties that are significantly different from soils and typical forms of clutter. For example, the shear wave velocity is approximately 20 times higher in the explosive and the plastics used in typical mines than in the surrounding soil. In addition, mines are complex mechanical structures with a flexible case, a trigger assembly, air pockets etc. The complex structure gives rise to structural resonances, non-linear interactions, and other phenomenology that is atypical for both naturally occurring and man made forms of clutter. Thus, this phenomenology can potentially be used to distinguish a mine from clutter. However, seismic detection techniques have fallen into disfavor because of practical system implementation issues.

A system has recently been developed at Georgia Tech that exploits the advantages of seismic techniques while overcoming many of the implementation issues [1-3]. The configuration of the system currently being studied is shown in figure 1. The system consists of an electromagnetic radar and an acoustic source. The source preferentially generates an acoustic surface (seismic) wave in the earth. The acoustic wave causes both the mine and the surface of the earth to be displaced. The displacement of the mine is different from the earth's, because the acoustic properties of the mine are quite different than those of soil. The displacement of the surface of the earth when a mine is present is different than when it is not present because of the local and propagating waves scattered by the mine. The electromagnetic radar is used to detect these displacements and, thus, the mine.

The technique is both complimentary to and compatible with existing mine detection techniques. It offers an additional queue for detection and classification because it senses different physical phenomena than conventional mine detection techniques such as ground penetrating radars (GPRs), metal detectors, and infrared sensors. There

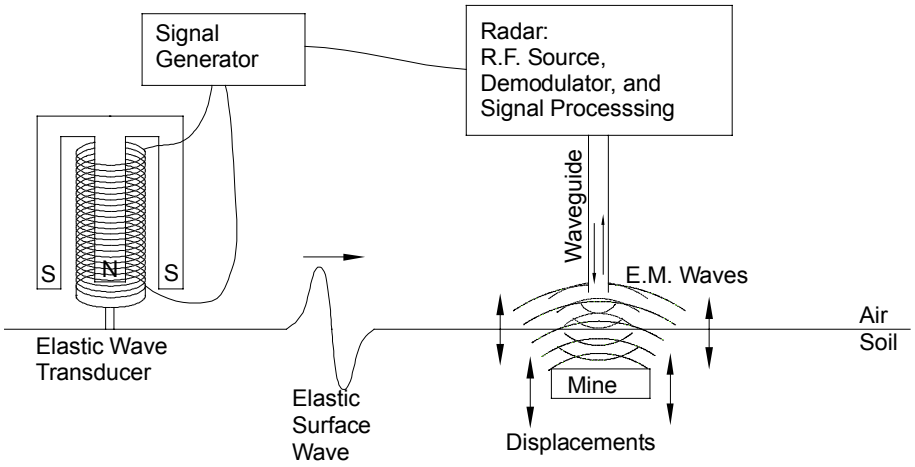


Figure 1. Acousto-electromagnetic mine detection system with the acoustic transducer placed on the surface of the earth.

is also no reason why the operation of the seismic source or sensor would preclude the simultaneous operation of these types of systems. This makes a seismic system an excellent candidate for sensor fusion with other techniques.

Experimental models for the system have been constructed and automated. The experimental models are being used to demonstrate the viability of the technique and to study the interactions of the acoustic and electromagnetic waves with various types of buried mines. The technique looks promising; detections have been demonstrated for both simulated antipersonnel mines and antitank mines buried in damp sand, and strong mine resonances have been observed in some of the experimental data which could serve to classify the detected objects. Additional investigation is needed to determine the potential capabilities of the technique in more varied conditions.

2. FIRST EXPERIMENTAL MODEL

Two experimental modes have been constructed; a drawing of the top view of the first experimental model is shown in figure 2. The model consists of a wooden tank that is filled with damp sand that has been packed to a relatively uniform density and cohesion. The tank is approximately 120 cm wide, 120 cm deep, and 240 cm long. A transducer is placed on the surface of the sand and is used to launch the elastic waves into the sand. The transducer is an electrodynamic shaker that is driven with a signal generator. The transducer has been coupled to the surface of the sand with a narrow base plate so that it preferentially excites elastic surface waves. These surface waves travel across the surface of the tank and interact with a mine that is buried in the sand.

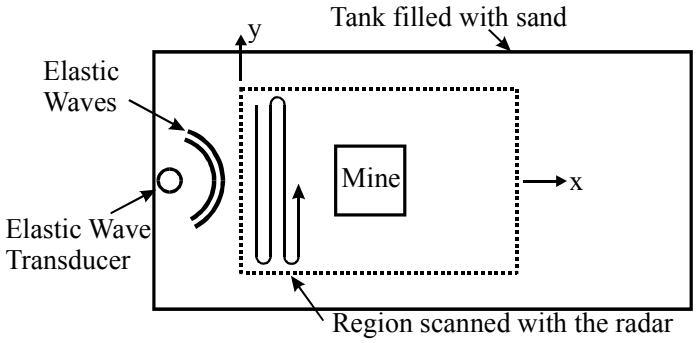


Figure 2. Top view of experimental model.

A radar has been designed and built to measure the acoustic vibrations of the soil and the mine. The two biggest challenges to make this radar perform adequately for the mine detection system are 1) to make it sufficiently sensitive to be able to detect small vibrations, and 2) to make the spot size (the area on the surface illuminated by the electromagnetic waves) sufficiently small to detect high wavenumber surface waves and localized wave fields. The radar can measure vibrations as small as 1 nm (10^{-9} m) as currently configured. To obtain this sensitivity, the radar was designed to minimize the effects of noise, such as the phase noise of the source and the electromagnetic interference from low-frequency magnetic fields. The spot size of the radar must be smaller than half of a wavelength of the acoustic waves at the highest operating frequency in order not to spatially integrate out the motion of interest. Currently, a small spot size is obtained by using an open-ended waveguide as the antenna for the radar. This antenna produces a sufficiently small spot size when the open end of the antenna is placed within a few centimeters of the surface. A digital beamforming array is being investigated as a possible method for obtaining a small spot size when the height of the antenna is increased [4]. All of the results in this report were obtained with the radar operating at 8 GHz and the end of the waveguide about 1 to 2cm from the surface.

A x-y positioner is used to scan the radar over the surface of the sand. This is controlled by a personal computer which automatically scans the radar over the surface and records the data. The region scanned with the radar is indicated on figure 2. Displacements are measured in this region as a function of frequency and position. The measurements are made on a uniform rectangular grid of discrete positions in the scanned region. The grid consists of 41 points in the y-direction that are spaced 2 cm apart and 121 points in the x-direction that are spaced 1 cm apart, for a total of 4961 points. Displacements are measured and recorded as a function of drive frequency at each of these points. Previously reported measurements [1-3] were made using short pulses in the time domain; the measurements reported here were made using a network analyzer in the frequency domain. This greatly improved the signal to noise ratio without increasing measurement times. The frequency domain data are transformed into the time domain in post processing, using a differentiated gaussian pulse with a center frequency of 400 Hz as the incident signal.

The current emphasis is on data integrity rather than scan time. The current scan time of 24 to 48 hours can be reduced greatly by reducing the integration time, increasing the distance between sample positions on the surface, or arraying the radar so that multiple positions can be measured simultaneously. For example, a scan could be completed in 5 to 10 minutes, if the integration time were reduced by a factor of 8, the spacing between the points were increased by a factor of 2 in each direction, and the radar were arrayed to take 10 measurement points simultaneously.

3. EXPERIMENTAL RESULTS

To date, the experimental model has been used to study the interaction of the elastic waves with seven different mines, both simulated mines and real mines that have been made inert by removing the explosive. A list of the mine with their approximate dimensions is shown in table I. The simulated anti-tank mine is made out of a solid piece of acrylic plastic, and the SIM9 mine is a case from a 9 cm diameter simulated antipersonnel mine partially filled with glazing putty to model the explosive. The results for the first two mines are presented in the previous papers [1-3], and the results for the remaining mines will be presented next.

A waterfall graph of the displacement of the surface is presented in figure 3 for the TS-50 mine when the top of the mine is 1.3cm below the surface of the sand. In this graph, the displacement is plotted as a function of time for the 121 points spaced along the x-axis with y=0 (along the center of the scanned region). The plot for each of these 121 points is shifted vertically from the previous one. The plot for x=0 is at the bottom of the graph and the plot for x=120 cm is at the top of the graph. The region in which the mine is located is indicated in gray. The incident pressure wave is seen to propagate toward and across the mine. The incident surface wave is also seen to travel across the mine. The larger displacements above the mine are due to a resonance in the mine. The resonance makes it much easier to detect the mine. The travelling waves reflected from the mine are relatively small indicating the difficulty that would be encountered in detecting this mine using a classical pulse echo

technique. A pseudo color graph of the frequency spectrum is shown in figure 4 as a function of x with y=0. The higher intensities near x=22cm are due to the resonance. The effect appears to be due to several resonances which occur in the 400 to 700 Hz frequency range. The specific spectral components of the effect may constitute a target signature for the TS-50 mine. It is apparent from figures 3 and 4 that the higher frequency waves attenuate more rapidly than do the lower frequency waves. In the absence of the resonance, a higher frequency incident would be required to image an object the size of the TS-50 mine. Thus, the resonance effect extends the distance from the source over which detections of this type of mine can be made.

Table I Mines investigated.

	Type	Description	Size
1	AT	Simulated anti-tank mine	30 cm wide, 30 long, 7.5 cm height
2	EM-3	Simulated anti-personnel mine	7.5 cm diameter, 2.5 cm height
3	TS-50	Inert anti-personnel mine	8 cm diameter, 4.5 cm height
4	VS-50	Inert anti-personnel mine	8 cm diameter, 4.5 cm height
5	Butterfly	Inert anti-personnel mine	4 cm wide, 12 cm long, 1.8 cm height
6	M-14	Inert anti-personnel mine	5 cm diameter, 4 cm height
7	SIM9	Simulated anti-personnel mine	9 cm diameter, 3 cm height

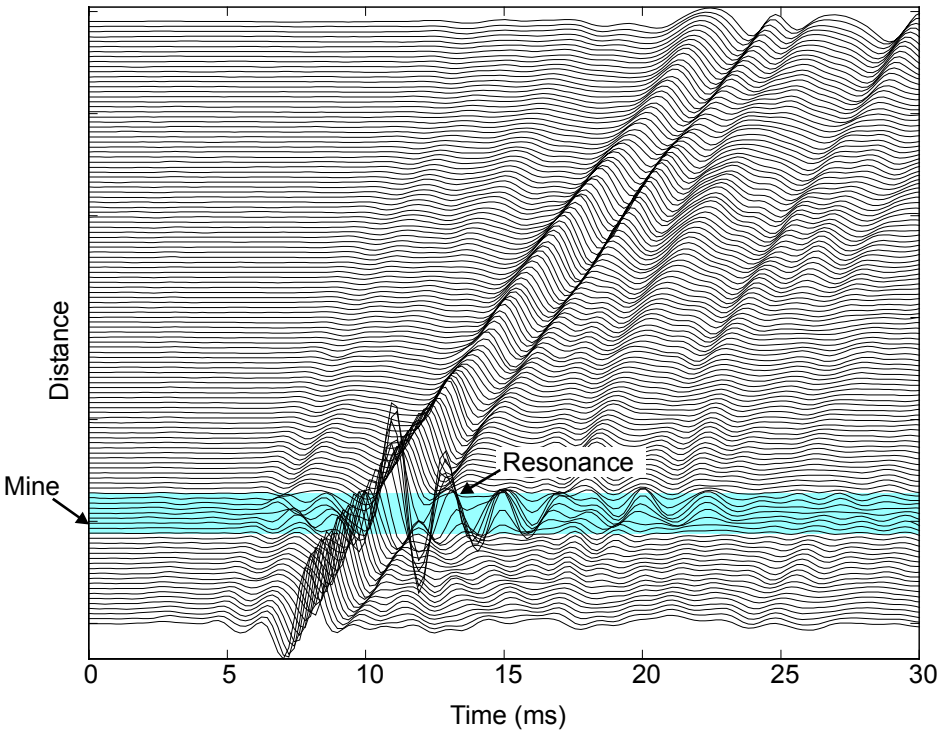


Figure 3. Waterfall graph of the displacement of the surface of the sand when the TS-50 anti-personnel mine is buried in the sand with y=0. The top of the mine is 1.3 cm below the surface.

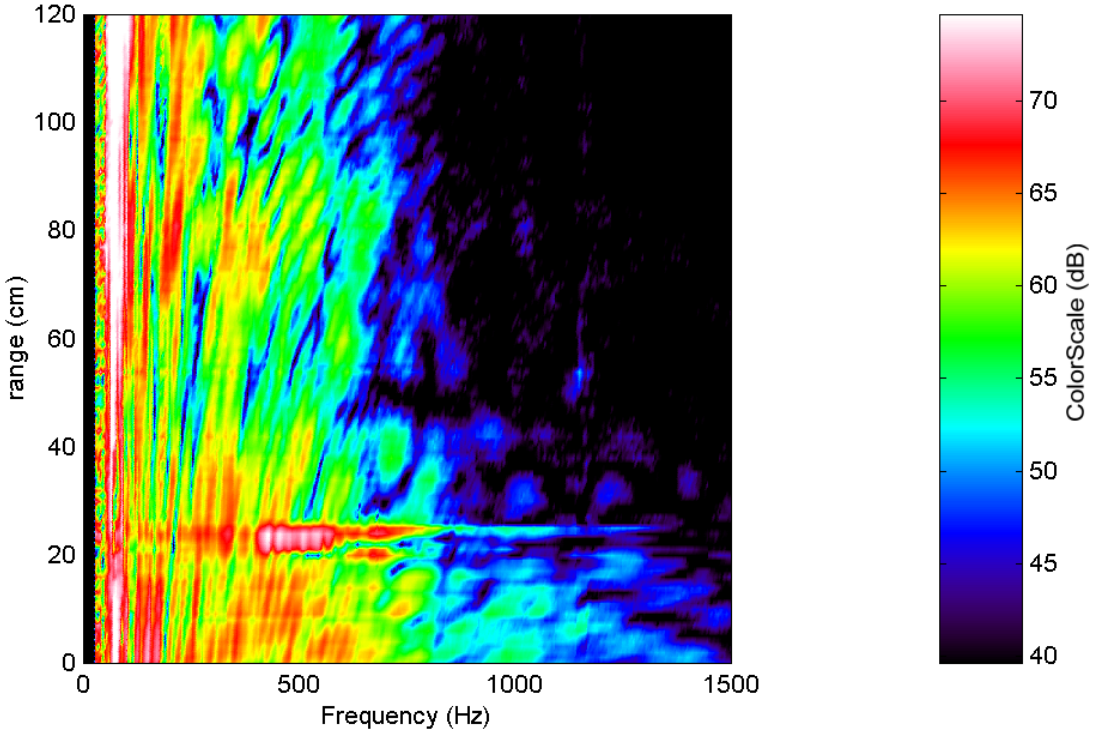
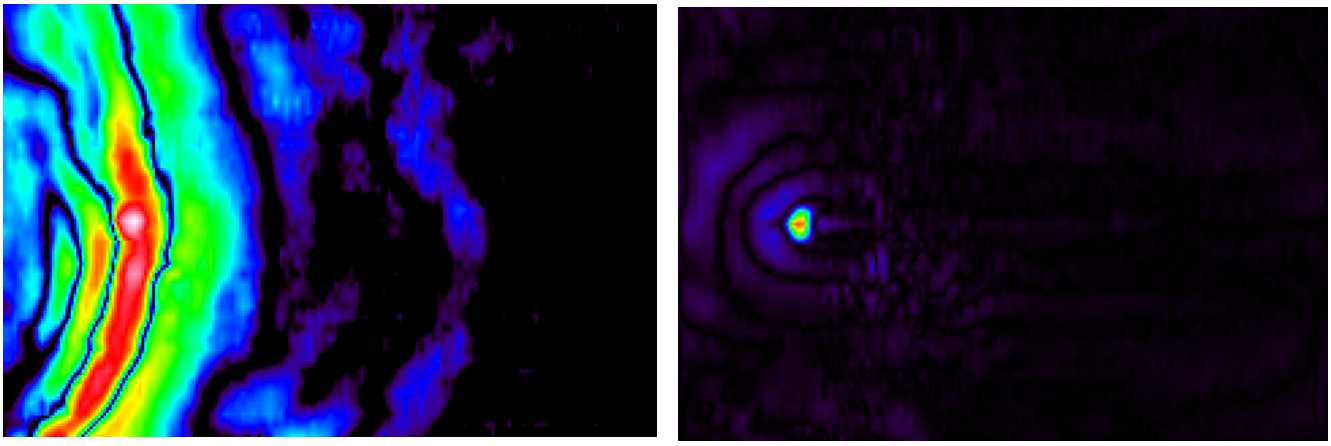


Figure 4. Pseudo color graph of the amplitude of the frequency spectrum for the surface displacement of the sand when the TS-50 anti-personnel mine is buried in the sand with y=0. The top of the mine is 1.3 cm below the surface.

Pseudo color graphs of the amplitude of the displacement of the surface are presented in figure 5 for the entire scanned region at a time at which the waves are interacting with the mine. The displacements are shown for the raw data and for the data with the forward traveling waves filtered out [1]. The mine can be seen in the graphs with and without the forward traveling waves filtered out. However, the effect of the mine is much more obvious in the graph with the forward traveling waves filtered out. The displacements due to the resonance and the reflected waves are clearly visible. A single image formed from the entire scanned region and the full



a) b)

Figure 5. Pseudo color graph of the amplitude of the displacement of the surface: a) raw measured data and b) data with the forward traveling waves filtered out. The top of the TS-50 mine is 1.3 cm below the surface.

measurement time by viewing the back propagating waves excited near the arrival time of the incident (thus rejecting the back wall reflection from the tank) is shown in figure 6. The location, shape and relative size of the mine can be seen in the image.

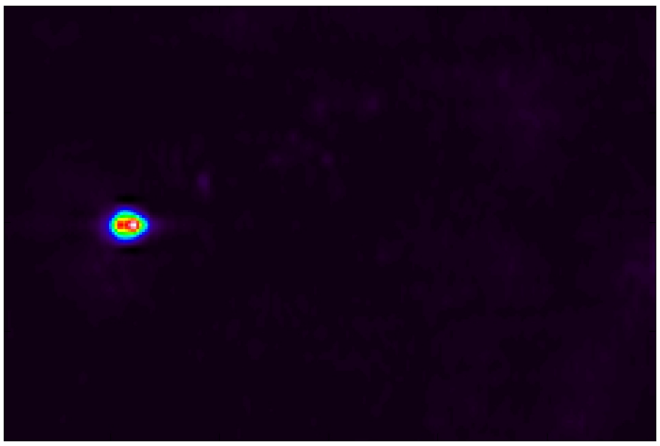
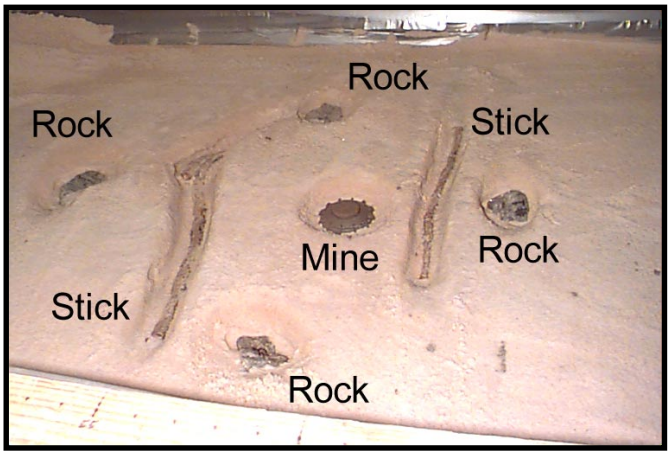


Figure 6. Image formed from the experimental results when the TS-50 anti-personnel mine is buried in the sand. The top of the mine is 1.3 cm below the surface.



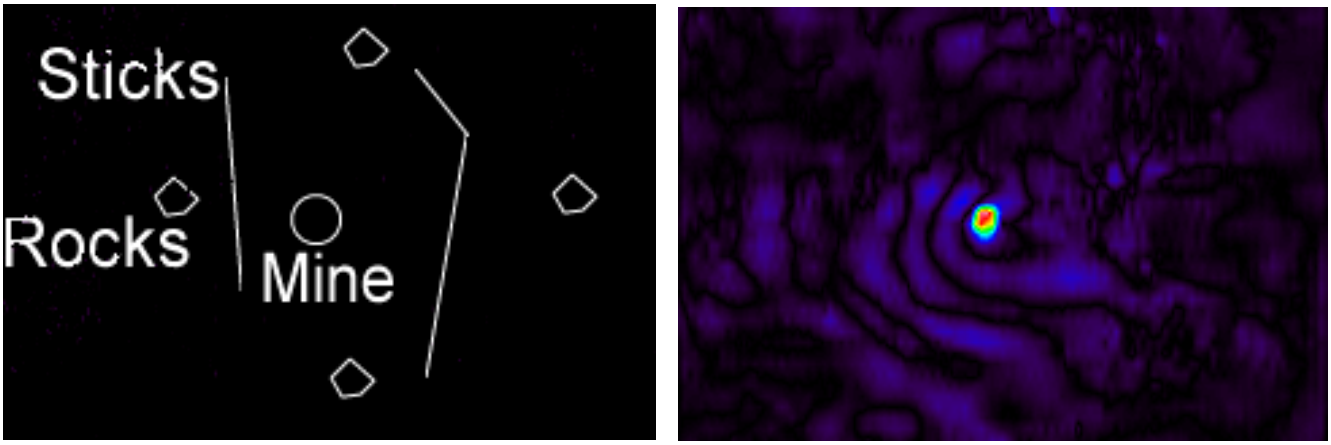
Photograph with the mine and the clutter uncovered.



Photograph showing the relative size of the mine and the clutter objects.

Figure 7. Photograph of the TS-50 mine surrounded by clutter. The top of the mine and the clutter is approximately 1.3 cm below the surface.

Another experiment was performed with this mine to see how well it could be differentiated from some common types of clutter. In this experiment, the mine was buried along with three buried rocks and two buried sticks; they were all buried approximately 1.3 cm deep. The mine and the buried clutter can be seen uncovered in the upper photograph in figure 7. The relative size of these objects can be seen in the lower photograph. Pseudo color graphs of the amplitude of the displacement of the surface are presented in figure 8 for the entire scanned region with the forward traveling waves filtered out. The time is chosen so that the incident pulse has reached the mine and the reflected pulses are clearly seen. An image formed from the data taken from the experiment is shown in figure 9. The location, shape and relative size of the mine can be clearly seen in the image; however, none of the clutter objects can be seen. Thus, the mine can be clearly differentiated from the clutter. The reason the mine is evident while the clutter is not is because the mine exhibits a resonance and the clutter objects do not. It is probable that all types of mines will exhibit similar resonances, but it is unclear if these will be manifest in the soil surface displacements or at what frequency they will occur. This is clearly an area for further investigations.



a) b)

Figure 8. a) Diagram showing the location of the mine and the clutter; b) Pseudo color graph of the amplitude of the displacement of the surface with the forward traveling waves filtered out. The top of the TS-50 mine is 1.3 cm below the surface.

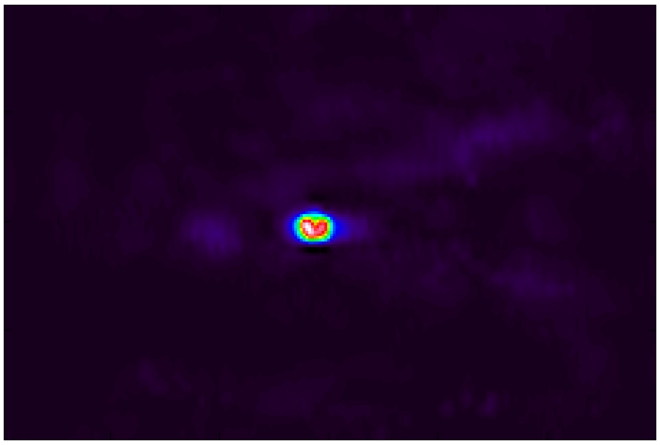


Figure 9. Image formed from the experimental results when the TS-50 anti-personnel mine is buried in the sand surrounded by clutter items (3 rocks and 2 sticks). The top of the mine and the clutter items is approximately 1.3 cm below the surface.

Waterfall graphs of the displacement of the surface are presented in figure 10a for y=0 (along the center of the scanned region). Graphs for 5 different mines as well as a graph without a mine are presented in the figure. The top of the mines is 6 mm below the surface of the sand. A time window of approximately 15 ms that follows along with the traveling wave is shown for each graph. The incident wave can be seen to propagate toward and across the buried mines. The wave in the no mine case is seen to propagate relatively undisturbed. Note, however, that it experiences significant dispersion as it travels through the sand. The wave can be seen to be disturbed in all of the cases in which a mine is buried. The disturbances are largest for the TS-50 and the butterfly mine, and the disturbances are smallest for the M-14 mine. Waterfall graphs of the same data with the forward traveling waves filtered out are presented in figure 10b. The waves seen in these graphs are due to both the reflected waves and local effects of resonances in the mines. The waves are seen to be significant for all of the mines, and insignificant for the no mine case. Thus, it is clearly possible to detect all of these mines.

4. NEW EXPERIMENTAL MODEL

The model presented above was designed with the primary consideration of the speed with which it could be built and instrumented. The model has performed well, but with the increasing refinement of the experiments several shortcomings of the model have become apparent. These problems were each addressed in the design of the new improved experimental model. A photograph of the new experimental model is shown in figure 11. The new model is much larger than the original one: the tank is approximately 4.5 m wide, 1.5 m deep and 4.5 m long; and is filled with 50 tons of packed damp sand. A larger faster positioner that can scan in the x, y, and z directions was incorporated into the model. The positioner is mounted 1.5 m above the surface of the sand with a large supporting frame to facilitate the larger standoff distances needed to test the beamforming arrays which are being investigated. The method for taking the data has also been changed. The data is taken in the time-domain using a swept frequency pulse ("chirp") that is temporally compressed in post processing. The new method is approximately 10 times faster while maintaining an adequate signal to noise ratio due to reduced ring-down time required at each measurement location. An additional gain in signal to noise ratio in the new model is provided by a larger source which drives the surface with 5 times the force of the acoustic source used in the first model.

The size of the first tank and the nature of its geometry and boundary conditions led to an artificially poor signal to reverberation ratio. This imposed as many limits as the signal to noise problem and resulted in restrictions on the time window for the measurement and repetition rate at which signals could be transmitted. Because of its increased size and lack of in band structural resonances the new tank greatly improves the signal to reverberation ratio. Also, the nature of the reverberation in the first tank varied with the depth of the water table in the sand. Since this could only be measured by disturbing the sand and could only by modified by the evaporation or addition of water there was considerable variation between measurements in the reverberation floor. This has been improved in the new tank with the addition of a drainage system that allows the depth of the water table to be measured and water to be both added and removed without disturbing the sand. Thus the reverberation floor is also more stable in the new experimental model. A comparison of a measured wave travelling in the old and new models is shown in figure 12. The wave in the new model is clearly cleaner. The reflections from the sides are not evident in the new model; they are the diagonal waves seen in the graph for the old model. The wave fronts are also straighter; this is due to the decreased measurement time and method of scanning.

In the past concerns had been raised over the validity of wet compacted sand as a surrogate for soil. Published data generally indicates higher shear wave speeds for soils than those that were measured in the damp compacted sand. This data, however, was taken under different conditions than those of interest for the mine detection problem and usually involves lower frequencies and deeper measurement locations. Prior to the construction of the new model field measurements around the Georgia Tech campus were made which confirmed that in the frequency range of interest soils have roughly the same surface wave speeds as those which have been observed in the sand (70-80 m/S). The only reasonable soil surrogate which could be shown to have a significantly higher surface wave speed (artificially crushed granite) cohered so tightly that mine burial would have been nearly impossible.

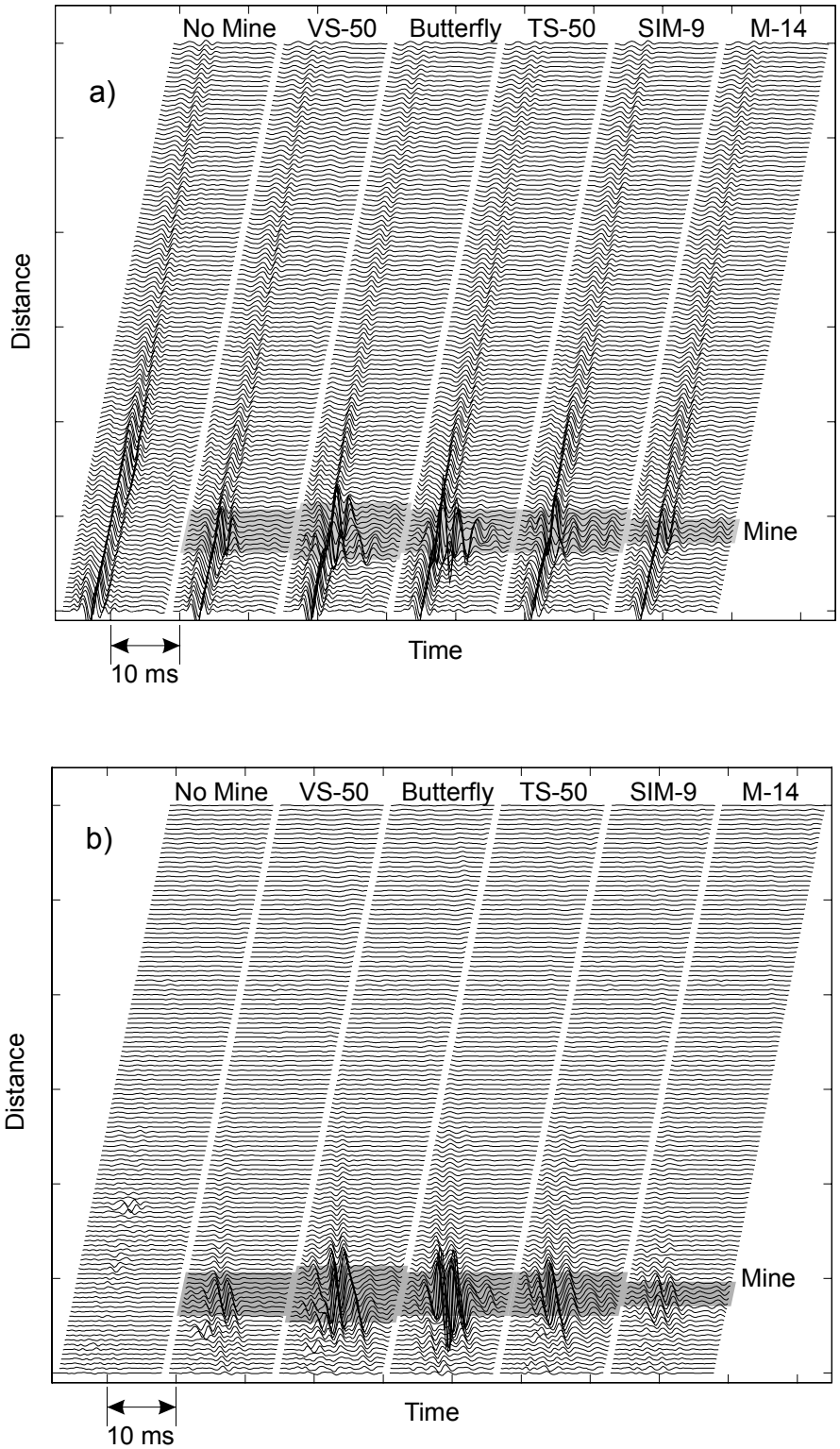
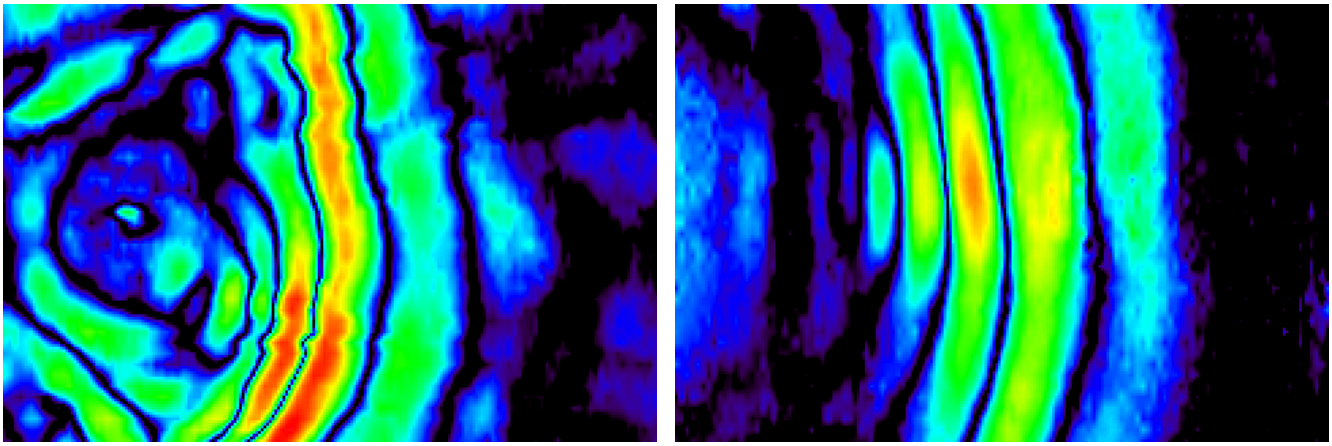


Figure 10. Waterfall graph of the displacement of the surface of the sand for a variety of anti-personnel mines are buried in the sand with y=0. The top of the mines are 0.6 cm below the surface.



Figure 11. Photograph of the new experimental model.



a) b)

Figure 12. Pseudo color graph of the amplitude of the displacement of the surface for a) the old experimental model and b) the new experimental model.

5. CONCLUSIONS

Experimental models for the acousto-electromagnetic sensor have been constructed and used to investigate the viability of this technique. The technique looks promising; it has been shown to be capable of detecting both simulated antipersonnel mines and antitank mines buried in damp sand. It has also been shown to be capable of distinguishing an antipersonnel mine from some types of naturally occurring clutter. However, much more work is needed to understand the capabilities of the sensor in more varied conditions.

6. ACKNOWLEDGEMENTS

This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0048

7. REFERENCES

1. Scott, W.R., Jr., Schroeder, C., and Martin, J.S., "An Acousto-Electromagnetic Sensor for Locating Land Mines," SPIE, AeroSense, *Detection and Remediation Technologies for Mines and Minelike Targets III*, Orlando, FL, pp. 176-186, April 1998
2. Scott, W.R., Jr. and Martin, J.S., "An Experimental Model of an Acousto-Electromagnetic Sensor for Detecting Land Mines," *Proceedings of the 1998 IEEE Antennas and Propagation Symposium,* Atlanta, GA, pp. 978-83, June 1998.
3. Scott, W.R., Jr., Schroeder, C., and Martin, J.S., "A Hybrid Acoustic/Electromagnetic Technique for Locating Land Mines," *Proceedings of the 1998 International Geoscience and Remote Sensing Symposium*, Seattle, Washington, July 1998
4. Lee, S.H. and Scott, W.R., Jr., "Beamforming Array for Detection Buried Landmines," SPIE, AeroSense, *Detection and Remediation Technologies for Mines and Minelike Targets III*, Orlando, FL, , April 1999

Finite-difference time-domain model for elastic waves in the ground

Christoph T. Schroeder and Waymond R. Scott, Jr.

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250
USA

ABSTRACT

A two-dimensional finite-difference model for elastic waves in the ground has been developed. The model uses the equation of motion and the stress-strain relation, from which a first order stress-velocity formulation is obtained. The resulting system of equations is discretized using centered finite-differences. A perfectly matched layer surrounds the discretized solution space and absorbs the outward traveling waves. The numerical model is validated by comparison to an analytical solution. The numerical model is then used to study the interaction of elastic waves with buried land mines. Results are presented for a buried antipersonnel mine. It can be seen, that an air-chamber within the mine is excited to resonant oscillations, which are clearly visible on the surface above the mine. The simulation results agree fairly well with experimental observations. Differences are mainly due to the numerical model being two-dimensional, whereas the experimental model is three-dimensional. Currently, the finite-difference model is being extended to three dimensions.

Keywords: land mine, elastic wave, acoustic, FDTD, finite-difference, numerical modeling

1. INTRODUCTION

A numerical finite-difference model for elastic waves in the ground has been developed and is described in this paper. The equation of motion and the stress-strain relation, together with a constitutive relation, form a set of first-order partial differential equations that completely describes the elastic wave motion in a medium. Introducing finite differences, this set of equations can be discretized and adapted to the finite-difference time-domain modeling scheme. Assuming that the field is known at one initial time t_0, this numerical scheme is used to determine the field values for all later times $t > t_0$.

The finite-difference model has been implemented in two dimensions. The solution space is discretized and a staggered finite-difference grid is introduced. The grid is surrounded by a perfectly matched layer, that absorbs the outgoing waves. The numerical model has been validated by comparison to an analytical solution. The analytical solution is obtained for elastic waves in a homogeneous half-space in form of an integral equation. Numerical integration must be used to determine the elastic waves. Excellent agreement is seen between the numerical model and the analytical solution.

The finite-difference model has been developed as part of a land mine detection project. In this project, a new technique is being investigated that synergistically uses both acoustic (elastic) and electromagnetic waves to detect land mines.[1] The numerical model has been very helpful for understanding the interaction of elastic waves with buried land mines.

The interaction of elastic waves with a buried antipersonnel mine has been explored. Two simple models for the TS-50 antipersonnel mine are investigated: one containing an air-filled chamber and one without an air-filled chamber. The interaction predicted by the simulation is in fairly good agreement with experimental results.[2] The differences are mainly due to the numerical model being two-dimensional, while the experimental model is three-dimensional. A three-dimensional numerical model is currently under development.

Electronic Mail and Telephone: christoph.schroeder@ee.gatech.edu, 404-894-3123, waymond.scott@ee.gatech.edu, 404-894-3048

2. TWO-DIMENSIONAL NUMERICAL MODEL

Elastic waves are launched into the earth and interact with a land mine that is buried in the ground. The interaction of the waves with the mine is to be investigated. For this problem, the earth can be modeled as an infinitely large half-space, bounded by a free surface. Under the conditions considered here, the earth can be approximated as isotropic and perfectly elastic and, thus, lossless.

2.1. Finite-difference Model

Figure 1 shows the two-dimensional finite-difference model. A normal point source is located on the free surface, exciting longitudinal (pressure) and transverse (shear) waves in the x-z-plane. The wave fields are invariant in the y-direction. At $z = 0$, a free surface bounds the solution space. A *perfectly matched layer* (PML) terminates the solution space at the remaining edges, absorbing all outward traveling waves. The space is discretized using a staggered finite-difference grid.

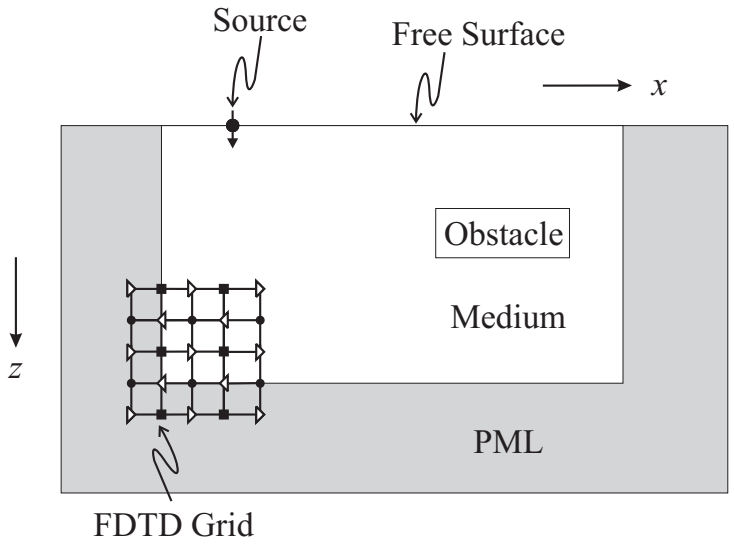


Figure 1. Two-dimensional finite-difference model.

A first-order velocity-stress formulation is chosen for the finite-difference model.[3] Here, the equation of motion and the stress-strain relation are discretized, leading to a system of first order partial differential equations. Since only field components in the x-z-plane are excited, the only non-zero field components are the three unknown stress components τ_{xx}, τ_{xz}, τ_{zz}, and the two unknown particle velocities v_x and v_z. Due to the invariance in the y-direction, all derivatives with respect to y vanish. The wave motion is then completely described by a system of five partial differential equations:

$$\rho\frac{\partial v_x}{\partial t} = \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xz}}{\partial z} \tag{1}$$

$$\rho\frac{\partial v_z}{\partial t} = \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{zz}}{\partial z} \tag{2}$$

$$\frac{\partial \tau_{xx}}{\partial t} = (\lambda + 2\mu)\frac{\partial v_x}{\partial x} + \lambda\frac{\partial v_z}{\partial z} \tag{3}$$

$$\frac{\partial \tau_{zz}}{\partial t} = (\lambda + 2\mu)\frac{\partial v_z}{\partial z} + \lambda\frac{\partial v_x}{\partial x} \tag{4}$$

$$\frac{\partial \tau_{xz}}{\partial t} = \mu\left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right), \tag{5}$$

where ρ is the material density and λ and μ are Lame's constants. These equations are discretized using centered finite-differences. Introducing the finite differences in space, Δx and Δz, and in time, Δt, Eq. (1) and (3), for

example, can be discretized as

$$\rho\frac{V_x^{k+0.5}(i-0.5,j-0.5)-V_x^{k-0.5}(i-0.5,j-0.5)}{\Delta t} = \frac{T_{xx}^k(i,j-0.5)-T_{xx}^k(i-1,j-0.5)}{\Delta x} + \frac{T_{xz}^k(i-0.5,j)-T_{xz}^k(i-0.5,j-1)}{\Delta z}, \quad (6)$$

$$\frac{T_{xx}^{k+1}(i,j-0.5)-T_{xx}^k(i,j-0.5)}{\Delta t} = (\lambda+2\mu)\frac{V_x^{k+0.5}(i+0.5,j-0.5)-V_x^{k+0.5}(i-0.5,j-0.5)}{\Delta x} + \lambda\frac{V_z^{k+0.5}(i,j)-V_z^{k+0.5}(i,j-1)}{\Delta z}. \quad (7)$$

Here, the capital letters mark the numerical value of the correspondent field component at a discrete location in space and time. For example, $V_x^{k+0.5}(i-0.5,j-0.5)$ stands for the numerical value of the particle velocity v_x at $(x,z)=((i-0.5)\Delta x,(j-0.5)\Delta z)$ at time $t=(k+0.5)\Delta t$. Knowing $V_x^{k-0.5}$, T_{xx}^k and T_{xx}^k, Eq. (6) can be solved for $V_x^{k+0.5}$, i.e. at the incremented time $t=(k+0.5)\Delta t$:

$$V_x^{k+0.5}(i-0.5,j-0.5) = V_x^{k-0.5}(i-0.5,j-0.5)+\frac{\Delta t}{\rho\Delta x}(T_{xx}^k(i,j-0.5)-T_{xx}^k(i-1,j-0.5)) + \frac{\Delta t}{\rho\Delta z}(T_{xz}^k(i-0.5,j)-T_{xz}^k(i-0.5,j-1)). \quad (8)$$

Similarly, T_{xx}^{k+1} is obtained from Eq. (7):

$$T_{xx}^{k+1}(i,j-0.5) = T_{xx}^k(i,j-0.5)+(\lambda+2\mu)\frac{\Delta t}{\Delta x}(V_x^{k+0.5}(i+0.5,j-0.5)-V_x^{k+0.5}(i-0.5,j-0.5)) + \lambda\frac{\Delta t}{\Delta z}(V_z^{k+0.5}(i,j)-V_z^{k+0.5}(i,j-1)). \quad (9)$$

In the same manner, discretized equations can be obtained for all field components.

Figure 2 depicts the position of the field components in the finite-difference grid. Note that the velocity components and the stress components are not known at the same position in time and space, but offset by $\frac{\Delta t}{2}$ in time and by $\frac{\Delta x}{2}$ and $\frac{\Delta z}{2}$ in space. This leads to the introduction of the staggered grid and the so-called leapfrog algorithm. In the leapfrog algorithm, the field components are updated sequentially in time: the velocity components are calculated first, then the stress components from the velocity components, the velocity components again using the stress components and so on. Thus, knowing the field components throughout the entire space at the time t_0 and $t_0+0.5\Delta t$, respectively, the field components can be determined for all later times $t>t_0$.

When implementing the finite-difference scheme, boundary conditions have to be treated in a special manner. Three different kinds of boundaries arise: the source point, internal boundaries (i.e. boundaries within the medium marked by a change in material properties), and external boundaries (i.e. the grid edges).

A normal point source is to be implemented on the free surface. This is done by forcing one field component to assume a specific time dependence. Here, the normal stress component τ_{zz} is excited by a differentiated Gaussian pulse.

The conditions at internal boundaries, i.e. at the interfaces between different media, are usually satisfied implicitly. However, to ensure numerical stability, the material properties have to be averaged for components on the boundary. While the material density ρ, appearing in the equation of motion, is averaged directly, the inverse of Lame's constants, λ and μ from the stress-strain relation, must be averaged. For transitions in between similar materials, the averaging may be omitted. However, it is absolutely indispensable at an interface between media with greatly different material properties (for example, at an air-solid-interface).

Four external boundaries arise at the four outer grid edges. At its upper edge, the half-space is bounded by a free surface. Due to the continuity of normal stress, the normal stress components vanish at a free surface. In order to satisfy this condition, an extra row must be inserted into the finite-difference grid one step beyond the free-surface boundary.

In order to model the infinite half-space, all waves that are reaching the three remaining outer grid edges must be perfectly transmitted and absorbed. The boundary condition that does this most accurately is the *Perfectly Matched*

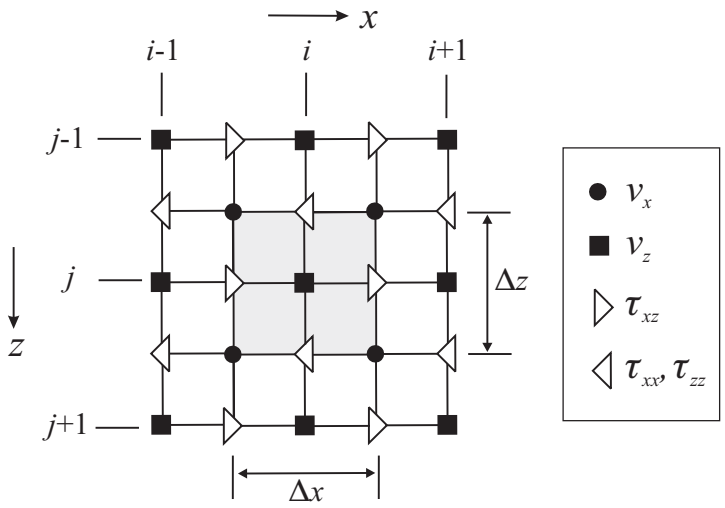


Figure 2. Part of the finite-difference grid. The field components are not known at the same locations, but offset by $\frac{\Delta x}{2}$ and $\frac{\Delta z}{2}$.

Layer (PML) boundary condition, first introduced by Berenger[4] and adapted to elastodynamics by Chew and Liu.[5] Here, a non-physical splitting of the wave fields allows the introduction of a lossy boundary layer that is perfectly matched to the solution space. It has been shown that an arbitrarily polarized wave incident on this PML medium is perfectly transmitted. The wave experiences the exact same phase velocity and wave impedance as in the solution space, while rapidly decaying along the axis normal to the PML-medium interface. The loss profile within the PML is chosen to be

$$\sigma(i) = \sigma_{\max} \cdot \left(\frac{i - i_{\mathrm{PML}}}{N_{\mathrm{PML}}} \right)^m , \tag{10}$$

where $m \approx 2.1$ and $\sigma_{\max} = \frac{0.1 N_{\mathrm{PML}}}{\Delta t}$; N_{PML} is the thickness of the PML in basis cells (here: $N_{\mathrm{PML}} = 20$), and $i - i_{\mathrm{PML}}$ indicates the position within the PML. This loss profile (with a slightly different $\sigma_{\max}$) has been found to yield best performance in electromagnetic finite-difference modeling.[6]

2.2. VALIDATION OF FINITE-DIFFERENCE RESULTS

An analytic solution for the particle displacement fields in an infinite homogeneous half-space subjected to a normal harmonic line-load of finite width can be derived.[7,8] Using a Fourier transform method, the particle displacement is obtained as an integral equation and can be determined using numerical integration.

The integral to be solved contains two poles and four branch cuts. Furthermore, the integration range goes from $-\infty$ to ∞. While the branch cuts do not impose a problem during numerical integration, the poles cannot be integrated numerically. To account correctly for the pole contributions, the poles must be subtracted from the integral and integrated analytically. To accelerate the integration, the infinite integration range may be truncated. However, for this to be possible, the integral has to converge sufficiently fast. In order to accelerate the convergence of the integral, an asymptotic approximation for the tails of the integrals is derived. This asymptotic approximation is subtracted from the integrals and integrated analytically.

To obtain the particle displacement in the time-domain, the displacement field must be Fourier transformed. Since the particle velocity is calculated in the FDTD simulation, the particle displacement is differentiated with respect to time, and the particle velocity is obtained.

2.2.1. Comparison to FDTD Results

In Fig. 3, the elastic waves due to a normal line load on a free surface are shown for both the analytic solution and the numerical simulation. A differentiated Gaussian pulse is launched at $t = 0$ ms from a source located at $(x, z) = (0, 0)$ cm. The magnitude of the particle velocity fields is shown at $t = 10$ ms. A pressure, a shear and a surface wave are seen to propagate. The pressure wave is the fastest of the three. The surface wave is propagating

slightly slower than the shear wave and contains the most energy. A plane wave arises at the free surface, induced by the pressure wave. This wave is called a head wave. Two more head waves can be seen that, however, do not appear to originate from any other wave.

Figure 3(a) shows the analytical solution for the particle velocity field, and Fig. 3(b) shows the finite-difference result. The analytic and numerical result are seen to be indistinguishable. In Fig. 4, the particle velocity components of numerical and analytical result are plotted along the x-axis and along the z-axis. The agreement of numerical and analytical results is very good. It can be seen that the finite-difference model predicts waves that travel slightly slower than those of the analytical solution. This is due to numerical dispersion within the finite-difference grid. By choosing a smaller grid spacing the effect of numerical dispersion can be reduced. The velocity component v_x must be zero along the z-axis at $x = 0$ cm. For the analytical solution, however, the v_x component is seen to have a nonzero value along the z axis, whereas it is seen to be zero for the finite-difference solution. This is due to an error introduced from an interpolation algorithm used for the analytical solution.

3. INTERACTION OF ELASTIC WAVES WITH THE TS-50 ANTIPERSONNEL MINE

The interaction of elastic waves with a TS-50 antipersonnel mine, buried in sand, is to be investigated. Two simple models for the TS-50 mine are shown in Fig. 5. In the first model, a small chamber filled with air is located on top of the mine's main chamber containing plastic explosives. The second model does not contain an air-filled chamber. The elastic properties of the materials used for the numerical simulation are summarized in Table 1.

Table 1. Parameters used for finite-difference simulation.

Sand	Shear wave velocity	$c_{s,\text{sand}}$	87 m/s
	Pressure wave velocity	$c_{p,\text{sand}}$	250 m/s
	Material density	ρ_{sand}	1400 kg/m^3
Mine	Shear wave velocity	$c_{s,\text{mine}}$	1100 m/s
	Pressure wave velocity	$c_{p,\text{mine}}$	2700 m/s
	Material density	ρ_{mine}	1200 kg/m^3
Air	Shear wave velocity	$c_{s,\text{air}}$	0 m/s
	Pressure wave velocity	$c_{p,\text{air}}$	330 m/s
	Material density	ρ_{air}	1.3 kg/m^3

Figure 6 shows the excitation as a function of time. The excitation is a differentiated Gaussian pulse with a center frequency of 800 Hz and launched at $t = 0$ ms. The source is located on the free surface at $x = 0$ cm. The left edge of the mine lies at $x = 45$ cm. Its upper edge is 2 cm underneath the surface of the ground.

Fig. 7 and 8 show the interaction of the excited waves with the antipersonnel mine. The mine in Fig. 7 contains an air-filled chamber, while the mine in Fig. 8 does not. The magnitude of the particle velocity is plotted on a logarithmic scale. The particle velocity field is shown at four different times: at $t = 2$ ms, $t = 5$ ms, $t = 6$ ms and $t = 10$ ms.

At $t = 2$ ms, a pressure, a shear and a surface wave have been launched. The pressure wave, the fastest of the three, just hits the mine. For the mine with the air-filled chamber, some energy is seen to be trapped in between the surface and the mine, while for the mine without the air-filled chamber no strong interaction occurs. At $t = 5$ ms and $t = 6$ ms, the surface wave hits the mine and is partially transmitted and partially reflected. At $t = 10$ ms, the waves reflected from the mine are clearly seen. For the mine without the air-filled chamber, the reflected waves are only weakly dispersed. However, for the case with the air-filled chamber, the reflected waves are seen to be strongly dispersed. Energy is still trapped in between the mine and the surface, though the incident surface wave has already passed by.

In Fig. 9, a waterfall graph of the particle velocity v_y at the surface is shown for a mine with and without the air-filled chamber. Here, v_y is plotted as a function of time and vertically offset by the distance from the source.

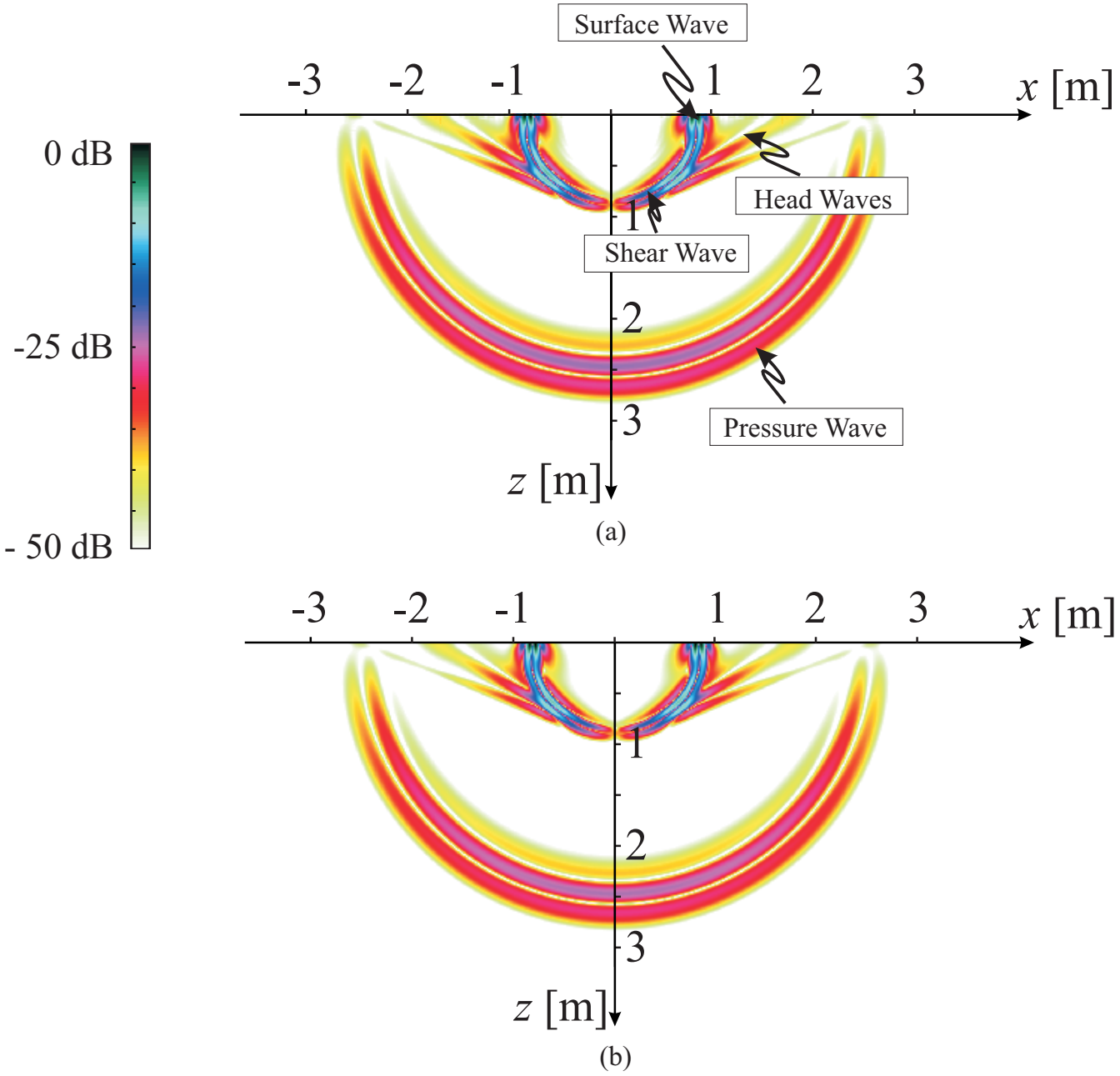


Figure 3. Plot of the magnitude of the elastic waves in a half-space subjected to a normal line-load at $t = 10$ ms: (a) analytical result; (b) finite-difference result.

The slope of the traveling waves in the graph indicates the wave speed. Thus, by looking at their slope, the different waves can be distinguished. A surface wave (S) and a pressure wave (P) are seen to be incident onto the mine. The pressure wave is reflected and transmitted by the mine. It converts into a reflected pressure wave (not visible in the graph), a reflected surface wave (rSP), a transmitted pressure wave (P) and a transmitted surface wave (SP). These waves are weak, due to the limited energy content of the pressure wave. The pressure wave is seen to travel faster over the mine than in the sand. This is due to the higher wave speed in the mine. The incident surface wave gives rise to a reflected pressure wave (rPS), a reflected surface wave (rSS), a transmitted pressure wave (PS) and a transmitted surface wave (S). For the mine with the air-filled chamber (Fig. 9(a)), a strong resonance can be seen at

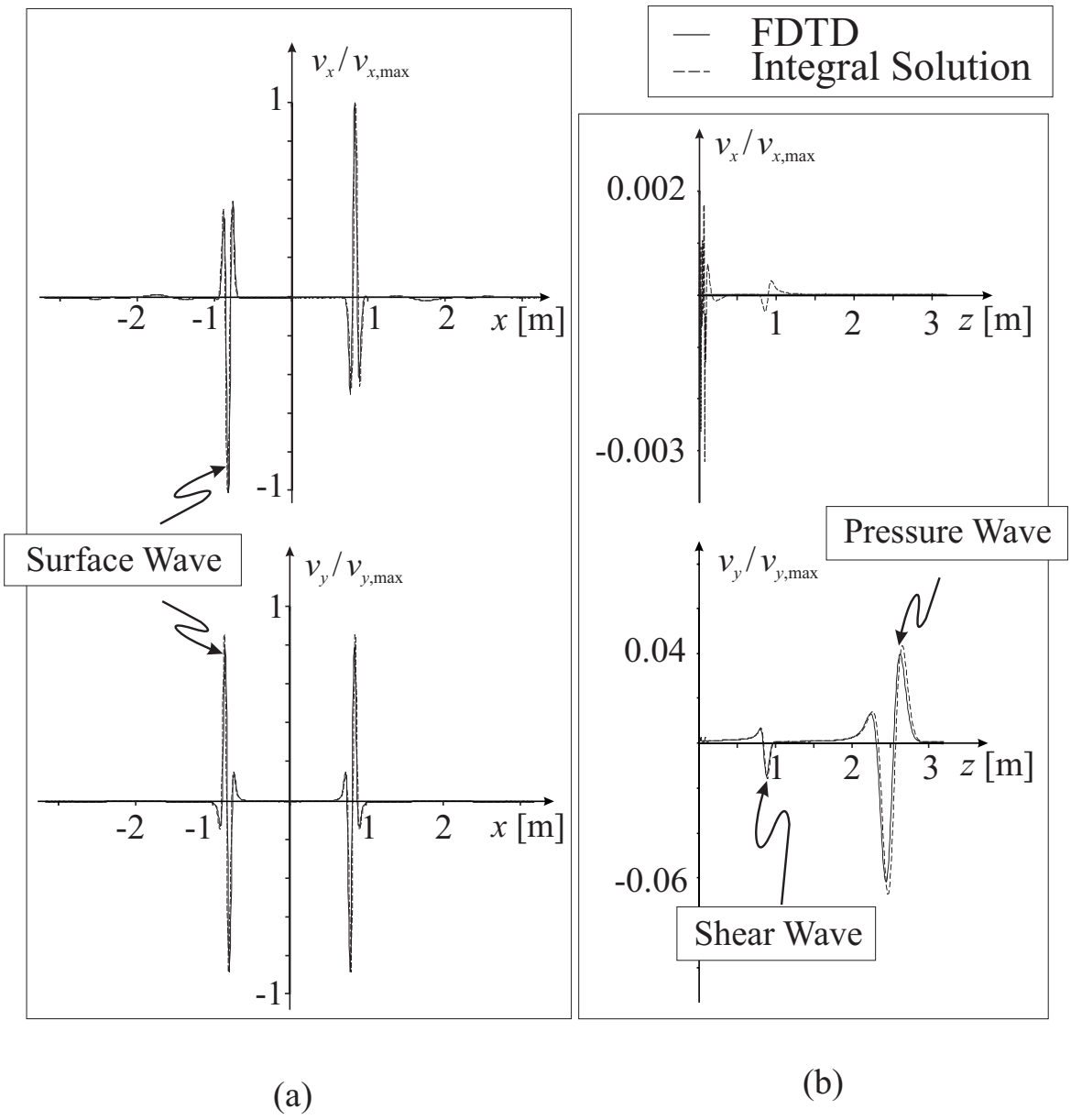


Figure 4. Comparison of Analytic and FDTD results: Displacement fields (a) along the x-axis, (b) along the positive z-axis. The particle velocity v_x along the y-axis is nonzero for the integral solution, due to an interpolation algorithm used during the Fourier transform.

the mine location. The resonance remains at the mine, even after the incident surface wave has passed by, and causes the surface wave to disperse. This resonance is due to the trapped energy in between the mine and the surface. For the mine without the air-filled chamber (Fig. 9(b)), no resonance occurs and the wave dispersion is weak.

The results obtained from the simulation are in fairly good agreement with experimental results.[2] A similar resonance for the mine with the air-filled chamber is observed in the experimental model. In the experimental model, the resonance is stronger than in the numerical model, whereas in the numerical model the reflections from the mine are much stronger. These differences are due to the numerical model being two-dimensional, whereas the experimental model is three-dimensional.

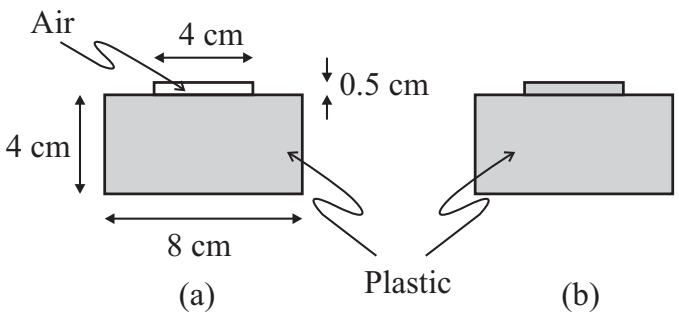


Figure 5. Simple model of the antipersonnel mine TS-50 (a) with and (b) without an air-filled chamber.

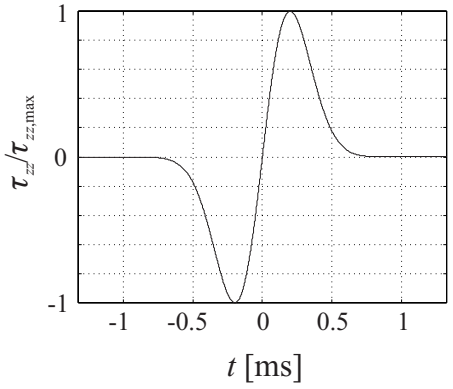


Figure 6. Source excitation: differentiated Gaussian pulse with a center frequency 800 Hz.

4. THREE-DIMENSIONAL NUMERICAL MODEL

The numerical finite-difference model is currently being extended to three-dimensions. A preliminary result is shown in Fig. 10. The particle velocity fields are plotted on the surfaces of a cube. The upper surface represents the surface of the ground and is modeled as a free surface. The cube faces on the left and on the right represent cross sections of the ground. A normal point source is located at $(x, y, z) = (0, 0, 0)$ cm. A differentiated Gaussian pulse is launched. Surface, shear and pressure waves are excited. As in the two-dimensional simulation, head waves arise at the free surface. On the cross sections, propagating surface, shear, pressure waves and head waves are visible, whereas on the surface only the surface waves and the head waves can be seen.

5. CONCLUSIONS

A finite-difference model has been developed and implemented in two dimensions. The model has been verified by comparison to an analytical solution. The finite difference model has been used to investigate the interaction of elastic waves with a buried antipersonnel mine. Results are obtained for a TS-50 antipersonnel mine with and without an air-filled chamber. For both cases, strong reflections are observed. For the mine with the air-filled chamber, a strong resonance at the mine location occurs. This resonance is caused by energy being trapped in between the mine and the surface. The results from the numerical model are found to be in fairly good agreement with experimental results. A three-dimensional model is currently being developed and will give further insight into the wave-mine interaction within the ground.

ACKNOWLEDGMENTS

This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0048.

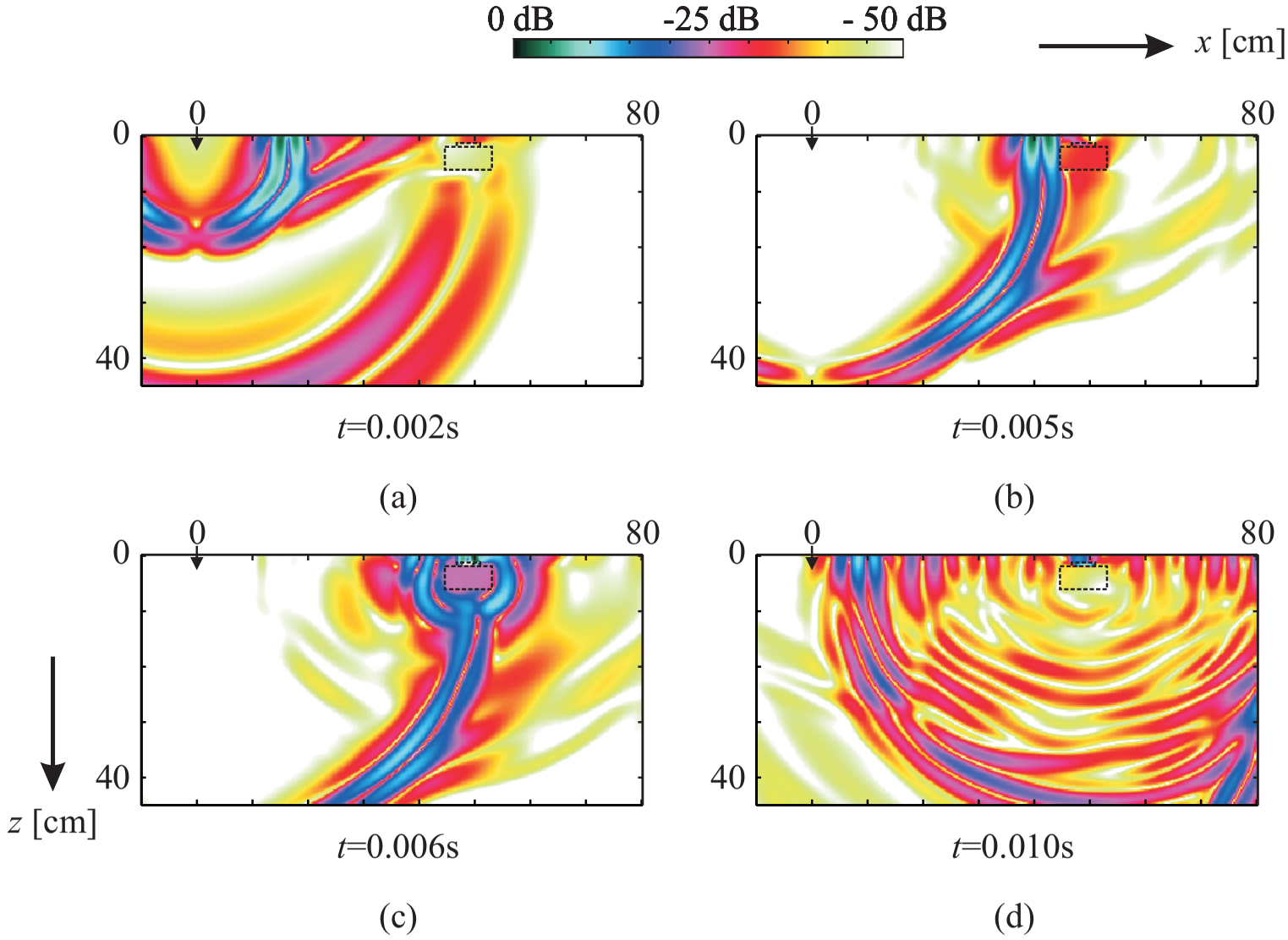


Figure 7.

Figure 7. Intensity plots of the interaction of elastic waves with a buried antipersonnel mine: mine with air-filled chamber.

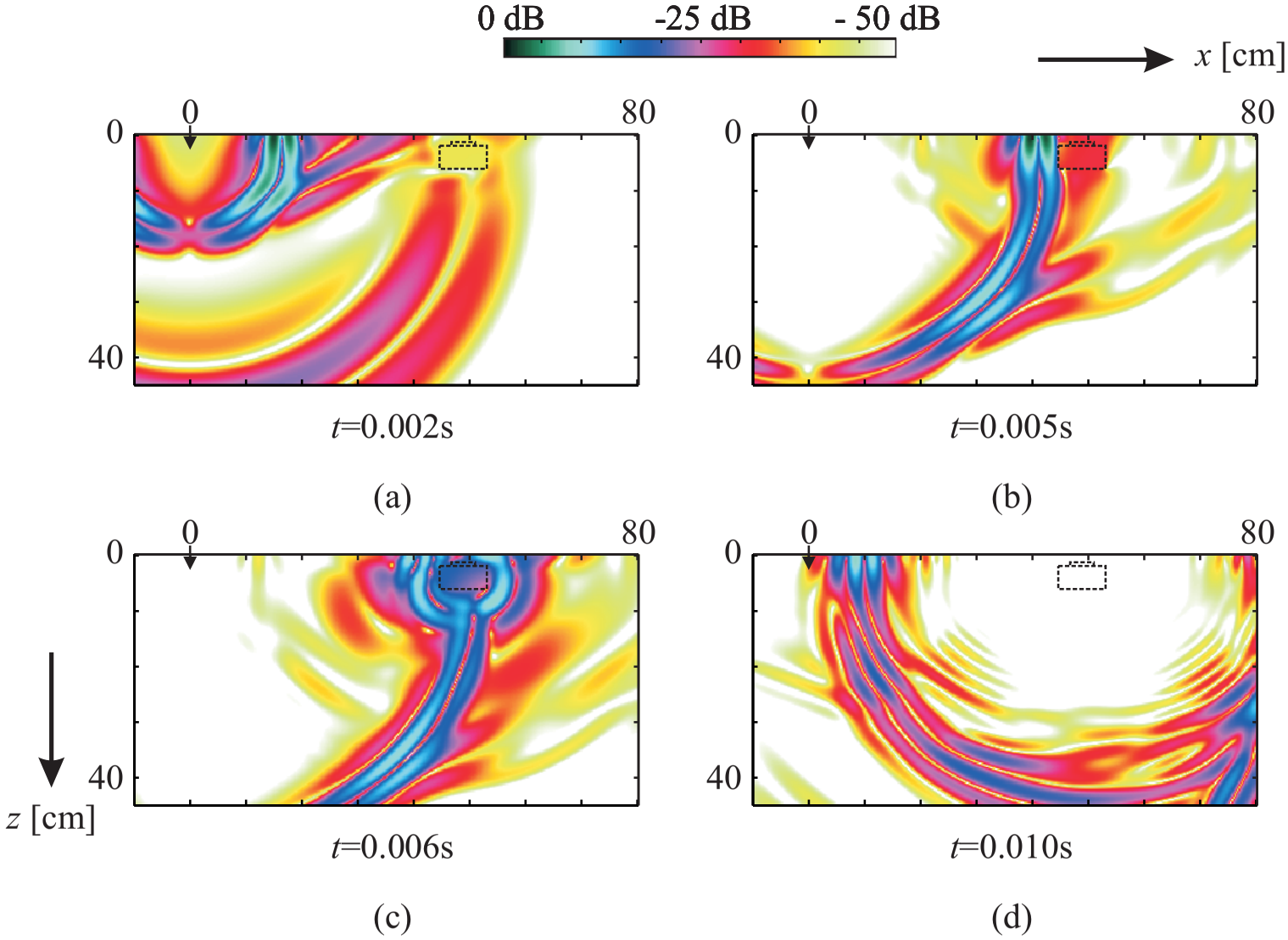


Figure 8.

Figure 8. Intensity plots of the interaction of elastic waves with a buried antipersonnel mine: mine without air-filled chamber.

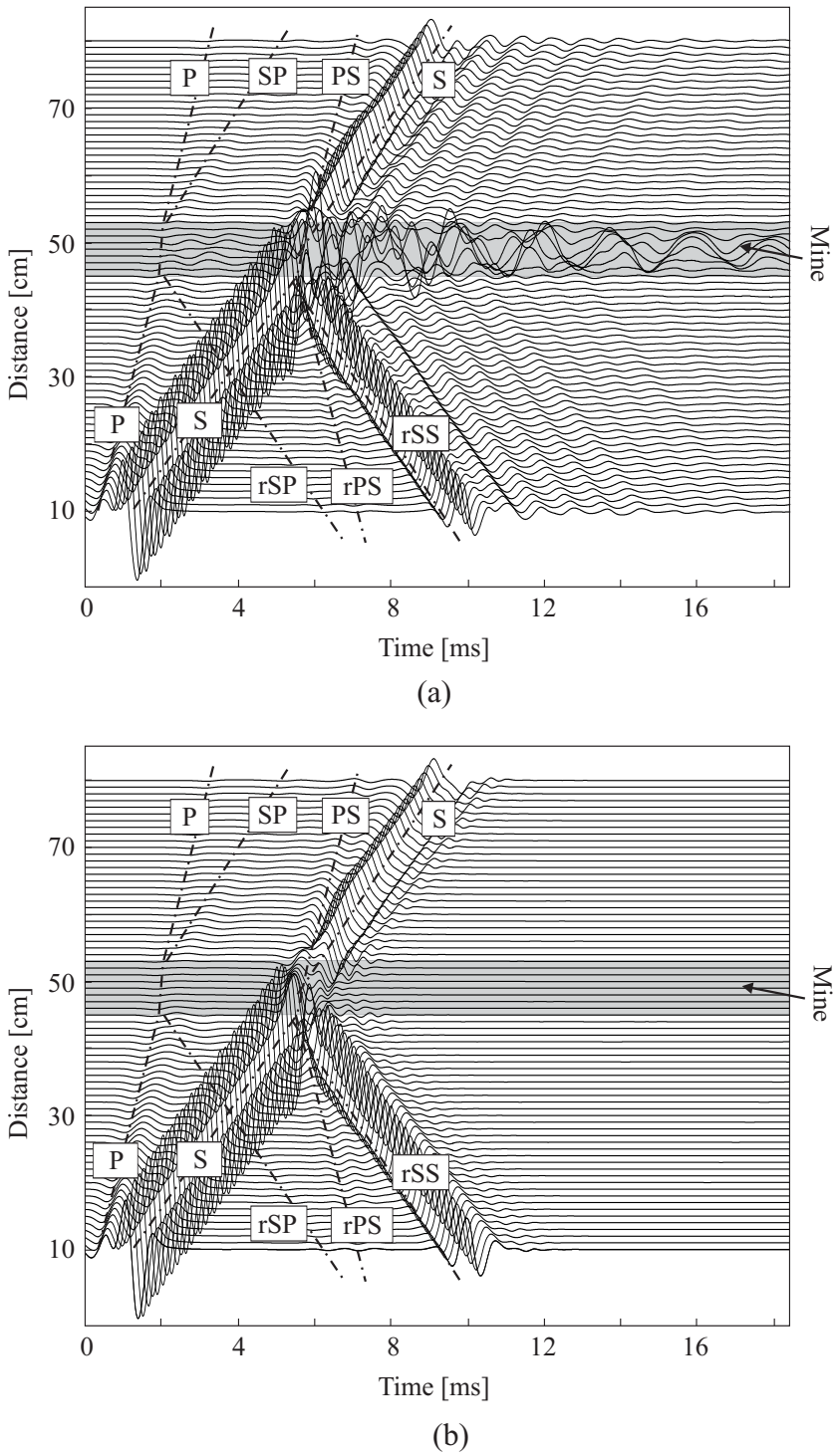


Figure 9. Interaction of elastic waves with a buried antipersonnel mine TS-50; waterfall graph of the particle velocity v_y at the surface for (a) a mine with the air-filled chamber and (b) a mine without the air-filled chamber. For the mine with the air-filled chamber, resonant vibrations can be seen at the mine location, which are due to energy being trapped in between the mine and the surface. P: pressure wave; S: surface wave; SP: surface wave due to pressure wave; PS: pressure wave due to surface wave; rSP: reflected surface wave due to pressure wave; rPS: reflected pressure wave due to surface wave; rSS: reflected surface wave due to surface wave. The vertical axis indicates the distance from the source.

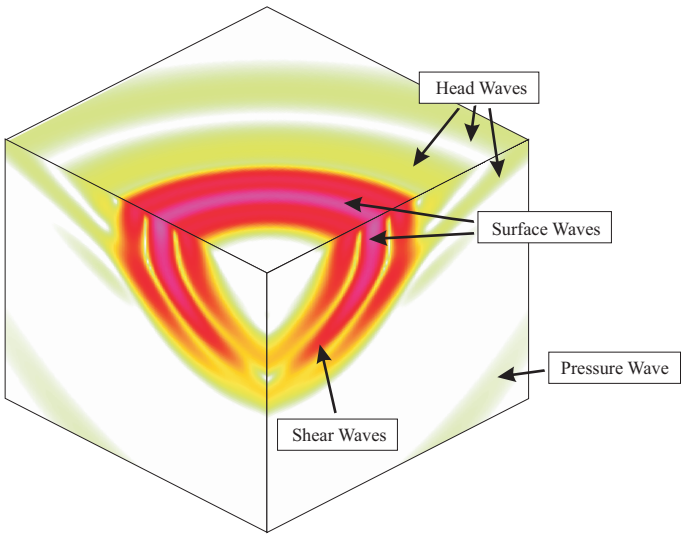


Figure 10. Propagating elastic waves in a three-dimensional half-space.

REFERENCES

1. W. R. Scott, Jr., C. T. Schroeder, and J. S. Martin, "An acousto-electromagnetic sensor for locating land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets III*, *Proc. SPIE* **3392**, pp. 176–186, 1998.
2. W. R. Scott, Jr. and J. S. Martin, "Experimental investigation of the acousto-electromagnetic sensor for locating land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, *Proc. SPIE* , 1999.
3. J. Virieux, "P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method," *Geophysics* **51**, pp. 889–901, 1986.
4. J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Physics* **114**, pp. 185–200, 1994.
5. W. C. Chew and Q. H. Liu, "Perfectly matched layer for elastodynamics; a new absorbing boundary condition," *J. Comput. Acoustics* **4**, pp. 341–359, 1996.
6. T. W. Hertel, *Pulse Radiation from an Insulated Antenna: An Analogue of Cherenkov Radiation from a moving Charge*, Master's Thesis, Georgia Institute of Technology, 1998.
7. G. F. Miller and H. Pursey, "The field and radiation impedance of mechanical radiators on the free surface of a semi-infinite isotropic solid," *Proc. R. Soc.* **A 223**, pp. 521–541, 1954.
8. K. F. Graff, *Wave Motion in Elastic Solids*, Dover Publications, Inc., 1975.

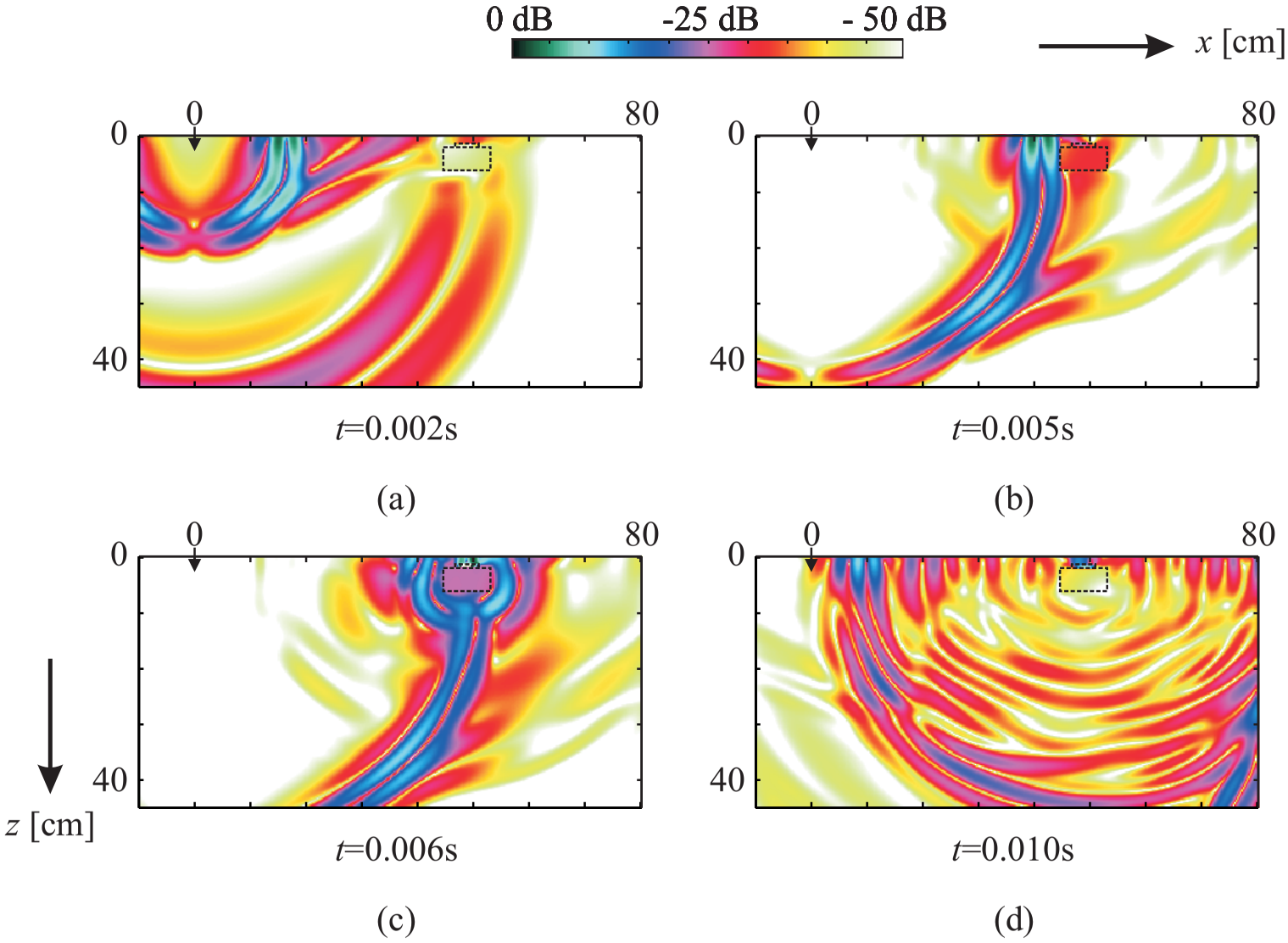


Figure 7. Intensity plots of the interaction of elastic waves with a buried antipersonnel mine: mine with air-filled chamber.

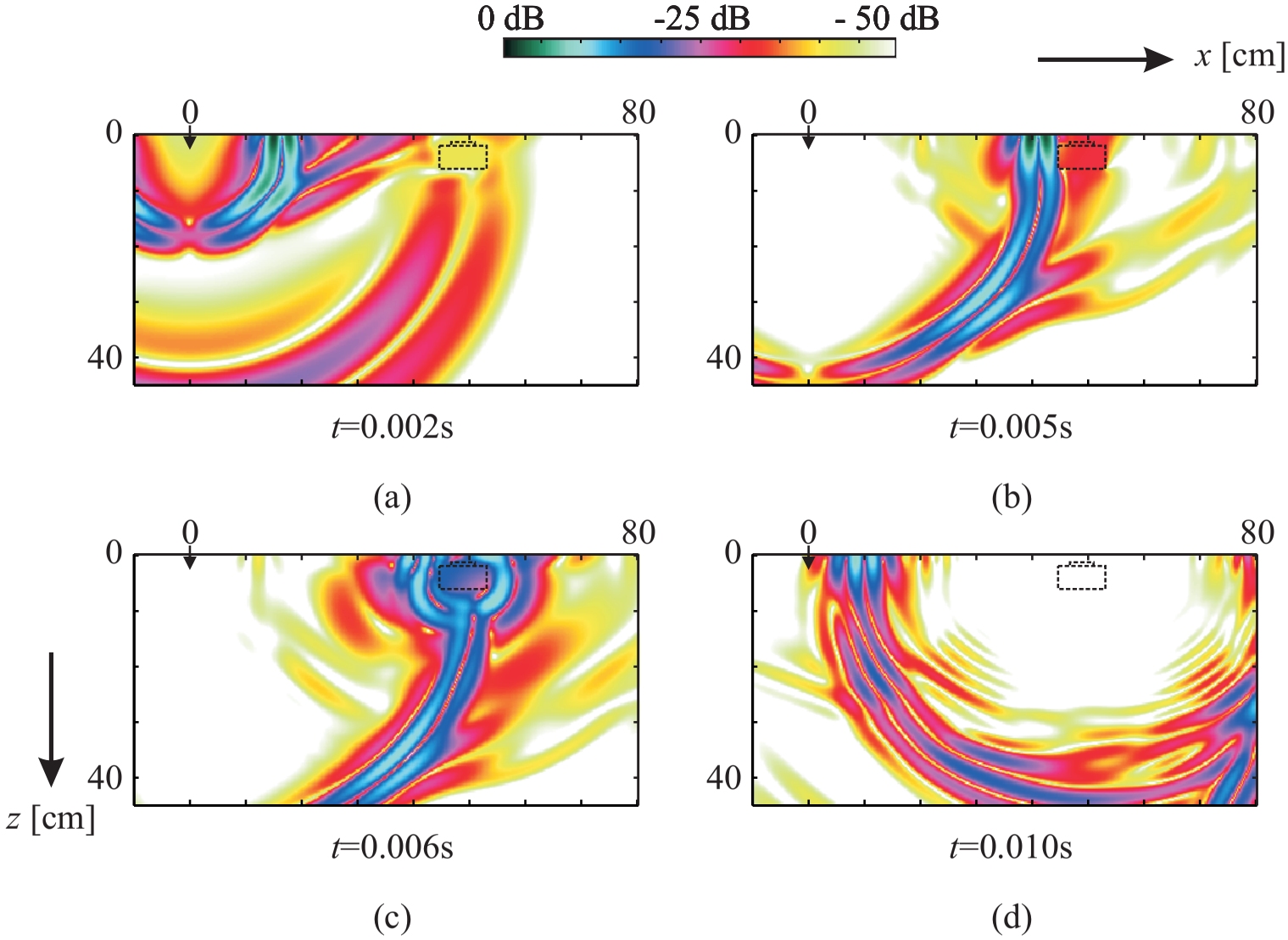


Figure 8. Intensity plots of the interaction of elastic waves with a buried antipersonnel mine: mine without air-filled chamber.

Near-Field Beamforming Array for Detecting Elastic Waves in the Earth

Seung-Ho Lee and Waymond R. Scott, Jr.

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

ABSTRACT

A near-field beamforming array is investigated for use in a radar system that is part of a hybrid elastic/electromagnetic technique for detecting land mines. The radar is used to measure the displacement of the surface of the earth and land mines due to elastic waves in the earth. The beamforming array is used to obtain a sufficiently small spatial resolution for the measurement of the displacement while allowing an adequate standoff distance for the radar. Both theoretical and experimental models are developed to investigate the viability of the beamforming array.

Keywords: Beamforming Array, Mine Detection, Displacement Measurement

1. INTRODUCTION

This work is part of a project in which a land mine detection system is being investigated that simultaneously uses both electromagnetic and elastic waves.[1] The configuration of the system currently being investigated is shown in Fig. 1. The system consists of an electromagnetic radar and an elastic wave source. An elastic wave is launched into the earth, and the elastic wave causes the surface of the earth and the mine to be displaced. The radar is used to measure the displacements and, thus, detect the mine. Currently, the radar uses an open-ended waveguide or a small horn as the antenna. The radar can detect displacements as small as 10^{-9} m and has a spatial resolution approximately 2 cm as currently configured. However, to obtain this spatial resolution, the antenna must be placed within a few centimeters of the surface of the earth.

In this work, we investigate the use of a near-field beamforming array so that the antenna can be placed farther from the surface while maintaining sufficiently small spatial resolution. The array can be implemented either by constructing an array of N elements and recording the signal from all of the elements simultaneously or by using a single element that is scanned to construct a synthetic array. The most practical solution will probably be to use a combination of these techniques and scan a subset of the N-element array to more quickly construct the synthetic array.

A theoretical and an experimental model has been developed for the beamforming array. These models are used to examine the feasibility of the beamforming array. The theoretical model is based on the integral equation method. It simulates the signals received by the antenna due to the displacement of the surface, and the beamforming array is implemented to reconstruct the displacement. The experimental model consists of an electrodynamic transducer to generate elastic waves, a tank filled with damp sand to simulate the earth, simulated mines, and a radar to measure the displacements of the surface of the sand. The beamforming algorithm is shown to improve the resolution of both the theoretical and experimental results.

2. THEORETICAL MODEL

A schematic diagram of the theoretical model is shown in Fig. 2. The array consists of I $\times$ J elements that are placed at the height h and parallel to the x-y plane. All elements are assumed to be identical. To make the model simple, the surface of the earth is modeled as a perfect electric conducting plane. The displacement of the surface is modeled as a circular bump with a diameter D and height Δz. For the i^{th} and j^{th} element, let us assume that electric and magnetic current densities $\vec{J}_s$ and $\vec{M}_s$ on the aperture of the element bounded by the surface s' are known, and are located at the position $\vec{R}'$. The position vector $\vec{R}'$ is used to locate the M $\times$ N observation points on the surface

E-mail and Tel.: S.H. Lee: gt2332b@prism.gatech.edu, 404-894-3123, W. R. Scott, Jr.: wrscott@ee.gatech.edu, 404-894-3048

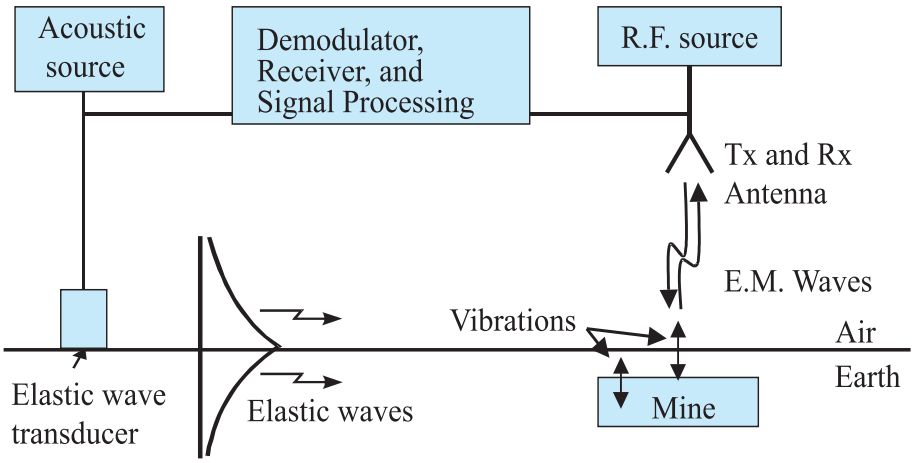


Figure 1. A system that simultaneously uses both elastic and electromagnetic waves to detect buried land mines.

of the conductor. The electric field radiated by the aperture, ignoring the effects of the conducting surface, can be expressed as[2]

$$\vec{E}^t_{e,ij}(\vec{R}) = \frac{-j\omega\mu_0}{4\pi}\iint_{s'}\{\vec{J}_s + \frac{1}{k_0}[-k_0^2(\vec{J}_s\cdot\hat{r})\hat{r} + \frac{3}{r}(jk_0+\frac{1}{r})(\vec{J}_s\cdot\hat{r})\hat{r} - \frac{\vec{J}_s}{r}(jk_0+\frac{1}{r})]\}\frac{e^{-jk_o r}}{r}\,ds', \tag{1}$$

$$\vec{E}^t_{m,ij}(\vec{R}) = -\frac{1}{4\pi}\iint_{s'}[\vec{M}_s\times\hat{r}(jk_0+\frac{1}{r})]\frac{e^{-jk_o r}}{r}\,ds', \tag{2}$$

where $\hat{r} = \frac{\vec{R}-\vec{R}'}{|\vec{R}-\vec{R}'|}$ and $r = |\vec{R}-\vec{R}'|$. The total electric field is then

$$\vec{E}^t_{ij}(\vec{R}) = \vec{E}^t_{e,ij}(\vec{R}) + \vec{E}^t_{m,ij}(\vec{R}). \tag{3}$$

Since the conducting surface is in the far-field of the source, the magnetic field is

$$\vec{H}^t_{ij}(\vec{R}) = \frac{1}{\eta_o}\hat{s}\times\vec{E}^t_{ij}(\vec{R}), \tag{4}$$

where $\hat{s} = \frac{\vec{R}}{|\vec{R}|}$(radial unit vector). Next, the surface current density is obtained from the tangential components of the magnetic field on the surface. Since the surface of the conductor is approximately planar, the surface current density can be approximated by

$$\vec{J}^{eq}_{ij}(\vec{R}) = 2\hat{z}\times\vec{H}^t_{ij}(\vec{R}). \tag{5}$$

Let $\Delta\vec{E}^r_{ij}(\vec{R})$ be the electric field at the center of the i^{th} and j^{th} element due to the current element at the location $\vec{R}$. Then, the voltage received by the i^{th} and j^{th} element is given by

$$V^r_{ij} \cong \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{N}\Delta\vec{E}^r_{ij}(\vec{R})\cdot\vec{l}^e_{ij}(\vec{R}), \tag{6}$$

where $\vec{l}^e_{ij}(\vec{R})$ is the vector effective height. Here the vector effective height is used to determine the voltage induced on the open-circuit terminals of the antenna from the received field. The effective height is determined from the radiated field(3) of the antenna[3]:

$$\vec{l}^e_{ij}(\vec{R}) = \frac{4\pi|\vec{R}|\vec{E}^t_{ij}(\vec{R})}{-j\eta_0 k_0 I_t e^{-jk_0|\vec{R}|}}. \tag{7}$$

For each element, the above steps must be repeated to synthesize the beamforming array.

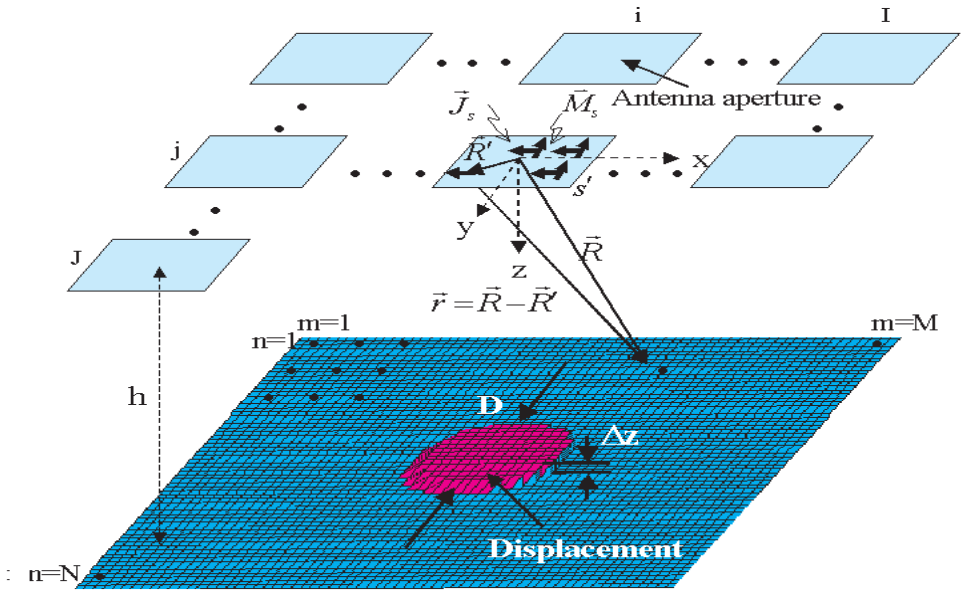


Figure 2. A schematic diagram of the theoretical model

3. BEAMFORMING ALGORITHM

After obtaining the received voltage at each element position, a beamforming algorithm can be applied. A schematic diagram of the beamforming array is shown in Fig. 3. The configuration of the array is the same as in Fig. 2 except for the coordinate system. The location of the element is indicated by $\vec{R}'$, and the received voltage at each element position is known. The array is focused at an arbitrary focal point located at the position $\vec{R}_f$. The received voltages of the elements are summed with the appropriate weighting function to focus the array at the focal point. The phase of the weighting function is chosen so that the voltages will add constructively by using the conjugate focusing method, and the amplitude of the weighting function is chosen to control the sidelobe levels of the array. The output of the array is then

$$V_A(\vec{R}_f) = \sum_{i=1}^{I}\sum_{j=1}^{J} V_{ij}^r(\vec{R}')W(\vec{R}_f,\vec{R}')e^{j\alpha_{ij}(\vec{R}_f)}, \tag{8}$$

where W is a window function used to adjust the amplitude of the weighting function, and $\alpha_{ij}(\vec{R}_f) = (-1)\{\text{phase of } \Delta V_{ij}^r(\vec{R}_f)\}$, where $\Delta V_{ij}^r(\vec{R}_f) = \Delta\vec{E}_{ij}^r(\vec{R}_f)\cdot\vec{l}_{ij}^e(\vec{R}_f)$. We have determined empirically that a Gaussian window function performs well:

$$W(\vec{R}_f,\vec{R}') = \exp\left(-\frac{1}{(0.6\tau)^2}|\hat{z}\times(\vec{R}_f-\vec{R}')|^2\right), \tag{9}$$

where τ is a parameter used to adjust the width of the window function at the 50% drop point from the peak. The window function is maximum directly above $\vec{R}_f$, $|\hat{z}\times(\vec{R}_f-\vec{R}')| = 0$, and it decreases with increasing distances from this point.

Parameters used in the example of the theoretical model are shown in Table 1. Using the procedure described above, the surface displacements are calculated from the outputs of the array to demonstrate the effectiveness of the beamforming array. Graphs of the reconstructed displacement of the conducting surface are shown in Fig. 4. The actual displacement and the displacements with and without beamforming are shown when the array is placed at 20 cm and 40 cm high. In these graphs, the displacements are plotted as a function of x with $y = 0$. Here the displacements are obtained by comparing the phase between the data obtained with and without the displacement in the model. The displacement obtained without beamforming is not a good replica of the actual displacement for either height. In order to solve this problem, the beamforming array technique was used to obtain the good image

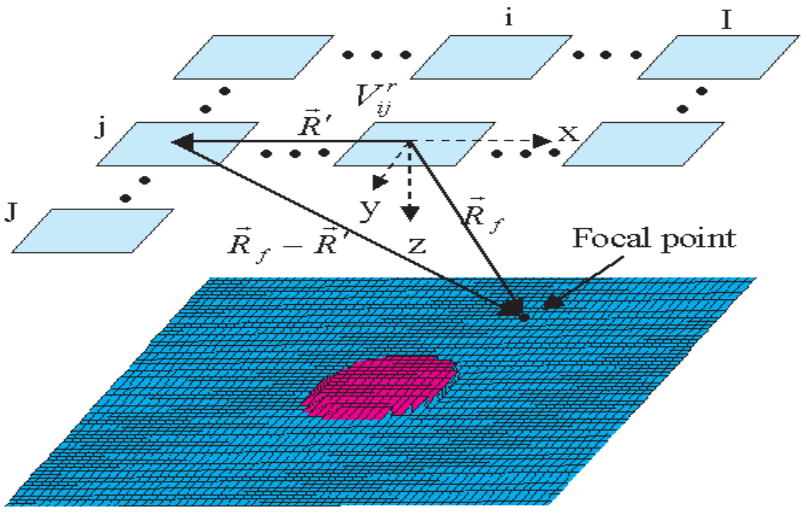


Figure 3. Beamforming array configuration

Table 1. Parameters used in the example

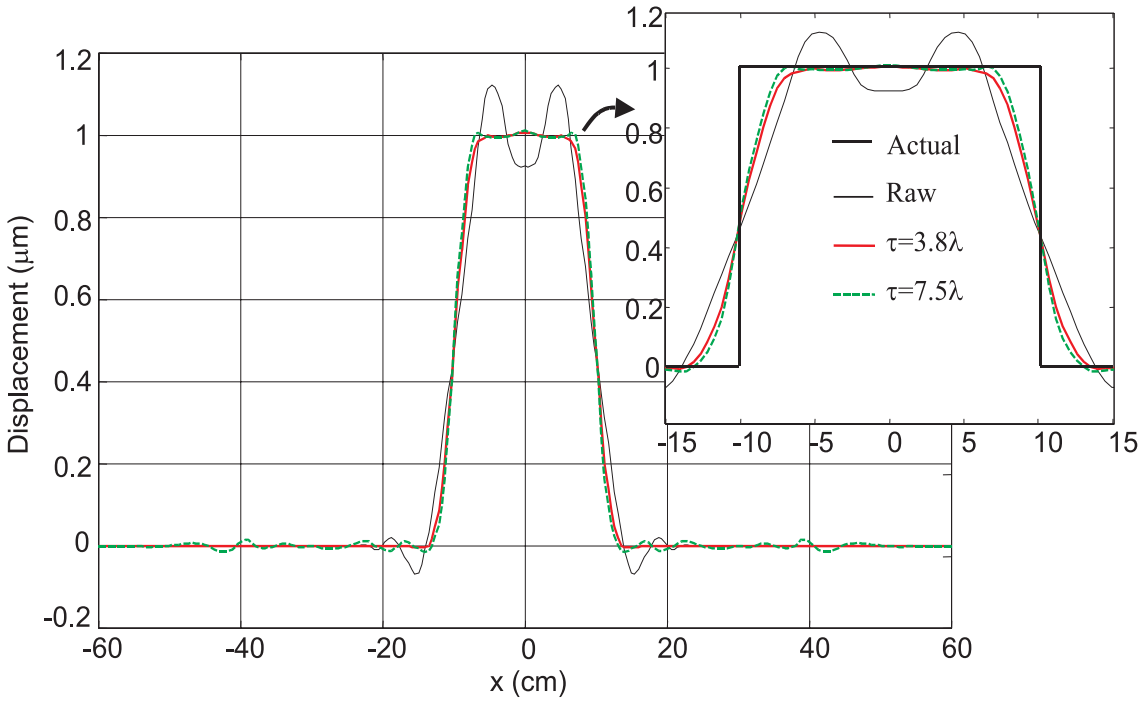
Array	65 × 65
Antenna	5 cm × 5 cm Pyramidal horn
Spacing between antennas	0.5λ
Frequency	8 GHz
Array height, h	20, 40 cm
Displacement	Circular with D=20 cm and Δz=1 μm
Plane	1.2 m × 1.2 m
Observation points on the surface	256 points/λ^2

of the displacements. The raw data were beamformed for several values of τ. When the raw data are beamformed with appropriate value of τ, the displacements are seen to be a better replica of the actual displacement than the results without beamforming. The spatial resolution is seen to increase with increasing τ, but the sidelobe levels also increase with increasing τ. Tradeoffs must be made between the spatial resolution and the sidelobe levels.[4]

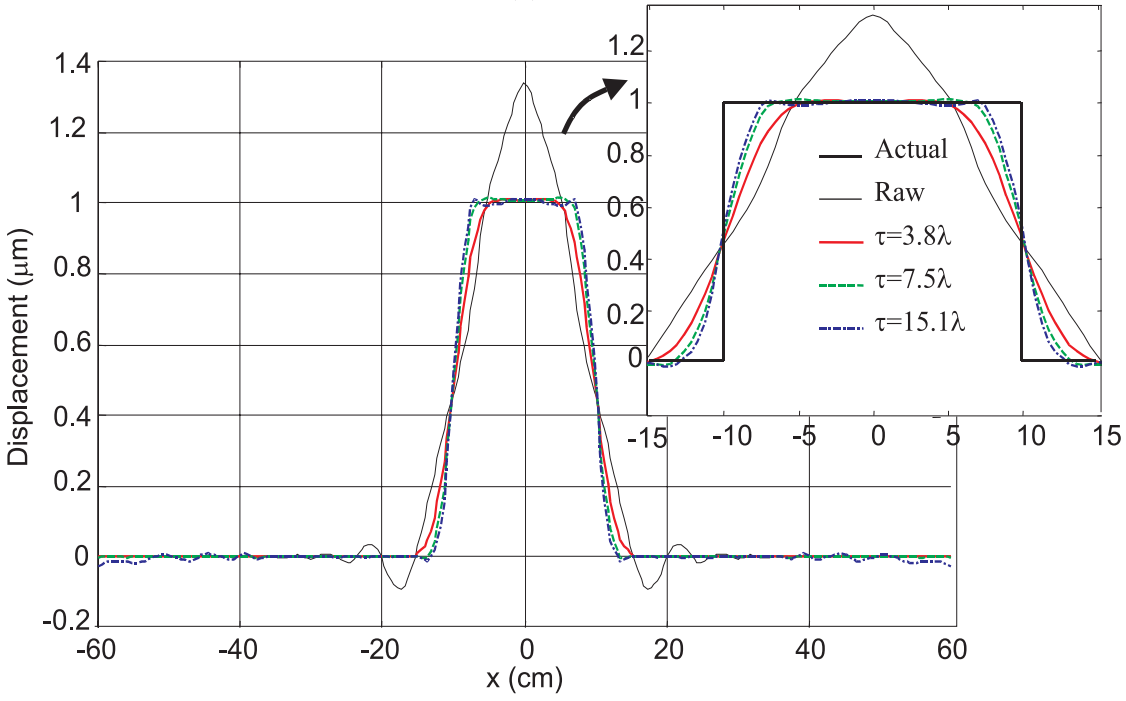
Pseudo color graphs of the amplitude of the displacement of the surface are presented in Fig. 5 for the entire conducting region. The displacements are shown for the data obtained with and without beamforming on a 40dB scale. The sidelobes are clearly visible around the displacement in the raw data for both h=20 cm and h=40 cm. The sidelobe levels increase when the height of the array increases. The top of the displacement is seen to be non-uniform in the raw data for both heights. After beamforming with the appropriate τ, the sidelobes are seen to be significantly reduced, and the top of the displacement is seen to be more uniform.

4. EXPERIMENTAL MODEL

A radar has been designed and built to measure the displacements of the surface of the sand and the mine due to the elastic waves.[1] Figure 6 is a block diagram of the radar system. The radar radiates electromagnetic waves that are reflected off of a vibrating boundary, and the reflected waves are received by the radar. A Homodyne system is used to demodulate the signals, and the signals from the I and Q channels are used to obtain the magnitude and phase of the reflected wave, which are then used to obtain the displacement of the surface. Currently, the radar is being used in a monostatic mode and uses a 5 cm × 5 cm Pyramidal horn as the antenna. The radar is operated at 8 GHz and is scanned using an x-y positioner. The radar is interfaced to a personnel computer to control the positioner and to store the measured data.



(a)



(b)

Figure 4. Reconstruction of the displacement of the surface : (a) h=20 cm, (b) h=40 cm.

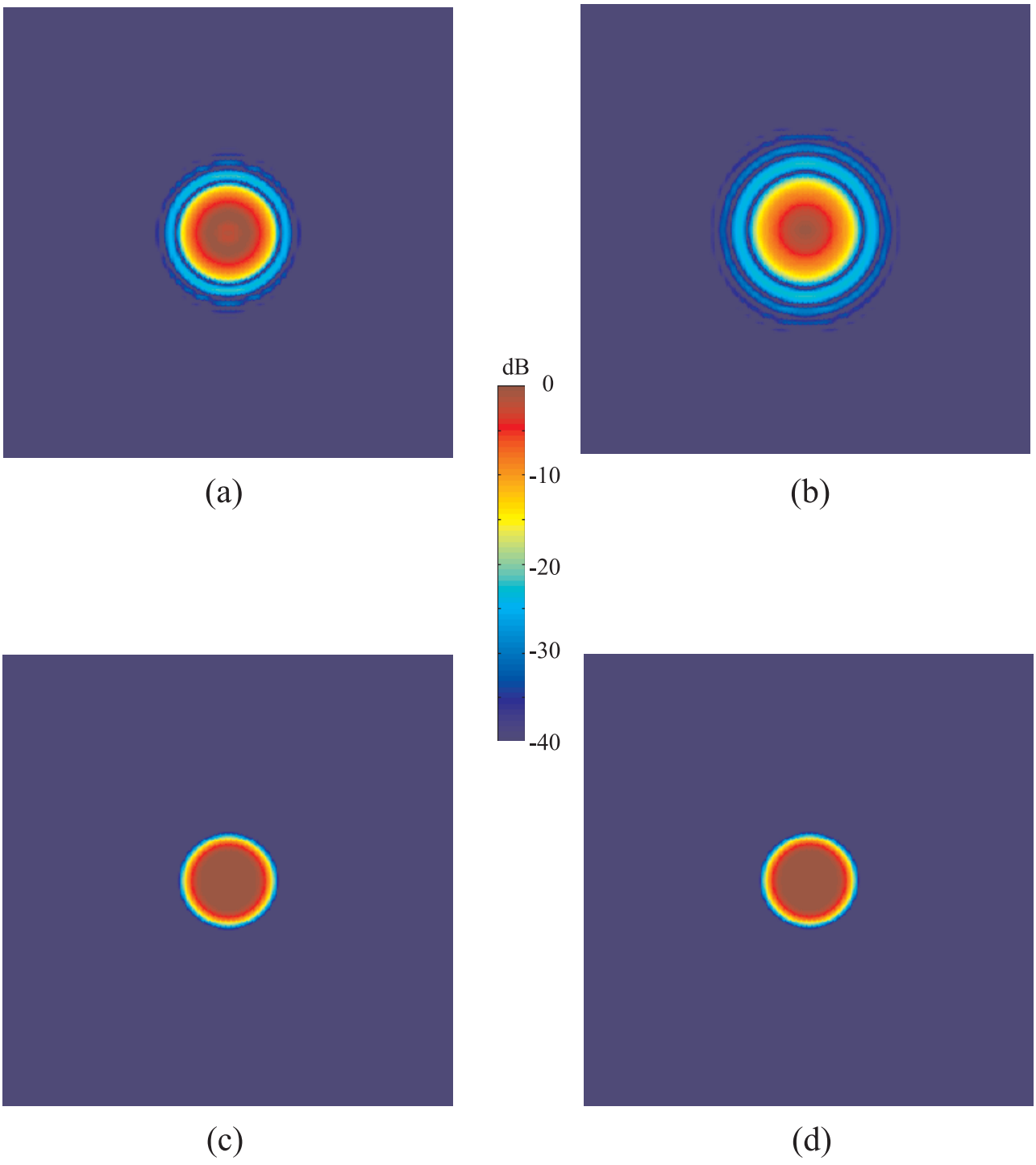


Figure 5. Pseudo color graphs of the displacement of the surface: (a) raw with h=20 cm, (b) raw with h=40 cm, (c) beamformed with $\tau = 3.8\lambda$ and h=20 cm, (d) beamformed with $\tau = 7.5\lambda$ and h=40 cm.

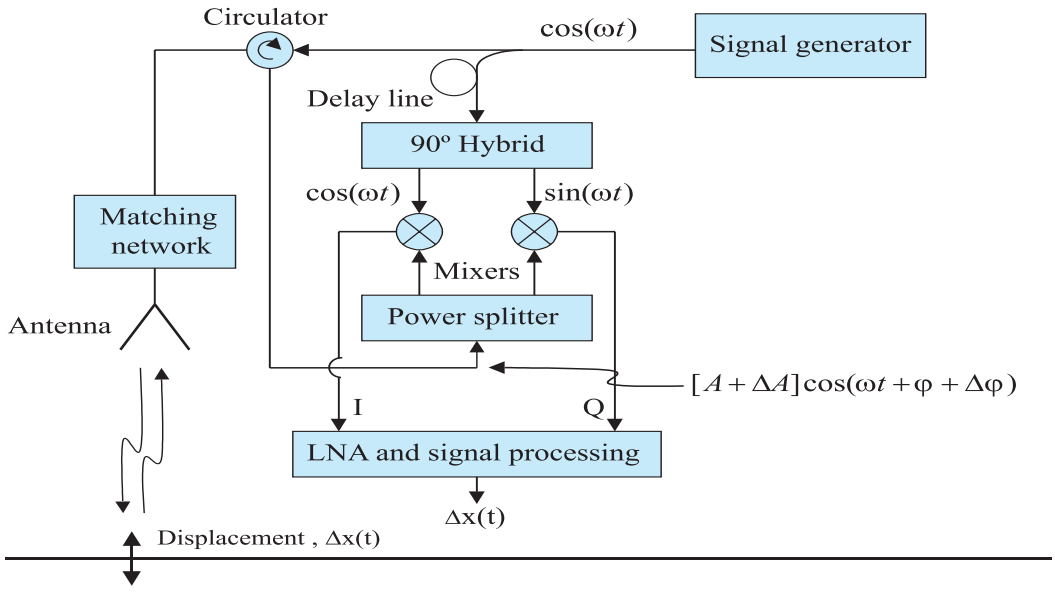


Figure 6. Radar system used in the experimental model

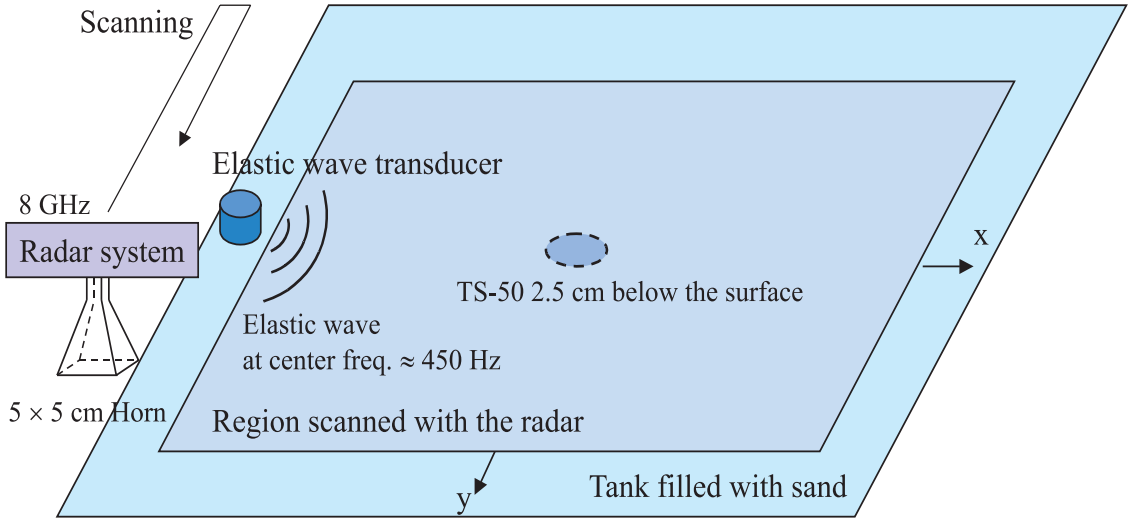


Figure 7. A schematic diagram of the experimental model

A schematic diagram of the experimental model is shown in Fig. 7. The model consists of a sandbox and a radar which is scanned over the surface of the sand to implement a synthetic beamforming array. The scanned area is 1.2 m square. A TS-50 anti-personnel mine is buried at the center of the area, and the top of the mine is 2.5 cm below of the surface of the sand. A transducer is placed on the surface of the sand and used to launch the elastic waves into the sand. These elastic waves travel across the surface of the sand and cause the surface of the sand and the mine to be displaced. The radar is used to measure the displacement of the surface of the sand due to the elastic waves. The response of the radar is measured on a uniform square grid of discrete positions in the scanned region. The grid consists of 65×65 equally spaced points that are spaced 1.875 cm apart. The displacements of the surface of the sand were measured and recorded as a function of time at each of these positions. Experiments have been performed for three different heights of the antenna. The antenna is placed at 5 cm, 20 cm, or 40 cm high above the surface of the sand.

Pseudo color graphs of the amplitude of the displacement of the surface are shown in Fig. 8 for the entire scanned region at a single instant in time. In these graphs, the elastic wave is traveling across the surface of the sand from the left hand side to the right hand side. These graphs show the displacements of the surface at a time at which the elastic wave has just passed the mine. A portion of the wave is seen to be scattered from the mine, and the displacements are seen to be biggest above the mine due to a resonance that occurs at the mine location. The resonance makes it much easier to detect the mine.[5] The mine is clearly visible at the center of the scanned region in all of the graphs. The displacements are shown for the raw data and for the beamformed data with $\tau = 3.8\lambda$ and $\tau = 7.5\lambda$ for h=20 cm and h=40 cm, respectively. Figures 8(a)-(c) show the raw measured data without beamforming for three different heights of the antenna. At a height of 5 cm, the raw data is a good representation of the actual displacement so that we can compare this result to the other cases. The spatial resolution becomes worse as the height of the radar is increased. Therefore, the mine is seen to appear much bigger than its actual size, and some artifacts(sidelobes) are observed around the mine.

The beamforming array has been used to improve the spatial resolution for the raw measured data. The beamformed displacements are also shown in Fig. 8(d) and (e) when h=20 cm and h=40 cm. The raw data was beamformed with $\tau = 3.8\lambda$ for h=20 cm and with $\tau = 7.5\lambda$ for h=40 cm. In both Fig. 8(d) and (e), the results are seen to be clearly better than the results without beamforming. Since the resolution has been improved, the location and the size of the mine are more evident in these graphs. Moreover, the artifacts around the mine have been significantly reduced after beamforming.

5. CONCLUSIONS

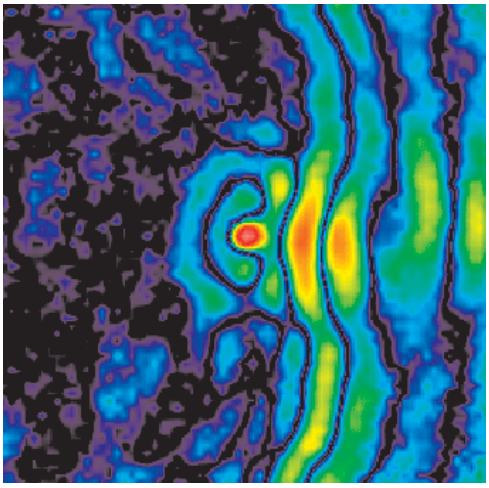
The feasibility of the near-field beamforming array has been investigated using a theoretical and an experimental model. By using the beamforming array, the spatial resolution is improved so that the displacement of the surface of the earth due to the elastic waves can be measured when the antenna is placed farther from the surface. However, more work is needed to understand the limitations of the beamforming array.

ACKNOWLEDGMENTS

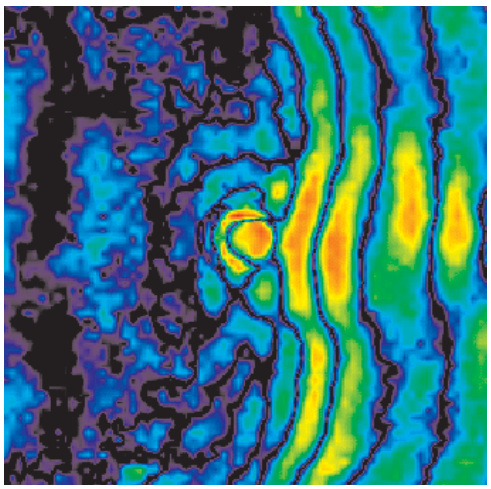
This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0448.

REFERENCES

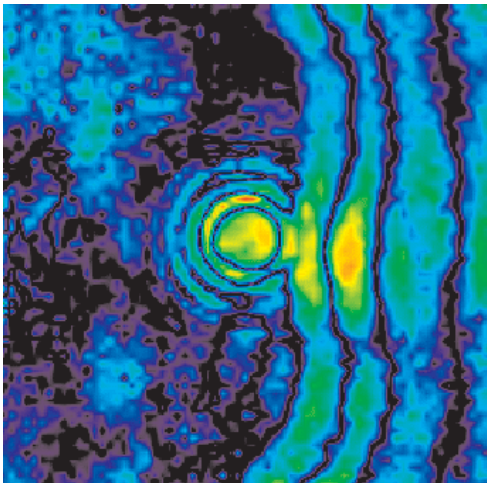
1. W. R. Scott, Jr., C. Schroeder, and J. Martin, "An acousto-electromagnetic sensor for locating land mines," *SPIE, AeroSense, Detection and Remediation Technologies for Mines and Minelike Targets III* , pp. 176–186, April 1998.
2. S. Silver, *Microwave Antenna Theory and Design*, McGraw-Hill, 1949.
3. C. A. Balanis, *Antenna Theory: Analysis and Design*, John Wiley and Sons, 1997.
4. S. H. Lee and W. R. Scott, Jr., "Beamforming array for detecting buried land mines," *SPIE, AeroSense, Detection and Remediation Technologies for Mines and Minelike Targets IV* , pp. 1343–1351, April 1999.
5. W. R. Scott, Jr. and J. Martin, "Experimental investigation of the acousto-electromagnetic sensor for locating land mines," *SPIE, AeroSense, Detection and Remediation Technologies for Mines and Minelike Targets IV* , pp. 204–214, April 1999.



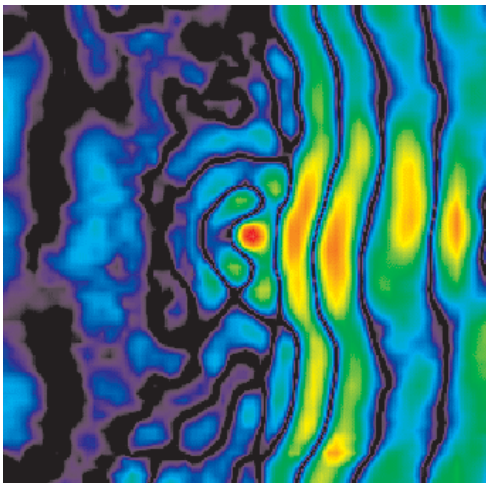
(a)



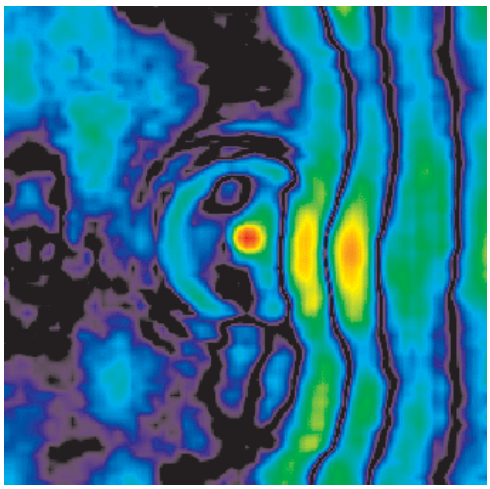
(b)



(c)



(d)



(e)

Figure 8. Pseudo color graphs of the displacement of the surface: (a) raw with h=5 cm, (b) raw with h=20 cm, (c) raw with h=40 cm, (d) beamformed with $\tau = 3.8\lambda$ and h=20 cm, (e) beamformed with $\tau = 7.5\lambda$ and h=40 cm.

Proceedings of SPIE, Vol. 4038, April 2000

Simultaneous Use of Elastic and Electromagnetic Waves for the Detection of Buried Land Mines

Waymond R. Scott, Jr.[a], Gregg D. Larson[b], and James S. Martin[b]

[a]School of Electrical and Computer Engineering
[b]School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

A hybrid technique has been developed that uses both electromagnetic and elastic waves in a synergistic manner to detect buried land mines. The system consists of a moving electromagnetic radar and a stationary elastic-wave source. The source generates elastic waves in the earth. These waves interact with the buried mine and cause both the mine and the earth to be displaced. Because the mechanical properties of the mine are different from those of the earth, the displacements in the region of interaction are distinct from those associated with the free-field propagation of the waves. The radar is used to detect these displacements and, thus, the mine. Initial investigations have demonstrated the feasibility of this scheme under controlled conditions. The current experimental effort is focused on understanding and overcoming the issues associated with using the system in field conditions.

Keywords: land mine, mine detection, acoustic, ground penetrating radar, GPR

1. INTRODUCTION

Seismic/elastic techniques show considerable promise for the reliable detection of all types of buried mines, even low-metal anti-personnel mines. The reason for this is that mines have mechanical properties that are significantly different from soils and typical forms of clutter. For example, the shear wave velocity is approximately 20 times higher in the explosive and the plastics used in typical mines than in the surrounding soil. In addition, mines are complex mechanical structures with a flexible case, a trigger assembly, air pockets, etc. The complex structure gives rise to structural resonances, non-linear interactions, and other phenomenology that is atypical for both naturally occurring and man-made forms of clutter. Thus, this phenomenology can be used to distinguish a mine from clutter.

A system has been developed at Georgia Tech that uses a radar based displacement sensor for the local measurement of seismic displacements without physically contacting the soil surface [1-3]. The non-contact nature of this sensor makes the system capable of interrogating the soil surface near or immediately above a mine. This substantially increases the measurable effects of the mine's presence over schemes which rely on elastic waves scattered by the mine to propagate to a remote sensor location. Figure 1 depicts the present system configuration. The system consists of the electromagnetic radar and the seismic source. The source (an electrodynamic shaker coupled to the ground by a narrow foot) preferentially generates an elastic surface (Rayleigh) wave in the earth. The Rayleigh wave causes both the mine and the surface of the earth to be displaced as it propagates past the mine. Since the amplitude of Rayleigh wave displacements decreases exponentially with depth, only the soil near the surface is interrogated for the presence of mines. The depth of soil which is examined is a function of the frequency of the source. For typical mine depths and sizes, this is in the l00 to 1,000 Hz range. The motion of the mine is different from the surrounding soil, because the elastic properties of the mine are quite different than those of soil. The displacement of the surface of the earth when a mine is present is different than when it is not present because of the local and propagating waves scattered by the mine. The

electromagnetic radar is used to detect these displacements and, thus, the mine. This idea of using elastic and electromagnetic waves synergistically has been proposed previously [4-6], but it has not been seriously investigated until now.

The radar based displacement sensor has been designed and built to measure small vibrations of the soil and the mine. The sensor radiates electromagnetic waves toward the earth. These waves are reflected from the surface of the earth and the mine, where they are amplitude and phase modulated by the transient displacements of the earth and the mine. The reflected waves are received and demodulated. The resulting demodulated signals are proportional to the surface displacement. The radar can measure vibrations as small as 1 nm (10^{-9} m) as currently configured. The end of the wave guide (which functions for both transmit and receive) illuminates an area on the earth's surface comparable to its own cross section (1cm x 2 cm) over which the displacement is integrated.

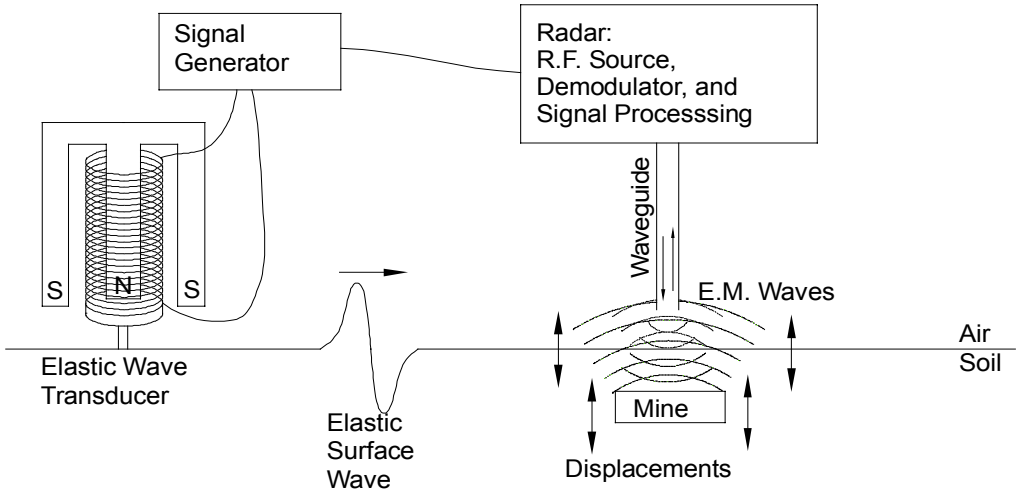


Figure 1: Diagram of the hybrid elastic and electromagnetic mine detection system.

The system is currently being studied in a laboratory scale experimental model. The model, which is depicted in figure 2, consists of a wedge shaped tank filled with over 50 tons of damp compacted sand to simulate soil. The seismic source is located near the tip of the wedge and is bi-directive toward the search area and the back wall. Simulated mines, inert mines, and clutter, such as rocks and sticks, are buried within a 2 m x 2m region in the center of the tank. The radar can be scanned above this region with a three degree of freedom positioner fixed above the tank.

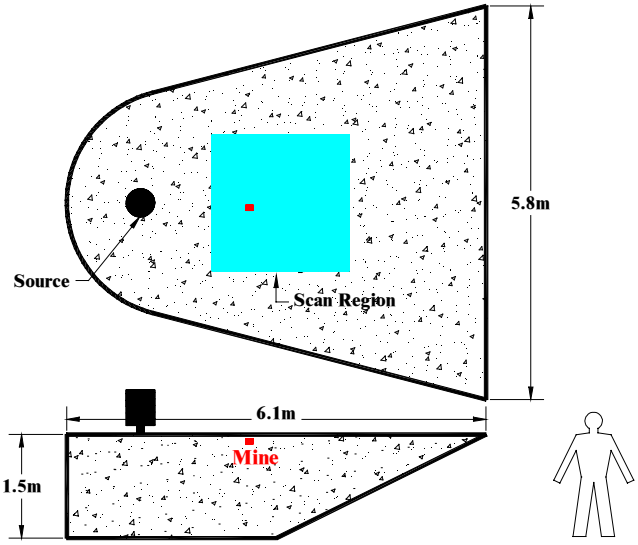


Figure 2: Experimental model for mine detection

Previously presented work with this experimental model and a previous experimental model has demonstrated the feasibility of the elastic / electromagnetic mine detection concept [7]. Current efforts with the model have focused on resolving critical issues for the refinement of the concept into a field operable mine detection system. Recent experimental work has sought to answer the following questions: Does the presence of clutter and other targets unacceptably degrade system performance? Does surface vegetation degrade the radar sensor's performance to make mine imaging impossible? Is the current 100 to 1000 Hz frequency range used in the system sufficient for the expected range of mine burial depths? Can the same system configuration and parameters be used for the detection of both anti-personnel (AP) and anti-tank (AT) mines? Can a source be constructed which can be translated along with the sensor for continuous rather than incremental scanning?

2. EFFECTS OF BURIED CLUTTER AND MULTIPLE TARGETS

The primary detection cues used for all of the AP mine types which have been studied in the experimental model have been resonances of the mine case, trigger mechanism, and overlying soil [6-8]. These are excited by the passage of the Rayleigh wave and characterized by large displacements which persist after the passage of the incident pulse. Although mines exhibiting resonances scatter a larger propagating wave field than similarly sized non-resonant objects, the most pronounced feature of the field scattered by these mines is its mostly local nature. This appears to result from a mostly reactive soil loading. For the mine types studied thus far, the localized resonant motion has been an excellent indicator of a mine's location and extent.

Imaging of mines from multiple surface displacement measurements can be done in many ways. The current imaging scheme involves a multi-step process which filters forward travelling waves (those components directed away from the source) out of the data in the wavenumber domain leaving the reflected waves and a portion of the non-propagating waves. The energy in these remaining waves at times near the time of arrival of the incident wave is assigned to each measurement point forming an image. While there is certainly room for refinement of this algorithm, it does provide an objective method for the generation of mine field images which seem to incorporate the subjective observations of resonant mine behavior. Unlike background subtraction, this algorithm does not rely on information which would be unavailable to a mine detection system operating in the field. It is therefore, a reasonable (but not a certain) candidate for use in a field ready system. Previous work has shown that this technique can be used to image several types of AP mines and distinguish them from non-resonant buried clutter [7,9].

An experiment was performed to address the issue of imaging a minefield containing multiple candidate mine targets, some actual and some false. For this study, the false targets were four mine-sized rocks. The actual targets were four inert AP mines of different types and one radar mine simulant. The 5 mines and 4 rocks were buried at depths of one to three centimeters within a 120 cm by 80 cm search area. The layout of this experiment and the relative scale of the buried objects is shown in figure 3. The goal was to determine whether the presence of the false targets or the scattered fields of the actual targets would obscure the apparent size or location of any of the actual targets in the image which could be formed of the minefield. Also of concern was the potential that multiple scattering would produce ghost images of some of the mines or that the signatures of the most identifiable AP mines (the TS-50 and VS-50) might dictate the dynamic range of the resulting image so as to mask the presence of the least identifiable mines (the M-14 and SIM-9). The image formed of this search area is shown in figure 4. It is clear from the figure that all 5 mine locations have been correctly depicted and that no false images were generated. The image of the VS-50 mine has been elongated in the direction of the butterfly mine which was buried behind it. This is an artifact of the technique used to filter out the forward traveling waves and can be fixed by improving the imaging algorithm. It is obvious from the nature of the elastic / electrodynamic system that false targets could have been selected that would have been problematic for this type of imaging. Rocks were selected because they are naturally occurring and ubiquitous. Resonant false targets are neither and should, therefore, not constitute an unacceptable false alarm rate.

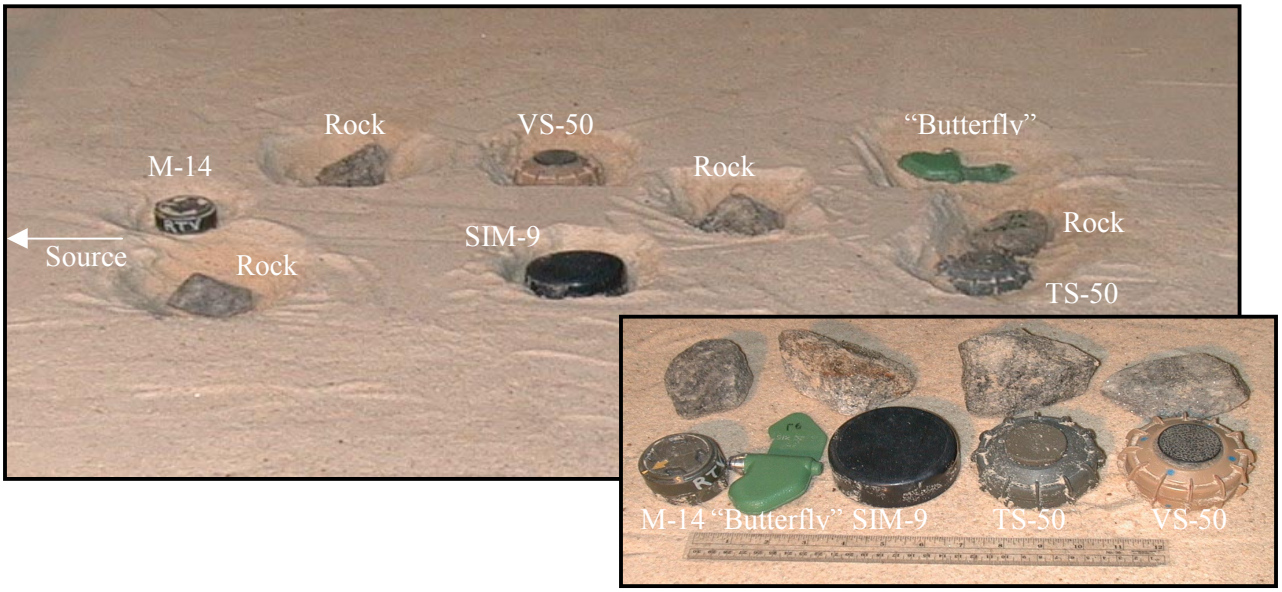


Figure 3: Mines and rocks in buried clutter experiment

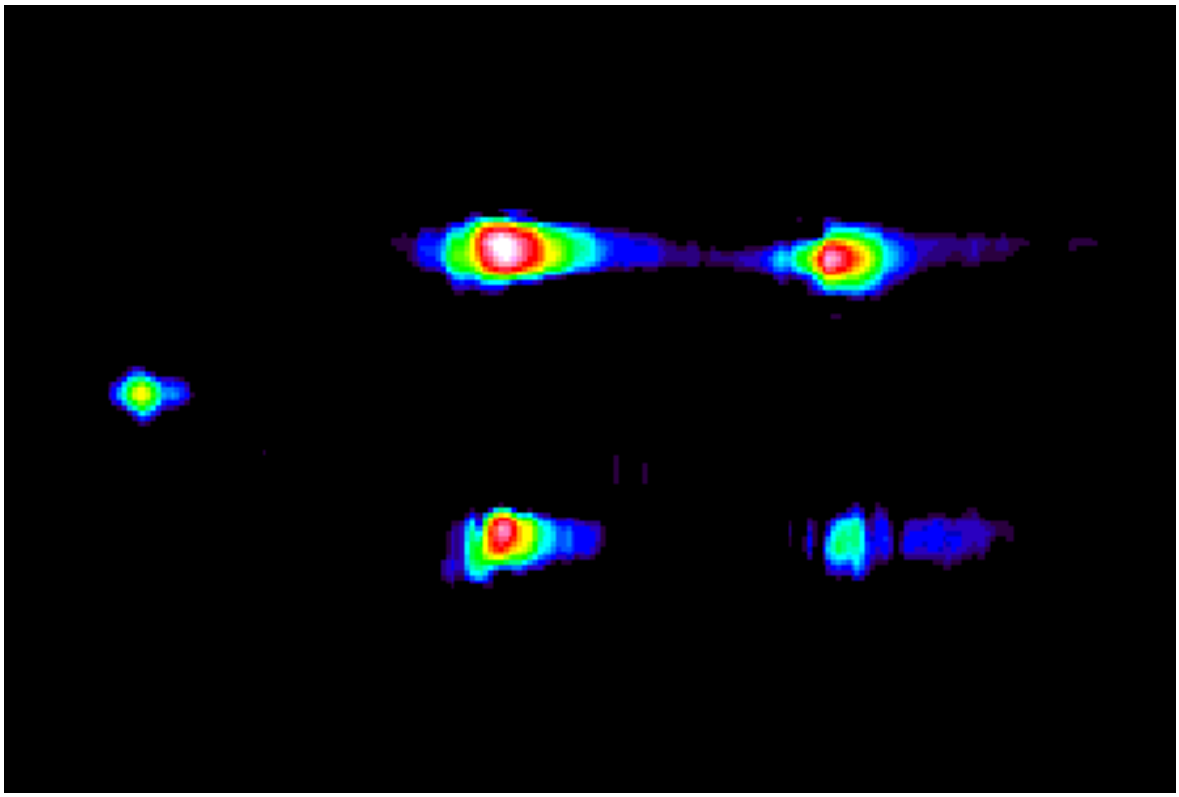


Figure 4: Image formed of minefield in buried clutter experiment

3. TRANSLUCENT SURFACE COVER

In early testing of the elastic / electromagnetic mine detection system, the surface of the soil surrogate used in the experiments was maintained level, smooth, and bare. This provided the radar sensor with a seismically modulated EM reflection from the surface that was nearly optimal (large and uncorrupted). In the field, the surface will be rough and lie under some type of ground cover. The ability of the system sensor to see through common surface cover such as grass or light vegetation is essential for practical system operation. Some types of ground cover, obviously, will be opaque to the radar's interrogation signal. These will require special consideration and possibly an alternate sensor design. Standing water is an example of such a problematic case. In general, vegetation is likely to be translucent to the radar sensor with a portion of the EM reflection originating in the vegetation and the remainder coming from the surface below. The signal which is reflected from the vegetation can corrupt the total return in two ways. First, it will reduce the electromagnetic signal component reflected from the surface. This will reduce the level of the measured displacement signal. Second, any motion of the vegetation will produce an additional modulation of the carrier unrelated to the motion of the underlying surface. This will increase the effective noise floor. To test the ability to penetrate surface cover, pine straw was selected both for the convenience of its application and because it is a commonly occurring ground covering. In the experiment, a 2.3 cm layer of pine straw was spread over the surface beneath which a TS-50 AP mine was buried along with 4 mine sized rocks and two sticks. The layout of this experiment can be seen in figure 5.

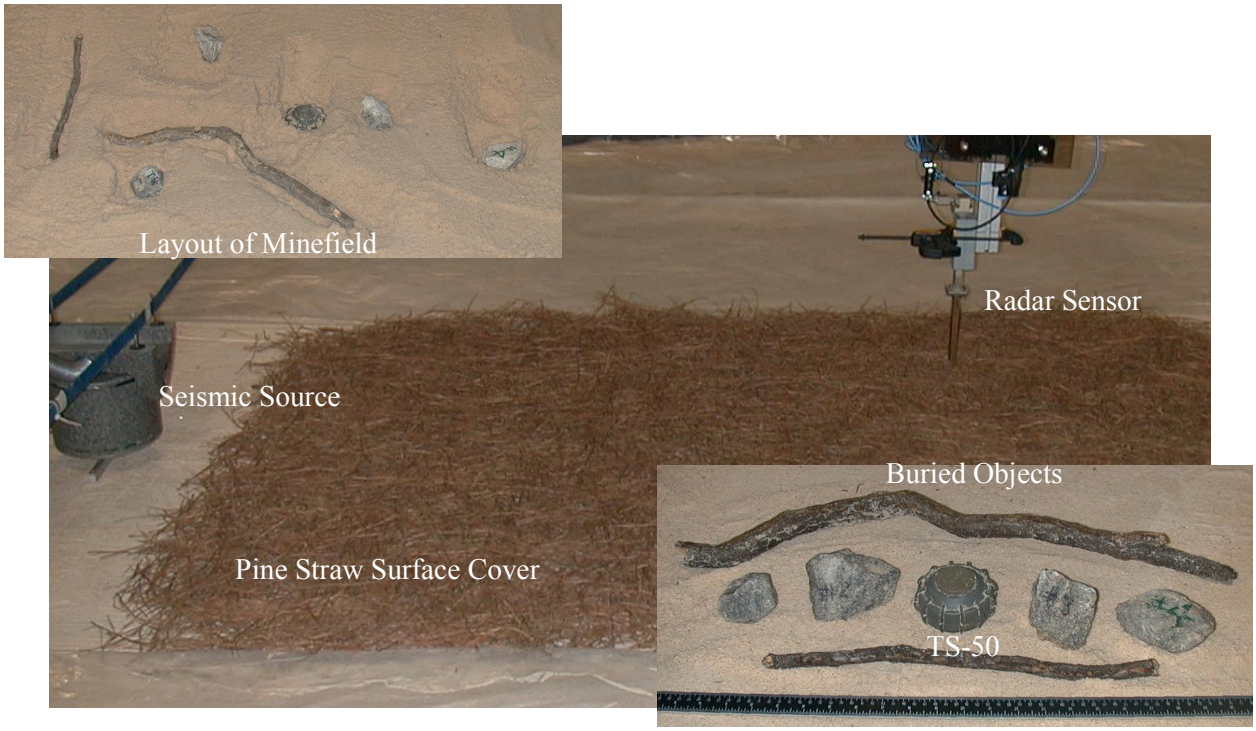


Figure 5: Surface Vegetation Experimental Setup

Figure 6 shows the image formed from the pine straw covered surface. The location and extent of the TS50 mine are apparent. There is less contrast in this image than for similar images formed in the absence of surface covering. An examination of the time domain signals which contribute to the image reveals that the dominant effect of the pine straw was to force an increase in the seperation between the antenna and the ground's surface which resulted in a reduced signal level and a loss of some spatial resolution. When this is corrected for, the only direct effect of the surface cover that can be observed is a slight increase in the spatial noise floor (signal variability with range). This is attributable to variations in the received EM surface reflection which is reduced by up to 6 dB when the pine straw is present.

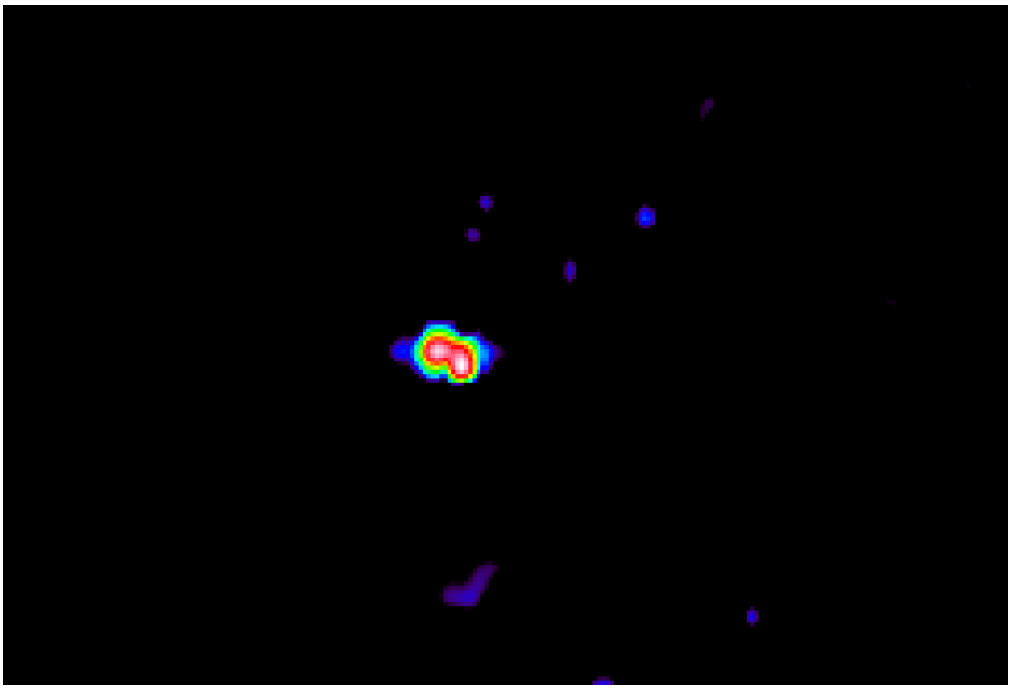


Figure 6: Image Formed of Pine Straw Covered Mine Field

4. MINE SIZE AND DEPTH RANGE OF DETECTABILITY

There are two ranges of mine size and burial depth which are of concern for the mine detection problem. The first of these is for AP mines which are typically a few hundred milliliters in volume and buried such that the mine trigger is a few centimeters below the soil surface. AP mines were the focus of most of the early work on the elastic/ electromagnetic mine detection system because they are usually perceived as a more difficult problem. The efficacy of a system like this for AP mine detection does not, however, automatically imply success in the detection of AT mines. These mines are typically several liters in volume and have burial depths up to a few decimeters. Since Rayleigh waves penetrate the surface to a depth proportional to their wavelength it is likely that lower frequencies would be required for the detection of AT mines. Since the detection of AP mines was greatly facilitated by resonances of the mines, it is unlikely that the AT mine problem is simply a direct scaling of the AP mine imaging technique. This would imply a system bandwidth down to 10 or 20 Hz which is not possible with the current experimental model, but should be possible in a system that operates in the field. Currently, the operating band of the system (100 – 1,000 Hz) is dictated by source response, ambient noise, and reverberation in the experimental model at low frequencies. Attenuation in the soil surrogate limits the bandwidth at high frequencies.

Two types of inert AT mines were tested: VS-1.6 and VS-2.2. These mines are shown in figure 7 along with a TS-50 AP mine for size comparison. Both AT mines have plastic cases and spring loaded piston triggers. The triggers are considerably stiffer than those on the AP mines which were previously tested. The results for both mines were quite similar. Figures 8 and 9 show seismograms generated by one-dimensional scans over the two AT mines as a function of mine burial depth. The data was processed in a slightly different way from the AP mine scans in that the pulse shape which was used had a center frequency that was an octave lower (225 Hz as opposed to 450 Hz for the AP mine scans). The experimental data used to reconstruct either pulse response are identical. The difference is in the post processing. Both processing techniques could be used simultaneously. The pulse with the lower center frequency was chosen for the AT mine in order to emphasize the effects of the spectral components which penetrated the soil to a sufficient depth to interact with the mine.

From data in figures 8 and 9 the presence of the mine is apparent up to a depth of about 10 cm. Above the mine, an amplification of the incident can be observed at 2, 4, and 6 cm depths. Propagating waves that have been reflected from the mines are clearly apparent at 4 and 6 cm depths. The strength of the reflected field at these depths appears to be inversely related to the duration of the localized motion above the mine which indicates that the radiation of Rayleigh waves may constitute a significant source of damping for the mine motion. There is also evidence of dispersion in the soil layer above the mine: the leading edge of the incident signal is clearly delayed as it passes above the mine and the pulse shape and arrival time well beyond the mine are quite different than they are at the same location for the no mine case.

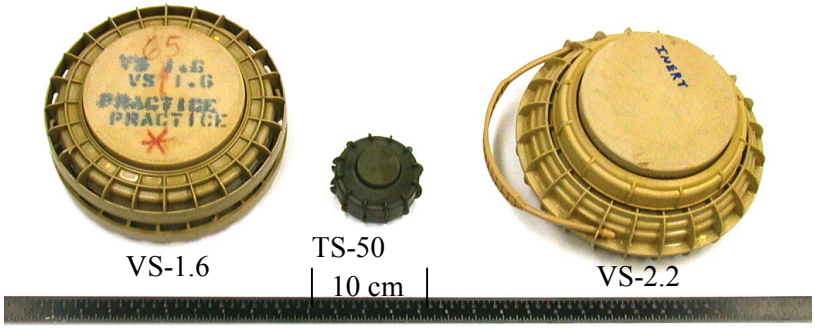


Figure 7: Size Comparison of AT Mines and AP Mine

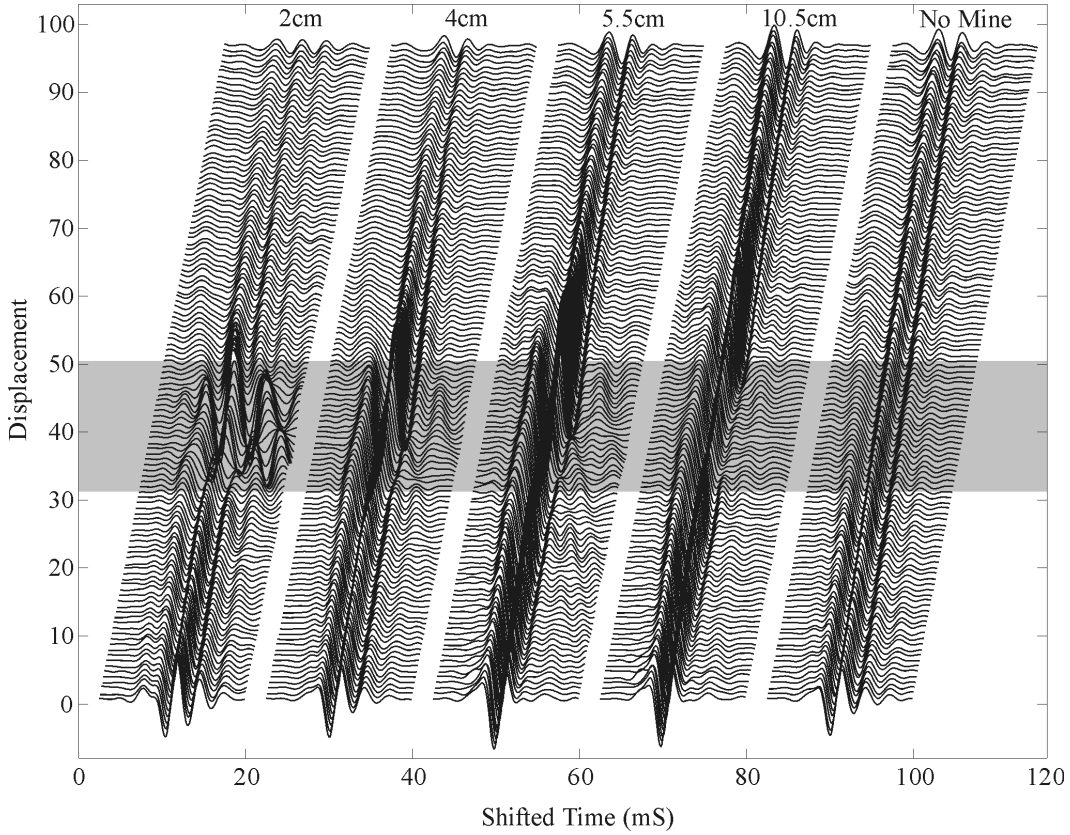


Figure 8: 1-D Scans of VS 2.2 AT Mine at Various Depths

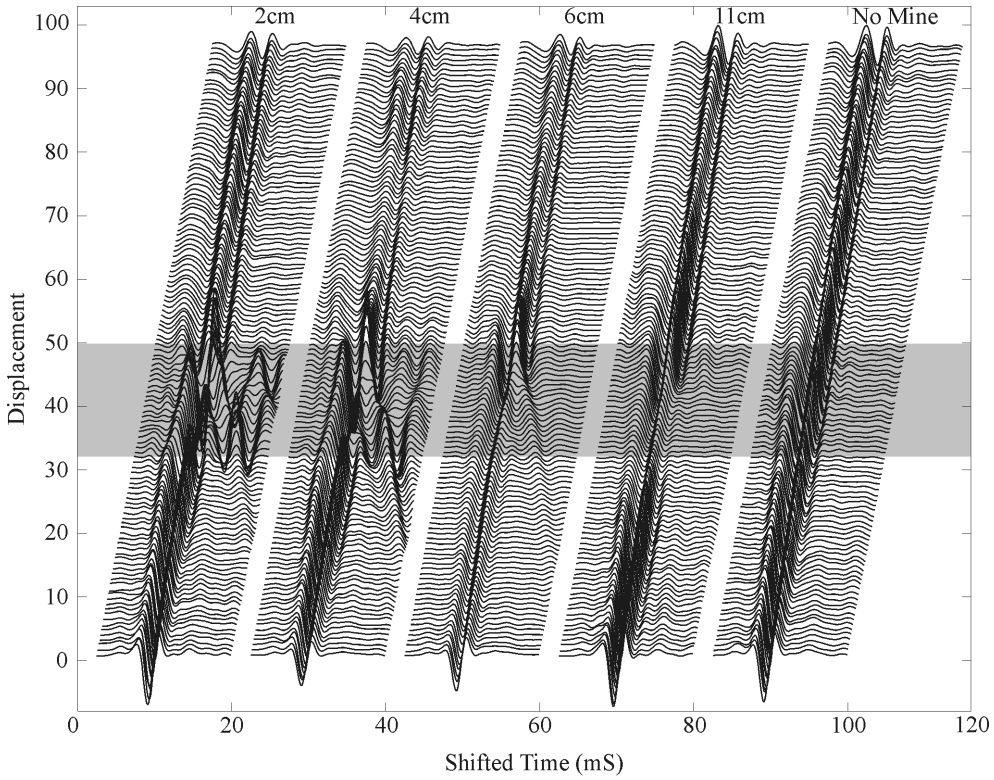


Figure 9: 1-D Scans of VS 1.6 AT Mine at Various Depths

In the seismograms of figure 8 there is very little sign of the AT mine at the 11 cm burial depth. Its presence is, however, much more apparent if the forward travelling waves are removed from the seismograms. This is because the forward travelling waves are the dominant component of the receive signal but are only slightly effected by the presence of the mine. In contrast, the smaller reverse propagating and standing waves are related almost entirely to the mine and other scatters in the search area. The seismograms of figures 10 and 11 demonstrate this. These have been generated by filtering the forward propagating components out of the data depicted in figures 8 and 9. A gain factor, which is proportional to mine depth, has been applied to force the traces onto comparable vertical scales. In the imaging process, where target depth is unknown, a logarithmic scale is used for this purpose. Where no mine was present the gain factor for the deepest corresponding burial was used to demonstrate the effective noise floor of this technique. These figures demonstrate that, using the current system and model, an AT mine could be imaged to a depth of 10 or 11 centimeters without substantial computational effort. This is, by no means a limit on the detection technique. The spectral content of the incident and the processing technique could be improved to extend this further. Acausal artifacts of the filtering operation which elongate the mine signature can be seen in the filtered seismograms. This is clearly one area in which the algorithm can be improved. The artifacts appear as wavefronts that originate at the mine but arrive instantaneously in the region behind the mine. Similar artifacts in the direction of the source are obscured by the Rayleigh wave reflected by the mine. Conceptually these should be precluded by the processing algorithm but the arise because of the spatial windowing of the scan region and the abrupt edge of the filtering function in the wavenumber domain.

In order to test the feasibility of imaging a deeply buried AT mine a 2-dimensional scan was performed with the VS-1.6 buried 11 cm deep. The imaging technique was identical to that used for the AP mines, but incorporated the lower center frequency pulse. The image which was formed is depicted in figure 12. The location and extent of the mine are apparent in this image. There is considerably more smearing of this image than for those formed of the AP mines. This is due, in part, to processing artifacts and, in part, to relatively large propagating wave reflections associated with the AT mine.

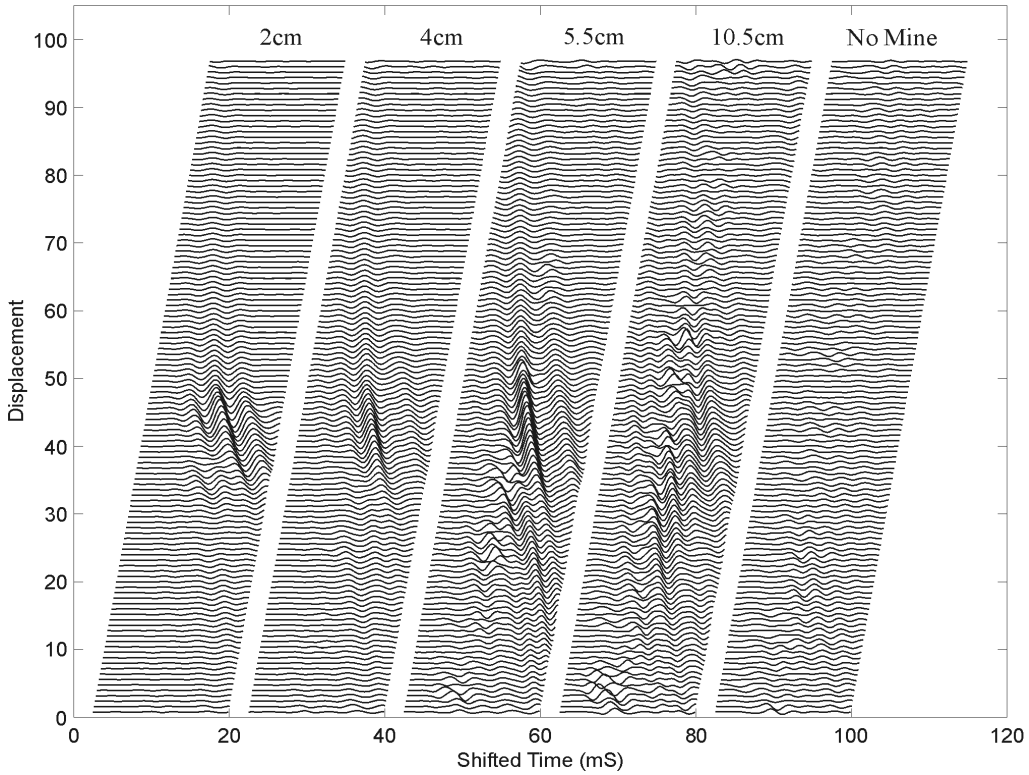


Figure 10: 1-D Scans over VS-2.2 AT Mine with Forward Waves Filtered Out

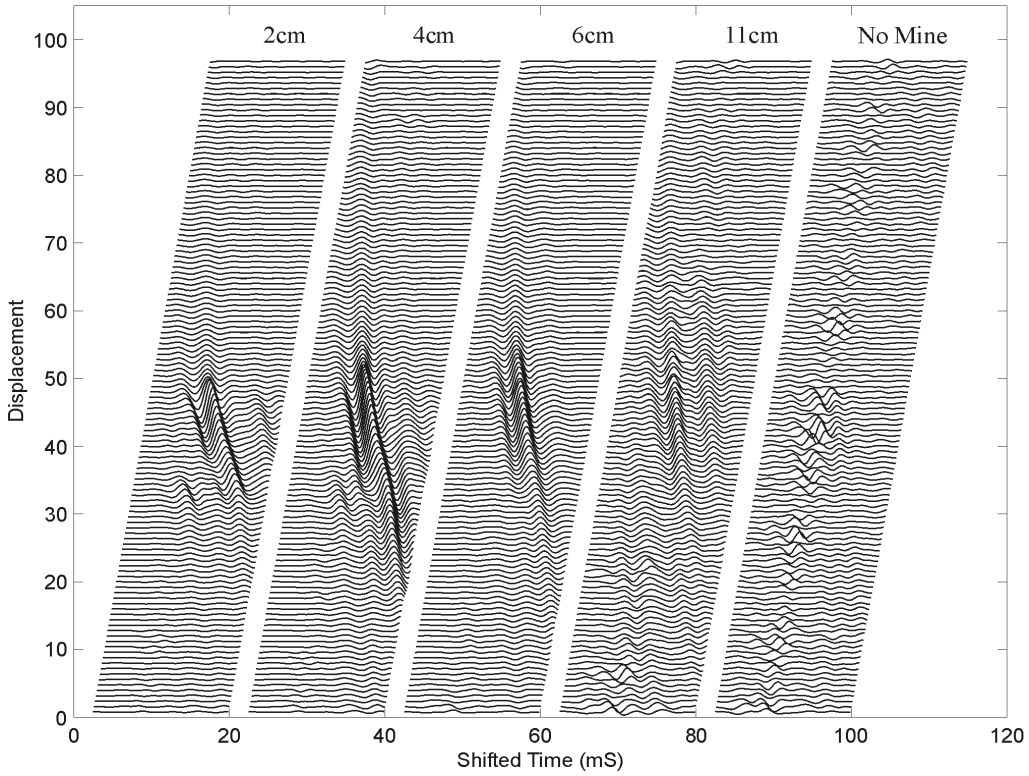


Figure 11: 1-D Scans over VS-1.6 AT Mine with Forward Waves Filtered Out

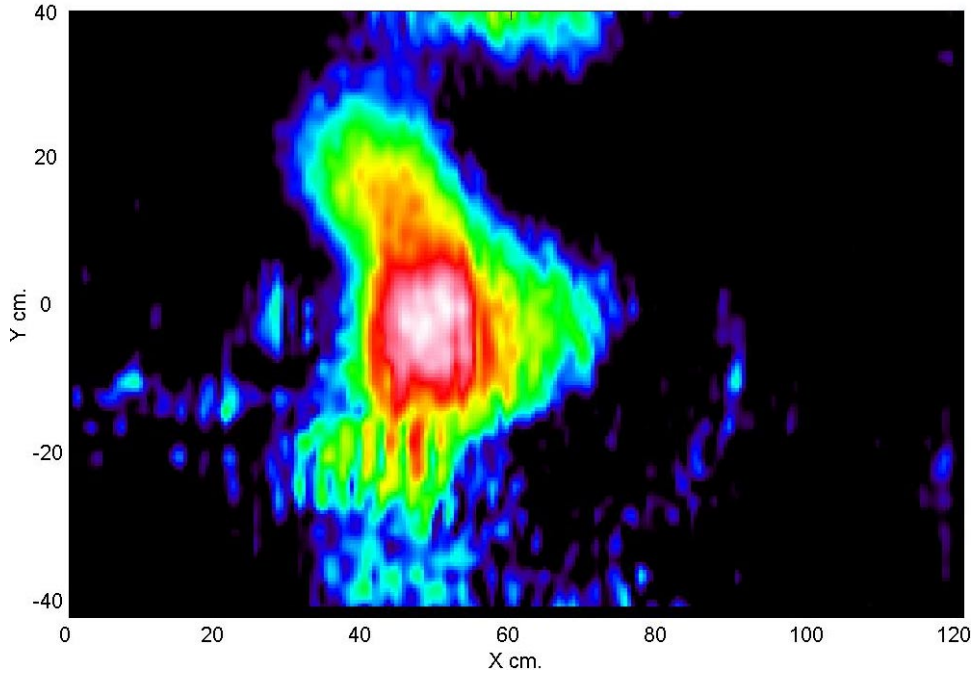


Figure 12: Image of VS-1.6 AT Mine Buried 11 cm Deep

5. SCANNABLE SEISMIC WAVE SOURCE

The electrodynamic shaker, which is currently being used as a seismic source, was selected for its compatibility with the experimental model. It provides a large broadband seismic excitation and does not couple energy into the sensor by any path other than the propagation of elastic waves through the ground. In this regard, it has been extremely well suited to the design of system components and the evaluation of system capabilities. It may, however, not be the best source design for a system which will operate in the field.

For a field operable mine detection system the seismic source needs to be relocated in either discrete increments or continuously during the search process. A reproducible source to surface coupling at each transmitting location is a desirable feature of such a source because it obviates the need to compute an additional transfer function in order to assemble a minefield image. The current seismic source is difficult and time consuming to move. It also does not generate a surface wave that is entirely reproducible when the source relocated. This has been observed even in the benign environment of the laboratory experimental model. The shaker design could be modified for use in a field operable system if the source were coupled to the ground through a rolling point of contact. This design has yet to be explored experimentally, but it is by no means the only option. Several alternative source designs have been proposed. The simplest of these are currently being tested in the experimental model.

Other authors have noticed that it is possible to couple roughly a thousand times more energy into soils from airborne sound than the large density mismatch at the air soil interface would indicate. This is, in part, because the elastic wave speeds in typical soils are considerably lower than the sound speed in air. Although, this observation seems promising for the system under development, an unfortunate consequence of the low propagation speeds is that soils excited by airborne sound respond in a locally reactive way to the excitation [10]. This, by implication, precludes the effective generation of Rayleigh waves (which propagate only along the soil surface) with airborne sound. In order to overcome this apparent limitation a surface excitation is required with substantial non-radiating wavenumber components in the air above the soil. This can be accomplished if the soil surface is in the near field of an aero-acoustic source.

The experimental aero-acoustic seismic source is a 10” moving coil loudspeaker which is suspended 1 to 3cm above the surface of the sand. Initial testing indicates that the speaker must be well within a seismic wavelength of the surface to effectively couple energy into the Rayleigh wave. At large distances from the surface there is no angle of incidence at which the airborne wave can match the phase velocity of the much slower Rayleigh wave. Measurable displacements can still be excited at these large separations, but they are due almost entirely to energy coupling into downward directed compressional waves. The trace velocity measured for the pulse along the surface under these circumstances is exactly the propagation speed in air. Figure 10 shows a seismogram of arrivals generated by translating the aero-acoustic source at a height 1 cm above the soil surface with respect to a fixed accelerometer receiver for a 1m transit. The spectrum of the received signals is comparable to that obtained using the surface-coupled shaker source although colored toward the lower frequencies. The displacements are within the measurable range of the existing radar based sensor and it is not unreasonable to believe that they could be made comparable to the current shaker source had a more powerful aero-acoustic source been used. Generally, the aero-acoustic source is more effective at lower frequencies because of the relatively large area beneath the speaker to which a uniform excitation can be applied. An accelerometer was used for these measurements rather than the radar sensor because it provides an absolute measure of displacement. Backside motion introduced by either mechanical or airborne acoustic coupling directly into the radar sensor is problematic for this sort of measurement and will be addressed in future work.

In addition to the two source designs mentioned above many other seismic sources are under consideration for the field operable system. These include thermoelastic laser and microwave sources, air and water jets, electrical arc sources, and ultrasonic parametric sources. Each of these options have intrinsic tradeoffs when compared with other designs. The laser, for example, would allow for very large standoff distances from the earth’s surface but would be substantially more expensive and less powerful than the sources considered here.

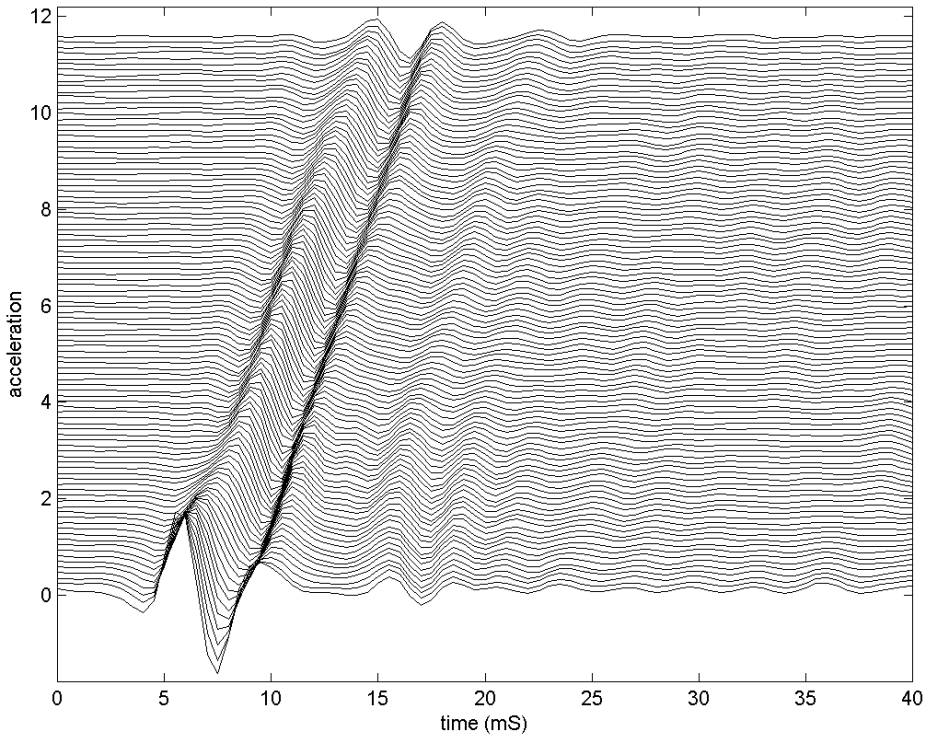


Figure 13: Rayleigh Wave Generated with an Aero-Acoustic Source

6. CONCLUSIONS

The elastic / electromagnetic land mine detection technique has been tested under laboratory conditions which mimic a variety of realistic mine detection scenarios. None of these conditions appear to be significant hurdles for the development of a field operable mine detection system.

A prototype mine detection system has been demonstrated which is capable of detecting inert AP and AT mines and distinguishing them from buried clutter. The efficacy of the system has been demonstrated over a range of anticipated burial depths with both bare ground and ground covered with thin vegetation. It has been shown that multiple mine targets in close proximity do not pose a significant imaging problem. A Rayleigh wave source has been successfully tested which does not require direct ground contact and can be scanned over the surface of the ground in field operations. Integration of the new source into the detection system, evaluation of alternative source and sensor designs, testing additional mine types, and addressing problematic operating environments will be the focus of future work on the development of the system.

7. ACKNOWLEDGEMENTS

This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0448

8. REFERENCES

1. Scott, W.R., Jr., Schroeder, C., and Martin, J.S., “An Acousto-Electromagnetic Sensor for Locating Land Mines,” SPIE, AeroSense, *Detection and Remediation Technologies for Mines and Minelike Targets III*, Orlando, FL, pp. 176-186, April 1998
2. Scott, W.R., Jr. and Martin, J.S., “An Experimental Model of an Acousto-Electromagnetic Sensor for Detecting Land Mines,” *Proceedings of the 1998 IEEE Antennas and Propagation Symposium,* Atlanta, GA, pp. 978-83, June 1998.
3. Scott, W.R., Jr., Schroeder, C., and Martin, J.S., “A Hybrid Acoustic/Electromagnetic Technique for Locating Land Mines,” *Proceedings of the 1998 International Geoscience and Remote Sensing Symposium*, Seattle, Washington, July 1998
4. C. Stewart, "Summary of Mine Detection Research," *U.S. Army Technical Report* 1636-TR, Vol. 1, May 1960.
5. G.S. Smith, "Summary Report: Workshop on New Directions for Electromagnetic Detection of Non-Metallic Mines," *Report for U.S. Army BRDEC and ARO*, June 1992
6. C. Stewart, Summary of Mine Detection Research, Vol. I, pp. 172-179, *Tech. Report* 1636-TR, May 1960, U.S. Army Engineering Res. and Devel. Labs, Corps. of Eng., Belvoir, VA.
7. Scott W.R., Jr. and Martin J.S., “Experimental Investigation of the Acousto-Electromagnetic Sensor for Locating Land Mines,” *Proceedings of the SPIE: 1999 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, Vol. 3710, pg. 204-14, April 1999.
8. Scott, W.R., Jr., Schroeder, C.T., and Martin, J.S., “An Acousto-Electromagnetic Method for Detecting Buried Objects,” *Proceedings of the XXVIth General Assembly of the International Union of Radio Science*, Toronto, Canada, pg. 724, August 1999
9. Behboodian, A., Scott, W.R., Jr. and McClellan, J.H. "Signal Processing of Elastic Surface Waves for Localizing Buried Land Mines," *Proceedings of the 33rd Assilomar Conference on Signals, Systems, and Computers*, Assilomar, CA, October 1999
10. Sabatier, J.M. et.al., “Acoustically Induced Seismic Waves” , Journal of the Acoustical Society of American Vol. 80 #2, pg. 646-9, August 1986

Proceedings of SPIE, Vol. 4038, April 2000

Three-Dimensional FDTD Model to Study the Elastic-Wave Interaction with Buried Land Mines

Christoph T. Schröder and Waymond R. Scott, Jr.

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250
USA

ABSTRACT

A three-dimensional finite-difference model for elastic waves in the ground has been developed and implemented. The model is used to investigate the interaction of elastic waves with buried land mines. When elastic waves interact with a buried mine, a strong resonance occurs at the mine location. The resonance can be used to enhance the mine's signature and to distinguish the mine from clutter. Results are presented for a single mine buried in the ground and several mines in the presence of clutter. The predictions of the numerical model are in fairly good agreement with experimental results.

Keywords: land mine detection, elastic, acoustic, FDTD, finite-difference

1. INTRODUCTION

A new technique is being investigated at the Georgia Institute of Technology, in which buried land mines are located by using both elastic (acoustic) and electromagnetic waves in a synergistic manner.[1–3] Here, elastic waves interact with a buried land mine and cause the mine and the surface above the mine to vibrate. An electromagnetic radar records the vibrations and, thus, detects the mine. During the process of developing the elastic/electromagnetic sensor, a major part has been the implementation of a numerical model which simulates the interaction of the elastic waves with the buried land mines.[4]

The numerical model is based on the finite-difference time-domain (FDTD) method. The equation of motion and the stress-strain relation, together with a constitutive relation, form a set of first-order partial differential equations that completely describes the elastic wave motion in a medium. Introducing finite differences, this set of equations can be discretized and adapted to the finite-difference time-domain modeling scheme. Assuming that the field is known at one initial time t_0, this numerical scheme is used to determine the field values at later times $t > t_0$.

The finite-difference model has been implemented in two and three dimensions. The solution space is discretized and a staggered finite-difference grid is introduced. The grid is surrounded by a perfectly matched layer, that absorbs the outward traveling waves and, thus, models the infinite extend of the solution space. The finite-difference model has been implemented in a fully parallel fashion. The computations for this paper's results have been performed both on a Cray T3E parallel supercomputer located at the ERDC Massively Shared Resource Center in Vicksburg, Mississppi, and on a Beowulf PC cluster located at the Georgia Institute of Technology. The Beowulf cluster has been developed and built especially for the model described in this paper.

The numerical model has been very helpful for explaining and understanding the experimental results. The numerical model bears several advantages over the experimental model. With the numerical model, it is possible to visualize the elastic wave motion within the ground, whereas with the experimental model only the wave fields on the surface can be observed. Furthermore, no noise is present in the numerical model, and material parameters can be adjusted easily.

Electronic Mail and Telephone: christoph.schroeder@ee.gatech.edu, 404-894-3123, waymond.scott@ee.gatech.edu, 404-894-3048

Results obtained with the three-dimensional model are presented in this paper. In Sec. 3.1, the propagation of a Rayleigh surface wave is shown as measured experimentally and computed numerically. The surface wave is seen to disperse as it propagates along the surface. The same effect can be also modeled numerically. In Sec. 3.2, the interaction of elastic waves with a buried anti-personnel mine is shown. In both experiment and numerical model, a strong resonance is observed at the mine location. Finally, in Sec. 3.3, the interaction of elastic waves with several buried mines in the presence of clutter is described.

2. THREE-DIMENSIONAL NUMERICAL MODEL

In a realistic setting, one or several mines are buried in the ground, surrounded and covered by various kinds of clutter. To approximate these conditions, experiments have been performed with mines buried in a large sand box.[1] In these experiments, elastic waves are launched by an electrodynamic transducer placed on the surface of the ground. The waves propagate along the surface and interact with the buried land mines. To study these mine-wave interactions, a three-dimensional finite-difference model has been developed.

2.1. Finite-Difference Model

Figure 1 shows the three-dimensional finite-difference model. To reasonably simplify the model, the ground is assumed to be linear, isotropic and lossless. The surface of the ground is modeled as a free-surface, a *Perfectly Matched Layer* terminates the solution space at the remaining grid edges and absorbs the outward traveling waves. The solution space is discretized using a staggered finite-difference grid.[5]

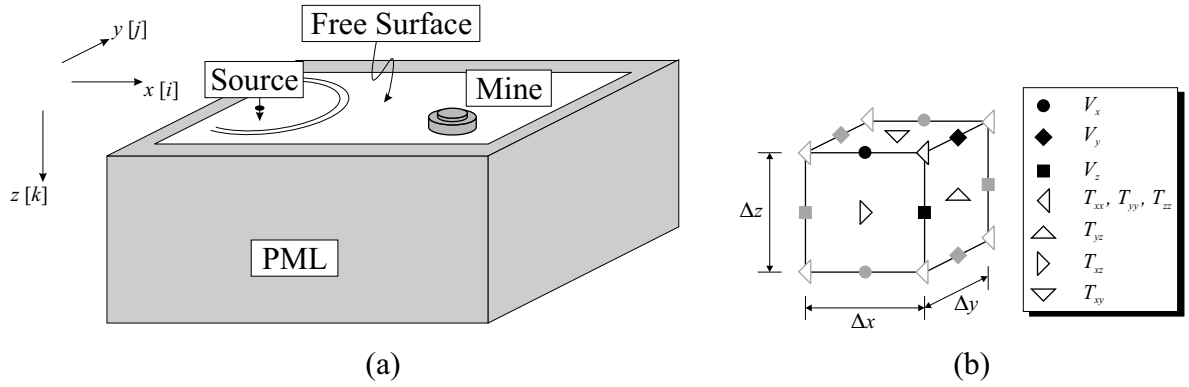


Figure 1. Three-dimensional finite-difference model; (a) lay-out, (b) finite-difference cell.

The elastic wave motion in solids is described by a set of fundamental partial-differential equations: the equation of motion relating the particle velocity vector and the mechanic stress tensor, the strain-velocity relation and the elastic constitutive relation. Combining these equations, a first-order system of equations is obtained describing the elastic wave fields entirely in terms of the particle velocity and the mechanic stress. In three dimensions, three unknown velocity components and six unknown stress tensor components arise.

For the numerical finite-difference model, the derivatives of the partial differential equations are approximated by finite-differences. The finite-difference algorithm shall be explained by deriving the update equations for the x-component of the particle velocity, v_x, (from the first component of the equation of motion) and for the longitudinal stress component, τ_{xx} (from the combined strain-velocity and constitutive relation). Update equations for the other field components can be derived in a similar manner.

The first-order system of equations consists of nine linear independent equations for the three unknown velocity vector components v_x, v_y and v_z, and the six stress tensor components τ_{xx}, τ_{yy}, τ_{zz}, τ_{yz}, τ_{xz} and τ_{xy}. The equations can be written out as

$$\rho \frac{\partial v_x}{\partial t} = \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} \tag{1}$$

$$\rho \frac{\partial v_y}{\partial t} = \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} \tag{2}$$

$$\rho\frac{\partial v_z}{\partial t} = \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{xz}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} \tag{3}$$

$$\frac{\partial \tau_{xx}}{\partial t} = (\lambda + 2\mu)\frac{\partial v_x}{\partial x} + \lambda\frac{\partial v_y}{\partial y} + \lambda\frac{\partial v_z}{\partial z} \tag{4}$$

$$\frac{\partial \tau_{yy}}{\partial t} = \lambda\frac{\partial v_x}{\partial x} + (\lambda + 2\mu)\frac{\partial v_y}{\partial y} + \lambda\frac{\partial v_z}{\partial z} \tag{5}$$

$$\frac{\partial \tau_{zz}}{\partial t} = \lambda\frac{\partial v_x}{\partial x} + \lambda\frac{\partial v_y}{\partial y} + (\lambda + 2\mu)\frac{\partial v_z}{\partial z} \tag{6}$$

$$\frac{\partial \tau_{yz}}{\partial t} = \mu(\frac{\partial v_y}{\partial z} + \frac{\partial v_z}{\partial y}) \tag{7}$$

$$\frac{\partial \tau_{xz}}{\partial t} = \mu(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}) \tag{8}$$

$$\frac{\partial \tau_{xy}}{\partial t} = \mu(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}), \tag{9}$$

where λ and μ are *Lame's constants* and describe the isotropic solid, and ρ is the material density of the medium.

Taking Eq. (1) and Eq. (4), the update equations for the particle velocity v_x and the longitudinal stress τ_{xx} are derived. Introducing finite-differences in space and time, the partial-differential equations are discretized:

$$\begin{aligned}
&\rho\frac{V_x^{l+0.5}(i,j-0.5,k-0.5) - V_x^{l-0.5}(i,j-0.5,k-0.5)}{\Delta t} = \\
&\quad\frac{T_{xx}^{l}(i+0.5,j-0.5,k-0.5) - T_{xx}^{l}(i-0.5,j-0.5,k-0.5)}{\Delta x} + \frac{T_{xy}^{l}(i,j,k-0.5) - T_{xx}^{l}(i,j-1,k-0.5)}{\Delta y} \\
&\quad+ \frac{T_{xz}^{l}(i,j-0.5,k) - T_{xz}^{l}(i,j-0.5,k-1)}{\Delta z},
\end{aligned} \tag{10}$$

$$\begin{aligned}
&\frac{T_{xx}^{l+1}(i+0.5,j-0.5,k-0.5) - T_{xx}^{l}(i+0.5,j-0.5,k-0.5)}{\Delta t} = \\
&\quad(\lambda+2\mu)\frac{V_x^{l+0.5}(i+1,j-0.5,k-0.5) - V_x^{l+0.5}(i,j-0.5,k-0.5)}{\Delta x} \\
&\quad+\lambda\frac{V_y^{l+0.5}(i+0.5,j,k-0.5) - V_y^{l+0.5}(i+0.5,j-1,k-0.5)}{\Delta y} \\
&\quad+\lambda\frac{V_z^{l+0.5}(i+0.5,j-0.5,k) - V_z^{l+0.5}(i+0.5,j-0.5,k-1)}{\Delta z}.
\end{aligned} \tag{11}$$

Here, the capital letters mark the numerical value of the correspondent field component at a discrete location in space and time. For example, $V_x^{l+0.5}(i,j-0.5,k-0.5)$ stands for the numerical value of the particle velocity v_x at $(x,y,z) = (i\Delta x,(j-0.5)\Delta y,(k-0.5)\Delta z)$ at time $t = (l+0.5)\Delta t$. Knowing $V_x^{l-0.5}$, T_{xx}^{l},T_{xy}^{l} and T_{xz}^{l}, Eq. (10) can be solved for $V_x^{l+0.5}$, i.e. for the particle velocity at the incremented time $t = (l+0.5)\Delta t$:

$$\begin{aligned}
&V_x^{l+0.5}(i,j-0.5,k-0.5) = V_x^{l-0.5}(i,j-0.5,k-0.5)+ \\
&\quad+\frac{\Delta t}{\rho}\left[\frac{T_{xx}^{l}(i+0.5,j-0.5,k-0.5) - T_{xx}^{l}(i-0.5,j-0.5,k-0.5)}{\Delta x} + \frac{T_{xy}^{l}(i,j,k-0.5) - T_{xx}^{l}(i,j-1,k-0.5)}{\Delta y}\right. \\
&\quad\left.+\frac{T_{xz}^{l}(i,j-0.5,k) - T_{xz}^{l}(i,j-0.5,k-1)}{\Delta z}\right].
\end{aligned} \tag{12}$$

Similarly, T_{xx}^{l+1} is obtained from Eq. (11):

$$\begin{aligned}
&T_{xx}^{l+1}(i+0.5,j-0.5,k-0.5) = T_{xx}^{l}(i+0.5,j-0.5,k-0.5)+ \\
&\quad+\Delta t\left[(\lambda+2\mu)\frac{V_x^{l+0.5}(i+1,j-0.5,k-0.5) - V_x^{l+0.5}(i,j-0.5,k-0.5)}{\Delta x}\right. \\
&\quad\left.+\lambda\frac{V_y^{l+0.5}(i+0.5,j,k-0.5) - V_y^{l+0.5}(i+0.5,j-1,k-0.5)}{\Delta y}\right.
\end{aligned}$$

$$+ \lambda \frac{V_z^{l+0.5}(i+0.5, j-0.5, k) - V_z^{l+0.5}(i+0.5, j-0.5, k-1)}{\Delta z} \Bigg] . \tag{13}$$

In the same manner, discrete update equations can be obtained for all field components.[6]

The discretization leads to the characteristic finite-difference grid. In this grid, the field components are staggered in space and time. The finite-difference grid can be thought of as being comprised of basis cells. The three-dimensional basis cell for the elastodynamic case resembles strongly the three-dimensional basis cell for electromagnetic finite-difference modeling, the so-called *Yee*-cell.[7] However, due to the stress being a tensor, more field components are present in the elastodynamic case. Figure 1 (b) shows the three-dimensional finite-difference basis cell. The position of the cell in space is labeled with i in the x-direction, with j in the y-direction and with k in the z-direction. Only the field components in black are assigned to the (i, j, k)-th cell. The other field components belong to adjacent cells and are gray in Fig. 1 (b). Note that the field components are not known at the same points in space and time. The grid is laid out such that each field component is surrounded by the field components it is dependent on. The finite-difference algorithm then works as follows. Knowing the field components at one initial time t_0 throughout the entire grid, the field values for later times can be determined. First, the velocity components on the grid are calculated at the incremented time $t_0 + 0.5\Delta t$ using Eq. (10) for v_x and equivalent equations for v_y and v_z. The stress components at $t_0 + \Delta t$ are determined from the velocity components using Eq. (11) and its equivalences. Then, the velocity components are updated using the stress values, the stress components are computed from the velocity components and so on. In this way, the field values can be determined up to any desired time.

When implementing the finite-difference scheme, boundary conditions have to be treated in a special manner. Three different kinds of boundaries arise: the source, internal boundaries (i.e. boundaries within the medium marked by a change in material properties), and external boundaries (i.e. the grid edges).

In the experimental model, an electrodynamic transducer placed on the surface launches the elastic waves. The transducer foot has the shape of a bar. In the numerical model, the transducer is approximated by exciting the particle velocity component normal to the surface, v_z, on an area equivalent to the area of the foot. The motion of the transducer foot has been measured using accelerometers and resembles closely the shape of a differentiated Gaussian pulse. In the numerical model, a differentiated Gaussian pulse is used as excitation, because the excitation must have a smooth shape to avoid numerical dispersion artifacts. To compare the experimental and numerical results, the transfer function of the numerical model at each point in space is determined and convolved with the excitation of the transducer. In this way, the elastic wave fields due to excitation with the real transducer foot motion are obtained.

The conditions at internal boundaries, i.e. at the interfaces between different media, are usually satisfied implicitly. However, to ensure numerical stability, the material properties have to be averaged for components on the boundary. While the material density ρ, appearing in the equation of motion, is averaged directly, the inverse of Lame's constants, λ and μ from the stress-strain relation, must be averaged.

Four external boundaries arise at the four outer grid edges. At its upper edge, the half-space is bounded by a free surface. Due to the continuity of normal stress, the normal stress components vanish at a free surface. In order to satisfy this condition, an extra row must be inserted into the finite-difference grid one step beyond the free-surface boundary.

In order to model the infinite half-space, all waves that are reaching the three remaining outer grid edges must be perfectly transmitted and absorbed. The boundary condition that does this most accurately is the *Perfectly Matched Layer* (PML) boundary condition, first introduced by Berenger[8] and adapted to elastodynamics by Chew and Liu.[9] Here, a non-physical splitting of the wave fields allows the introduction of a lossy boundary layer that is perfectly matched to the solution space. In continuous space, it has been shown that an arbitrarily polarized wave incident on this PML medium is perfectly transmitted. The wave experiences the exact same phase velocity and wave impedance as in the solution space, while rapidly decaying along the axis normal to the PML-medium interface. In discrete space, however, the PML will not be matched perfectly to the solution space. To keep reflections at the interface small, a tapered loss profile is chosen for the PML.[3]

2.2. Parallelization

The finite-difference model has been implemented in a fully parallel fashion. When implemented on a parallel computer, the three-dimensional finite-difference grid is divided into several sub-grids, and each sub-grid is assigned

to one processor of the parallel machine. The processors compute the wave fields on their sub-grids and share only the field values on the interfaces with their neighbors. The finite-difference model has been implemented to run both on a Cray T3E supercomputer with 544 processors (located at the ERDC Massively Shared Resource Center in Vicksburg, Miss.) and on a PC computer cluster located at the Georgia Institute of Technology. The PC cluster contains 50 Pentium III processors and has been especially built for this project.

3. INTERACTION OF ELASTIC WAVES WITH BURIED ANTIPERSONNEL MINES

The interaction of elastic waves with anti-personnel mines, buried in sand, is to be investigated. First, the propagation of a Rayleigh surface waves along the surface is compared as measured experimentally and as computed numerically. Then, the interaction of elastic waves with a single anti-personnel mine is described. Finally, three anti-personnel mines are buried in the ground together with some rocks and a wooden stick.

In the experiment, the mines are buried in a large sand-filled box. The properties of sand, together with the properties of plastic, air, rocks, and wood, are depicted in Table 1. Note that the shear wave speed in sand is assumed to be depth-dependent. Note also that wood is usually anisotropic, but is assumed to be isotropic to simplify the numerical model. This assumption is justified by considering the small dimensions of the wooden stick used here.

Table 1. Parameters used for finite-difference simulation.

Sand	Shear wave velocity	$c_{s,\text{sand}}$	(depth-dependent)
	Pressure wave velocity	$c_{p,\text{sand}}$	250 m/s
	Material density	ρ_{sand}	1400 kg/m^3
Plastic	Shear wave velocity	$c_{s,\text{plastic}}$	1100 m/s
	Pressure wave velocity	$c_{p,\text{plastic}}$	2700 m/s
	Material density	ρ_{plastic}	1200 kg/m^3
Air	Shear wave velocity	$c_{s,\text{air}}$	0 m/s
	Pressure wave velocity	$c_{p,\text{air}}$	330 m/s
	Material density	ρ_{air}	1.3 kg/m^3
Granite	Shear wave velocity	$c_{s,\text{granite}}$	3500 m/s
	Pressure wave velocity	$c_{p,\text{granite}}$	5500 m/s
	Material density	ρ_{granite}	4100 kg/m^3
Wood	Shear wave velocity	$c_{s,\text{wood}}$	100 m/s
	Pressure wave velocity	$c_{p,\text{wood}}$	300 m/s
	Material density	ρ_{wood}	650 kg/m^3

The space step for the numerical model is chosen to be $\Delta x = \Delta y = \Delta z = 0.5$ cm; the time step is $\Delta t = 11.54$ μs and, thus, fulfills the Courant condition (the necessary condition for stability of the finite-difference algorithm). The excitation has the shape of a Gaussian pulse. To be able to compare the numerical and experimental results, the transfer function of the system is determined in a post-processing step for each point in space and convolved with the transducer motion as measured in the experiment.

3.1. Surface Wave Dispersion

A Rayleigh surface wave is seen to disperse as it propagates along the surface of the ground. This effect has been observed experimentally[1,2] and is believed to be mainly due to the shear wave speed in sand increasing with depth. The pressure within the sand grows with depth, thus enhancing the cohesion between the grains of the sand. The

increased cohesion causes the medium to be stiffer and the shear wave speed to become larger. The shear wave speed as a function of depth is empirically found to be

$$c_s = 150\left(1 - \frac{3}{10z^{0.3}+5}\right) \text{ m/s}, \tag{14}$$

where z is the depths in meters. The depth profile is graphed in Fig. 2.

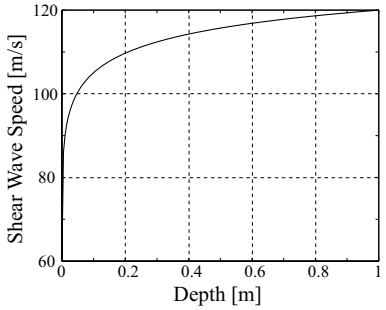


Figure 2. Shear-wave-speed depth profile.

Figure 3 shows waterfall graphs of the vertical particle displacement at the surface as measured experimentally (Fig. 3 (a)) and as determined numerically (Fig. 3 (b)). In these graphs, the displacement at a number of points along a straight line on the surface is plotted as a function of time and offset by the distance from the source. The slope of the traveling waves in the graph indicates the wave speed. Thus, by determining the slope, the different wave types can be distinguished.

Both in the experiment and in the numerical model, a pressure wave (P) and a Rayleigh surface wave (R) are seen to propagate. The pressure wave is faster than the Rayleigh wave and, thus, its slope is larger. The pressure wave is seen to decay quickly, due to the pressure wave being a volume wave rather than a surface wave. As the surface wave propagates, the pulse disperses and changes its shape. Dispersion arises, if wave components with different frequencies travel with different speeds. Due to the shear-wave speed increasing with depth, high-frequency components, which are more confined to the surface, will experience a lower shear-wave speed than low-frequency components, which reach deep into the ground. The agreement of experiment and numerical simulation is fairly good. The differences are mostly due to non-linear effects occurring in the sand. The non-linearities dampen the high-frequency components and, in this way, cause the pulse to broaden and to change in shape. By decreasing the amplitude of the excitation in the experiment, the agreement of experiment and numerical simulation can be improved. However, the signal-to-noise ratio of the radar output in the experiment will also decrease and, thus, the noise contents of the data will increase. Thus far, the non-linear effects are not incorporated into the numerical model.

Figure 4 shows the the vertical particle displacement on the surface and on a cross section through the ground at two times, T_1 and T_2, corresponding to the vertical lines in Fig. 3 (b). The wave fields have been computed numerically and are plotted on a logarithmic scale, in which black corresponds to the largest magnitude (0 dB) and white to the smallest (-60 dB). At T_1, a Rayleigh surface wave (R), a shear wave (S) and a pressure wave (P) are seen to propagate. The shear wave and the pressure wave are visisble only on the cross section through ground. The pressure wave induces head waves (H) as it propagates along the surface. Head waves are downwardly directed plane shear waves generated by the passage of bulk waves along the free surface. They usually decay rather quickly. Both the head waves and the surface wave are visible on the cross section and on the surface. At T_2, the waves have propagated farther. The pressure wave has left the range of the plot. The surface wave (R) and the shear wave (S) have separated. The head waves (H) have decayed and are not visible any more on the surface.

3.2. Interaction with a Buried Anti-Personnel Mine

To investigate the interactions with a buried land mine, a simple model for an anti-personnel mine is inserted into the numerical model. Figure 5 shows the simple model, together with a simplified cross-sectional drawing and a picture of a real TS-50 anti-personnel mine. The simple mine model consists of a main chamber containing plastic

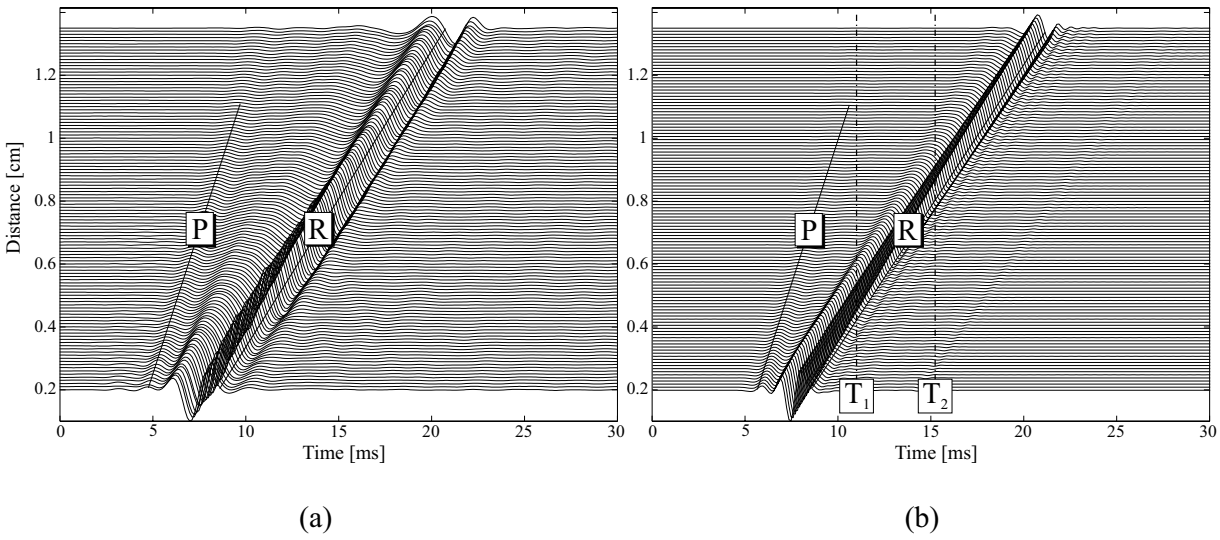


Figure 3. Waterfall graphs of the vertical particle displacement on the surface according to (a) experiment and (b) numerical simulation.

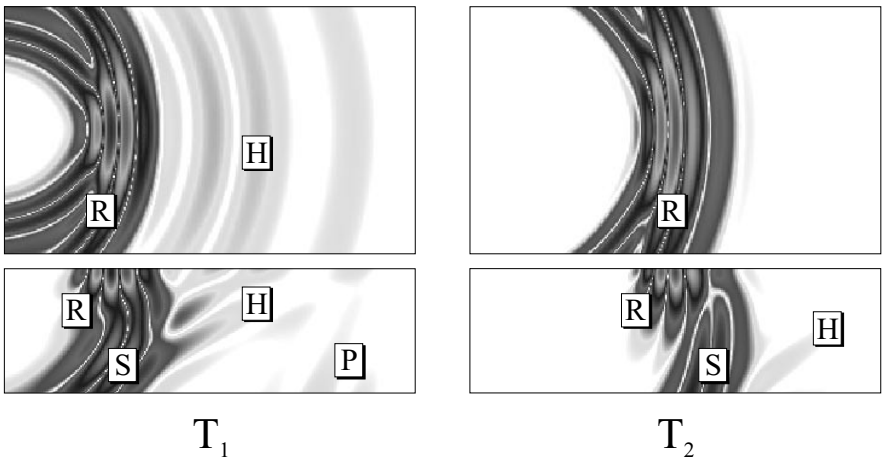


Figure 4. Normal particle displacement on the surface (top) and on a cross section through the ground (bottom). T_1 and T_2 correspond to the vertical lines in Fig. 3 (b).

explosives, and a smaller chamber on top of the mine's main chamber filled with air. In the cubic finite-difference grid, the round shape of the mine is approximated by cubes, leading to the stair-case form evident in Fig. 5. The air-filled chamber is inserted into the model to approximate the structure of a real TS-50 mine. A real TS-50 mine contains plastic explosives, a trigger mechanism, several chambers and is enclosed in a plastic case (see Fig. 5).

Figure 6 shows waterfall graphs of the mine-wave interaction for both experiment and numerical simulation. The mine is buried 2 cm beneath the surface, at a distance of 70 cm from the source. A pressure wave (P) and a Rayleigh surface wave (R) arise. The waves hit the mine and are reflected (rR) and transmitted. While the interaction of the mine with the pressure wave is weak, the surface wave strongly interacts with the mine. In both experiment and numerical simulation, resonant oscillations occur at the mine location and remain even after the waves have passed the mine. For the numerical model, it can be shown that the incident waves couple into flexural waves which arise in the thin soil layer above the mine.[4] These flexural waves are confined to the thin layer and form a standing wave pattern, giving rise to the resonant oscillations. While this explains the resonance in the numerical model, it gives only one possible cause for the resonance in the experiment. A real TS-50 mine has several chambers, it has a flexible case that can support both flexural and longitudinal waves, and it contains springs that can also give rise to resonances. The authors are currently working on refining the numerical model to incorporate more details of the mine.

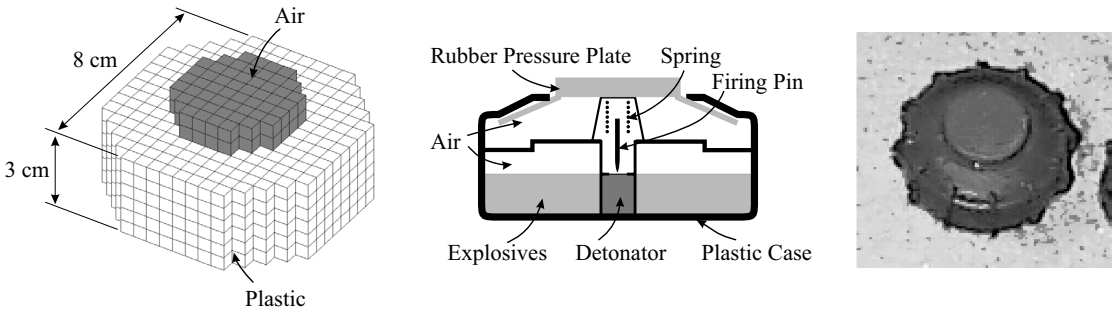


Figure 5. Simple model for the TS-50 anti-personnel mine; cross-sectional drawing and photograph of a real TS-50 mine.

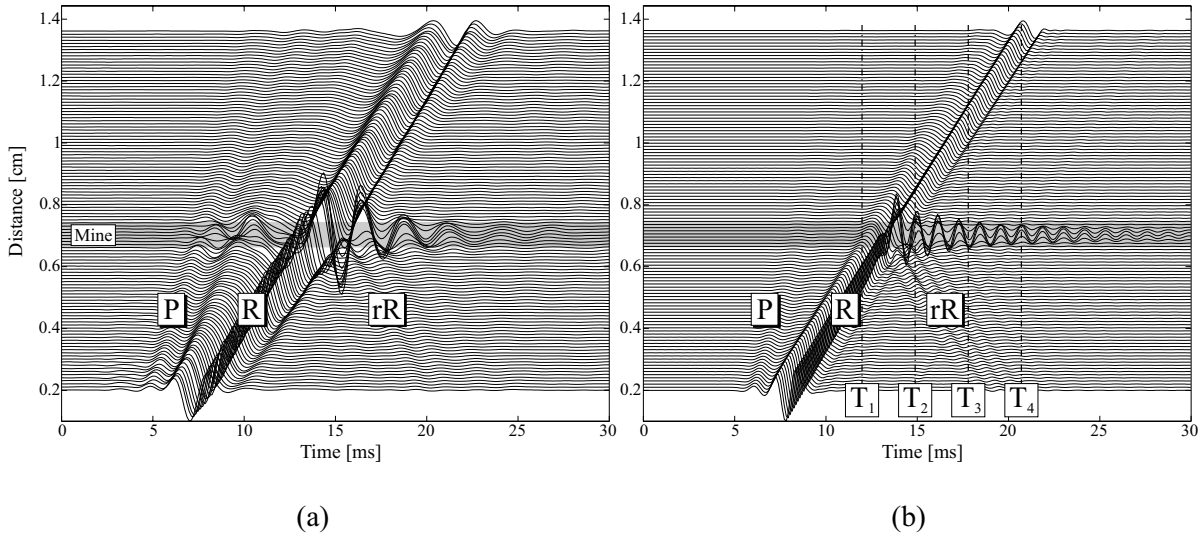


Figure 6. Interaction of elastic waves with a buried anti-personnel mine; waterfall graphs of the vertical particle displacement on the surface according to (a) experiment and (b) numerical simulation.

In Fig. 7, the vertical particle displacement on the surface and on a cross section through the ground is plotted. The fields are determined using the numerical model. The dynamic range of the plots is 50 dB. The wave fields are shown at four different instances in time, corresponding to the vertical lines marked with T_1, T_2, T_3 and T_4 in Fig. 6 (b). In the first plot, the surface wave is seen to just hit the mine. While only the surface wave (R) is visible on the surface, both the surface wave (R) and the shear wave (S) appear on the cross section. In the second plot, the surface wave has just passed the mine. The interaction of the surface wave with the mine gives rise to reflected surface waves (rR) and reflected shear waves (rS), which are clearly visible on the cross section. Pressure waves are also induced by the surface wave, but they are weak and not visible. In the third and fourth plot, the surface wave has passed the mine. Some energy, however, remains at the mine and causes the mine to vibrate and to radiate.

3.3. Mine-Wave Interaction in the Presence of Clutter

Next, it is investigated how the presence of clutter affects the mine-wave interaction. For this, three mines are buried in the ground together with a wooden stick and four rocks of various sizes and shapes. The material properties of the rocks (granite) and wood can be found in Table 1. Figure 8 shows how the mines, the rocks and the stick are arranged in the ground. The mine closest to the source is smaller (6 cm in diameter) than the other two mines (8 cm). The mines are buried 2 cm beneath the surface. The first mine lies at a distance of 50 cm from the source, the second at 70 cm and the third at 80 cm. The rocks and the wooden stick are located at different depth and have different shapes. The rock closest to the source is a tilted cuboid, the second rock is a sphere, and the other two consist of rectangular plates of random size. The wooden stick is modeled as a cylinder with a bent axis.

Figure 9 shows the vertical particle displacement on the surface in two series of pseudo color plots as obtained with the numerical model. The first series, Fig. 9 (a), has a dynamic range of 40 dB, whereas the second series,

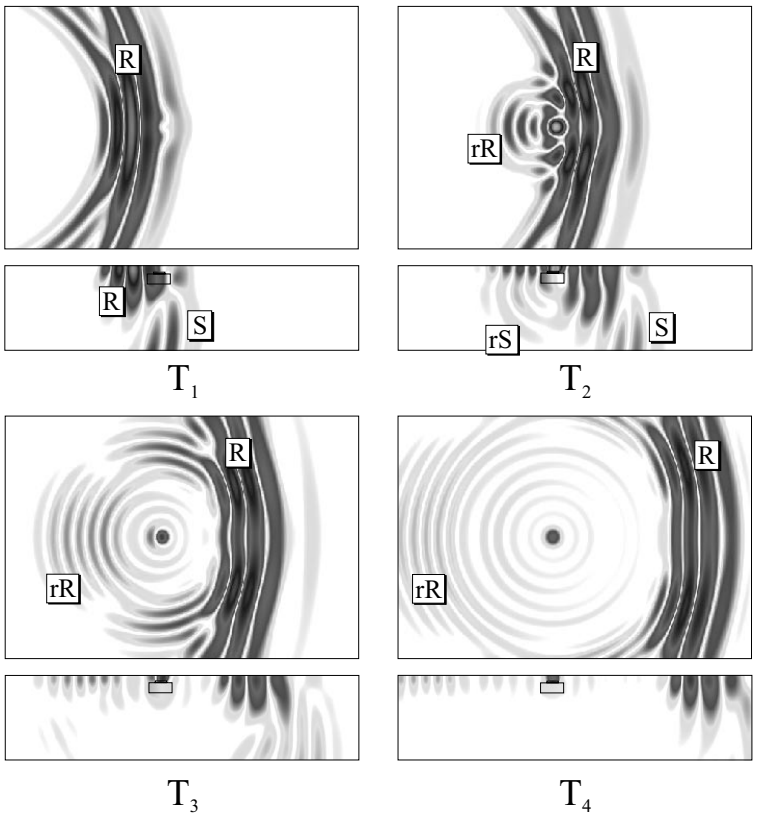


Figure 7. Interaction of elastic waves with a buried anti-personnel mine; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instances in time, corresponding to the vertical lines in Fig. 6 (b).

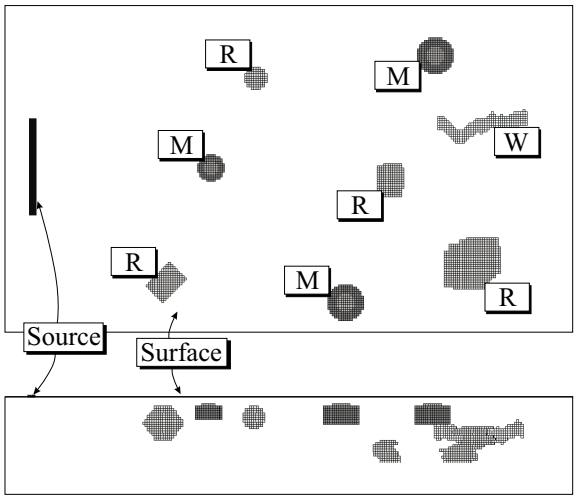


Figure 8. Three mines (M), 4 rocks (R) and a wooden stick (W) buried in the ground; top view and side view.

Fig. 9 (b), has a range of 70 dB. The dynamic range of the first series (40 dB) corresponds to the dynamic range of the experimental measurements. In the 70 dB range, however, more details of the wave-interaction with mines and clutter will be visible. In all plots, black represents the highest magnitude (0 dB) and white the lowest (-40 and -70 dB, respectively). The wave fields are shown at four instances in time.

At the first instance, i.e. the first row in Fig. 9, the surface wave just hits the mine closest to the source. In front

of the surface wave, head waves are propagating. The head waves decay quickly and, therefore, appear only at the earliest instance. On the 40 dB scale, only one faint head wave is visible. However, on the 70 dB scale, three head wave fronts are evident. These head waves have already reached the other two mines. At the second instance, the surface wave has passed the first and second mine and has just reached the third mine. A similar resonance to the one described in Sec. 3.2 is observed at the location of the first mine. Energy is trapped at the mine, and the mine radiates continuously. At the third and fourth time, the surface wave has passed all mines and all three are excited to resonant oscillations. The mines are clearly visible on the surface, whereas the rocks and the stick cannot be seen. The interaction with all rocks is rather weak. Very similar results have been observed also experimentally using real anti-personnel mines. Strong resonances could be seen at the mine locations, making it easily possible to distinguish the mines from the clutter.

4. CONCLUSIONS

A three-dimensional finite-difference model for elastic waves in the ground has been developed and implemented in a fully parallel fashion. Results are obtained and compared to experimental results. Experimentally, a Rayleigh surface wave is seen to disperse as it travels along the surface of the ground. This dispersion effect can be modeled also numerically by assuming a depth-dependent shear-wave speed profile. When an anti-personnel mine interacts with elastic waves, a strong resonance occurs at the mine location. This resonance has been observed in both experiment and numerical simulation. When clutter is present, this resonance can be used to distinguish mines from objects like rocks and wooden sticks, making it easy to locate the mines in the ground.

ACKNOWLEDGMENTS

This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0448, by a grant from the US Office of Naval Research under contract N00014-99-1-0995, and by an equipment grant from the Intel Corporation.

REFERENCES

1. W. R. Scott, Jr., C. T. Schröder, and J. S. Martin, "An acousto-electromagnetic sensor for locating land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets III, Proc. SPIE*, vol. 3392, pp. 176–186, 1998.
2. W. R. Scott, Jr. and J. S. Martin, "Experimental investigation of the acousto-electromagnetic sensor for locating land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proc. SPIE*, 1999.
3. C. T. Schröder and W. R. Scott, Jr., "Finite-difference time-domain model for elastic waves in the ground," in *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proc. SPIE*, 1999.
4. C. T. Schröder and W. R. Scott, Jr., "A finite-difference model to study the elastic-wave interactions with buried land mines," *IEEE Trans. on Geophysics and Remote Sensing* , to be published.
5. C. T. Schröder, *A Finite-Difference Model for Elastic Waves in the Ground*, Diploma Thesis, Technische Universität Braunschweig, 1999.
6. J. Virieux, "P-SV wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophysics* **51**, pp. 889–901, April 1986.
7. A. Tavlove, *Computational Electromagnetics: The Finite-Difference Time-Domain Method*, Artech House, 1995.
8. J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Physics* **114**, pp. 185–200, 1994.
9. W. C. Chew and Q. H. Liu, "Perfectly matched layer for elastodynamics; a new absorbing boundary condition," *J. Comput. Acoustics* **4**, pp. 341–359, 1996.

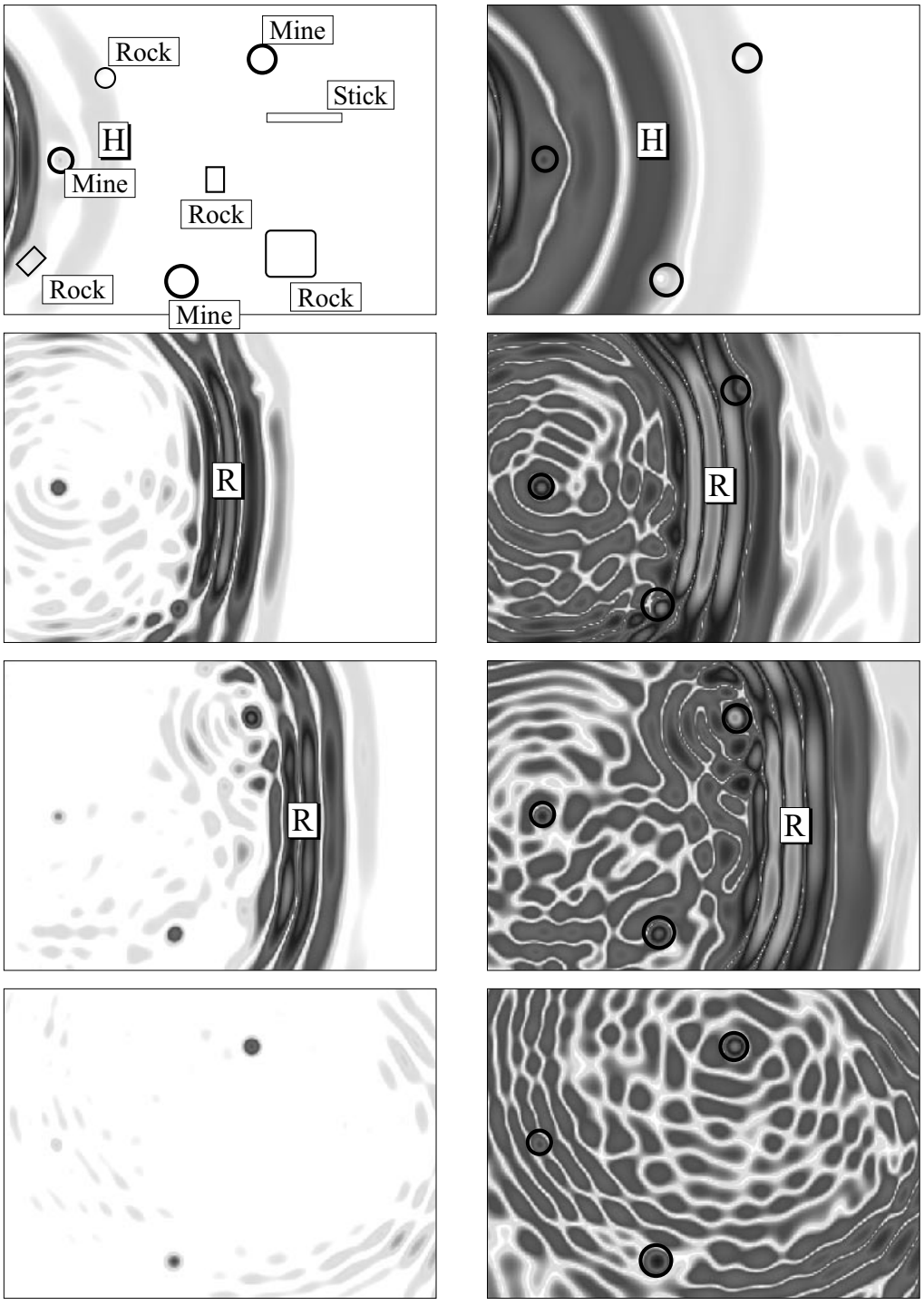


Figure 9. Three mines, 4 rocks and a wooden stick buried in the ground; pseudo color plots of the normal particle displacement on the surface with a dynamic range of (a) 40dB and (b) 70 dB.

Proceedings of SPIE, Vol. 4394, April 2001

Environmental Factors that Impact the Performance of a Seismic Land Mine Detection System

Gregg D. Larson [a], James S. Martin[a], Waymond R. Scott, Jr. [b] and George S. McCall II [c]
[a] School of Mechanical Engineering
[b] School of Electrical and Computer Engineering
[c] Georgia Tech Research Institute
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

A system has been developed that uses high frequency seismic waves and non-contacting displacement sensors for the detection of land mines. The system consists of a moving displacement sensor and a stationary elastic-wave source. The source generates elastic waves in the earth. These waves propagate across the minefield where they interact with buried mines. The sensor measures the displacements at the earth's surface due to the passage of the waves and the interactions of the waves with mines. Because the mechanical properties of the mine are different from those of the earth, the surface displacements caused by the interaction are distinct from those associated with the free-field propagation of the waves. This provides the necessary cue for mine detection. The system has been demonstrated in a controlled laboratory environment, and efforts are currently underway to transition this work into field tests. Moving the experimental effort into the outdoor environment is a critical milestone toward the ultimate goal of this research effort, which is the design of a field-operable mine detection and classification system. There are many issues associated with this transition. Foremost among these is the propagation characteristics of seismic waves in the field environment and, particularly, the mechanisms that limit the energy which can be coupled into the seismic signal that is used to search for mines. To investigate this, a field site was selected which reasonably emulates the existing laboratory experimental model, and a series of measurements was undertaken to determine the effects of environmental factors at both sites on the generation and propagation of seismic waves. At both sites, strong non-linearity was observed which limited the energy content of the incident signal.

Keywords: land mine, mine detection, non-linearity, seismic waves

1. INTRODUCTION

Seismic/elastic techniques show considerable promise for the reliable detection of all types of buried mines, even low-metal anti-personnel mines. The reason for this is that mines have mechanical properties that are significantly different from soils and typical forms of clutter. For example, the shear wave velocity is approximately 20 times higher in the explosive and the plastics used in typical mines than in the surrounding soil [1]. In addition, mines are complex mechanical structures with a flexible case, a trigger assembly, air pockets, etc. The complex structure gives rise to structural resonances, non-linear interactions, and other phenomenology that is atypical for both naturally occurring and man-made forms of clutter. Thus, this phenomenology can be used to distinguish a mine from clutter.

A system has been developed at Georgia Tech that uses a radar-based displacement sensor for the local measurement of seismic displacements without physically contacting the soil surface [2-3]. The non-contact nature of this sensor makes the system capable of interrogating the soil surface near or immediately above a mine. This substantially increases the measurable effects of the mine's presence over schemes which rely on elastic waves scattered by the mine to propagate to a remote sensor location [4]. Figure 1 depicts the present system configuration. The sensor shown in the figure is bistatic. Both bistatic and monostatic configurations have been

studied. The system consists of the non-contact sensor and the ground coupled seismic source. The source (an electrodynamic shaker resting on a narrow foot) excites an incident signal composed primarily of a Rayleigh surface wave with measurable energy also propagating in bulk shear and compressional waves. The incident signal travels across the minefield and excites motion of soil near the surface that includes the buried mines. The sensor measures this motion over an array of measurement locations. Sensor data are processed in both space and time domains to yield an image of the minefield [5]. To interrogate the earth over the range of typical mine sizes and burial depths, the incident signal has a spectrum from 50Hz to 1,000 Hz.

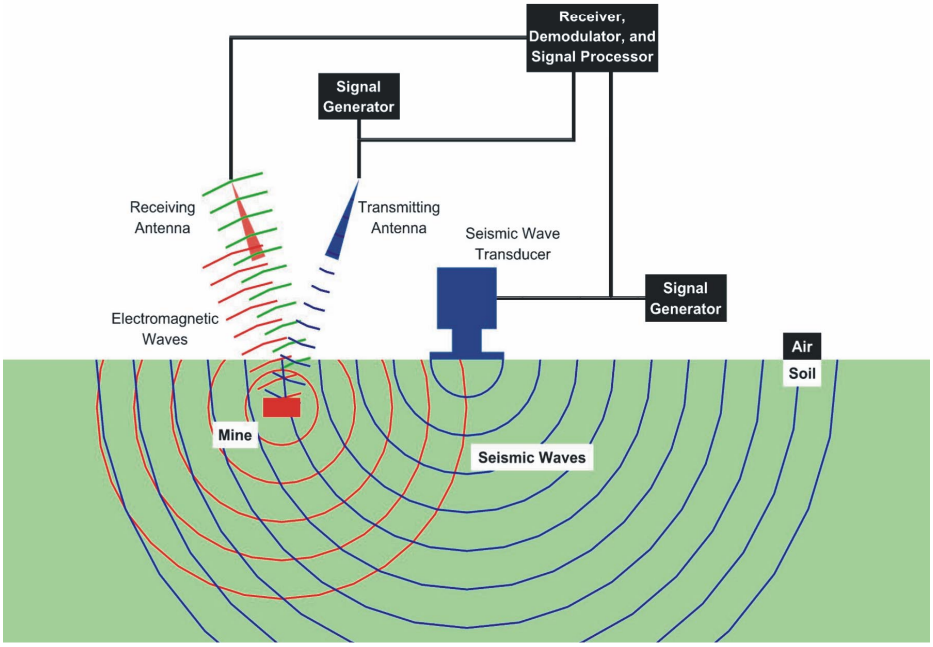


Figure 1: Schematic of a Seismic Mine Detection System

The seismic mine detection system has been investigated experimentally over the past 3 years using a series of laboratory models of increasing size and complexity. Damp compacted sand was selected as the medium for these models because of the ease and repeatability it offered for the burial of mines, mine simulants, and clutter objects. Anecdotal evidence exists in the literature to suggest that this is a reasonable model for soils [6,7]. The published data are sparse on the range of frequencies and depths that are of interest for this problem and are limited to the characterization of soils within the linear regime. Scaling of published experimental data and theoretical results involving very different frequency ranges and material parameters requires a sufficient experimental database to justify the comparison. Additionally, results generated with the laboratory model suggest phenomena with no obvious precedent in the literature. These non-linear effects can be observed at relatively low drive levels (displacements on the order of 1 μm). They include dispersion of surface guided waves, an inverse relationship between amplitude and group velocity, and amplitude-dependent loss. Whether these are typical of the natural environment or peculiarities of the laboratory model is yet to be determined. The work presented here represents the first attempt to address some of these questions.

The radar-based displacement sensor is not presently configured for field tests. Since the topic of current interest is environmental characteristics rather than sensor development, it was not deemed necessary that the system sensor be used for the field measurements. The baseline sensors that were used for the laboratory-to-field comparisons were, therefore, ground coupled accelerometers. These have the advantage of having comparable sensitivity to the radar sensor over most of the band of interest while being waterproof and robust in a comparatively hostile environment. The accelerometers also offer a simple method for the construction of a

physical array, whereas the radar sensor is currently a single sensor that is scanned to form a synthetic array. The synthetic array cannot easily distinguish between long-timescale changes, such as tidal motions of the water table, and short-timescale changes, such as dispersion of a propagating waveform. Unlike the laboratory environment, these factors cannot be controlled in the field. Vibra-Metrics VMI-9002A accelerometers encapsulated in syntactic foam with 100 mV/g sensitivity were used in conjunction with custom-built preamplifiers and power supply decouplers for their internal charge amplifiers. Data were acquired without low-pass filters on 24 channels sampled at 8 KHz with 12-bit resolution in all of the accelerometer measurements. Anti-aliasing filters were not used in the data acquisition system because previous experiments in the linear regime had demonstrated that the need for these was obviated by the rapid attenuation of high frequencies with propagation distance in the sand. Although the Nyquist frequency was 2 octaves above the highest transmit frequency (1 KHz) and well above the dominant spectral components of the background noise, aliasing problems introduced by the lack of filters complicated the analysis of the data. Aside from this problem, the accelerometer system offered a signal to noise ratio (SNR) equivalent to that achieved by the radar sensor in the laboratory at 100 Hz. Since the radar sensor responded to displacement rather than acceleration, there was a 12 dB per octave SNR improvement associated with the accelerometers at high frequencies and a corresponding degradation at low frequencies. An additional relative improvement in the accelerometers' high frequency SNR is offered by its small integration area on the ground's surface (3 cm^2) compared with approximately 9 cm^2 illuminated by the radar sensor (depending on the sensor's height).

2. LABORATORY AND FIELD MEASUREMENTS IN THE LINEAR REGIME

Figure 2 shows the laboratory model and the field site. The laboratory model, which has been described in detail in a previous paper [2], is a 1.5 m deep wedge-shaped tank filled with 50 tons of masonry sand that is kept damp and compacted with a water table maintained 50 cm below the surface. The field test site is at the Navy beach adjoining the campus of the Naval Postgraduate School (NPS) in Monterey, California. This has been the site of ongoing micro-seismic experiments undertaken at NPS and is quite near a beach location extensively characterized by Bachrach and Nur in 1995 [8]. Several different accelerometer array configurations were used at both sites. The cross configuration depicted in Figure 2 represents the largest of these. Tides varied the water height at the field site both during and between experiments. The dimensions shown are typical, although depth changes of up to 50 cm occurred during the experiments. This translated into horizontal changes of up to 10 m. Some of the experiments were performed at low tide on the harder sand below the high water mark. Others were performed at high tide on the softer dry sand above the high water mark. Similar propagation speeds were observed in both cases, although the wet sand results more closely matched those of the laboratory model. Figure 3 depicts waterfall plots generated by pulse compression of data taken at both the laboratory and field sites. The data shown here were not measured directly but were reconstructed from measured transfer functions. For each of these data sets, the raw data, which were generated using a 4 second linearly-swept-frequency chirp covering the entire band of interest, were used to synthesize the response of the system to a temporal discrete waveform with a spectrum centered in the middle of the band of interest. In this case, the compressed drive signal was the first derivative of a Gaussian pulse centered at 400 Hz. The graphs in Figure 3 indicate the similarity between the laboratory model and field site. Two distinct arrivals are observed at each site. The first of these (any apparent earlier arrivals are caused by noise-induced processing artifacts) arrives with a group velocity of 170 m/s to 183 m/s, which is consistent with observed P-wave velocities [8] and with observations of leaky surface waves [9] at similar sites. The series of small arrivals immediately following the first arrival may account for the latter observation, although both wave types may not exist with measurable amplitude in the data. The second and largest pulse arrival propagates with a speed of 87 m/s to 90 m/s and is due to the Rayleigh surface wave. The peaks of this waveform have been clipped in the figure in order to display both wave types in a single graph. This speed is consistent with the other reported observations of Rayleigh waves in both sand and hard soil [7,9]. The coda associated with the Rayleigh wave arrival is longer and more complex in the field data than in the laboratory because of the variety of surface and subsurface inhomogeneities at the field site. These include footprints, buried flotsam, and a water table that was inclined with respect to the ground surface and non-uniform due to wave

action. Although identical pulse shapes were used with both data sets, there are also some differences in the shapes of the received signals: the data from the laboratory model have lower spectral content than the field data. This is due to a difference in the frequency response of the two different transducers that were used in these tests. A MB Electronics 20-lb. shaker was used as a source for the field tests, but this was damaged in shipping and replaced by a 100-lb. VTS shaker for the later laboratory measurements. Although the Rayleigh wave pulse has been clipped in these plots, dispersion is apparent in both the laboratory and field data. In a homogenous, isotropic, semi-infinite medium, these waves would be non-dispersive. The most likely cause for this dispersion is a vertical inhomogeneity in the sand. This may be due to the depth dependence of water content (as hypothesized in [8]) or to the lithostatic force on the lower strata (as in the earth's crust on large length scales [10]).

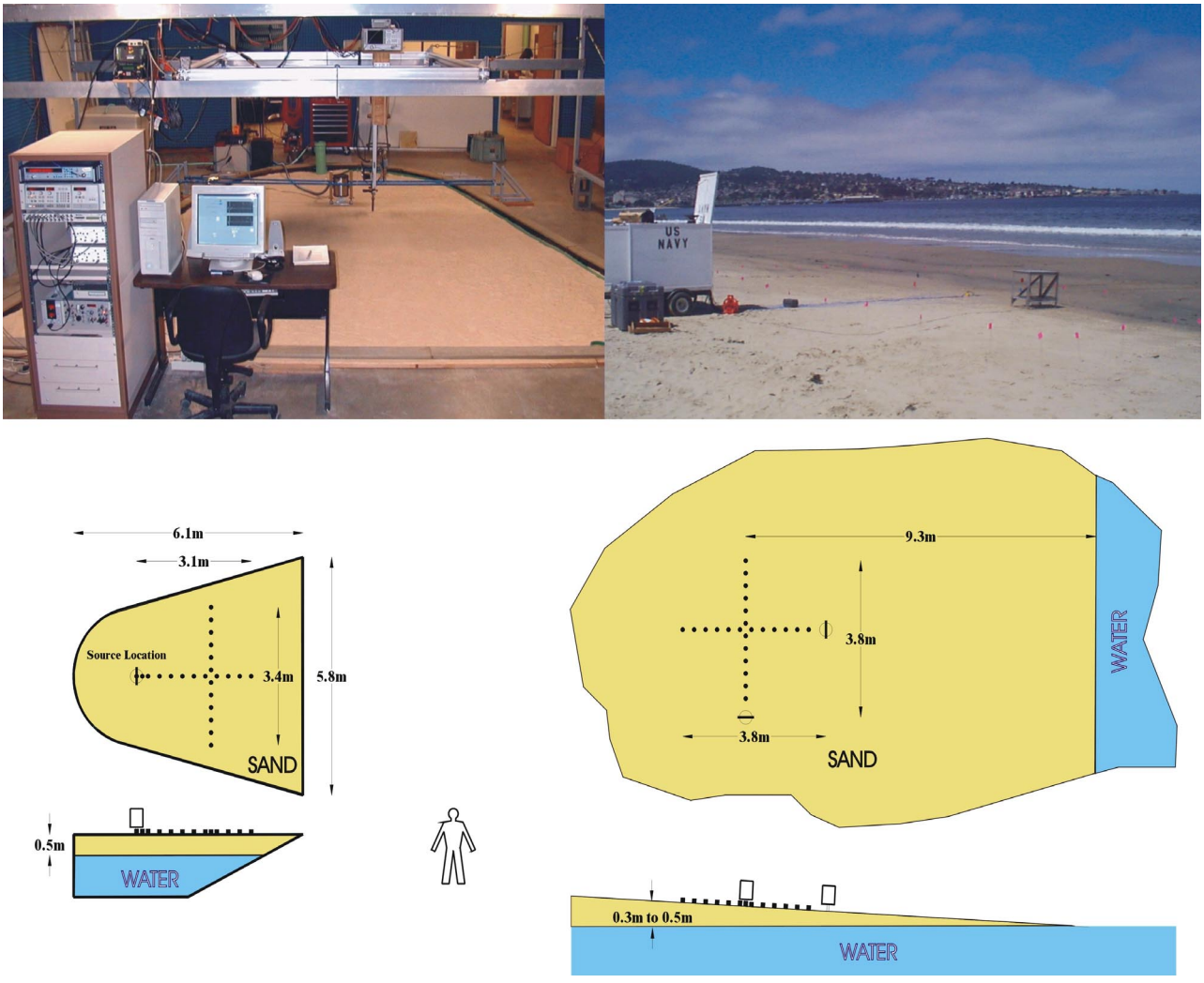


Figure 2: Experimental Models, (A) the Laboratory Model and (B) the Field Test Site

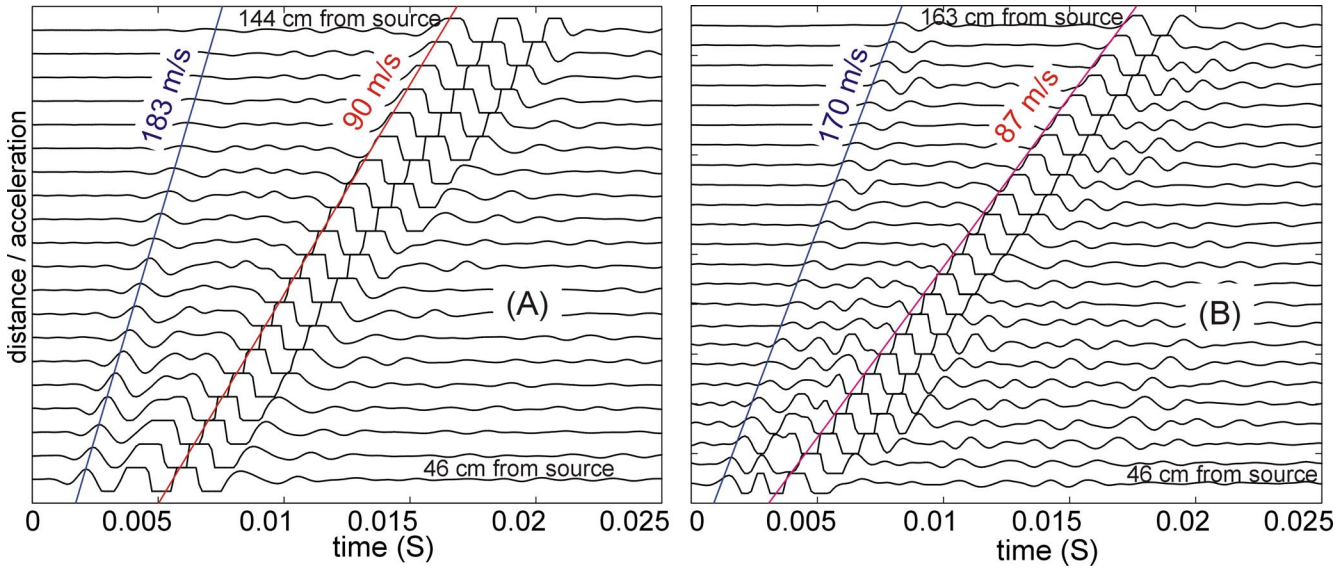


Figure 3: Propagation of a Seismic Pulse, in the Laboratory Model (A) and at the Field Test Site (B)

A second difference between the field site and the laboratory model is the horizontal anisotropy in the wave propagation introduced by the incline of the water table. Figure 4 shows a comparison of measurements made perpendicular to the direction of wave propagation for two cases from the field site: first, for a wave propagated along the shoreline, and second, for a wave propagated directly inland. The waveforms represent only the surface wave pulse (the faster arrivals are time gated out) and are not clipped, but they have been normalized by their peak values to correct for the directivity of the source. The inland propagation measurement shows a very strong top/bottom symmetry along the array as one would expect for an isotropic medium. The along-shore measurement, however, shows a distinctly different pulse shape arriving at the shoreward end of the array. This dispersion may be due to the narrowing dimension of wave guide between the surface and the water table or to a lowering of the wave speed with water content below complete saturation. This was observed by Bachrach and Nur [8] for P-waves in a similar environment.

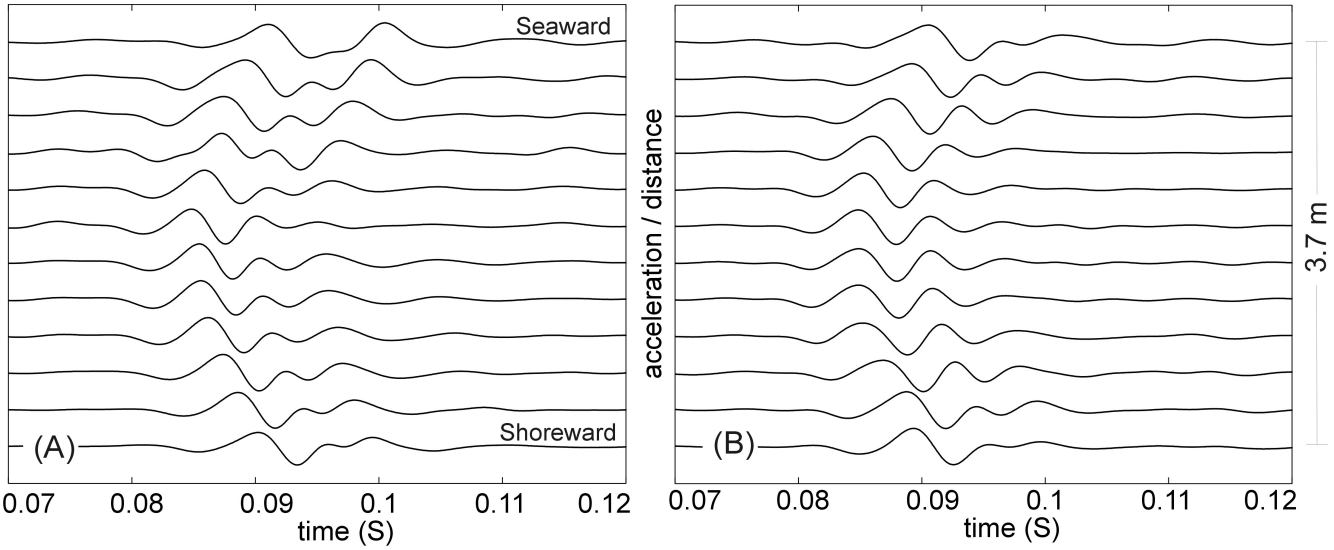


Figure 4: Comparison of Wave-Front Transects: Propagation Parallel to the Beach (A) and Shoreward Propagation (B)

It is also interesting to note that a third difference between the laboratory model and the field site was in the observed ambient noise fields. This difference, surprisingly, favored the field site. In spite of wave action, wind, and foot traffic, the noise floors measured at the field site were entirely due to the instrumentation noise of the measurement system below 500 Hz. In contrast, the noise floor in the laboratory model has several distinctive features that can be attributed to artificial sources and couple into the model through the foundation of the building that houses it. On average, the laboratory measurements were made over a noise floor which was 4 to 6 dB higher than the field site's noise floor in the 50 Hz to 1000 Hz band.

3. HARMONIC GENERATION AND NON-LINEAR STEEPENING

As mentioned previously, the incident signals used for these experiments were linearly-swept-frequency chirps. This makes it possible to separate non-linear effects in the long-time domain using a short-time Fourier transform as a function of long times (i.e. a spectrogram). Figure 5 depicts such a spectrogram for the signal received at an accelerometer located 50 cm from the seismic source on dry sand (above the high water mark) at the field test site. The lowest of the positively sloped ridges on the contour plot represents the fundamental component of the drive signal. The ridges with steeper slopes represent each of the higher harmonics generated in the source to sensor path as indicated by the numbers on the plot. Thus, it is a relatively simple matter to distinguish the 200 Hz, 300 Hz, and 400 Hz components generated by non-linear mechanisms from the same spectral components contained in the original drive signal because these effects are not concurrent on a long timescale. It is clear from the figure that the system is highly non-linear with more than 8 harmonics distinguishable above the noise floor. The aliasing problem mentioned earlier is also clear in this figure. Although the drive signal contains no energy above 1 KHz, the harmonics of the drive signal higher than the second contain energy above the Nyquist frequency of 4 KHz. Thus, the ridges appear to *reflect* from the upper frequency bound of the figure. At the times when these aliased signals are coincident with the drive or other harmonics, the two effects cannot be separated in post-processing. Both plots in Figure 5 depict the same signal spectrogram. In Figure 5B, the dynamic range has been compressed and converted to a gray scale, which uses only the bottom 20 dB of the 60-dB dynamic range available above the noise floor. From this figure, it is apparent that sub-harmonics of the drive signal are also generated at very low levels in the wave propagation. This may be a feature unique to the dry sand or may not have been manifest above the noise floor in the data acquired below the high water line, which showed a similar content of overtones.

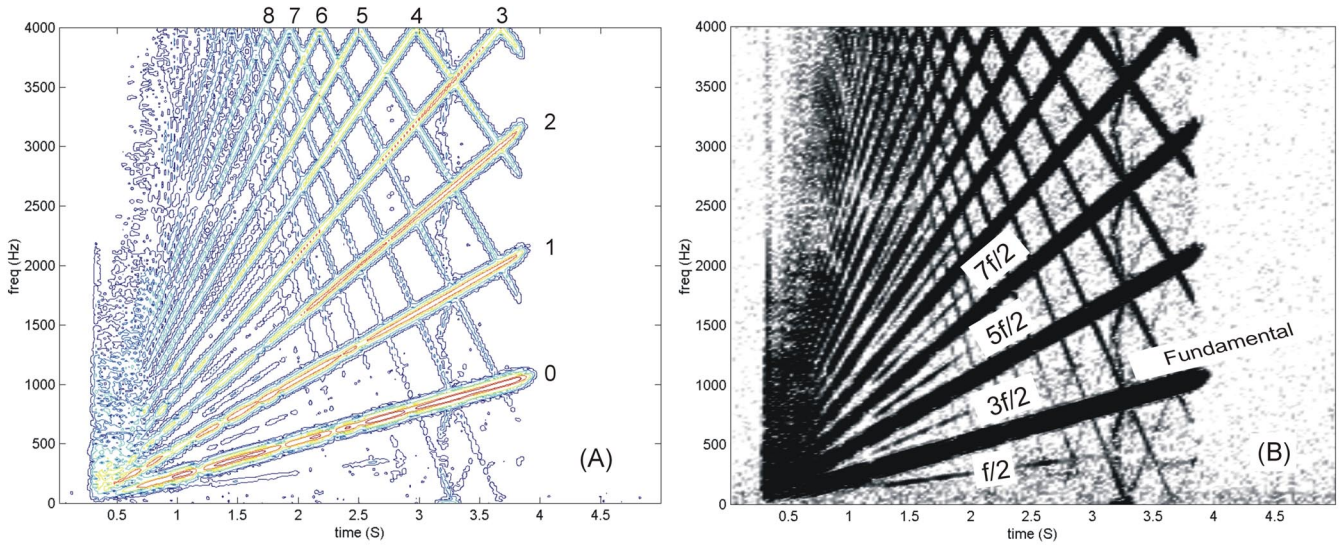


Figure 5: Spectrogram of Acceleration Signal Measured 50 cm from the Seismic Source at the Field Site on Dry Sand, Depicted on a 60 dB Scale (A) and on a 20 dB Scale at the Bottom of the Dynamic Range Measured (B).

It is difficult to attribute the generation of the higher harmonics and sub harmonics solely to the non-linearity of the sand. Source non-linearity and non-linearity in the source-to-sand coupling may also play a role in the observed phenomena. Spectral analysis of multiple points along the propagation path shows that, across most of the measurement region ranging from 0.35 m to 5.0 m from the source, the propagation appears to be linear in the data acquired at the field site. The question remains, however, as to whether or not there is a distinguishable exchange of energy between fundamental frequencies and overtones within the measured region. This is somewhat difficult to address because geometrical spreading reduces the signal with distance, and the higher frequencies are attenuated very quickly in the sand. Harmonics, therefore, cannot be seen to grow along the propagation path in the field test data unless normalized by the fundamental and then only over the nearest 3 to 4 measurements. However, an examination of a similar experimental setup from the laboratory reveals evidence to support propagation path non-linearity as the cause of the harmonics observed in the field tests. Figure 6 shows a comparison between spectra that were observed in the sand and at the source. Although harmonics of the drive signal exist at the source, they have clearly been amplified relative to the fundamental by the first 35 cm of propagation through the sand. Only three overtones are observable in the source motion and six can be seen in the sand. Further evidence in support of this is available from an examination of waveforms from the laboratory model in the time domain. Figure 7 depicts measurements made using a 396 Hz CW drive signal. A single cycle of motion has been plotted at the source and at distances of 30.5 cm and 244 cm away from the source in the sand. All the signals have been shifted in time and have been scaled to a common peak amplitude in order to emphasize changes in the waveform shape over delay and attenuation. Although very little distortion is observable in the source motion, the signal has converted to a series of sharp high frequency impulses of downward motion separated by broad low frequency positive phases. This transition occurs over the first 30 cm of propagation. Losses, then, strip the high frequency energy from the signal, and it has nearly returned to the shape of the drive at the 244 cm range, which is effectively in the "old age" region for the evolution of this waveform. The distortion observed at 30 cm is not consistent with the formation of shock fronts predicted by the Fubini solution for a one-dimensional compressional wave [11]. It is, however, consistent with the numerical predictions of Zabolotskaya for the vertical displacement component of a non-linear Rayleigh wave in a homogeneous solid (steel)[12].

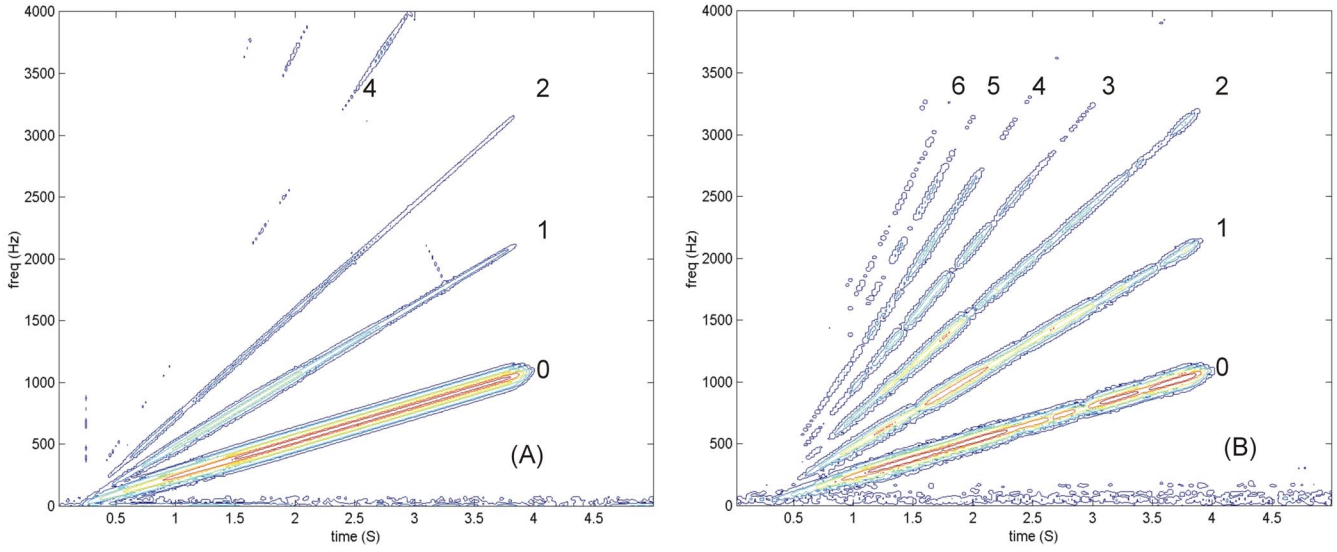


Figure 6: Spectrograms of Measured Acceleration on a 60 dB Dynamic Range Measured at the Source (A) and in the Sand 30.5 cm from the Source (B) in the Laboratory Model Experimental Model

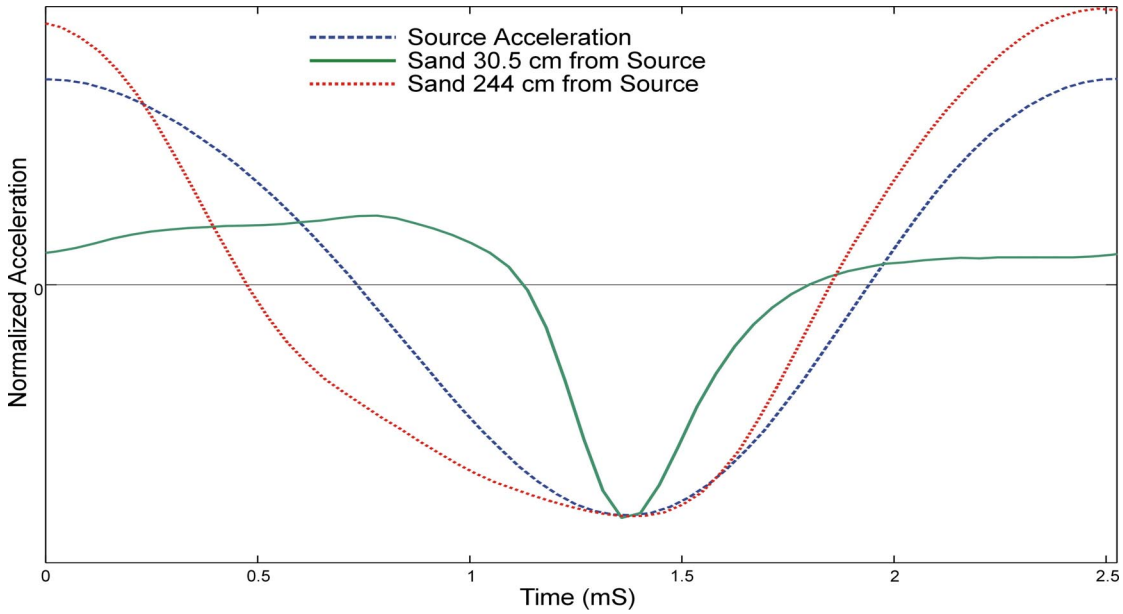


Figure 7: Steepening of 396 Hz CW Waveform with Propagation Distance

4. SATURATION AND SPALL

There are two obvious limits to the transfer of energy into seismic waves: saturation and spall. Under saturation, marginal level additions to the drive signal are entirely converted into higher harmonics and attenuated so that they make no contribution to the signal level received far from the source where the fundamental component dominates. Thus, an increase in drive signal does not produce a corresponding increase in the amplitude of the propagating elastic wave. Detachment of the source from the sand surface, known as spall, is the equivalent of cavitation at the face of an acoustic transducer immersed in a fluid. This occurs when either the biasing force on the shaker contact (the weight of the shaker itself for the data presented here) or the cohesion within the sand below that contact point is inadequate to maintain contact between the source and the sand.

Experimental evidence indicates that each of these mechanisms may limit energy transfer under different circumstances both in the laboratory model and at the field test site. Also, there is some evidence to suggest that the two effects may not be entirely separable in that spalling on a granular scale may be involved in the saturation limit. Spall at the source occurred more readily at the field test site in the tests performed on the dry sand above the high water mark. It was audible during the tests as a rumble midway through some of the frequency sweeps. Figure 8 depicts a spectrogram containing a spall event measured 51 cm from the shaker source on dry sand; the signals from closer accelerometers clipped in the analog electronics during this event. The event is seen midway through the signal and is easily distinguishable from the harmonic generation, which precedes and succeeds it, because of its broadband nature. Both before and after the event, energy is confined to discrete bands around the drive frequency and its overtones, whereas the rupture of the sand distributes energy evenly over the entire bandwidth of the measurement system and probably beyond this as well. Although the force applied to the contact surface was constant through the drive signal, the contact has obviously been restored after the spall event. This is probably because the sand surrounding the contact was fluidized by the spall event, and the settling of the shaker into the surface improved the contact sufficiently to end the event.

All of the field measurements were performed with transient signals at empirically maximized drive levels; unfortunately, the point at which individual spectral components of the drive signal went into saturation could not be determined from this limited data set. The experiment was therefore repeated in the laboratory model using a pure tone drive at 396 Hz. This is near the middle of the band of interest. The results of this experiment are plotted in Figure 9. Here the measured levels of the fundamental component and the first two harmonics have been plotted for the acceleration measured in the sand at 4 different ranges against that measured at the source. All

of the curves depict a nominally linear amplitude ratio for source accelerations up to about 1g. Beyond this, the received signals begin to saturate. Interestingly, the signals measured at 61 cm and 122 cm from the source enter a phase where incremental additions to source level result in an actual decrease in measured acceleration of the sand. This may be due to an interference, which occurs when portions of the source contact have fluidized while others have not. It is also possible that the combined effects of dispersion and regeneration of the fundamental frequency by mixing of overtones produces a destructive interference.

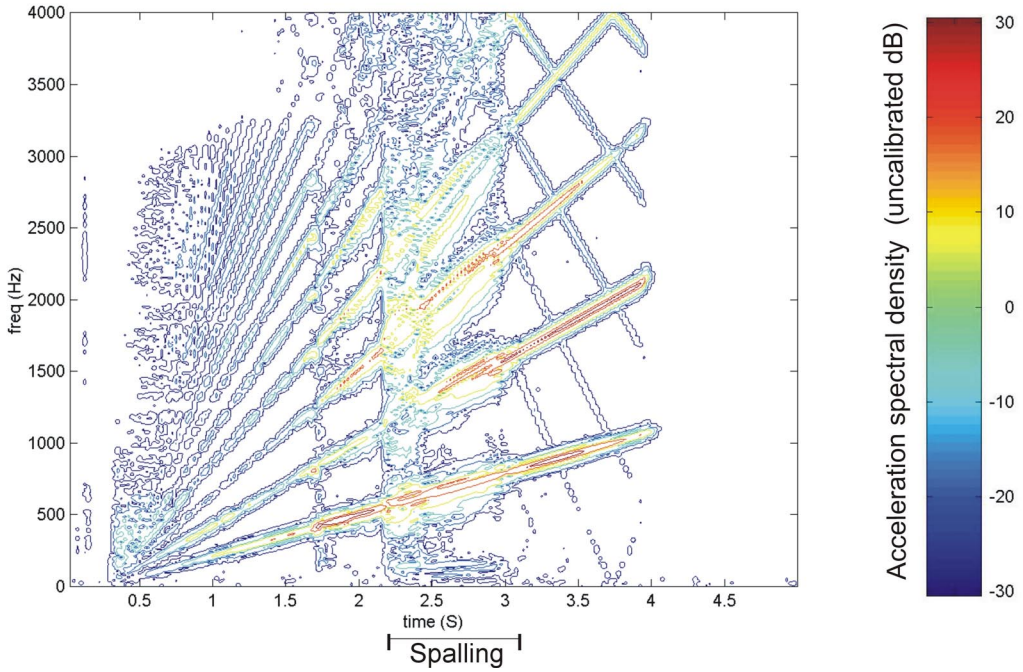


Figure 8: Spectrogram of a Spall Event Measured at the Field Site 46 cm from the Source on Dry Sand

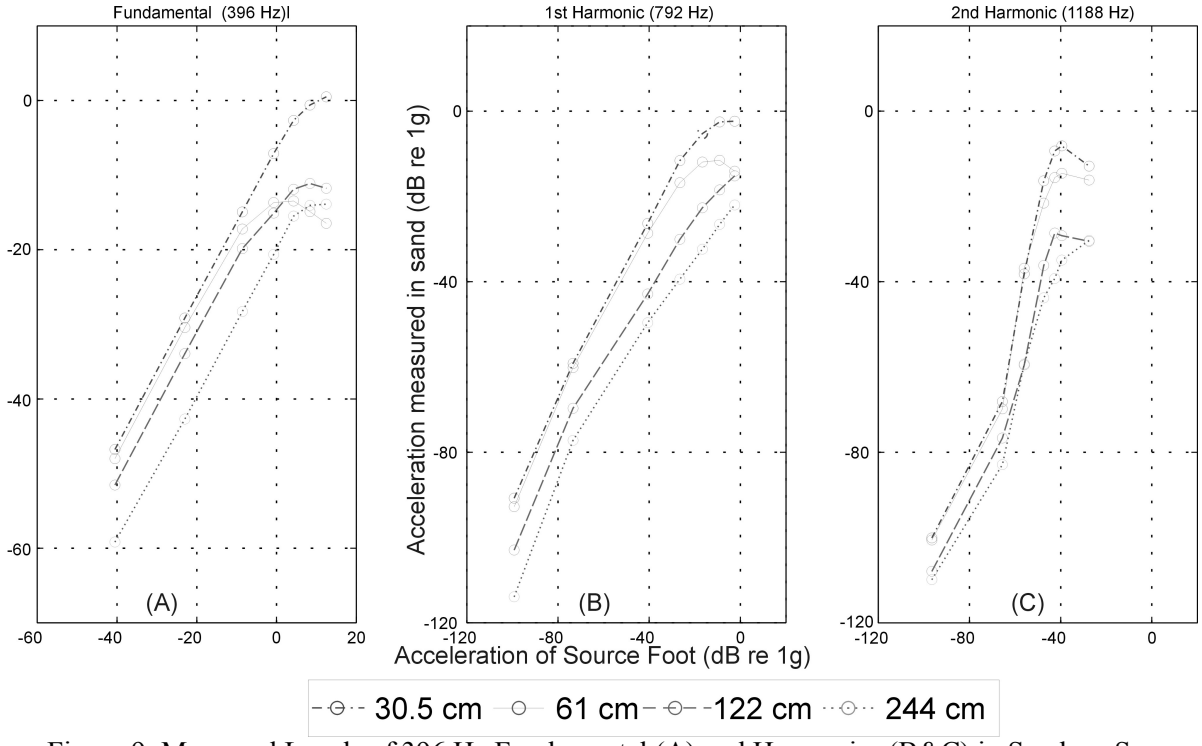


Figure 9: Measured Levels of 396 Hz Fundamental (A) and Harmonics (B&C) in Sand vs. Source

5. NON-LINEAR DISPERSION

Dispersion can be observed in measured surface waveforms in most of the data from both the laboratory model and the field site. It is apparent in the change in pulse shape as a function of location shown in Figures 3 and 4. The nature of this dispersion has significant implications for the post-processing of experimental data. Linear dispersion of Rayleigh waves, which can occur as a consequence of the vertical gradients in material properties, is fully accounted for in the pulse compression routine that generated the waveforms depicted in Figures 3 and 4. Non-linear dispersion, however, is not accounted for, since sum and difference frequencies occur at times dependent on pulse shape and cannot be accounted for with a simple impulse response. The differences between these two types of dispersion are illustrated by the laboratory results shown in Figure 10. Here, the acceleration was measured at four separate locations using a differentiated Gaussian pulse with a center frequency of 450 Hz as the original drive signal rather than reconstructing the pulse response in post-processing. What is depicted are raw data, although 1000 pulses have been averaged together to build up the signal-to-noise ratios for each measurement. The horizontal separation in Figure 10 indicates the four different measurement locations in the model at 30.5 cm, 61 cm, 122 cm, and 244 cm along a line away from the source. The magnitude of the offset is arbitrary. Propagation delays of 3 ms, 6 ms, 16 ms, and 31 ms have been subtracted before the waveforms were stacked so that all the data could be displayed in a single graph. Each vertical separation between traces indicates an increase of 3 dB in drive signal amplitude. The signals have all been normalized by their peak acceleration amplitude in order to emphasize changes in pulse shape over acceleration amplitude differences which are approximately three orders of magnitude for the data depicted.

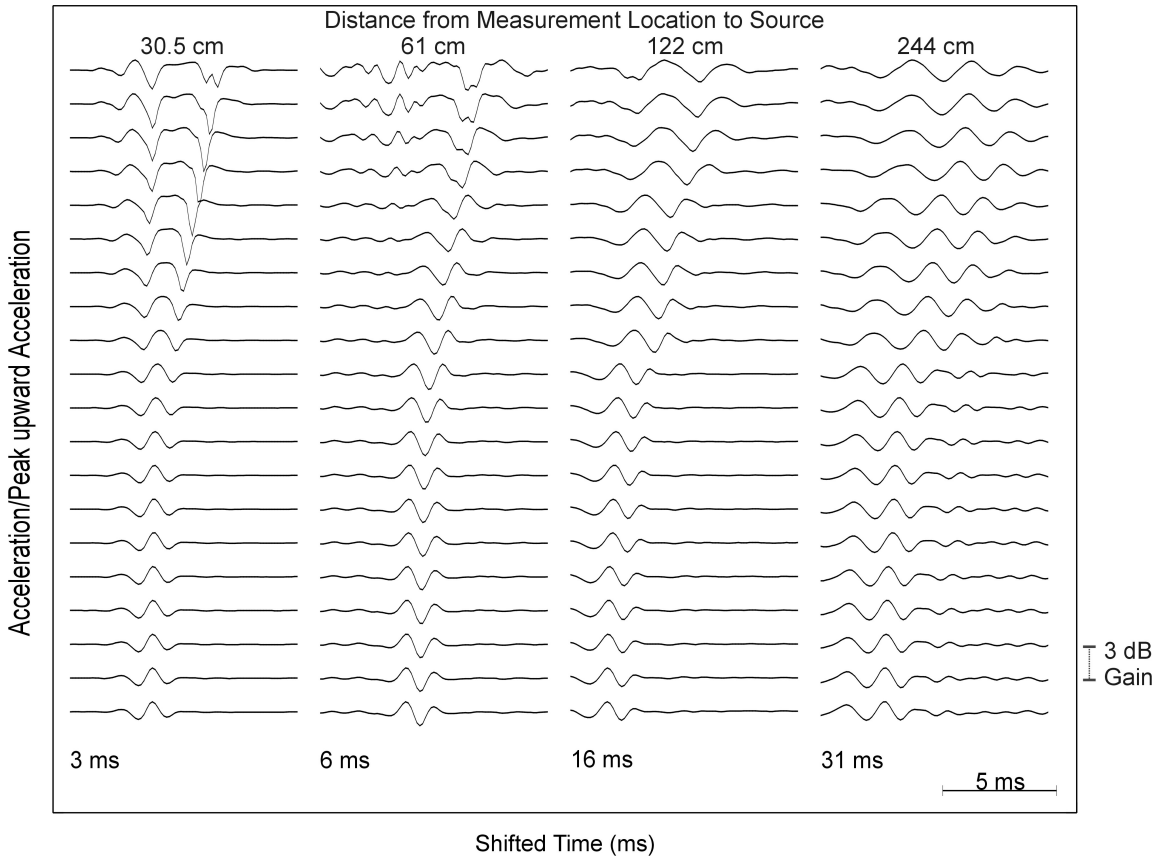


Figure 10: Changes in Shape of Rayleigh Wave Pulses with Increasing Drive Level and Propagation Distance

The apparent inversion of the signal between the first and second measurement points at very low drive levels is an indication of linear dispersion. Viewed as a function of position, the peaks and troughs of the signal

can be seen to propagate through the pulse envelope as this envelope propagates across the sand surface. The obvious change in pulse shape and the delays associated with the pulse arrival at high drive levels are the effect of non-linear dispersion. The troughs of the original pulse can be seen to form two downward acceleration impulses 30 cm from the source, similar to those depicted in Figure 7 for a continuous drive. These impulses separate in time as drive level is increased, much like the broadening of shock waves observed in air [11]. At greater distances there is, however, a preference toward not only the low frequency components of the original pulse, but also the energy that arrives late in time. This is particularly apparent in the data taken at 61 cm where the arrival of the pulse is visibly delayed with increasing drive. This inverse relationship between the group velocity and the drive level appears to have no precedent in the literature and will be a focus of further research in this area. The physical processes that mediate this phenomenon are not well understood. Its nature is further obfuscated by the fact that it was originally observed in completely linear pulse compression of an FM sweep in the non-linear regime and was believed to be an artifact of this clearly invalid processing scheme. This implies that the effect is mediated by a dependence of phase velocity on both drive amplitude and frequency, but not by the generation of harmonics, which are time gated out of the pulse compression. It may involve small spatial regions (slip planes or un-compacted zones) where stress concentrations cause local fluidization and lower the average shear strength along the propagation path in a way that recovery cannot occur on the timescale of the excitation. It is also possible that energy preferentially leaks out of the leading portion surface wave into the bulk of the medium. Since all bulk waves propagate with higher velocities than the surface wave, this form of energy loss might preferentially attack the leading edge of the waveform and create the impression of the pulse's slowing down.

6. CONCLUSIONS

Measurements have been conducted at a field test site and in a laboratory model that qualify the role played by system non-linearity in limiting the potential range and signal level that can be exploited by a seismic mine detection system. Non-linearity in both the source coupling and the propagation path have been observed and demonstrated to be possible signal limitations. Observed steepening and non-linear dispersion pose potential problems for the post-processing of data that may also limit useable signal levels. Work is currently underway to understand and quantify all of these effects and their potential impact on system operations.

7. ACKNOWLEDGEMENTS

This work is supported in part by the Office of Naval Research under contact number N00014-99-1-0995.

8. REFERENCES

1. Patel, D.L., "Handbook of Land Mines and Military Explosives for Countermine Exploitation," *US Army Belvoir RD&E Center Technical Report No. 2495*, March 1992.
2. Scott W.R., Jr., Larson, G.D. and Martin J.S., "Simultaneous Use of Elastic and Electromagnetic Waves for the Detection of Buried Land Mines," *Proceedings of the SPIE: 2000 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, Vol. 4038, April 2000.
3. Scott W.R., Jr., Martin J.S., and Larson G.D., " Experimental Model for a Seismic Landmine Detection System," *IEEE Trans on Geoscience and Remote Sensing*. To appear July 2001
4. "Feasibility of Acoustic Landmine Detection: Final Technical Report," *BBN Technical Report No. 7677*, May 1992.

5. Behboodian, A., Scott, W.R., Jr. and McClellan, J.H., "Signal Processing of Elastic Surface Waves for Localizing Buried Land Mines," *Proceedings of the 33rd Assilomar Conference on Signals, Systems, and Computers*, Assilomar, CA, October 1999.

6. Sabatier, J.M., Bass E.H., Bolen, L.N., and Attenborough, K., "Acoustically Induced Seismic Waves," *Journal of the Acoustical Society of America,* Vol. 80, no. 2, pp. 646-9, August 1986.

7. Westebbe, M., Bohme, J.F., Krummel, H., and Matthews, M.B., "Model fitting and testing in near surface seismics using maximum likelihood in frequency domain," *Conference Record of thirty-second Asilomar Conference on Signals, Systems and Computers*, Vol. 2, pp.1311-15, November 1998.

8. Bachrach, R, Nur, A., and Dvorkin, J., "High-Resolution Shallow-Seismic Experiments in Sand," "Part I" and "Part II," *Geophysics,* Vol. 63, No. 4, pg. 1225-40, July 1998.

9. Smith, P., Wilson P.S., Bacon F.W., Manning, J.F., Behrens, J.A., and Muir T.G, "Measurement and Localization of Interface Wave Reflections from a Buried Target," *Journal of the Acoustical Society of America*, Vol. 103, no. 5, pg. 2333-43, May 1998.

10. Lay T. and Wallace, T.C., "*Modern Global Seismology*." NY: Academic Press, 1995, pp. 252-63.

11. Beyer, R.T., "*Nonlinear Acoustics*." NY: Acoustical Society of America Press, 1997, pp. 91-109 and pp. 189-202.

12. Zabolotskaya, E.A., "Nonlinear Propagation of Plane and Circular Rayleigh Waves in Isotropic Solids," *Journal of the Acoustical Society of America*, Vol. 91, no. 5, pp. 2569-75, May 1992.

Use of High-Frequency Seismic Waves for the Detection of Buried Land Mines

Waymond R. Scott, Jr.[a], Seung-Ho Lee[a], Gregg D. Larson[b],
James S. Martin[b], and George S. McCall II[c]

[a]School of Electrical and Computer Engineering
[b]School of Mechanical Engineering
[c]Georgia Tech Research Institute
Georgia Institute of Technology
Atlanta, GA 30332-0250

ABSTRACT

Over the past three years a system has been under development at Georgia Tech that utilizes a seismic interrogation signal in combination with a non-surface-contacting, radar-based displacement sensor for the detection of buried landmines. Initial work on this system investigated the workability of the system concept. Pragmatic issues regarding the refinement of the current experimental laboratory system into a system which is suitable for field testing and, in turn, one which would be suited to field operations have been largely ignored until recently. Both field operations and realistic field testing require a system that is different from the original laboratory system in two crucial ways. One of these is that a field system needs a sensor standoff from the ground surface larger than the original 1 to 2 cm. This is necessary in order to account for small-scale topography, to avoid ground cover such as grass, and to minimize the risk to the operator. A second difference is that the scanning speed of a field system must be substantially greater than that of the original laboratory system, which takes several hours to image 1 m^2 of ground surface. From an operational standpoint, the reason for this is obvious. From an experimental standpoint, it is also important because ambient conditions are difficult to control on long time scales outdoors. Both of these new requirements must be met within the design parameters that were established empirically during the development of the laboratory system. One of these is that the system must be capable of measuring peak displacements as small as 1 nm with a 1 Hz resolution band in the low audio frequency range (30 Hz to 2 KHz). This is necessary because of limits placed by the medium on the amplitude of seismic signals that can be generated. Another parameter is that the measurement must be integrated over a region with a diameter of about 2 cm to 5 cm on the ground surface. This is necessary in order to preserve spatial structure of interest for the imaging of small mines. An array of radar-based displacement sensors configured with focused antennas meets all these requirements. A sensor was designed to test this concept. The new sensor was equipped with an antenna consisting of a conical corrugated horn and a bifocal dielectric lens that permitted a 20-cm surface-standoff distance. The sensor was tested in a laboratory experimental model in scenarios similar to those used to evaluate the original system sensor and was found to satisfy all of the new system requirements.

Keywords: land mine, mine detection, seismic waves, focused antenna, radar lens

1. INTRODUCTION

Seismic/elastic techniques for the detection of buried landmines have been under investigation by various groups for more than 30 years.[1,2] These are promising for the reliable detection of all types of buried mines, even low-metal anti-personnel mines. The reason for this is that mines have unique mechanical properties that can be used to distinguish them from soils and many forms of clutter. Typically, mine components are made from materials that are very stiff in comparison to the surrounding soil. They are also complex mechanical structures with a flexible case, a trigger assembly, air pockets, etc. This complex structure gives rise to structural resonances, non-linear interactions, and other phenomenology that is atypical for both naturally occurring and most types of man-made clutter. Using seismic techniques, this phenomenology can be exploited to distinguish mines from clutter.[1]

A seismic mine detection system has been developed at Georgia Tech that uses a radar-based displacement sensor for the local measurement of seismic displacements without physically contacting the soil surface. The non-contact nature of this sensor makes the system capable of interrogating the soil surface near or immediately above a mine.

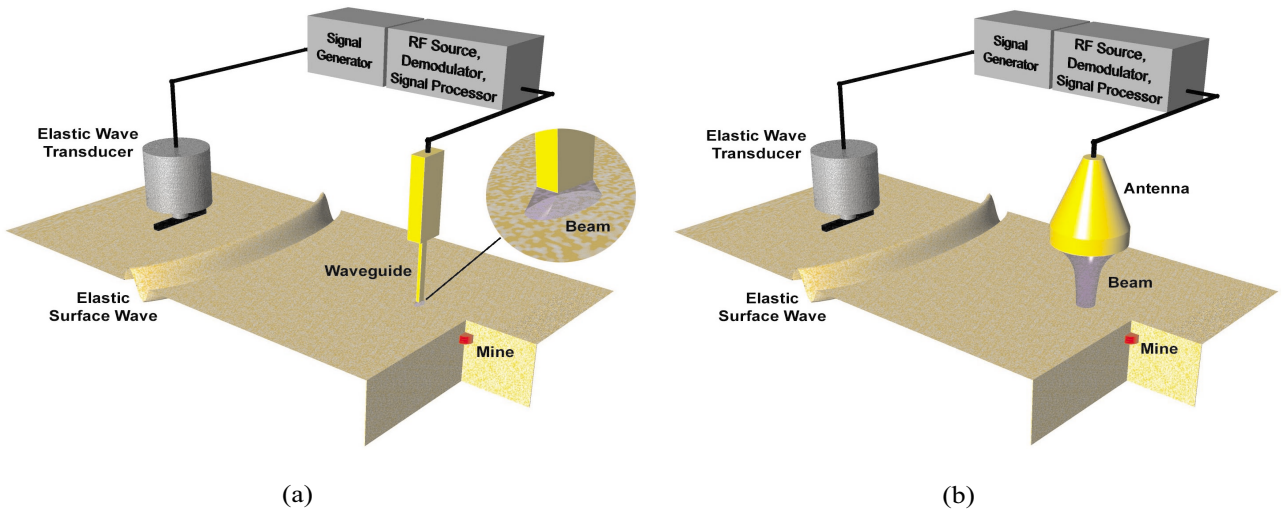


Figure 1. Configuration of laboratory experimental mine detection system: (a) original system and (b) new system.

This substantially increases the measurable effects of the mine's presence over schemes which rely on elastic waves scattered by the mine to propagate to a remote sensor location.[2] As it is normally configured, this system is depicted in Fig. 1(a). For comparison, the modified system that is currently under investigation is depicted in Fig. 1(b). In both configurations, the source (an electrodynamic shaker resting on a narrow foot) excites an incident signal composed primarily of a Rayleigh surface wave in the 30 Hz to 2 KHz frequency range. This travels across the minefield and excites motion of soil near the surface that includes the buried mines. Deeper soil is not excited by this signal because the Rayleigh wave motion decays exponentially away from the surface of the ground. The transmitted radar signal from the sensor antenna illuminates a spot on the ground surface above the minefield. The reflected radar signal from this spot is phase and amplitude modulated by the seismic surface motions. This modulation, which is proportional to the ground displacement integrated over the spot size, is extracted from the reflected radar signal by analog mixing with a reference signal after it is received. In operation, the sensor is translated to measure surface motion over an array of measurement locations. Sensor data are then stored and post-processed in both space and time domains to yield an image of the minefield.[3] In both of the configurations shown in Fig. 1, moving a single sensor forms a synthetic array. A physical array could also be used. This would provide a time saving in proportion to the number of sensors that populate the array. The original sensor (shown in Fig. 1(a)) utilizes an aperture element (an open-ended waveguide) as both a transmitting and a receiving antenna. This is effectively omnidirectional at the 8 GHz operating frequency of the radar sensor. The spot size is kept small (on the order of the 1-cm by 2-cm waveguide dimensions) because the waveguide is in close proximity to the ground (1cm to 2 cm above the surface). The new sensor (Fig. 1(b)) achieves a similar spot size using a focused antenna.

The mine detection system has been tested with the waveguide antenna in laboratory experiments and using numerical models for seismic propagation and scattering. Test conditions have included a variety of inert mines, mine-like targets, realistic clutter objects, and thin organic surface cover. Tests have also been performed in which realistic burial scenarios were simulated. These included trenching from recent mine burials, weathering, and multiple antipersonnel (AP) and antitank (AT) mines buried in close proximity. The results of these experiments, which have been reported in previous papers,[1,4] show that the system can be used for the detection and imaging of both AT and AP mines over a fairly wide range of realistic conditions. Typical images have shown a contrast between mines and clutter of 20 to 30 dB. Studies have been conducted with this sensor in which the separation of the waveguide and the ground surface was increased well beyond the 1 to 2 cm that was normally used.[5] In these experiments there was significant degradation in the spatial resolution of the image which could be formed from the experimental data. Digital beamforming techniques were shown to provide some improvement in the spatial resolution with large standoff distances. The problem was also mitigated somewhat by the fact that the image artifacts after beamforming were symmetrically distributed about the actual mine location. They could, therefore, be used to enhance the mine's image in post-processing, provided that there was an a priori knowledge that only one mine was present. Further

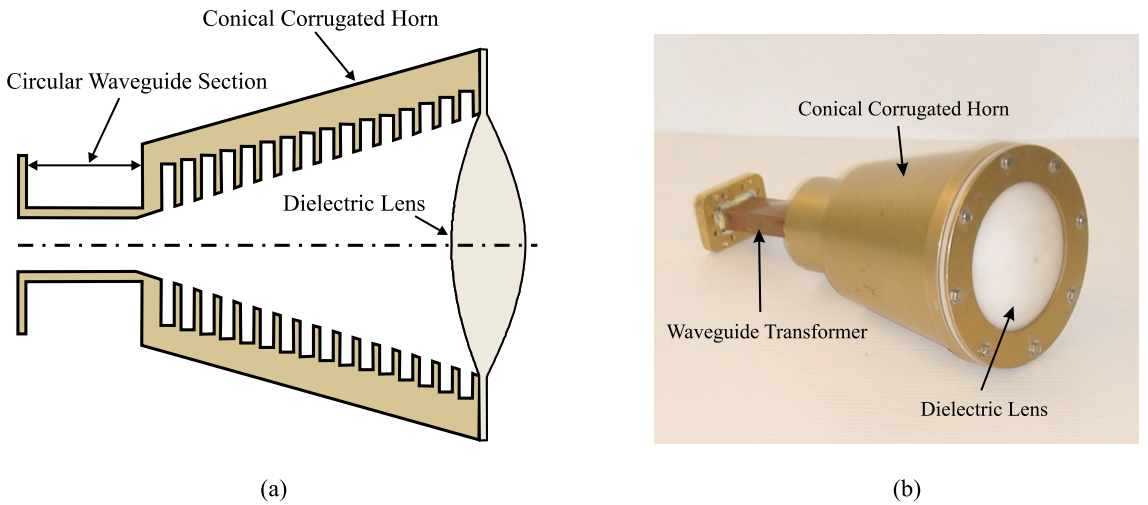


Figure 2. The FRD sensor antenna in (a) cutaway and (b) photograph.

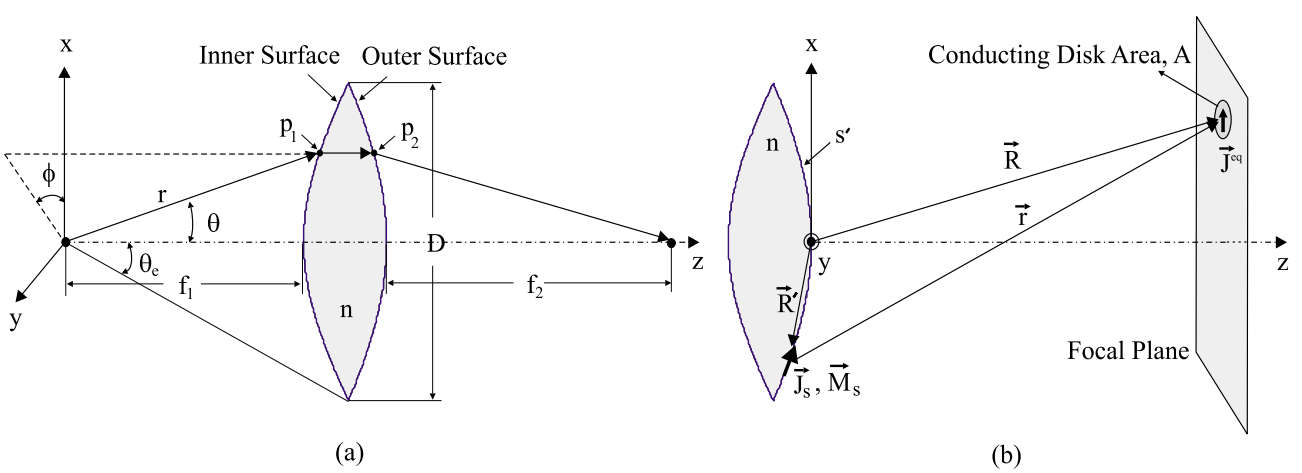


Figure 3. Geometry of computational model for FRD antenna lens: (a) lens geometry and (b) two-way beam pattern.

substantial improvements in image quality required an unrealistically accurate model for the existing sensor, an impracticably large array of sensors, or an array of more directional sensors. There are many ways in to construct a directional sensor. These include the use of ellipsoidal reflectors, passive microstrip antennas, slotted waveguide antennas, or focusing dielectric lenses. The last of these designs is the one that was pursued and is reported in this paper. It is expected that similar results could be achieved with other design concepts, provided that they were correctly implemented.

2. SENSOR DESIGN

Figure 2 depicts the design of the antenna for the experimental focused radar displacement (FRD) sensor. The antenna consists of a dielectric lens with foci at the surface of the ground and at the open end of a circular waveguide. The lens and the waveguide are connected by a conical metal horn with a corrugated interior. The FRD sensor is monostatic, as its predecessor was, so that this antenna functions as both a transmitter and a receiver.

A theoretical model was constructed for the function of the new antenna. The geometry for this model is shown in Fig. 3. Here the dielectric lens with a refraction index n and a diameter D has two focal lengths, f_1 and f_2, the inner and the outer focal length, respectively. The diameter of the lens is confined by the semi flare angle (θ_e) of the horn. The inner focal point is assumed to be the same as the phase center of the horn. The inner and outer contours of the lens are hyperbolic curves. The lens profile is rotationally symmetric and is a function of D, f_1, f_2, and n. The geometry of the inner surface of the lens is defined as[6]

$$x = r\sin\theta, \qquad 0 \le \theta \le \theta_e, \tag{1}$$

$$z = \frac{a_1 n + \sqrt{(a_1 n)^2 - (n^2-1)(a_1^2 - x^2)}}{n^2 - 1}, \tag{2}$$

where $a_1 = (n-1)f_1$.

The inner surface of the lens is illuminated by a spherical wave front emanating from the horn. The horn is excited by the TE_{11} mode propagating in the circular waveguide shown in Fig. 2. The TE_{11} mode is converted into the hybrid HE_{11} mode by the discontinuity between the circular waveguide and the horn and facilitated by the corrugations in the horn.[6] The electric field on the inner surface of the lens can then be described in terms of the geometric parameters by:

$$\vec{E}(\theta,\phi,r) = \left(\frac{P_\nu^1(\cos\theta)}{\sin\theta} + \frac{dP_\nu^1(\cos\theta)}{d\theta}\right)(\cos\phi\hat{\theta} - \sin\phi\hat{\phi})\frac{e^{-jk_0 r}}{r}, \tag{3}$$

where $P_\nu^1(\cos\theta)$ is the associated Legendre function of the first kind.[7] ν is determined by the solution of the equation:

$$\frac{P_\nu^1(\cos\theta_e)}{\sin\theta_e} + \frac{dP_\nu^1(\cos\theta_e)}{d\theta_e} = 0. \tag{4}$$

The field on the outer surface of the lens is found by ray tracing. First order reflections from both inner and outer surfaces are considered in this computation by using the Fresnel reflection coefficient. The electric and magnetic surface current densities ($\vec{J}_s$ and $\vec{M}_s$) on the outer surface can then be determined by an application of boundary conditions and from these the radiation pattern of the lens can be found. The electric and magnetic fields at any point on the focal plane may be represented in terms of the lens's outer surface current densities by the following equations:

$$\vec{E}(\vec{R}) = \frac{-j\omega\mu_0}{4\pi}\iint_{s'}[\vec{J}_s\Psi + \frac{1}{k_0^2}(\vec{J}_s\cdot\nabla')\nabla'\Psi)]ds' - \frac{1}{4\pi}\iint_{s'}(\vec{M}_s\times\nabla'\Psi)ds', \tag{5}$$

$$\vec{H}(\vec{R}) = \frac{-j\omega\epsilon_0}{4\pi}\iint_{s'}[\vec{M}_s\Psi + \frac{1}{k_0^2}(\vec{M}_s\cdot\nabla')\nabla'\Psi)]ds' + \frac{1}{4\pi}\iint_{s'}(\vec{J}_s\times\nabla'\Psi)ds', \tag{6}$$

where $\Psi = \frac{e^{-jk_o|\vec{r}|}}{|\vec{r}|}$ and $\nabla' = \frac{\partial}{\partial x'}\hat{x} + \frac{\partial}{\partial y'}\hat{y} + \frac{\partial}{\partial z'}\hat{z}$. Here, x', y', and z' are defined by the coordinate system of the source. The radiated power pattern of the lens-focused corrugated horn is then given by $|\vec{E}(\vec{R})|^2$.

Rather than the radiation pattern of the antenna, it is the two-way beam pattern involving the reflection from a scatterer back to the antenna that is of interest for the determination of the FRD sensor's spot size. This is because the antenna functions as both the radar transmitter and the receiver for the sensor. In order to compute this two-way pattern, a small conducting disk of area A is placed on the focal plane at the location $\vec{R}$. The surface current on the disk is approximated from the magnetic field defined by Eq. 6. When the disk is electrically small, the electric field at the antenna radiated by the disk is then

$$\vec{E}^r(\vec{R}) = \frac{-j\omega\mu_0}{4\pi}[\vec{J}^{eq}\Psi + \frac{1}{k_0^2}(\vec{J}^{eq}\cdot\nabla)\nabla\Psi)]A, \tag{7}$$

where $\vec{J}^{eq}(\vec{R}) = -2\hat{z}\times\vec{H}(\vec{R})$, $\Psi = \frac{e^{-jk_0|\vec{R}|}}{|\vec{R}|}$, and $\nabla = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}$.

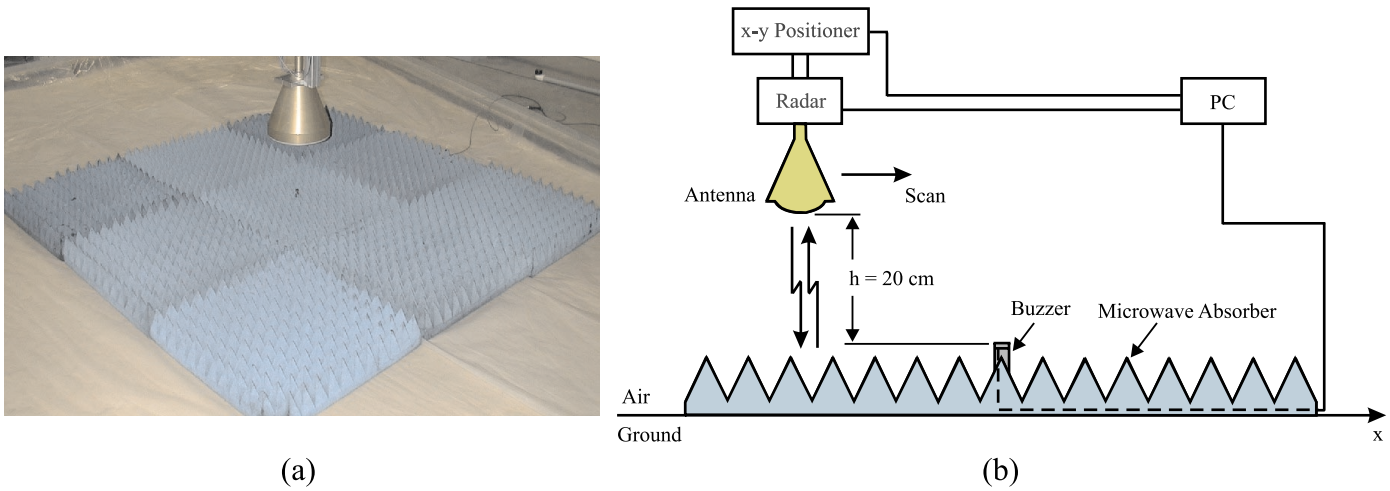


Figure 4. FRD sensor directivity measurement: (a) photograph and (b) schematic.

The voltage received by the antenna from each of the secondary sources is found to be:

$$V^r(\vec{R}) = \vec{E}^r(\vec{R}) \cdot \vec{l}^e(\vec{R}), \tag{8}$$

where $\vec{l}^e$ is the effective height of the antenna, which can be determined by the antenna's radiated field as computed in Eq. 5 to be[8]:

$$\vec{l}^e(\vec{R}) = \alpha \frac{4\pi|\vec{R}|\vec{E}(\vec{R})}{-j\eta_0 k_0 e^{-jk_0|\vec{R}|}}, \tag{9}$$

where α is a proportional constant that depends on the details of how the antenna is fed. The two-way pattern is then obtained from $|V^r(\vec{R})|^2$.

If the lens were illuminated by a spherical wavefront with uniform amplitude, strong side lobes would be expected in its two-way beam pattern. The corrugations on the interior of the conical horn were designed so as to ameliorate these by tapering the amplitude of the electric field near the edges of the lens. The corrugations were designed using standard procedures for this type of device.[6] The width of the corrugations is less than $\frac{\lambda_o}{2}$. The depth of the slots is between $\frac{\lambda_o}{4}$ and $\frac{\lambda_o}{2}$. The slot depths taper from $\frac{\lambda_o}{2}$ near the throat of the horn to $\frac{\lambda_o}{4}$ at the aperture with the variation occurring on a scale which is large compared to the wavelength. As it was designed and built, the prototype antenna had the same inner and outer focal distances, both of which were 20 cm. This created a plane of symmetry for the lens and provided some fabrication convenience but was not required for the design.

3. SENSOR EVALUATION

In order to measure the spot size of the FRD sensor, an experiment was conducted in which a small piezoelectric buzzer was placed in the center of a 1.8 m by 1.8 m anechoic surface. The buzzer was driven at 1KHz, and the sensor output was recorded at this frequency using a two-second-integration time to build dynamic range above the noise floor. The antenna was then translated linearly across the surface. From a system standpoint, an ideal sensor under these circumstances would have measured only the displacement of the buzzer when it was directly above the buzzer and zero elsewhere. The anechoic treatment assured that the measurement could not be contaminated by ambient vibration and by single and multiple reflected signals from the ground and the buzzer. The configuration of this experiment is depicted in Fig. 4.

The results of the directivity experiment are shown in Fig. 5. The 3dB beam width on the ground was found to be about 4.7 cm. The beam pattern was well predicted by theory for horizontal sensor offsets up to about 15 cm. Beyond 15 cm the FRD sensor consistently outperformed model predictions. This result is difficult to explain. The antenna may actually perform better than predicted due to inaccuracies in the model, or it may be an inaccuracy in the measurement. The measurement inaccuracy could be due to abnormal motion of the buzzer, such as horizontal

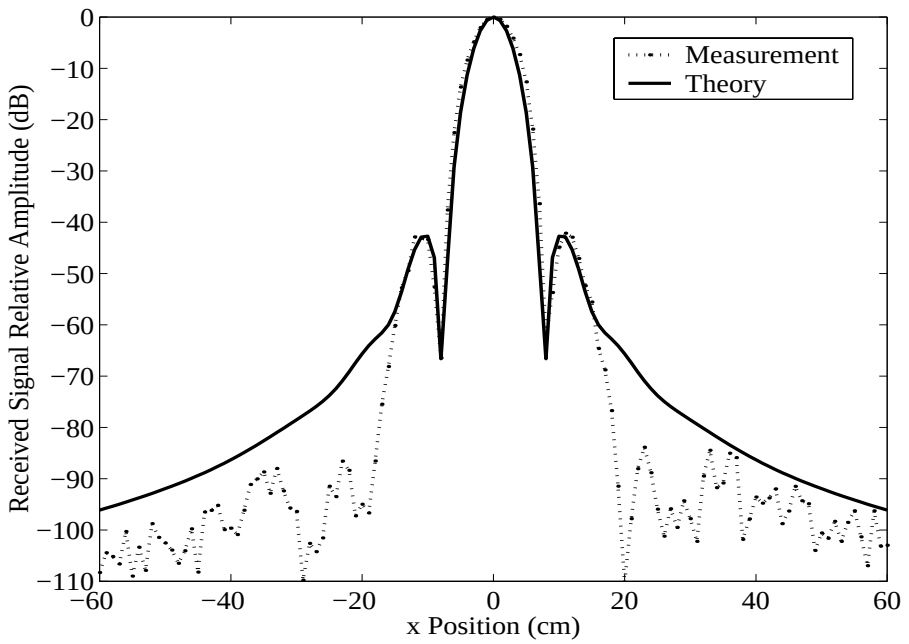


Figure 5. Results of FRD sensor directivity measurement.

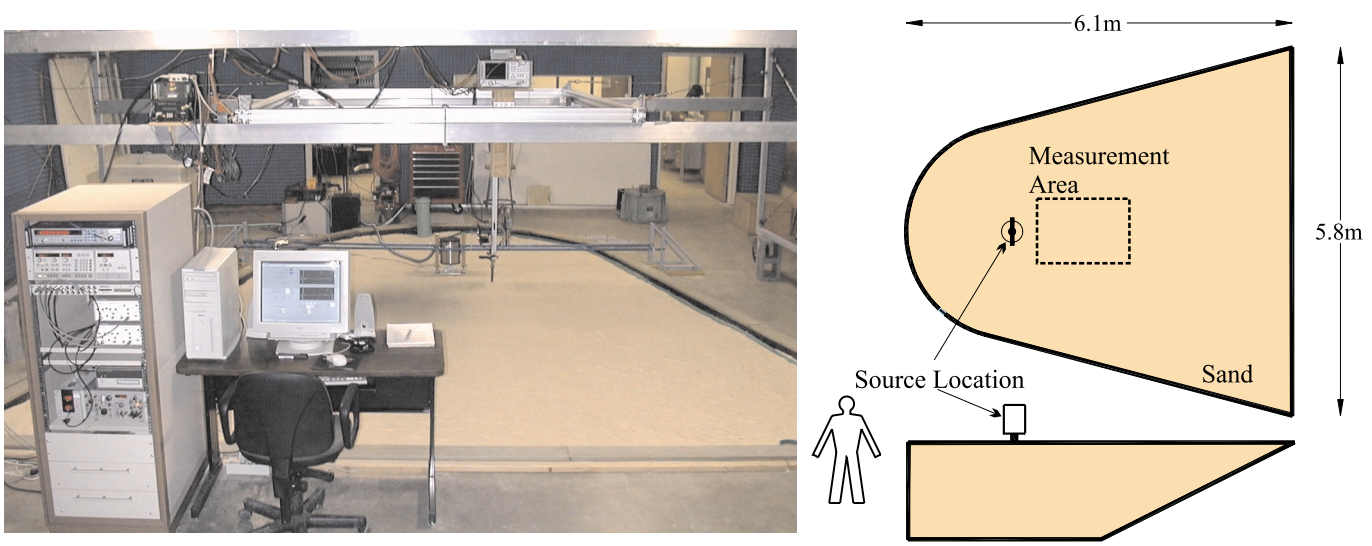


Figure 6. The laboratory experimental model.

motion of its case, which adds destructively with the motion sensed from the radiating surface on the top of the buzzer. Since the FRD sensor only responded to motion in the direction of the EM propagation path, it would be insensitive to this lateral motion when it was directly above the buzzer. The relative contribution of this motion to the sensor output would have increased as the antenna was translated away from the buzzer, and the EM path became more horizontal.

4. LABORATORY EXPERIMENTS

The FRD sensor was tested in the laboratory model that was used for the development of the mine detection system. This model is depicted in Fig. 6. It is a wedge-shaped tank filled with over 50 tons of damp compacted sand to simulate soil. The seismic source is located near the tip of the wedge and is bi-directive toward the search area and the back wall. Simulated mines, inert mines, and clutter, such as rocks and sticks, are buried within a 2-m-by-2-m region in the center of the tank. The FRD sensor was scanned above this region with a three-degree-of-freedom positioner fixed to a framework above the tank. This was done using a typical two-dimensional scan that covers an 80-cm by 120-cm boundary-inclusive area. It is sampled in 1-cm increments along the direction of seismic wave propagation, from 30 cm to 150 cm out from the source. 2-cm increments are sampled across the seismic wave fronts (out to 40 cm on either side of the plane of symmetry, apparent in the top view of the model in Fig. 6). A total of 4,961 measurement locations were sampled in this scan before an image was formed. These scan parameters were arrived at during work with a smaller experimental model. The search dimensions have not been increased in most of the work with the larger model in order to save time. Since there is currently a four-second dwell over each measurement location used to build up the signal-to-noise ratio in the received signal, even the current 1 m^2 scan requires 7 hours to complete. In addition to building an array of sensors, reducing the dwell time and increasing the scan increments are being considered to speed a field-operable system. The 1 cm and 2 cm scan increments used in the measurements are clearly smaller than necessary since the antenna averages the displacement over a region of the ground with a diameter of about 4.7 cm. These scan increments were used to make comparisons with previous data easier.

Four-second displacement time histories are acquired with 8-KHz sampling and 12-bit resolution for each measurement location and stored as frequency domain transfer functions between the source's drive signal and the sensor's received signal. This must ultimately be reduced to a single image which depicts the probability of a buried mine below each location on the planar surface. There are many ways in which this sort of post processing can be performed. The scheme that is currently employed has been described in a previous paper.[3] The technique involves a multi-step process that exploits observed features of mine response, such as localized resonance effects, to enhance the background contrast in the final image. First the response of the system to a differentiated Gaussian pulse, which typically has a center frequency of 450 Hz, is synthesized from the recorded transfer functions. The data are then filtered temporally and spatially to remove low and high frequency noise. The forward-propagating waves (waves travelling away from the source) are then filtered out in k-space, leaving only the back-scattered waves. Two data sets are created from the remaining time histories. The first data set is the energy at each spatial position in the back-scattered waves at times of arrival near the incident wave. The second data set is the energy that propagates back toward the seismic source from each spatial position. Taking the product of these two data sets forms the final image. Nothing about the displacement sensing mechanism is presumed in this scheme other than that the measurements represent true point displacements. Thus, no modifications to the imaging scheme were made to accommodate the FRD sensor. This provided the additional advantage of allowing direct comparison with existing images in the evaluation of sensor performance.

5. AP MINE DETECTION AND IMAGING

Two mine-imaging experiments were performed to evaluate the FRD sensor. In the first of these a single TS-50 AP mine was buried with its top-most surface 1 cm below the sand in the center of the scan region. The sand surrounding the mine was carefully repacked after the burial to simulate the effects of weathering. This is a baseline configuration, which has been tested many times with other sensor types and in other experimental models for comparison purposes.[1,9] The minefield was then covered with a 15-cm layer of pine straw to simulate realistic organic ground cover. This is similar to earlier experiments in which 2-cm to 3-cm layers of pine straw were used in testing with the waveguide antenna. The thickness of pine straw covering the scan region was increased to 15 cm because the greater standoff of the FRD sensor permitted this. Figure 7 shows a comparison between a scan performed with the waveguide antenna 1 to 2 cm high (Fig. 7(a)) and the two scans performed with the FRD antenna 20 cm high: one without the pine straw (Fig. 7(b)) and one with the pine straw (Fig. 7(c)). It is clear from the data presented in the figure that the FRD sensor images this relatively small AP mine with an image contrast comparable to the sensor that is equipped with the waveguide antenna. Unlike earlier attempts to implement digital beamforming arrays of omnidirectional sensors,[5] the image shows no artifacts that are remote from the actual mine location. However, the spatial resolution is somewhat less than with the waveguide antenna; this is expected since the FRD antenna averages the displacement over a somewhat larger area than does the waveguide antenna. In either

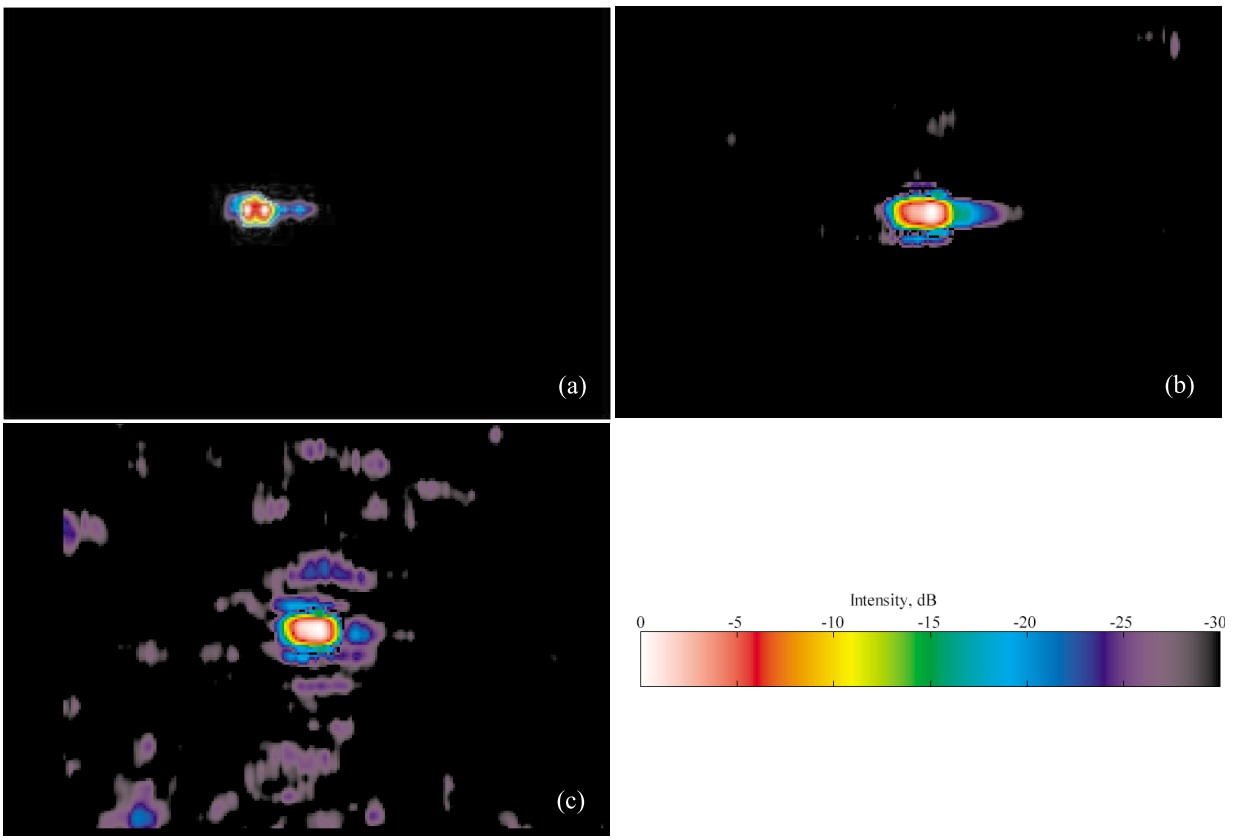


Figure 7. Images formed from two-dimensional scans above a TS-50 mine buried 1cm deep: (a) using the waveguide antenna, (b) using the FRD sensor over a clean surface, and (c) using the FRD sensor over 15 cm of pine straw.

event, the image shows more than adequate spatial resolution for operational purposes when one considers the 120-cm by 80-cm scale on which it is depicted. Further research is planned to determine the appropriate spatial resolution for the sensor: higher spatial resolution yields a higher fidelity image but at the cost of slower measurement times because the higher spatial resolution will necessitate smaller scan increments. The ground cover seems to introduce some clutter into the image, but this is more than 25 dB below the level of the mine image. The source of this clutter is not currently known: It may be due to pine straw or it may be due to an increase in environmental noise for this particular measurement. Previous attempts to quantify the effects of ground cover were confounded by the small standoff distance of the waveguide sensor.[1] There is, therefore, no appropriate data for comparison with the FRD sensor's performance. The fact that this measurement was even possible makes a favorable statement regarding the FRD sensor's design.

6. MULTIPLE MINE IMAGING

The most challenging burial scenario that can be simulated in the experimental model is the detection of both AP and AT mines in close proximity to each other and in the presence of buried clutter. This is a realistic scenario, since it is a common practice of mine warfare to plant multiple AP mines in close proximity to AT mines. The AP mines thereby protect the AT mine from sappers who can more easily detect the larger object and remove it with little personal danger. This poses a unique detection problem in that it requires a system to operate with sensitivity appropriate to both mine types simultaneously. Also, the system must be capable of distinguishing individual targets and rejecting ghost images formed by multiple-scattering effects. This scenario has been previously investigated using the waveguide antenna, which was found to offer adequate spatial resolution to image all of the buried mines with a reasonable set of parameters selected for the installation.[1] The FRD sensor was tested over a similar burial arrangement. The layout of this experiment and the relative scale of the buried objects can be seen in Fig. 8. The

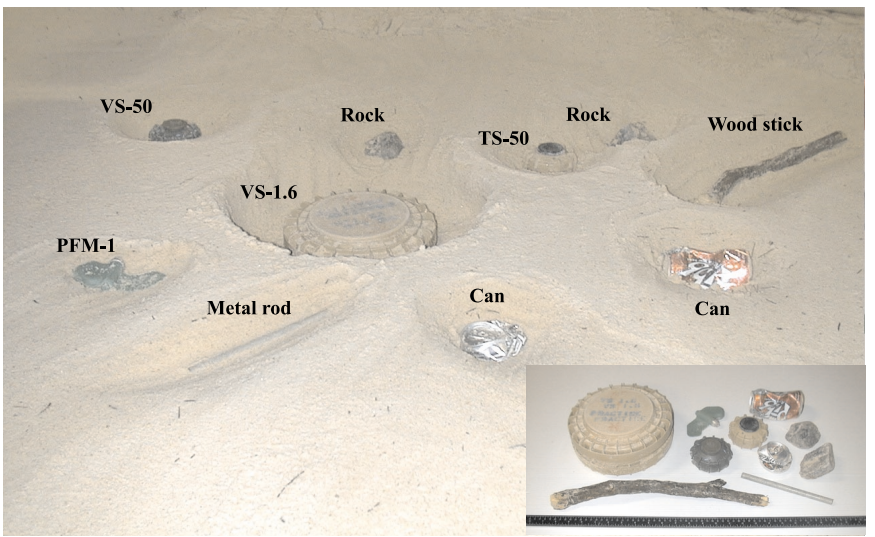


Figure 8. Photograph showing the multiple mine burial uncovered and the relative scale of the buried objects (insert).

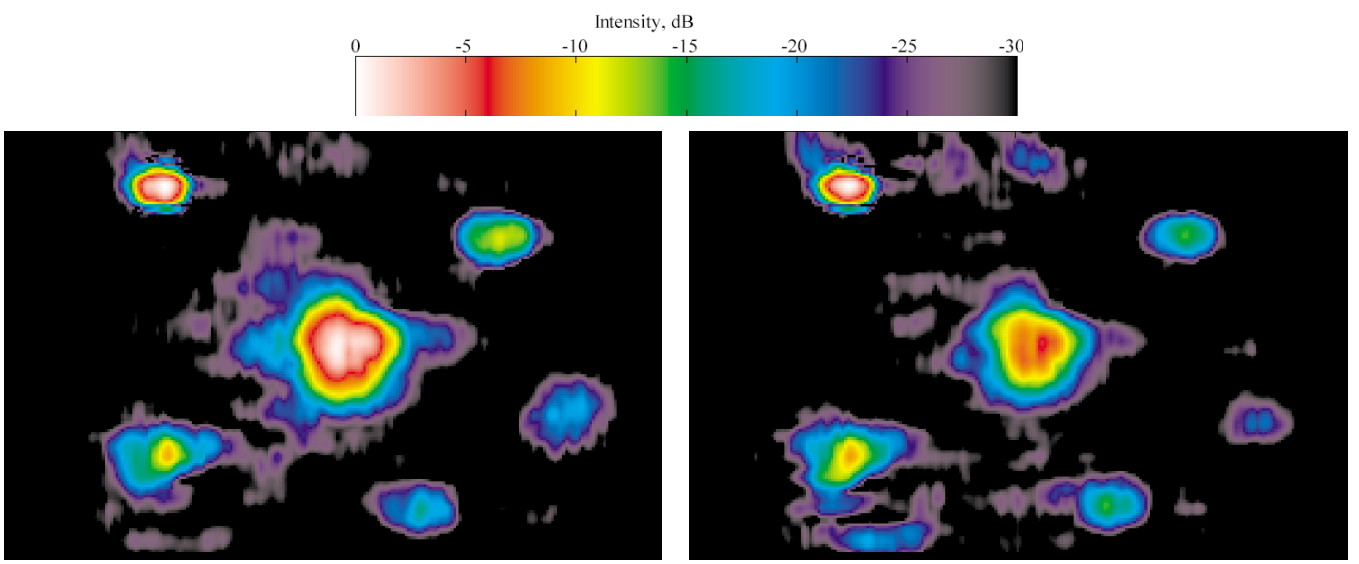


Figure 9. Images formed of multiple mine burial using the FRD sensor: (a) clean scan and (b) with 15 cm of pine straw.

AT mine was a VS-1.6. Three AP mines surrounded it: a TS-50, a VS-50, and a PFM-1. The clutter objects included two AP mine-sized rocks, two aluminum beverage cans, a steel rod, and a large stick. As before, the minefield was covered with a layer of pine straw 15 cm thick, and the sand was carefully packed around all of the buried objects. Figure 9 shows the image that was formed of this burial with and without the 15 cm of pine straw. All of the mines are clearly visible in this image. Note that the image with the pine straw is comparable to that without the pine straw for this case: it does not have the increased clutter visible in Fig. 7. In addition to the mines, the beverage cans also show up well in the image. This is because they are air-filled and therefore exhibit low-frequency structural resonances similar to those that identify the mines. This type of clutter could easily appear as a mine to the current seismic imaging algorithm much as it would fool an individual using a metal detector. In contrast, the steel rod

would also have fooled the metal detector but is not imaged using the seismic system. This provides an illustration of how multiple sensor types might be exploited for the reduction of false alarms in a field-operational system.

7. CONCLUSIONS

A seismic mine detection system has been tested in a laboratory experimental model. The system is similar to one that has been described in previous papers[1,3,4,9] but was outfitted with a focused-radar displacement (FRD) sensor. The FRD sensor replaced an earlier sensor equipped with an open-ended waveguide antenna that needed to be located 1 to 2 cm above the ground in order to achieve sufficient spatial resolution for mine detection. The FRD sensor, which is based on a dielectric lens and corrugated horn design, allows the system to stand off 20 cm from the ground surface while achieving detection and imaging performance comparable to the earlier design. This sensor addresses the standoff requirements of a field operable system. In conjunction with other planned hardware and software improvements, an array of these sensors should also address the scanning speed requirements of such a system. Future plans for the development of the system involve building a small array of these sensors and transitioning the experimental work into outdoor testing.

ACKNOWLEDGMENTS

This work was supported in part by the US Army Night Vision Electronic Systems Directorate, S&T Division, Countermine Technology Team and by the OSD MURI program by the US Army Research Office under contract DAAH04-96-0448.

REFERENCES

1. W. R. Scott, Jr., J. Martin, and G. Larson, "Experimental model for a seismic landmine detection system," *IEEE Transactions on Geoscience and Remote Sensing* , to appear July 2001.
2. *Feasibility of Acoustic Landmine Detection: Final Technical Report*, BBN Technical Report NO. 7677, May 1992.
3. A. Behboodian, W. R. Scott, Jr., and J. H. McClellan, "Signal processing of elastic surface waves for localizing buried land mines," *Proceedings of the 33rd Assilomar Conference on Signals, Systems, and Computers* , Assilomar, CA, October 1999.
4. W. R. Scott, Jr., G. Larson, and J. M. and, "Simultaneous use of elastic and electromagnetic waves for the detection of buried land mines," *Proceedings of the SPIE: 2000 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls* , Orlando, FL, Vol. 4038, pp. 667-678, April 2000.
5. S. H. Lee and W. R. Scott, Jr., "Near-field beamforming array for detecting elastic waves in the earth," *Proceedings of the SPIE: 2000 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls* , Orlando, FL, Vol. 4038, pp. 691-699, April 2000.
6. Y. T. Lo and S. W. Lee, *Antenna Handbook: Antenna Theory*, Van Nostrand Reinhold, New York, 1993.
7. G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, Academic Press, 1995.
8. C. A. Balanis, *Antenna Theory: Analysis and Design*, John Wiley and Sons, 1997.
9. W. R. Scott, Jr. and J. Martin, "Experimental investigation of the acousto-electromagnetic sensor for locating land mines," *Proceedings of the SPIE: 1999 Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls* , Orlando, FL, Vol. 3710, pp. 204-214, April 1999.

Proceedings of SPIE, Vol. 4394, April 2001

Resonance Behavior of Buried Land Mines

Christoph T. Schröder and Waymond R. Scott, Jr.

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250
USA

ABSTRACT

A three-dimensional finite-difference time-domain model for elastic waves in the ground has been developed and implemented on a massively parallel computer. The numerical model has been developed as part of a project in which elastic and electromagnetic waves are used synergistically to detect buried land mines. The numerical model is used to study the interaction of elastic waves with buried land mines. As a first approach, a simple model for a TS-50 antipersonnel mine has been developed, and the interaction of elastic waves with the buried mine has been investigated. In both experimental results and numerical simulations, resonant oscillations occur at the location of the buried mine. To further explore the resonant behavior of a buried land mine, a refined mine model has been developed that includes more details of the actual mine. Using the refined mine model, the nature of the resonance is explained, and the parts of the mine that influence the resonant oscillations are identified. Results are presented which describe the resonance as a function of burial depth and soil parameters.

Keywords: land mine, elastic, acoustic, FDTD, finite-difference, resonance

1. INTRODUCTION

At the Georgia Institute of Technology, a system is being developed in which buried land mines are located by using both elastic (acoustic) and electromagnetic waves.[1,2] In this system, elastic waves are excited that interact with a buried land mine. The waves cause the mine and the surface above the mine to vibrate. An electromagnetic radar records the vibrations and, thus, detects the mine. To investigate the characteristics of elastic wave propagation in the ground and to explore the elastic response of buried land mines, a numerical model has been developed.[3,4]

In experimental studies, resonant oscillations have occurred at the location of a buried land mine, when the mine is excited by elastic waves. These resonant effects have been observed for mines of different types, shapes and sizes, including antipersonnel and antitank mines. The resonance significantly enhances the signature of a buried mine and clearly distinguishes it from surrounding clutter objects. In this paper, the resonant behavior of buried antipersonnel mines will be described.

To compute the interaction of elastic waves with a buried land mine on a large scale, a simple model for an antipersonnel mine has been used. Although the model is simple, it predicts resonant oscillations at the location of the buried land mine that are qualitatively very similar to the ones observed in experiments. To investigate the resonant behavior in more detail, a refined mine model has been developed, which more closely resembles the structure of an actual TS-50 antipersonnel mine. The detailed mine model incorporates the parts of the mine which are believed to have an impact on the elastic behavior of the mine. With this model, the characteristics of the resonance can be explored.

In the subsequent section, Sec. 2, the numerical model is explained briefly. In Sec. 3, results obtained with the numerical model are described. First, the interaction of elastic waves with a buried mine is depicted. Then, the resonant behavior of a buried land mine is explained. For this, a parametric study is conducted in which the resonance is studied as a function of the soil properties and the burial depth of the mine.

Electronic Mail and Telephone: christoph.schroeder@ece.gatech.edu, 404-894-3123, waymond.scott@ece.gatech.edu, 404-894-3048

2. THREE-DIMENSIONAL NUMERICAL MODEL

In the course of the development of the elastic/electromagntic sensor, experiments have been performed with land mines buried in a large sand-filled box.[1,2] In the experiments, elastic waves are launched by an electrodynamic transducer placed on the surface of the ground. The waves propagate along the surface and interact with the buried land mines. To investigate the characteristics of the elastic wave propagation in the ground, a numerical model has been developed.

2.1. Finite-Difference Model

A schematic drawing of the model is shown in Fig. 1 (a). The numerical model is a three-dimensional finite-difference model that computes the elastic wave fields in the time-domain. For the model, the ground is approximated as a linear, isotropic, lossless medium. A first-order formulation in terms of the particle velocity and the mechanical stress is employed. The governing partial-differential equations are discretized by central finite differences, and the characteristic finite-difference grid is introduced. In the finite-difference grid, the field components are not known at the same points in space and time, but they are staggered, as indicated in Fig. 1 (b). The air-ground interface is modeled by a free-surface boundary condition. The solution space is surrounded by a *Perfectly-Matched Layer* (PML) absorbing boundary condition, which absorbs the outward traveling waves and, thus, prevents artificial reflections at the outer grid faces. A detailed description of the numerical model can be found in the literature.[4]

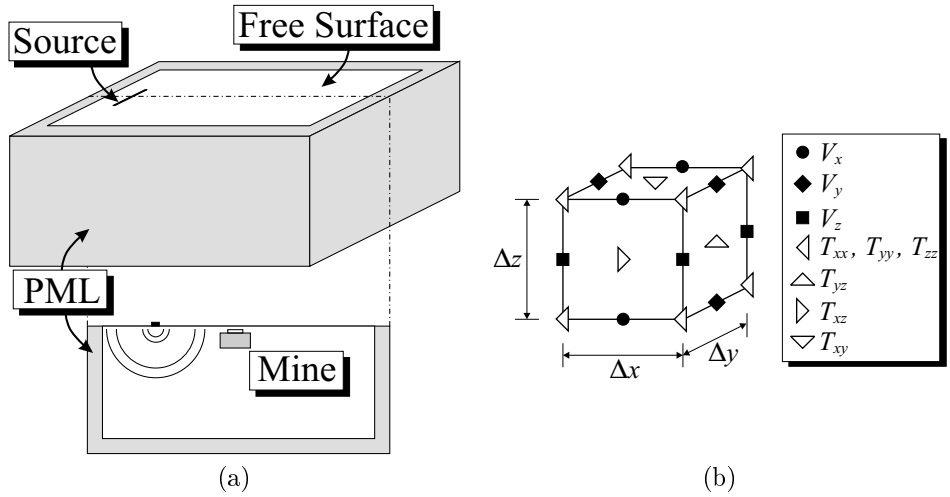


Figure 1. Three-dimensional finite-difference model; (a) lay-out, (b) finite-difference cell.

2.2. Material Properties

In the experiments, the mines are buried in a large sand-filled box. If sand is assumed to act in a linear and isotropic manner, its material properties can be described by three independent quantities: the material density ρ, the pressure wave speed c_p, and the shear wave speed c_s. In order to model the sand and its behavior accurately within the numerical model, these quantities have been measured as a function of depth for the sand box used in the experiments.

The material density, the pressure wave speed and the shear wave speed as a function of depth are shown in Fig. 2. All three quantities have been found to be depth dependent. The material density has a value of roughly 1400 kg/m^3 at the surface and increases to about 1600 kg/m^3 at a depth of half a meter. The variation in the material density is due to changes in the water content in the soil. At the surface, the sand is fairly dry. Within the ground, the water content increases steadily, until the water table is reached at a depth of about half a meter. The pressure wave speed and the shear wave speed are both also found to be depth-varying. At the surface, a layer with fairly slow wave speeds exists. Beyond this layer, the pressure wave speed and the shear wave speed increase rapidly. At a depth of half a meter, the pressure wave speed and the shear wave speed have values of 210 m/s and 130 m/s,

respectively.* The depth variation of the wave speeds is influenced by several mechanisms. The cohesion between the grains of the sand, for example, is affected by the water content, because the water binds the grains of the sand. The cohesion is also altered by the pressure in the ground, which, due to gravity, increases with depth. The increase in the cohesion will cause an increase in the wave speeds. On the other hand, the increase in the material density with depth would, if the stiffness was constant, reduce the wave speeds. It must be noted that, although the measured depth profiles are fairly good approximations of the actual depth profiles in the soil, they are only approximations and the actual depth profile might be more complex.

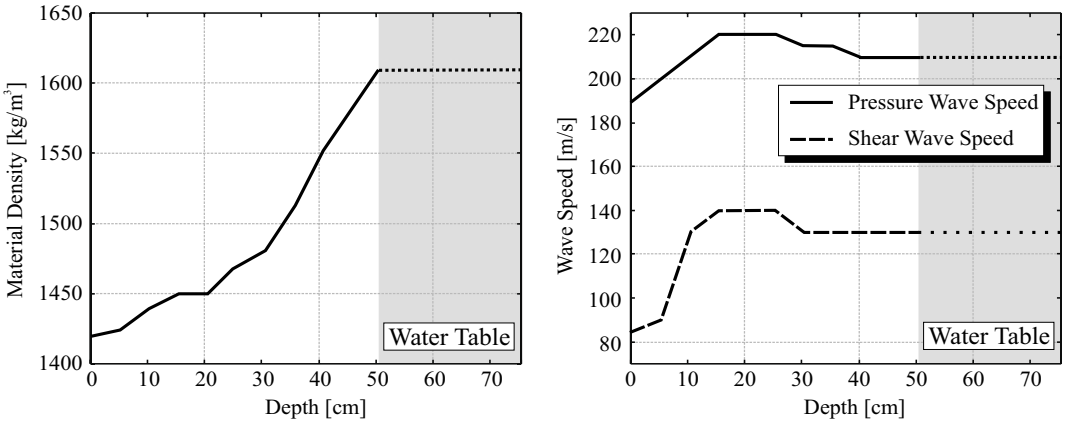


Figure 2. The material density and the elastic wave speeds as a function of depth.

3. RESULTS

Using the numerical model, the interaction of elastic waves with buried land mines is investigated. First, results are presented that describe the interaction on a large scale using a simple model for the buried mine. Then, the resonant oscillations that are observed at the location of the buried mine are explored by employing a mine model that incorporates many features of an actual mine.

3.1. Interaction of Elastic Waves with a Buried Land Mine

Fig. 3 shows an actual TS-50 antipersonnel mine. Antipersonnel mines are fairly small and have a complex structure. Parts of the mine that might influence its elastic behavior are the explosives, the triggering mechanism, the mine's case, the rubber pressure plate and the various air chambers. As a first approach, the mine is approximated by a simple model. The simple model has roughly the same dimensions as the TS-50 mine. Due to the cubic structure of the discrete finite-difference grid, the mine is discretized by cubes. The model consists of two parts: a large plastic chamber and a small air-filled chamber on top of the plastic chamber. By inserting the simple mine model into the numerical model, the interaction of elastic waves with a buried mine is computed.

Results are shown in Fig. 4. The interaction of elastic waves with a mine buried 2 cm beneath the surface of the ground is depicted. The vertical particle displacement on the surface plane (top) and on a cross section through the ground (bottom) is shown at four instants in time. A logarithmic color scale is employed, ranging from white (0 dB) to black (-60 dB). The plots describe a surface area of 1.0 m by 0.8 m, and a cross sectional area with a depth of 0.7 m. The source is placed on the surface, to the left of the plots. The distance from the source to the mine is 0.8 m. The excitation is a differentiated Gaussian pulse with a center frequency of 450 Hz.

At the first instant in time, T_1, a Rayleigh surface wave (R) and a shear wave (S) are seen to propagate. A pressure wave has also been launched, but has already left the range of the plot. On the cross section it is evident

*The material properties are measured only up to a depth of half a meter and are assumed constant deeper in the ground. This is a vague assumption and probably not true within the water table, but is justifiable, because, within the frequency range of interest, the variation of the material properties deeper in the ground will not have a significant impact onto the wave propagation at and close to the surface.

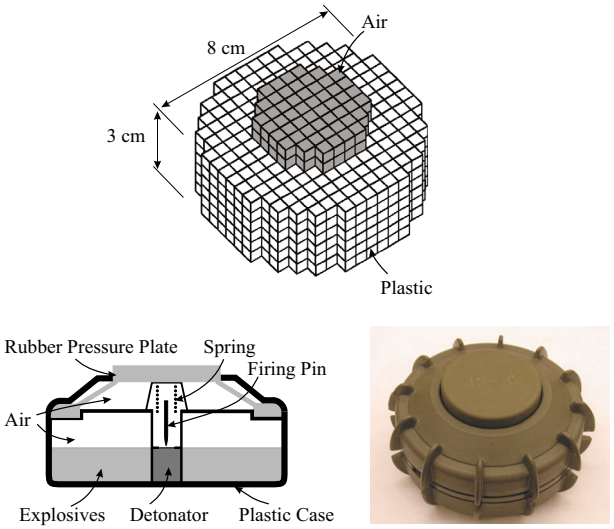


Figure 3. A simple model for a buried antipersonnel mine.

that, due to the variation of the wave speeds in the ground, the shear wave is refracted back towards the surface, and a guided shear wave propagates along the surface. The guided shear wave is ahead of the Rayleigh wave and is clearly visible on the surface. At T_2, the Rayleigh surface wave just hits the mine. At T_3, scattered Rayleigh surface waves (rR) and shear waves (rS) can be distinguished. At T_4, resonant oscillations are visible at the location of the buried mine. Due to the resonant oscillations, the location of the buried mine can be clearly identified on the surface plane. When looking at the wave fields on the cross section, it is evident that the resonant oscillations are confined to the thin soil layer above the mine. It has been shown that, within the numerical model, the resonant oscillations are due to flexural waves coupling into the thin soil layer and forming a pattern of standing waves.[3] A similar resonance has been observed also experimentally for an actual TS-50 mine. For an actual mine, however, the resonant oscillations are due to the soil-mine system, and the resonance is strongly influenced by the resonant behavior of the mine. Thus, for an actual mine the resonant behavior is certainly more complex and cannot be attributed solely to flexural modes in the soil layer above the mine.

3.2. Resonance Behavior of a Buried Land Mine

To investigate the resonant oscillations at the location of a buried land mine in detail, a mine model has been developed that closely resembles the structure of an actual mine. The new mine model is shown in Fig. 5. The mine model includes the details of the mine that are believed to have an impact on the mine's resonant behavior, i. e., the explosives, the plastic case, the rubber pressure plate and the air chambers.

To incorporate the features of an actual mine, a much finer discretization must be chosen for the detailed mine model. Due to limited computer resources, the size of the discretized solution space is reduced to the area immediately surrounding the mine. Because the dimensions of the solution space for this model are small (10 cm by 10 cm by 10 cm), the medium can be approximated as homogeneous. Thus, for all of the following results, the soil properties are assumed constant throughout the solution space.

Results obtained with the detailed mine model are shown in Fig. 6. The vertical particle displacement on the surface plane (top) and on a cross section through the ground (bottom) is shown at four instants in time using the same color scale as earlier. The mine is buried 2 cm beneath the surface. Instead of modeling the transducer on the surface, a vertically polarized plane shear wave is injected from the left. The excitation is a differentiated Gaussian pulse with a center frequency of 1 kHz and, thus, has a wider bandwidth than the excitation used previously. The solution space is surrounded by a PML. For the results in Fig. 6, the material density is chosen to be $\rho = 1400$ kg/m^3, the pressure wave speed is $c_p = 250$ m/s and the shear wave speed $c_s = 80$ m/s. The same values for the pressure wave speed and the material density will be used for all results presented in the following.

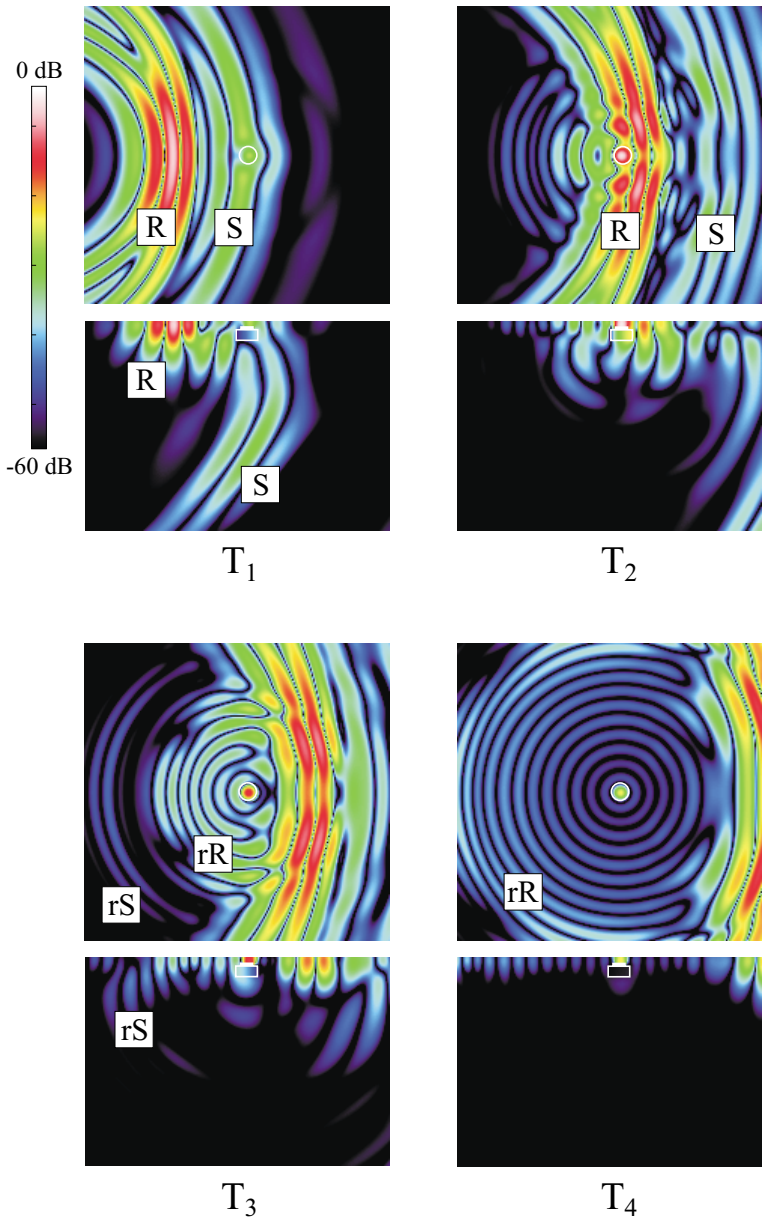


Figure 4. Interaction of elastic waves with a buried antipersonnel mine; mine buried 2 cm beneath the surface. The vertical particle displacement on the surface (top) and on a cross section through the ground (bottom) is shown at four instants in time. R: Rayleigh wave; S: shear wave; rR: scattered Rayleigh wave; rS: scattered shear wave.

In Fig. 6, resonant oscillations are seen to be excited at the buried mine. The oscillations appear to be mainly due to the motion of the thin soil layer above the mine and the rubber pressure plate of the mine. The oscillations decay as time progresses. The surface motion seems to be composed of different modes. At T_1 and T_3, the motion is dominated by a mode describing an upward and downward motion of the surface above the mine. At T_4, a mode describing a rocking motion of the surface is apparent. A null in the vertical particle displacement along the diameter of the soil layer is clearly visible on the surface and on the cross section. At T_2, the surface motion seems to be composed of the superposition of at least these two modes.

From the results obtained with the detailed mine model, it becomes clear that the resonance is not only due to flexural modes of the thin soil layer, but rather must be attributed to the resonant behavior of the soil-mine system.

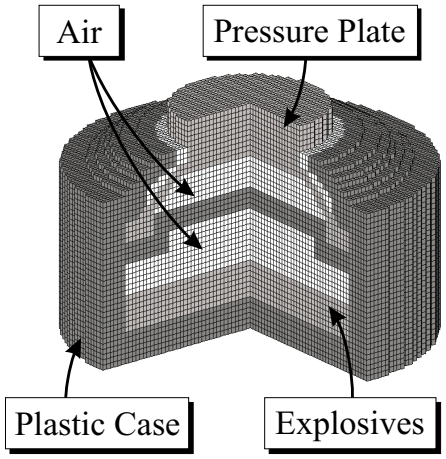


Figure 5. Refined mine model that includes details of an actual TS-50 mine such as the explosives, the case, the rubber pressure plate, and the air chambers.

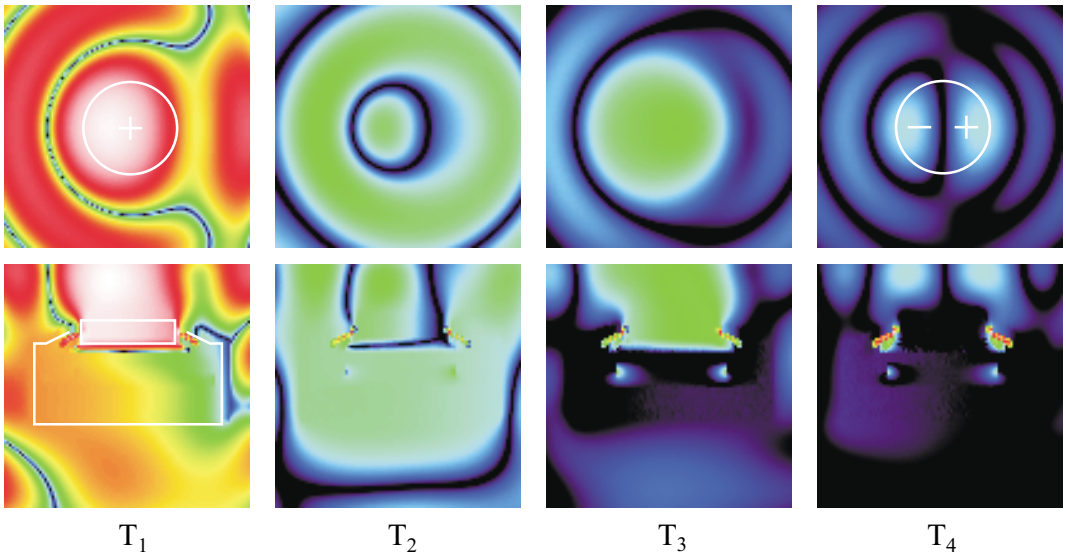


Figure 6. Resonant oscillations at the location of a buried land mine; mine buried 2 cm beneath the surface. The vertical particle displacement in the surrounding of the mine as computed with the detailed mine model is shown at four instants in time on the surface plane (top) and on a cross section through the ground (bottom).

The resonant oscillations are caused by the motion of both the soil layer and the rubber pressure plate. This motion will be influenced by various factors. On the one hand, it will strongly depend on the stiffness and thickness of the soil layer above the mine. On the other hand, it will be affected by the mine and its structure. For example, below the rubber pressure plate lies a closed air chamber (see Fig. 5). This air chamber behaves like a spring and has a significant impact on the motion of the pressure plate.

To analyze the resonant behavior at the location of a buried mine in detail, a parametric study of the resonance as a function of the burial depth and the shear wave speed in the ground has been conducted. For this, the vertical motion at one point on the surface above the mine is computed in the time domain. A point is picked that lies 1 cm off the vertical axis of the mine. The motion is Fourier transformed from the time domain into the frequency domain, and, by dividing by the Fourier transform of the incident pulse, the transfer function is determined. From

the transfer function the resonant frequencies of the particle motion can be identified. For the parametric study, shear wave speeds in the range from 40 m/s up to 150 m/s are considered in order to determine the response for fairly loose as well as rather stiff soils. These values are representative for the shear wave speeds at the surface of many common soil types.

Fig. 7 shows the transfer function for a mine buried 1 cm beneath the surface. The shear wave speed in the ground is assumed to be 40 m/s. Below 2.1 kHz, two resonant peaks can be distinguished. The two peaks correspond to two modes of the surface motion. The first mode has a resonant frequency of about 630 Hz and corresponds to an up- and downward motion of the surface. The second mode, at about 1.4 kHz, describes a rocking motion. The behavior of the modes is illustrated in Fig. 7. The two modes that are excited within the soil-mine system are similar in shape to the first two modes of a thin circular disc with clamped edges.[5,6] However, the analogy to the circular disc cannot be used to accurately predict the resonant frequencies of the soil-mine system, because along its edge the soil layer in the model is not rigidly fixed and the structure of the mine has a significant impact on the resonance.

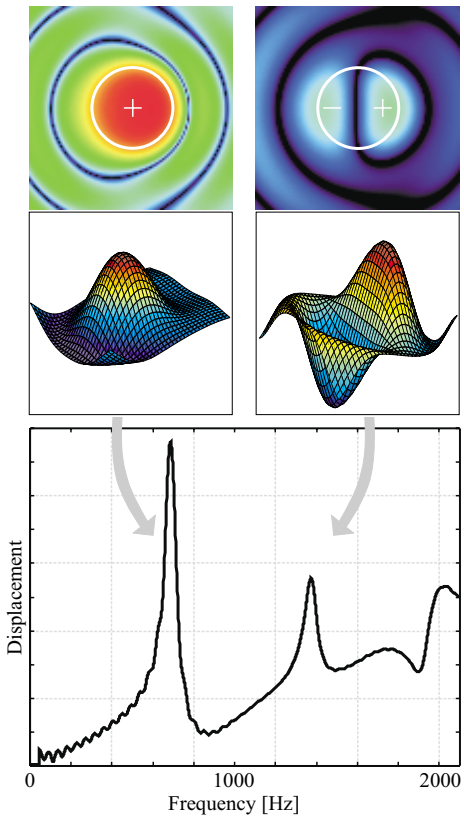


Figure 7. Transfer function of a point on the surface above the mine; mine buried at 1cm; $c_s = 40$ m/s. The two resonant peaks correspond to two modes of the surface motion.

To explore the resonance for mines in different burial scenarios, the resonant behavior is determined as a function of the shear wave speed in the ground and the burial depth of the mine. Fig. 8 shows the transfer function on the left as a function of the shear wave speed in the ground with the burial depth of the mine fixed at 2 cm, and on the right as a function of the burial depth with the shear wave speed kept constant at 40 m/s. It is apparent that the resonant frequencies of the first and second peak are shifted upwards in frequency with increasing shear wave speed. At the same time, their amplitudes drop, and the quality of the resonance is degraded. For increasing burial depth, the resonant frequencies are shifted slightly downwards, and the amplitudes drop drastically. For the mine at 4 cm, the second resonant peak is not distinguishable any more.

To summarize the results, the resonant frequency and the quality factor of the first resonant peak are plotted in

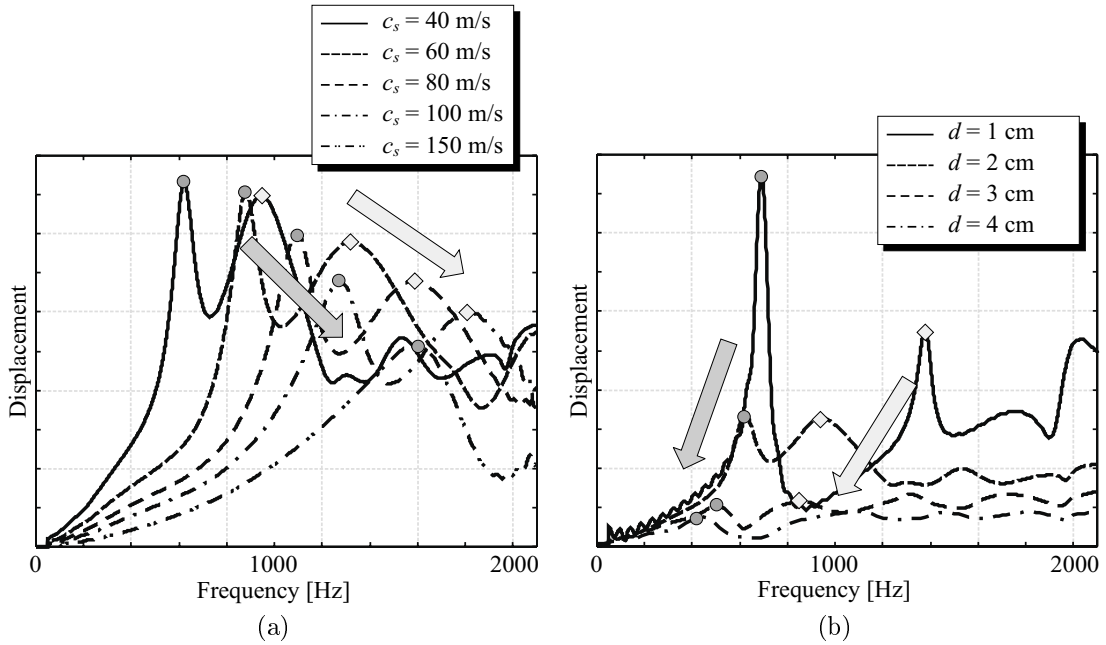


Figure 8. Surface motion at a point above the mine in the frequency domain. (a) Mine buried 2 cm beneath the surface, variable shear wave speed; (b) shear wave speed $c_S = 40$ m/s, variable depth.

two parametric graphs as a function of the shear wave speed in the ground with the burial depth as a parameter. Note that, by varying the shear wave speed in the ground while keeping the material density constant, effectively the stiffness in the ground is being varied. Thus, in these plots the resonant behavior is shown as a function of the stiffness in the ground. In Fig. 9, the resonant frequency and the quality factor are plotted. The top line in each plot corresponds to the mine at a burial depth of 1 cm. Qualitatively, it can be seen that the resonant frequency increases with shear wave speed and decreases with burial depth. The quality factor of the resonance decreases with both shear wave speed and burial depth. The change in the resonant frequency is mainly due to the stiffness in the soil, whereas the change in the quality of the resonance can be attributed to radiation damping. The dominant damping mechanism for the resonance within the model is the radiation of energy into the surrounding soil. The coupling of the thin soil layer above the mine to the surrounding soil increases with the stiffness and the thickness of the layer. Thus, the energy that is radiated also increases, and the damping of the resonance is enhanced.

Resonant oscillations have been observed also experimentally. Fig. 10 shows experimental results obtained with an actual TS-50 mine buried 1 cm beneath the surface. The solid curve indicates the measured transfer function, whereas the dashed gray curve is the transfer function as computed with the numerical model for a mine at 1 cm and with a shear wave speed in the soil of 40 m/s. The curves are seen to be quite different. The second resonant peak in the experimental results is very weak. The authors believe that these differences are mainly due to non-linear effects occurring in the sand. The non-linear effects strongly dampen the high frequency components of the elastic waves and effectively suppress frequency components above about 500 Hz.

4. CONCLUSIONS

If a buried land mine interacts with elastic waves, resonant oscillations are excited at the location of the buried mine. In this paper, these resonant oscillations are explored as a function of burial depth and soil parameters using a numerical model. It has been found that, in general, the resonant behavior is strongly dependent on both the burial depth and the shear wave speed in the soil. The resonant frequency is seen to increase with shear wave speed and to decrease with burial depth. At the same time, the quality factor of the resonance decreases with both shear wave speed and burial depth. The results presented in this paper show conclusively that resonant oscillations do exist at the location of a buried mine for a wide range of different soil types and burial depths.

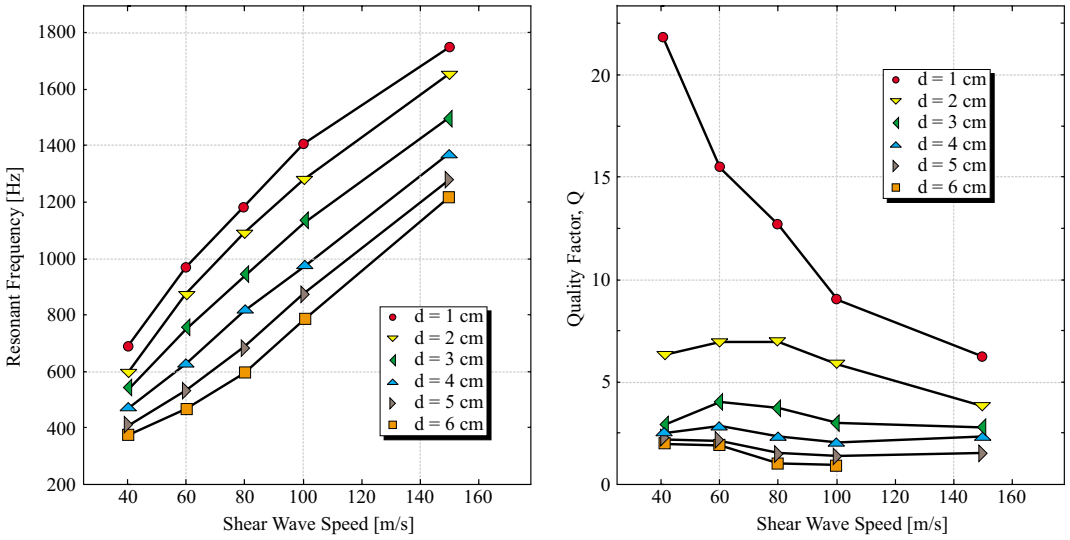


Figure 9. Parametric graph of the resonant frequency and the quality factor of the first resonant peak as a function of the shear wave speed in the ground with the burial depth as a parameter.

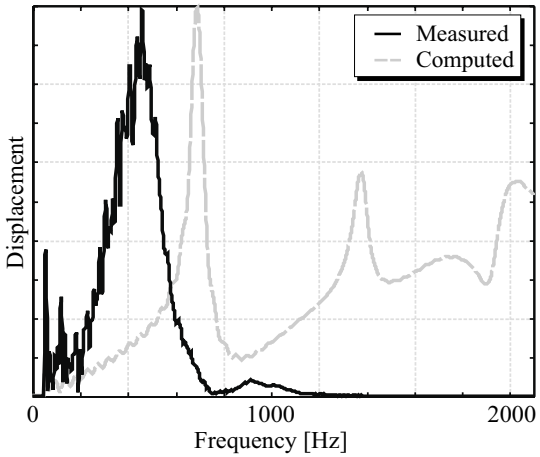


Figure 10. Experimental results for a mine buried at 1 cm; transfer function.

ACKNOWLEDGMENTS

This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0448, by a grant from the US Office of Naval Research under contract N00014-99-1-0995, and by an equipment grant from the Intel Corporation.

REFERENCES

1. W. R. Scott, Jr., G. D. Larson, and J. S. Martin, "Simultaneous use of elastic and electromagnetic waves for the detection of buried land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets V, Proc. SPIE*, vol. 4038, pp. 667–678, Apr. 2000.
2. W. R. Scott, Jr., J. S. Martin, and G. D. Larson, "Experimental model for a seismic landmine detection system," *IEEE Trans. on Geoscience and Remote Sensing* , to appear July 2001.

3. C. T. Schröder and W. R. Scott, Jr., "A finite-difference model to study the elastic-wave interactions with buried land mines," *IEEE Trans. on Geoscience and Remote Sensing* **38**, pp. 1505–1512, Jul. 2000.
4. C. T. Schröder and W. R. Scott, Jr., "Three-dimensional FDTD model to study the elastic-wave interaction with buried land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets V, Proc. SPIE*, vol. 4038, pp. 680–690, Apr. 2000.
5. K. F. Graff, *Wave Motion in Elastic Solids*, Dover Publications, 1975.
6. N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, Springer-Verlag, New York, 1st ed., 1991.

A poroelastic model for acoustic landmine detection

Y. Q. Zeng[a] and Q. H. Liu[a]

[a]Department of Electrical and Computer Engineering,
Duke University, Durham, North Carolina 27708, U.S.A.

ABSTRACT

Acoustic waves can be a viable tool for the detection and identification of land mines and unexploded ordnance (UXO). Design of acoustic instruments and interpretation and processing of acoustic measurements call for accurate numerical models to simulate acoustic wave propagation in a heterogeneous soil with buried objects. Compared with the traditional seismic exploration, high attenuation is unfortunately ubiquitous for shallow surface acoustic measurements because of the loose soil and the fluid in its pore space. To adequately model such acoustic attenuation, we propose a comprehensive model to simulate the acoustic wave interactions with land mines and soils based on the Biot theory for poroelastic media. The finite-difference time-domain method is then used to solve the Biot equations. For the truncation of the computational domain in the FDTD method, we extend the acoustic and elastic perfectly matched layer (PML) to poroelastic media. Numerical experiments show that, with only 10 cells of PML medium, a high attenuation of about 50 dB can be achieved for outgoing waves. The numerical model is validated by comparison with analytical solutions. Unlike the pure elastic wave model, this efficient PML-FDTD model for poroelastic media incorporates the interactions of waves and the fluid-saturated pore space. The difference between elastic model and poroelastic model is investigated by studying surface wave amplitude variation with offset (AVO) in three different ground media: dry sand, fully water saturated sand and partly water saturated sand. The interaction of elastic wave with a plastic mine buried in dry sand is simulated. The results show that the surface wave is significantly influenced by the existence of a mine-like object. The diffraction of the surface wave can serve as an acoustic target signature.

Keywords: poroelastic, land mine detection, staggered-grid, elastic wave, PML, finite-difference

1. INTRODUCTION

Recently acoustic methods have received considerable attention in the characterization of shallow objects. It has been shown that acoustic waves can be a viable tool for the detection and identification of land mines and unexploded ordnance (UXO). Design of acoustic instruments and interpretation and processing of acoustic measurements call for accurate numerical models to simulate acoustic wave propagation in a heterogeneous soil with buried objects. Conventional methods for acoustic modeling include the popular finite difference time-domain (FDTD) solution of pure elastic media. Unfortunately, the pure elastic model cannot adequately incorporate the attenuation mechanism in its governing equations, although some approximate models are possible.

Compared with the traditional seismic exploration, high attenuation is unfortunately ubiquitous for shallow surface acoustic measurements because of the loose soil and the fluid in its pore space. To adequately model such acoustic attenuation, we propose a comprehensive model to simulate the acoustic wave interactions with land mines and soils based on the Biot theory.[1–3] for poroelastic media. The finite-difference time-domain method is then used to solve the Biot equations.[3] Like velocity-stress FD method,[4] a velocity-strain, finite-difference method is developed in a staggered gird for heterogeneous poroelastic media. In this method, Biot's equations[3] are reformulated into first-order equations to arrive at a leap-frog system in a staggered grid both in time and space domains. Numerical solutions have been validated by analytical solutions.

In order to simulate an unbounded medium, an absorbing boundary condition (ABC) must be implemented to truncate the computational domain in numerical algorithms. There are many kinds of ABCs developed for numerical simulation of wave propagation. Cerjan *et al.*[5] introduces a simple damping taper to the boundaries that attenuates the outgoing waves. Since this lossy layer is not perfectly matched to the interior region, however, it requires a substantial number of layers for the taper to be effective. Clayton and Engquist[6] (CE) use the paraxial approximation to the wave equation to make the boundary transparent to outgoing waves. The CE ABC can lead

Correspondence to Q.H. Liu, Email: qhliu@ee.duke.edu; http://www.duke.edu/∼qhliu.

to instability when the Poisson's ratio is greater than two.[7] Since Berenger[8] proposed the highly effective perfectly matched layer (PML) as an absorbing boundary condition for electromagnetic waves, it has been widely used for finite-difference and finite-element methods. Chew and Liu[9] first proposed the PML for elastic waves in solids, and proved the zero reflections from PML to the regular elastic medium. Hastings *et al.*[11] have independently implemented the PML ABC for two-dimensional problems by using potentials. The PML has also been extended to model acoustic waves and electromagnetic waves in lossy media.[12] PML has been applied to the second-order Biot's equation for fluid-saturated poroelastic media,[13] which requires a complicated convolution. In this paper, as the PML is developed for the first-order system, incorporation of PML becomes much simpler. The performance of PML is investigated. The effectiveness of this ABC is confirmed by examining the reflection from the boundary.

The difference between elastic model and poroelastic model is investigated by studying surface wave amplitude variation with offset (AVO) in three different ground media: dry sand, fully water saturated sand and partly water saturated sand. The interaction of elastic wave with a plastic mine buried in dry sand is simulated. The results show that the wave responses is significantly influenced by the existence of a mine-like object. After processing, the target can be detected by using surface acoustic measurements.

2. FORMULATION

2.1. The governing equations

The propagation of acoustic waves in porous and fluid-saturated media is different from that in single phase elastic media. In addition to the regular P waves and S waves in solid elastic media, a slow P wave which results from the relative motion between solid frame and fluid may be present in porous media. Thus the pattern of energy dissipation in porous media is different from that in solid elastic media. Based on continuum mechanics and macroscopic constitutive relationship, Biot[1–3] developed a theory of wave motion in a porous elastic solid saturated with a viscous compressible fluid. Biot's theory was confirmed by Burridge and Keller[14] based on the dynamic equations which govern the behavior of medium on a microscopic scale. Plona[15] also confirmed Biot's theory through experiments.

In an isotropic, heterogeneous porous elastic medium, the parameters describing the physical properties of the medium are as follows,

μ	shear modulus of dry porous matrix
λ_c	Lame constant of saturated matrix
ϕ	porosity
η	viscosity
κ	permeability
ρ	the overall density of the saturated medium determined by $\rho_f\phi + (1-\phi)\rho_s$
ρ_s	density of solid material
ρ_f	density of fluid
a	tortuosity of the matrix
K_s	bulk modulus of the solid
K_f	bulk modulus of the fluid
K_b	bulk modulus of the dry porous frame

The macroscopic displacements and strains are defined as

u_i	ith component of displacements of solid particle
U_i	ith component of displacements of fluid particle
w_i	ith component of relative displacement, $w_i = \phi(U_i - u_i)$
e_{ij}	ij component of strain tensor in porous medium, $e_{ij} = [(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j})/2]$
e	$e = \sum_{i=1,2,3} e_{ii}$
ξ	dilatation for the relative motion, $\xi = -\sum_{i=1,2,3} \frac{\partial w_i}{x_i}$

For a three-dimensional isotropic, heterogeneous and porous elastic medium, wave propagation is governed by Biot's equations[3]:

$$2\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) + \frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) = \frac{\partial^2}{\partial t^2}(\rho u_i + \rho_f w_i) \tag{1}$$

$$\frac{\partial}{\partial x_i}(\alpha M e - M\xi) = \frac{\partial^2}{\partial t^2}(\rho_f u_i + m w_i) + \frac{\eta}{\kappa}\frac{\partial w_i}{\partial t} \tag{2}$$

where $m = a\rho_f/\phi$ and

$$M = \frac{1}{\frac{\phi}{K_f} + \frac{(\alpha-\phi)}{K_s}}$$

$$\alpha = 1 - \frac{K_b}{K_s}$$

Let $\mathbf{v^s}$ be the velocity of the solid particle, $\mathbf{v^f}$ be the velocity of the pore fluid relative to the solid frame. Then the second-order equations (1) and (2) can rearranged as the first-order equations,

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial v_i^s}{\partial t} &= 2m\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) + m\frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) \\ &\quad -\rho_f\frac{\partial}{\partial x_i}(\alpha M e - M\xi) + \rho_f\frac{\eta}{\kappa}v_i^f \qquad (3)\\ (m\rho - \rho_f^2)\frac{\partial v_i^f}{\partial t} &= \rho\frac{\partial}{\partial x_i}(\alpha M e - M\xi) - \rho\frac{\eta}{\kappa}v_i^f \\ &\quad -2\rho_f\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) - \rho_f\frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) \qquad (4)\end{aligned}$$

The time derivatives of strains e_{ij} and ξ can be expressed as

$$\begin{aligned}\frac{\partial e_{ij}}{\partial t} &= [(\partial v_j^s/\partial x_i + \partial v_i^s/\partial x_j)/2] \qquad (5)\\ \frac{\partial \xi}{\partial t} &= -\nabla\cdot\mathbf{v^f} \qquad (6)\end{aligned}$$

In the explicit first-order finite-difference schemes, equations (3) to (6) consist of a leap-frog system for the strain field e_{ij}, ξ and velocity field $\mathbf{v^s}$ and $\mathbf{v^f}$. With proper absorbing boundary conditions, these equations can be solved numerically for the wave field in an unbounded medium.

Equations (3) to (6) predict the existence of three different waves in fluid-saturated poroelastic media: a shear wave and two compressional waves with a faster and a slower propagation velocities.

2.2. Equations for the PML absorbing boundary condition

In this paper, the perfectly matched layer (PML) will be used to truncate the unbounded medium, absorbing all outgoing waves. The artificial absorptive medium is introduced in the regular medium by modifying equations (3) to (6) with complex coordinates.[10,11] In the frequency domain, a complex coordinate variable is chosen as

$$\tilde{x_i} = \int_0^{x_i} e_i(x_i')dx_i' \tag{7}$$

$$e_i = a_i + j\frac{\omega_i}{\omega} \qquad (i = 1, 2, 3) \tag{8}$$

where $a_i \geq 1$ is a scaling factor, and $\omega_{x_i} \geq 0$ is an attenuation factor. The operator $\frac{\partial}{\partial x_i}$ can be expressed in terms of the regular coordinate,

$$\frac{\partial}{\partial \tilde{x_i}} = \frac{1}{e_i}\frac{\partial}{\partial x_i} \tag{9}$$

The PML formulation is to replace x_i in (4)-(6) by the corresponding complex coordinate $\tilde{x_i}$. In a PML region, the real part a_i is scaling factor, and the imaginary part ω_i represents a loss in the PML. In a regular non-PML region, $a_i = 1$ and $\omega_i = 0$.

In order to make equations with PML simple, the equations and variables are split as the following,

$$v_i^s = \sum_{k=1}^{3} v_i^{s(k)}$$

$$v_i^f = \sum_{k=1}^{3} v_i^{f(k)}$$

where $v_i^{s(k)}$ and $v_i^{f(k)}$ represent the split field variables containing space derivative $\frac{\partial}{\partial x_k}$ only . For example, equation (3) for $i = 1$ can be split into the following 3 equations,

$$\begin{aligned}
(m\rho - \rho_f^2)\frac{\partial v_1^{s(1)}}{\partial t} &= 2m\frac{\partial}{\partial x_1}(\mu e_{11}) + m\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M\xi) - \rho_f\frac{\partial}{\partial x_1}(\alpha M e - M\xi) \\
(m\rho - \rho_f^2)\frac{\partial v_1^{s(2)}}{\partial t} &= 2m\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f\frac{\eta}{\kappa}v_1^f \\
(m\rho - \rho_f^2)\frac{\partial v_1^{s(3)}}{\partial t} &= 2m\frac{\partial}{\partial x_3}(\mu e_{13})
\end{aligned}$$

The diagonal strain components e_{ii} need not be split. However, other strain components have to be split as $e_{ij} = \sum_{k=1}^{3} e_{ij}^{(k)}$ for $i \neq j$, and $\xi = \sum_{k=1}^{3} \xi^{(k)}$, for example,

$$\begin{aligned}
\frac{\partial e_{ii}}{\partial t} &= \frac{\partial v_i^s}{\partial x_i} \\
e_{12} &= e_{12}^{(1)} + e_{12}^{(2)} \\
\frac{\partial e_{12}^{(1)}}{\partial t} &= \frac{1}{2}\frac{\partial v_2^s}{\partial x_1} \\
\frac{\partial e_{12}^{(2)}}{\partial t} &= \frac{1}{2}\frac{\partial v_1^s}{\partial x_2} \\
\xi &= \xi^{(1)} + \xi^{(2)} + \xi^{(3)} \\
\frac{\partial \xi^{(1)}}{\partial t} &= -\frac{\partial v_1^f}{\partial x_1} \\
\frac{\partial \xi^{(2)}}{\partial t} &= -\frac{\partial v_2^f}{\partial x_2} \\
\frac{\partial \xi^{(3)}}{\partial t} &= -\frac{\partial v_3^f}{\partial x_3}
\end{aligned}$$

In frequency domain PML formulation, after x_i is replaced by $\tilde{x_i}$, equations for v_1^s, for example, can be rewritten as

$$\begin{aligned}
(m\rho - \rho_f^2)(-j\omega)(1 + j\frac{\omega_{x_1}}{\omega})v_1^{s(1)} &= 2m\frac{\partial}{\partial x_1}(\mu e_{11}) + m\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M\xi) - \rho_f\frac{\partial}{\partial x_1}(\alpha M e - M\xi) \\
(m\rho - \rho_f^2)(-j\omega)(1 + j\frac{\omega_{x_1}}{\omega})v_1^{s(2)} &= 2m\frac{\partial}{\partial x_2}(\mu e_{12}) + (1 + j\frac{\omega_1}{\omega})\rho_f\frac{\eta}{\kappa}v_1^f \\
(m\rho - \rho_f^2)(-j\omega)(1 + j\frac{\omega_1}{\omega})v_1^{s(3)} &= 2m\frac{\partial}{\partial x_3}(\mu e_{13})
\end{aligned}$$

By taking inverse Fourier transform, the above equations yield the time-domain PML equations,

$$\begin{aligned}
(m\rho - \rho_f^2)(\frac{\partial v_1^{s(1)}}{\partial t} + \omega_1 v_1^{s(1)}) &= 2m\frac{\partial}{\partial x_1}(\mu e_{11}) + m\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M\xi) \\
&\quad -\rho_f\frac{\partial}{\partial x_1}(\alpha M e - M\xi)
\end{aligned} \tag{10}$$

$$(m\rho - \rho_f^2)(\frac{\partial v_1^{s(2)}}{\partial t} + \omega_1 v_1^{s(2)}) = 2m\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f \frac{\eta}{\kappa}(v_1^f + \omega_2 \int_{-\infty}^{t} v_1^f d\tau) \quad (11)$$

$$(m\rho - \rho_f^2)(\frac{\partial v_1^{s(3)}}{\partial t} + \omega_1 v_1^{s(3)}) = 2m\frac{\partial}{\partial x_3}(\mu e_{13}) \quad (12)$$

$$\frac{\partial e_{11}}{\partial t} = \frac{\partial v_1}{\partial x_1} - \omega_1 e_{11} \quad (13)$$

$$\frac{\partial e_{12}^{(1)}}{\partial t} = \frac{1}{2}\frac{\partial v_2^s}{\partial x_1} - \omega_1 e_{12}^{(1)} \quad (14)$$

$$\frac{\partial e_{12}^{(2)}}{\partial t} = \frac{1}{2}\frac{\partial v_1^s}{\partial x_2} - \omega_2 e_{12}^{(2)} \quad (15)$$

$$\frac{\partial \xi^{(1)}}{\partial t} = -\frac{\partial v_1^f}{\partial x_1} - \omega_1 \xi^{(1)} \quad (16)$$

$$\frac{\partial \xi^{(2)}}{\partial t} = -\frac{\partial v_2^f}{\partial x_2} - \omega_2 \xi^{(2)} \quad (17)$$

$$\frac{\partial \xi^{(3)}}{\partial t} = -\frac{\partial v_3^f}{\partial x_3} - \omega_3 \xi^{(3)} \quad (18)$$

Similar equations can be obtained in the same way for other components v_2^s, v_3^s, v_1^f, v_2^f, v_3^f, and e_{13}, e_{23}. Within PML region, equations (3) and (4) are split into 18 equations for 3D. Equation (5) is split into 9 equations and equation (6) is split into 3 equations. So the total number of the equations is 30 for the PML region, compared to 13 for the regular interior region. Thus, the memory requirement within the PML region is about two and half times that required by a regular medium for three-dimensional problems. This extra memory requirement in PML region is offset by the effectiveness of PML in absorbing the outgoing waves.

2.3. Finite difference implementation

The governing equations for the PML absorbing boundary condition, such as equations (10)-(18) are first-order partial differential equations for particle velocity and strain. They can be solved with different numerical methods. For the evaluation of seismic and acoustic responses of specific models, the accuracy and convenience of the numerical method are of primary concern. The finite-difference method is widely used in wave modeling because of its flexibility and accuracy. For the first-order, leap-frog system of equations (3) to (6), the explicit finite-difference method is used.

To implement a 3-D finite-difference solution to the equations with PML, the material parameters and unknown field components are discretized on a regular 3-D grid at the intervals $\triangle x_1$, $\triangle x_2$ and $\triangle x_3$. The time domain is also discretized with time step $\triangle t$. There are two kinds of operators to approximate the first-order derivatives,*i.e.,* the centered grid and staggered grid. Because the centered-grid operator to perform first derivatives are less accurate than staggered grid operators,[16] a staggered grid is used in this paper. For the FD implementation of Biot's equations on a staggered grid in Fig. 1, the velocity field components are located at the cell's face centers, while material parameters and normal strains are located at the center of the cell and shear strains are located the at six edge centers. Strain field is computed at $n\triangle t$ and velocity field is computed at $(n + \frac{1}{2})\triangle t$. This staggered grid is similar to that for elastic waves in a solid.[17,18]

With this discretization, the leap-frog system can be written in a time-stepping form. In order to make the layout of the formulas simple, the governing equations with PML boundary conditions can be generalized as the first-oder differential equation. For examples, (11), (13) and (16) can be rewritten as

$$\frac{\partial v_1^{s(2)}}{\partial t} + c_0 v_1^{s(2)} = c_1 \int_{-\infty}^{t} v_1^f d\tau + c_2[2m\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f \frac{\eta}{\kappa}(v_1^f)] \quad (19)$$

$$\frac{\partial e_{11}}{\partial t} + \omega_1 e_{11} = \frac{\partial v_2^s}{\partial x_1} \quad (20)$$

$$\frac{\partial \xi^1}{\partial t} + \omega_1 \xi^1 = -\frac{\partial {v_1}^f}{\partial x_1} \quad (21)$$

where c_0, c_1 and c_2 space-dependent coefficients. The time-stepping equations can be written as,

$$v_1^{s(2)}[j_1, j_2, j_3, (n+\frac{1}{2})] = f_1 v_1^{s(2)}[j_1, j_2, j_3, (n-\frac{1}{2})] + f_2 R_1 \quad (22)$$

$$e_{11}[j_1, j_2, j_3, (n+1)] = g_1 e_{11}[j_1, j_2, j_3, n] + g_2 R_2 \quad (23)$$

$$\xi^1[j_1, j_2, j_3, (n+1)] = g_1 \xi^1[j_1, j_2, j_3, n] - g_2 R_3 \quad (24)$$

$$f_1 = \frac{c_0/2 - 1/\triangle t}{1/\triangle t + c_0/2}$$

$$f_2 = \frac{1}{1/\triangle t + c_0/2}$$

$$g_1 = \frac{\omega_1/2 - 1/\triangle t}{1/\triangle t + \omega_1/2}$$

$$g_2 = \frac{1}{1/\triangle t + \omega_1/2}$$

where R_1, R_2 and R_3 are right-hand sides of equations (19),(20) and (21) respectively. It should be noted that the material parameters in the above equations must be properly averaged in order to arrive at a higher accuracy.[17] In order to incorporate the PML boundary condition, the computational domain is divided into a PML region and an interior region. The absorption of outgoing waves is achieved by the PML region, which consists of several cells of mathematically defined materials with a quadratically tapered ω_j profile to increase the attenuation toward the outer boundary. In this paper, ω_j of the PML region is chosen as

$$\omega_i(j_i) = \frac{(M - 1/2 - j_i)^2}{(M - 1/2)^2} \omega_{i,max} \quad (25)$$

where $\omega_{i,max}$ is the value at the center of the cell at outermost boundary. At the outer boundary, the velocity components and shear strain are forced to be zero. For convenience, $\omega_{i,max}$ can be expressed in term of dominant frequency and normalized coefficient. Then equation (25) becomes

$$\omega_i = \begin{cases} 2\pi a_0 f_0 (l_{x_i}/L_{PML})^2, & \text{inside PML,} \\ 0, & \text{outside PML} \end{cases} \quad (26)$$

where f_0 is the dominant frequency of the source, and L_{PML} is the thickness of the PML region, and l_{x_i} is the distance from the interface between the interior region and PML region.

3. NUMERICAL RESULTS

The finite-difference algorithm on a staggered grid has a higher accuracy than on a centered grid. Thus, the PML region can be made thiner with a staggered grid with a better absorption than with a centered grid. In this paper, the length of PML region is chosen to be 10 cells. The FD algorithm is illustrated by its two-dimensional implementation.

In the following numerical simulations, the source will be a monopole source for a homogeneous medium and a dipole source for a half space. The source time function is the first derivative of a Gaussian function

$$S(t) = (t - t_0) e^{-\pi^2 {f_0}^2 (t - t_0)^2}$$

where f_0 is the predetermined dominant frequency, and t_0 the time shift.

The source energy is partitioned linearly between the solid and the fluid phases with factors

$$W_f = \phi$$
$$W_s = (1 - \phi)$$
$$W_r = \phi |W_f - W_s|$$

where W_f is the weighting factor for the fluid motion, W_s is for the solid motion and W_r is for the relative motion between solid frame and pore fluid.

3.1. PML performance

The effectiveness of the absorbing boundary condition is an essential factor for the successful numerical simulation. A good absorbing boundary condition has the characteristics of effective absorption of outgoing waves without requiring a large memory. In this paper, the performance of the PML is investigated on a homogeneous, fluid-saturated poroelastic medium by comparing numerical solutions of PML model with reference solutions that do not have reflections. The optimized a_0 can be obtained by examining the attenuation of the wave field in the PML region.

The homogeneous model for PML performance test is a sandstone fully saturated with water and has the properties of $\rho_s = 2650$ kg/m^3, $\rho_f = 1040$ kg/m^3, and porosity $\phi = 0.3$. This model has the wave properties as $v_{fast} = 2365.7$ m/s for fast P wave, $v_{slow} = 776.95$ m/s for slow P wave and $v_{shear} = 960.5$ m/s for shear wave.

We first study the performance of the PML ABC for a homogeneous medium. Fig. 2 compares a numerical solution at a location 5-cell away from PML interface with a reference solution. (The reference solution is obtained from a much larger model with Dirichlet boundary in which reflections have not arrived within the time window of interest.) No apparent reflections are present in the numerical solution with the PML boundary. Compared to the incident signal, the the reflection is about 50dB down. Another advantage of PML absorbing boundary condition is its superior stability. For this particular model, a high ratio of $v_{fast}/v_{shear} = 2.46$ will cause instability problem for CE boundary condition[10].

3.2. Validation of numerical results

An analytical solution for the particle velocity field in a homogeneous, fluid-saturated poroelastic medium subject to a point source in 3-D space or a line source in 2-D space can be derived.[4] The particle velocity is obtained in a closed form via potential functions.

For Biot's equations, it is convenient to solve for the particle velocity through potential functions. The velocity of particles and body force at source can be expressed in term of potentials as

$$\mathbf{u} = \nabla\psi_s + \nabla \times \Psi_s$$

$$\mathbf{U} = \nabla\psi_f + \nabla \times \Psi_f$$

$$\mathbf{f} = \nabla\Phi + \nabla \times \Psi$$

where $\nabla \cdot \Psi_f = 0$, $\nabla \cdot \Psi_s = 0$ and $\nabla \cdot \Psi = 0$ and describe the rotational potentials, while ψ_s, ψ_f and Φ describe the dilatational potentials. If the source is purely dilatational, then its rotational component disappears. In the time domain, for a purely P-wave point source and ideal non-viscous fluid, the potential can be expressed as

$$\psi_s(r,t) = \frac{\alpha s(t - \frac{r}{V_f}) + \beta s(t - \frac{r}{V_s})}{4\pi r}$$

$$\psi_s(r,t) = \frac{\alpha A_f s(t - \frac{r}{V_f}) + \beta A_s s(t - \frac{r}{V_s})}{4\pi r}$$

where r is the distance and $s(t)$ is the source time function. A_f and A_s represent the ratios between the solid and fluid motion for the fast P-wave and the slow P-wave. The coefficients α and β are determined by the regularity conditions. V_f and V_s are the velocities of the fast P-wave and the slow P-wave, respectively.

In two dimensions, for a pure P-wave line source along y-axis, the solution can be obtained by integrating the point source solution in y direction. In $x - z$ plane, the dilatational potentials are

$$\psi_s(x,z,t) = \frac{\alpha H(t - \frac{r}{V_f})}{2\pi} \int_{r/V_f}^{t} \frac{s(t-\tau)}{\sqrt{\tau^2 - r^2/{V_f}^2}} d\tau + \frac{\beta H(t - \frac{r}{V_s})}{2\pi} \int_{r/V_s}^{t} \frac{s(t-\tau)}{\sqrt{\tau^2 - r^2/{V_s}^2}} d\tau$$

$$\psi_f(x,z,t) = \frac{\alpha A_f H(t - \frac{r}{V_f})}{2\pi} \int_{r/V_f}^{t} \frac{s(t-\tau)}{\sqrt{\tau^2 - r^2/{V_f}^2}} d\tau + \frac{\beta A_s H(t - \frac{r}{V_s})}{2\pi} \int_{r/V_s}^{t} \frac{s(t-\tau)}{\sqrt{\tau^2 - r^2/{V_s}^2}} d\tau$$

where $H()$ is the Heaviside step function and $r = \sqrt{x^2 + z^2}$. Once the potential functions are available, the velocity can be easily obtained by taking gradient of potential functions.

The validation of the numerical method can be done by comparing the numerical results with the above analytical solution. A homogeneous model whose parameters are the same as the previous model is considered. A P-wave line source of the first derivative Gaussian time function with $f_0 = 40$Hz is located at (0,0). Then the solution at (30 m,30 m) is calculated. The numerical result and analytical solution in Fig. 3 have an excellent agreement.

3.3. Comparison between elastic and poroelastic models

In the conventional elastic model, a single-phase medium is considered. There is no energy dissipation in such an elastic medium. In poroelastic model, however, a multi-phase medium (usually solid frame and pore fluid) is considered. The in-phase motion between solid and fluid leads to the regular P wave and S wave, and the out-phase motion between solid and fluid leads to the slow P wave. Because the fluid is always viscous, there is energy dissipation between the solid frame and the pore fluid. In Biot's theory, the ratio of viscosity and permibility ($b = \eta/\kappa$) determines this dissipation. The difference between an elastic model and a poroelastic model is investigated by studying the influence of b on AVO response in three different ground media: dry sand, fully water saturated sand and partly water saturated sand (70% air and 30% water). For a half-space problem, since the source and receivers are usually on the free surface and the surface wave is the dominant signal, only surface wave is investigated.

Fig. 4, Fig. 5 and Fig. 6 show the influences of b on the surface wave AVO for dry sand, fully water saturated sand and partly water saturated sand models, respectively. The difference between elastic and poroelastic models for dry sand is obvious, but b has very little influence on AVO. For the fully water saturated sand model, b has a significant effect on AVO. When b is very large, the AVO response is very close to that of elastic model. For partly water saturated sand model, the difference between elastic and poroelastic model is obvious and b also has an obvious effect on AVO.

3.4. Applications

In this section, this method is used to model a plastic antipersonnel mine that is difficult to detect with the more conventional electromagnetic induction sensors (*i.e.*, metal detectors). The geometry of the model of a buried mine is shown in Fig. 7.

The parameters for the plastic mine are chosen as $\rho = 1200$ kg/m^3, S-wave velocity $v_{sh} = 1100$ m/s and P-wave velocity $v_p = 2700$ m/s. The material in the ground may be considered as a two-phase composite material consisting of granular solid and pore fluid. The nature of this composite varies with environment, geographic location, and with depth below the surface of ground. In this model, the soil is chosen as dry sand, and the fast P wave velocity is 250 m/s and S wave velocity is 87 m/s. The peak frequency of the source is 450 Hz.

Fig. 8 shows the four snapshots at 4.55 ms, 10.4 ms, 11.05 ms and 15.6 ms. The obvious surface wave reflections are shown on the snapshot at 15.6 ms. Fig. 9 shows the seismograms of the vertical velocity of solid frame at the surface. Fig. 10 shows the scattered fields at 4.55 ms, 10.4 ms, 11.05 ms and 15.6 ms. Fig. 11 shows the seismograms of the scattered fields. Compared to the surface wave, the P wave is very weak. It is the surface wave reflection from the mine that is useful for land mine detection at a relatively low frequency.

4. CONCLUSION

A comprehensive model is developed to simulate the acoustic wave interactions with land mines and soils based on the Biot theory for poroelastic media. The velocity-strain, finite-difference method is then used to solve the Biot equations. For the truncation of the computational domain in the FDTD method, we extend the acoustic and elastic perfectly matched layer (PML) to poroelastic media. Numerical experiments show that, with only 10 cells of PML medium, a high attenuation of about 50 dB can be achieved for outgoing waves. The numerical model is validated by comparison with analytical solutions. Unlike the pure elastic wave model, this efficient PML-FDTD model for poroelastic media incorporates the interactions of waves and the fluid-saturated pore space.

The difference between elastic model and poroelastic model is investigated by studying surface wave amplitude variation with offset in three different ground media: dry sand, fully water saturated sand and partly water saturated sand. For dry sand, b has little influence on surface wave AVO. But for wet sand, b has a profound effect on surface AVO. This numerical method has been used to simulate the interaction of elastic waves with a buried mine-like object in dry sand. The results show that the surface wave response is significantly influenced by the existence of a mine-like object. This numerical model provides a useful tool for system design and data interpretation in acoustic demining.

ACKNOWLEDGMENT

This work is supported by the Army Research Office under Grant DAAH04-96-1-0448 (Demining MURI).

REFERENCES

1. M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. 1.Low-frequency range," *J. Acoust. Soc. Am.* **28**, pp.168–178, 1956.
2. M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. 1.High-frequency range," *J. Acoust. Soc. Am.* **28**, pp.179–191, 1956.
3. M. A. Biot, "Mechanics deformation and acoustic propagation in porous media," *J. Appl. Phys.* **33**, pp.1482–1498, 1962.
4. N. Dai, A. Vafidis and E. R. Kanasewich, "Wave propagation in heterogeneous, porous media: A velocity-stress, finite-difference method," *Geophysics* **60**, pp.327–340, 1995.
5. C. Cerjan, D. Kosloff, R. Kosloff, and M. Reshef, "A non-reflecting boundary condition for discrete acoustic and elastic wave equations," *Geophysics* **50**, pp.705–708, 1985.
6. R. Clayton and B. Engquist, "Absorbing boundary conditions for acoustic and elastic wave equations," *Bull. Seism. Soc. Am.* **67**, pp.1529-1540, 1977.
7. K.D. Mahrer, "An empirical study of instability and improvement of absorbing boundary conditions for elastic wave equation," *Geophysics* **51**, pp.1499–1501, 1986.
8. J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.* **114**, pp.185–200, 1994.
9. W. C. Chew and Q.H. Liu, "Perfectly matched layers for elastodynamics: A new absorbing boundary condition," *J. Comp. Acoust.* **4**(4), pp.72–79, 1996.
10. Q. H. Liu, "Perfectly matched layers for elastic waves in cylindrical and spherical coordinates," *J. Acoust. Soc. Am.* **105**, pp.2075–2084, 1999.
11. F. D. Hastings, J. B. Schneider, and S. L. Broschat, "Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation," *J. Acoust. Soc. Am.* **100**, pp.3061–3069, 1996.
12. Q. H. Liu and J. Tao, "The perfectly matched layer for acoustic waves in absorptive media," *J. Acoust. Soc. Am.* **102**, pp.2072–2082, 1997.
13. Y. Q. Zeng, J. He and Q. H. Liu, "The application of the perfectly matched layer in numerical modeling of wave propagation in poroelastic media," Research report, Duke University (1999).
14. R. Burridge and J. B. Keller, "Poroelasticity equations derived from microstructure," *J. Acoust. Soc. Am.* **70**, pp.1140–1146, 1981.
15. J. Plona, "Observation of the second bulk compressional wave in a porous medium at ultrasonic frequencies," *App. Phys. Lett.* **36**, pp.259–261, 1980.
16. G. Kneib and C. Kerner, "Accurate and efficient seismic modeling in random media," *Geophysics* **58**, pp.576–588, 1993.
17. Q. H. Liu, F. Daube, C. Randall, E. Schoen, H. Liu, and P. Lee, "A 3D finite difference simulation of sonic logging," *J. Acoust. Soc. Am.* **100**, pp72–79, 1996.
18. Q. H. Liu, "Some current trends in numerical methods for transient acoustic and elastic waves in multidimensional inhomogeneous media," *Current Topics in Acoustical Research*, **2**, pp.31–42, 1998.

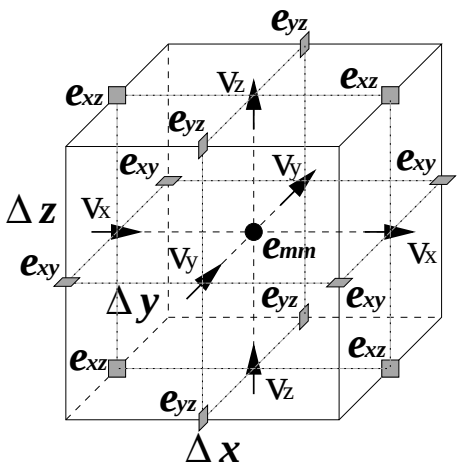


Figure 1. The relative locations of field components in a unit cell of staggered grid.

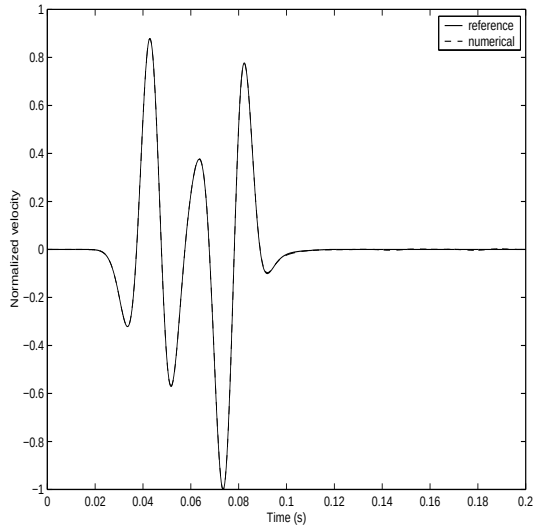


Figure 2. Comparison of a numerical result with PML ABC with a reference solution.

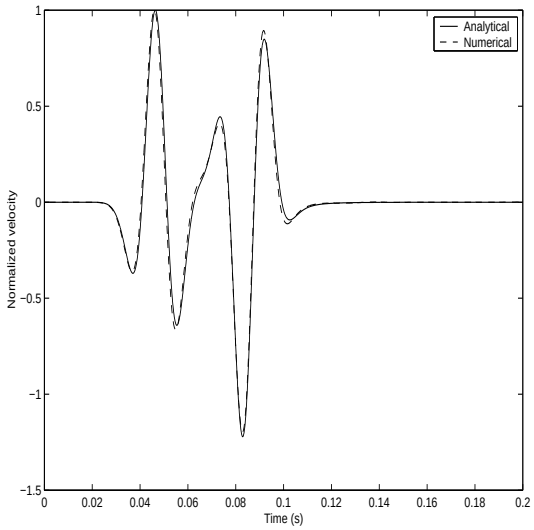


Figure 3. Comparison of analytical and numerical solutions for the vertical velocity component in a homogeneous solid.

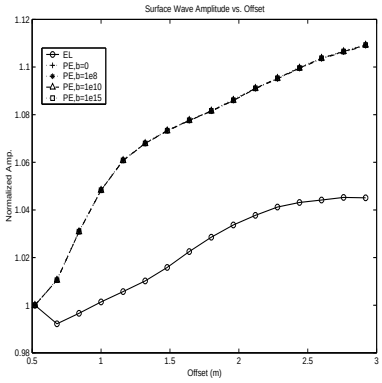


Figure 4. Comparison between elastic and poroelastic model for dry sand.

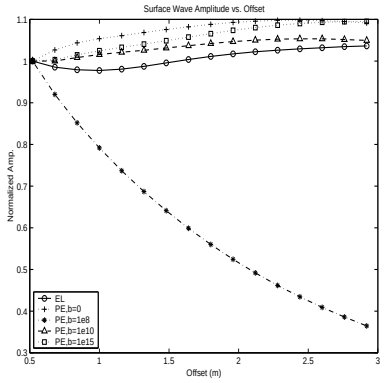


Figure 5. Comparison between elastic and poroelastic model for fully water saturated sand.

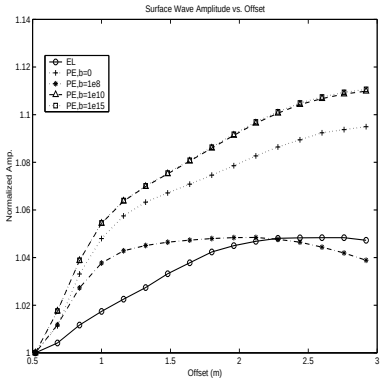


Figure 6. Comparison between elastic and poroelastic model for partly water saturated sand.

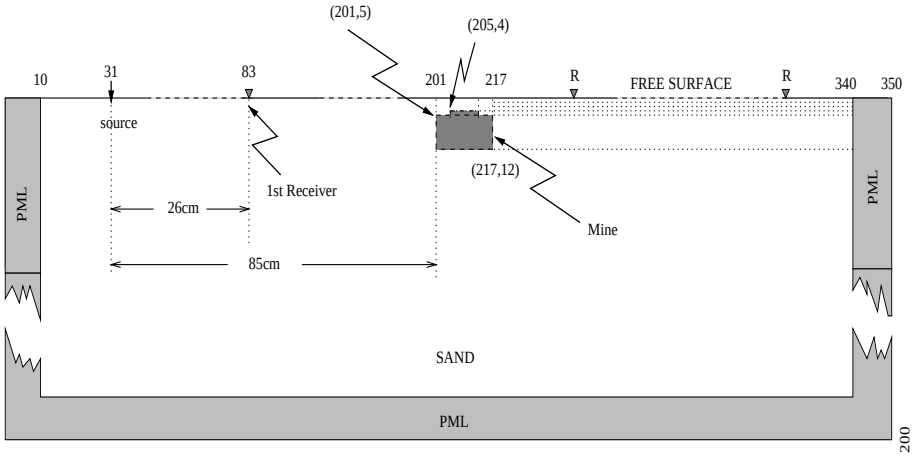


Figure 7. Geometry of a buried land mine model.

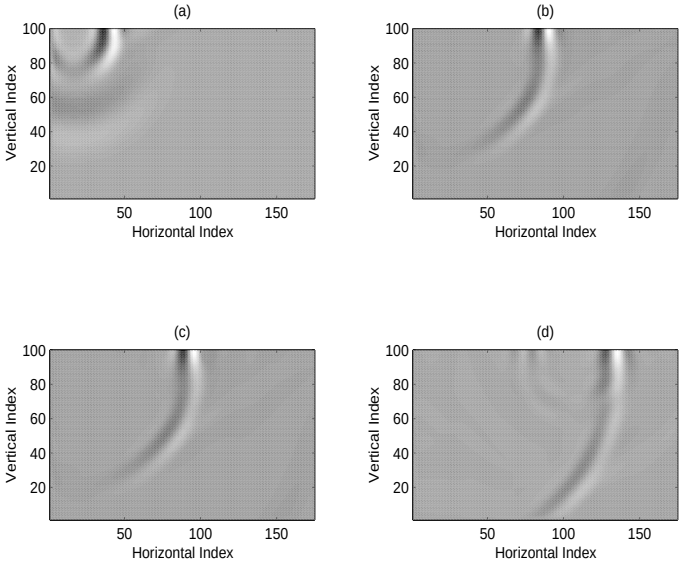


Figure 8. Snapshots at (a) 4.55 ms; (b) 10.4 ms; (c) 11.05 ms; (d) 15.6 ms.

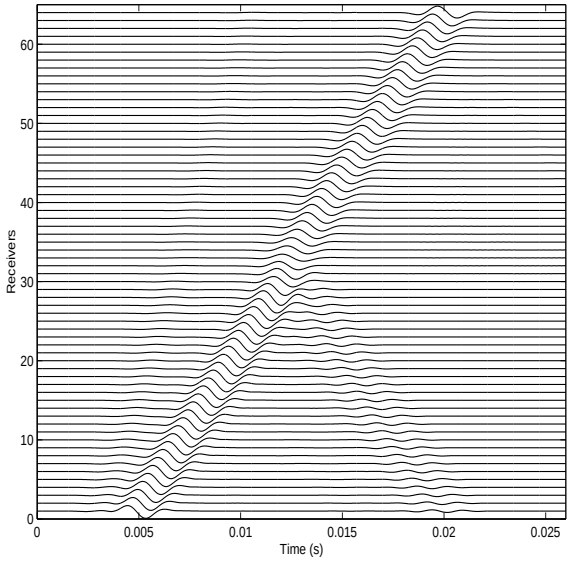


Figure 9. Seismograms of vertical velocity of solid frame.

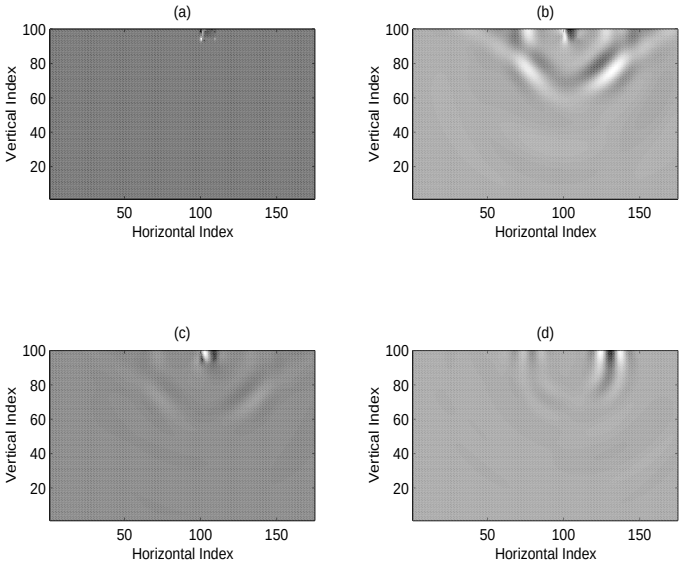


Figure 10. Scattered field at (a) 4.55 ms; (b) 10.4 ms; (c) 11.05 ms; (d) 15.6 ms.

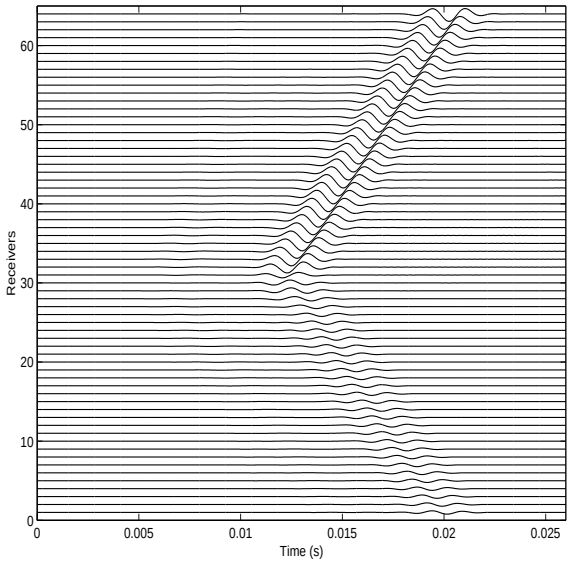


Figure 11. Seismograms of scattered vertical velocity field of solid frame.

Theoretical Study of Microwave Radiometry for Buried Object Detection

Joel T. Johnson

The Ohio State University ElectroScience Laboratory

1320 Kinnear Rd., Columbus, OH 43212-1191

johnson@ee.eng.ohio-state.edu

ABSTRACT

An analytical study of environmental and clutter effects on microwave radiometers used for the detection of buried objects is presented. To simplify the analysis, it is assumed that the soil/target medium has a constant physical temperature versus depth, so that Kirchhoff's law can be applied to determine emissivities, and a simple layered medium geometry is used to model a buried target. Changes in brightness temperatures which result due to the presence of a buried target are illustrated for varying soil dielectric properties, radiometer frequencies, and target depths, and are contrasted with changes in brightness temperatures which can occur when no target is present due to slight soil moisture or soil temperature variations. Brightness temperature clutter due to a small surface roughness is also analytically modeled, through application of the small slope approximation for the homogeneous medium case and the small perturbation method in the presence of a subsurface layer, and it is shown that surface clutter effects can be mitigated through proper choice of sensor polarization and observation angle. Particular attention is given to the relationship between passive (radiometer) and active (ground penetrating radar) microwave sensors; results demonstrate that these two can provide complementary information. Finally, the use of wideband radiometric measurements are discussed as a means for reducing environmental clutter effects and improving detection algorithms.

Keywords: Thermal emission, microwave radiometry, landmine detection

1. INTRODUCTION

Microwave and millimeter wave radiometers are currently being considered for use in military and humanitarian demining applications.[1–7] A radiometer measures the microwave portion of Planck blackbody radiation from the object under view and is sensitive to emissivity variations at greater depths than IR passive sensors due to the greater penetration depths obtained as the frequency is decreased. Previous analytical and experimental studies initially emphasized millimeter wave frequencies[1–4] due to the favorable spatial resolutions that can be obtained, but more recent studies have proposed the use of microwave frequencies to obtain greater sensitivities for deeper targets and for higher moisture content soils.[5–7] Previous efforts with microwave frequencies have focused on laboratory demonstrations,[5] including demonstrations of the potential of synthetic aperture microwave radiometry to improve spatial resolutions,[6] but a more recent study has described field tests performed with a 5 GHz system.[7] Although these tests have shown the detection of buried objects using microwave radiometers, as with any device proposed for humanitarian demining applications, the utility of the sensor should be examined under a wide range of environmental conditions and including potential clutter sources in order to optimize sensor design procedures and to determine conditions under which the sensor would be most effective.

A study of environmental and clutter effects on microwave radiometer systems is presented in this paper. As in previous modeling studies, a layered medium model is used to describe the soil/target medium, and Kirchhoff's law is applied to calculate brightness temperatures under the assumption that the medium has a constant physical temperature versus depth. While these assumptions greatly reduce the complexity of the model, many of the dominant physical mechanisms of emission from a buried target should be retained, as studies with more sophisticated models (including variations in layered medium temperature versus depth) in the soil moisture remote sensing community have shown.[8] Emphasis is placed on the detection of plastic targets, and soil moisture changes in this case are shown to modify target/no-target brightness temperature differences through both dielectric contrast and attenuation effects. Clutter effects are then investigated through consideration of brightness temperature sensitivities to changes in surface physical temperature or moisture content. In addition, the effect of surface roughness is analytically

Figure 1. Geometry of layered medium

modeled and the potential for reducing sensitivity to surface roughness described through choice of sensor observation angle and polarization. The potential of multi-frequency radiometric measurements for reducing surface clutter is also described.

2. EMISSION FROM A LAYERED MEDIUM

The layered medium model to be applied is illustrated in Figure 1 and consists of a boundary between free space and a soil medium with relative permittivity ϵ_1. A target if present is located at depth d below the soil medium boundary and is modeled as a layer of thickness t with relative permittivity ϵ_2. Note the no-target case is obtained when $\epsilon_2 = \epsilon_1$. The entire medium is assumed to have physical temperature T_s, which is set to 290 K throughout the study. To reduce the number of parameters involved, a one inch thick target ($t = 0.0254$ m) with a relative permittivity representative of plastic ($\epsilon_2 = 3.00 + i0.03$) will be used in all cases.

Kirchhoff's law for this configuration states that the brightness temperature measured at polar observation angle θ and in polarization $\gamma = H$ or V is

$$T_{B,\gamma}(\theta) = T_s \left(1 - |\Gamma_\gamma(\theta)|^2\right) \tag{1}$$

where $\Gamma_\gamma(\theta)$ is the reflection coefficient of the layered medium at polar angle θ, and H and V represent horizontal and vertical polarizations, respectively. Note the term inside the parentheses is defined as the medium emissivity. Radiometer measured brightnesses can also obtain a contribution from sky brightnesses reflected off the layered medium,[1] but for microwave systems observing plastic targets this is not a significant contribution. When the layered medium model is applied to describe emission from a buried target, a "beam fill" multiplicative factor should also be included[5] if the target does not completely fill the antenna pattern; such factors depend on the antenna used and thus are not considered in this study.

Figure 2 illustrates vertically (plots (a) and (c)) and horizontally (plots (b) and (d)) polarized brightness temperatures versus observation angle obtained at 1 GHz for a homogeneous medium and for a medium with a target layer at depths of 2.5, 5, 7.5, and 10 cm. Plots (a) and (b) consider a sand medium with volumetric moisture content $m_v = 20$ percent, while plots (c) and (d) consider $m_v = 5$ percent; soil permittivities as a function of soil type, frequency, and moisture content are obtained from an empirical model.[9] Homogeneous medium brightnesses show the distinct variations that occur versus observation angle for horizontal and vertical polarizations, with the maximum vertically polarized brightness obtained at the pseudo-Brewster angle of the medium. These results also illustrate the strong dependence of brightness temperatures on soil moisture with differences of more than 50 K observed between the 20 ($\epsilon_1 = 12.12 + i0.43$) and 5 percent ($\epsilon_1 = 4.32 + i0.08$) moisture cases. Results in plots (a) and (b) including a target layer show the large changes (tens of K) that can occur at 1 GHz even for the deepest target in the large moisture content case. Note both positive and negative deviations from homogeneous brightness temperatures are observed, as will be discussed later. Plots (c) and (d), however, show greatly reduced variations in the low moisture content

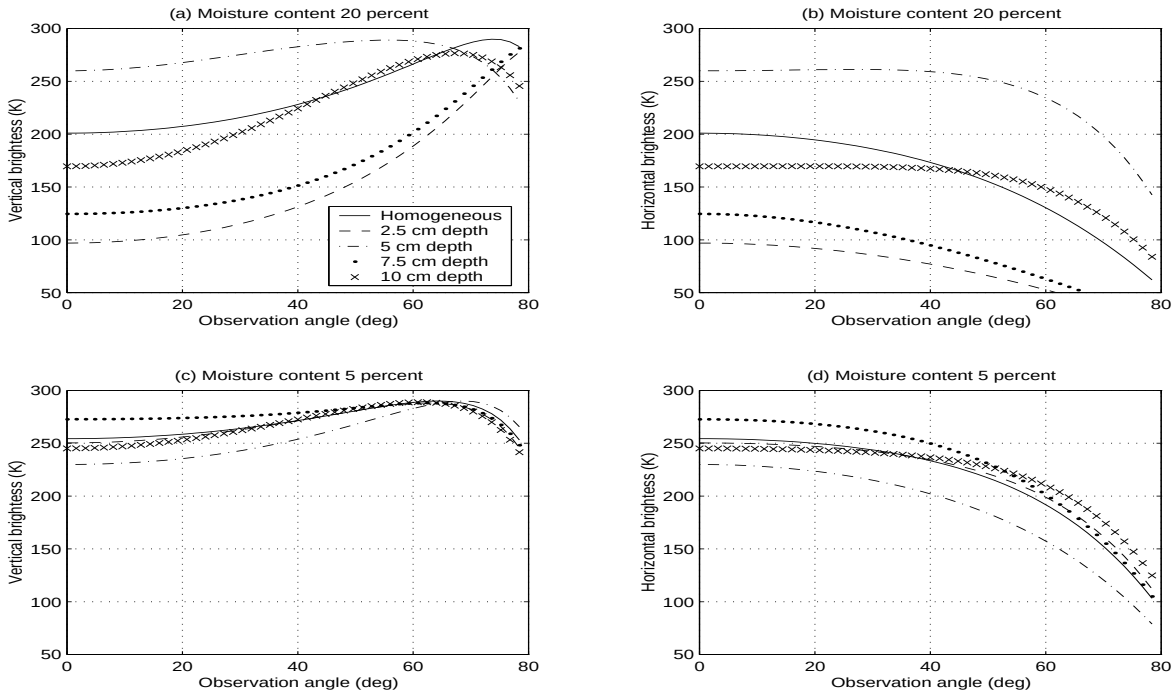


Figure 2. Layered medium brightness temperatures at 1 GHz

case, due to the small dielectric contrast between the target ($\epsilon_2 = 3.00 + i0.03$) and background media. Figure 3 illustrates the same comparisons at 5 GHz ($\epsilon_1 = 11.64 + i2.0$ for 20 percent moisture, $4.24 + i0.36$ for 5 percent moisture), and demonstrates the importance of attenuation in determining the detectability of a buried target. The smaller penetration depths obtained at 5 GHz result in much smaller brightness contrasts in the 20 percent moisture case (plots (a) and (b)), particularly for deeper targets, while dielectric contrast issues again remain a problem in the 5 percent moisture case.

2.1. Relationship with GPR observations

Note if observation at polar angle zero is considered, the dependence on the reflection coefficient is identical to that which a nadir observing ground penetrating radar system would obtain, since the radar system would measure the specular reflection off the layered medium. Thus the basic physics of both radiometer and GPR measurements is identical for nadir observations of layered media, and attenuation and dielectric contrast effects are the major physical factors at work.[10] A coherent GPR system however would also measure scattered field phases that are not captured by radiometric measurements; this information is highly advantageous, since time domain responses can then be constructed to provide additional insight into layered medium structure. A coherent GPR system therefore is clearly superior to a radiometric system for nadir observations of perfectly layered media (i.e., a clutter free environment).

When observation at oblique angles is considered, however, GPR and radiometric measurements respond to significantly different physical phenomena. In the case of a perfectly layered medium excited by an ideal antenna, a GPR system would measure no scattered field returns since no scattering objects are present to generate backscattered fields. In more realistic models, GPR responses would be obtained due to discontinuities in the layered medium structure or due to diffraction from a finite size target or surface features. Radiometric measurements, however, continue to follow Kirchhoff's law at oblique observation angles for an ideal layered medium, and therefore remain measurements of the layered medium reflection coefficient at oblique incidence. Significant sensitivities to a subsurface layer can therefore still be obtained with radiometric measurements at oblique incidence as observed in Figures 2 and 3 without requiring scattering or diffraction effects. The recent increasing interest in "forward-looking" sensors, which operate at oblique incidence, motivates careful consideration of radiometric systems. GPR

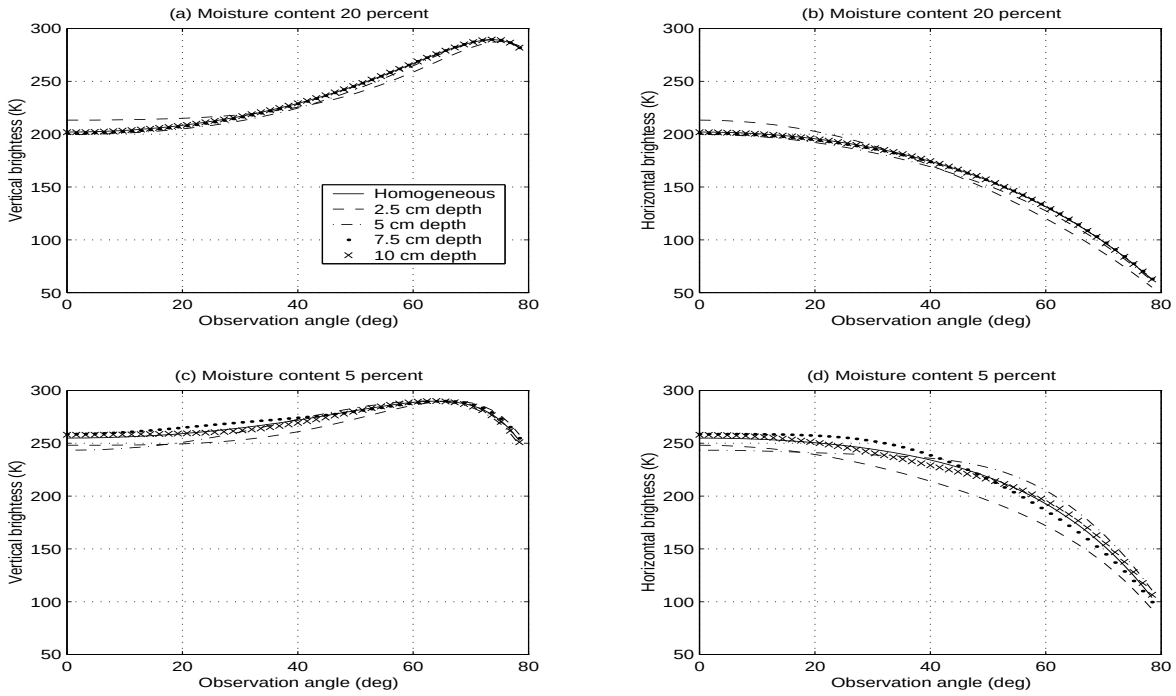


Figure 3. Layered medium brightness temperatures at 5 GHz

and radiometric systems will also obtain differing clutter effects at both nadir and oblique incidence, as discussed in the next section.

2.2. Brightness temperature clutter for a layered medium

It is clear from Figure 3 that in some situations brightness temperature contrasts between the target and no target cases can become small, on the order of a few Kelvin or smaller. The detectability of such small contrasts will be dependent on both system parameters and the level of variations in brightness temperatures ("clutter") that are expected in the homogeneous medium case with no target present. Current microwave technology has made extremely sensitive radiometric measurements possible,[11] so the dominant limiting factor is likely to be the signal to clutter rather than signal to noise ratio. If the ideal layered medium geometry is retained, brightness temperature variations with a homogeneous medium can arise from two sources only: variations in the surface physical temperature (T_s) or in the medium volumetric soil moisture content. Studies of soil surface temperatures[12] demonstrate that variations in T_s of more than 1% are unlikely, providing a corresponding approximate 1% variation in medium brightness temperatures, on the order of $2-3$ K. Variations in soil moisture are clearly even more significant, as observed in Figures 2 and 3. Figure 4 plots sensitivities in Kelvin per percent moisture variation for $m_v = 5$ percent and 20 percent at frequencies 1 and 5 GHz. This figure shows the large sensitivities to soil moisture obtained by a microwave radiometer, particularly for smaller soil moisture contents. This fact makes the radiometer an advantageous sensor for the remote sensing of soil moisture,[9] but also represents a serious limitation for radiometric systems observing low contrast or highly attenuated targets. Note that sensitivity to soil moisture can be reduced by observing in vertical polarization at angles between 60 to 70 degrees (near the pseudo Brewster angle); however at these angles Figures 2 and 3 also show reduced contrasts in the presence of a buried target. Assuming an ideal layered medium with variations in temperature and soil moisture of only 1 percent, targets having brightness temperature contrasts less than approximately 10 Kelvin would be difficult to separate from environmental false alarms with a single frequency radiometer.

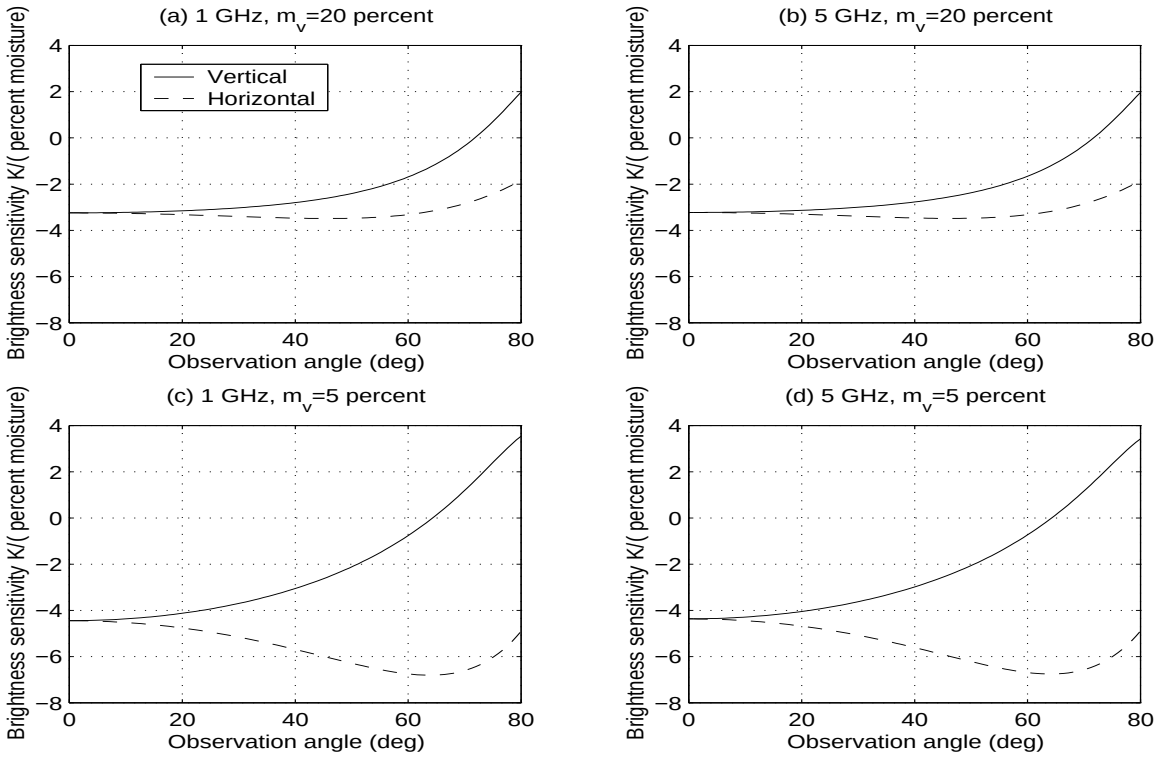


Figure 4. Layered medium brightness temperature sensitivities to changes in soil moisture

2.3. Multi-frequency observations

To address clutter issues for the layered medium geometry, radiometric measurements over a range of frequencies are proposed. Figures 5 and 6 plot brightness temperatures as a function of frequency between 1 and 5 GHz for the homogeneous medium and including a buried target at observation angles of 0 (nadir) and 50 degrees, respectively. Note the oscillatory patterns of brightness temperatures obtained in the presence of a subsurface layer, due to the coherent interference effects that occur in the reflection coefficient with the subsurface layer. The reduced oscillation amplitude observed in the 3 and 4.5 GHz range occurs in the region where the thickness of the target is approximately one half of a wavelength. Although an incoherent model was found most suitable for millimeter wave frequencies in a previous study,[1] coherent effects are more likely to be observed at microwave frequencies (for targets which fill a large percentage of the antenna pattern) due to reduced scattering effects in the background medium at lower frequencies; coherent reflection processes have been observed even at millimeter wave frequencies with radiometry systems for the detection of sea surface oil slicks.[13] The oscillatory pattern versus frequency occurs only in the presence of a subsurface layer; homogeneous medium brightness temperature variations are within 1.5 K including permittivity variations with frequency from the empirical model,[9] and show a near linear dependence on frequency. The source of the positive and negative single frequency target/no target contrasts observed in Figures 2 and 3 is now apparent, due to the oscillatory pattern which varies with target depth, as well as the fact that targets at some depths will produce no contrast for single frequency observations. Note that multiple frequency observations eliminate sensitivities to soil temperature and soil moisture variations, since in all cases homogeneous medium brightness show only slight variations versus frequency. Multiple frequency brightness temperature measurements also allow some of the "phase" information that would be obtained in a GPR measurement to be recovered, since the rapidity of the oscillatory pattern depends on target depth. A detection algorithm based on multiple frequency brightness measurements could simply search for any oscillatory pattern in the data (for example through a Fourier transform based algorithm), and in the ideal layered medium case would eliminate almost all false alarms due to environmental clutter even with small target/no target contrasts or heavily attenuated targets.

Of course, such near perfect performance will not be obtained when more realistic target and clutter emission models are included. However, the basic success of multiple frequency radiometric measurements for the simple

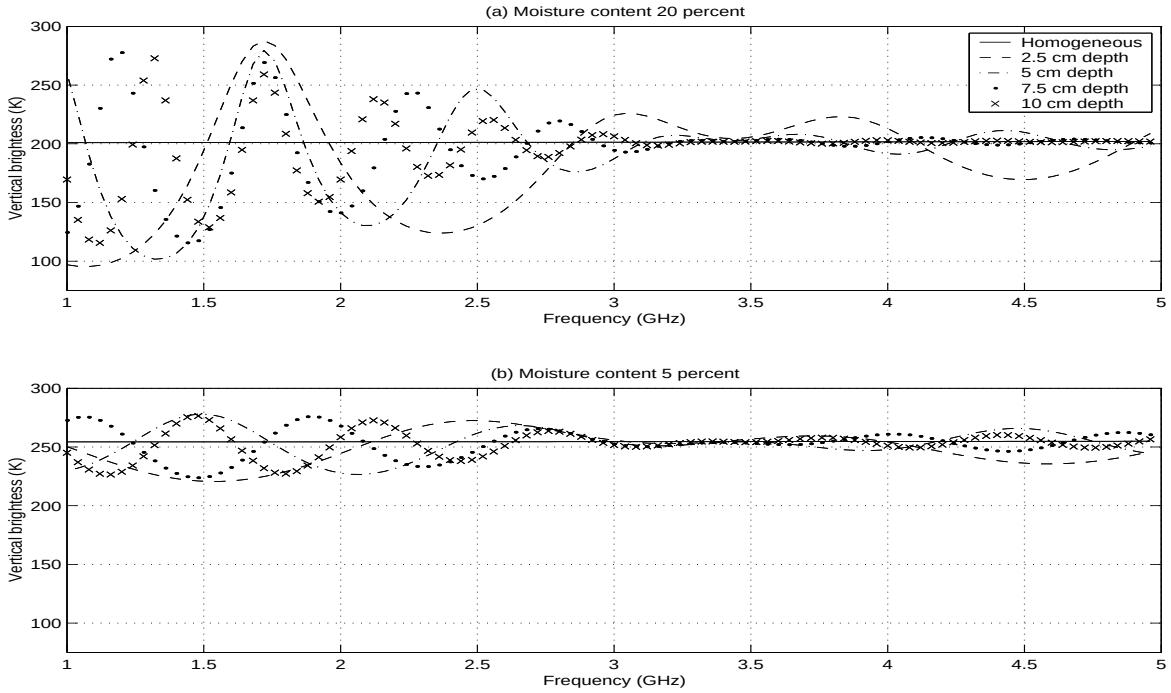


Figure 5. Layered medium brightness temperatures versus frequency for nadir observation

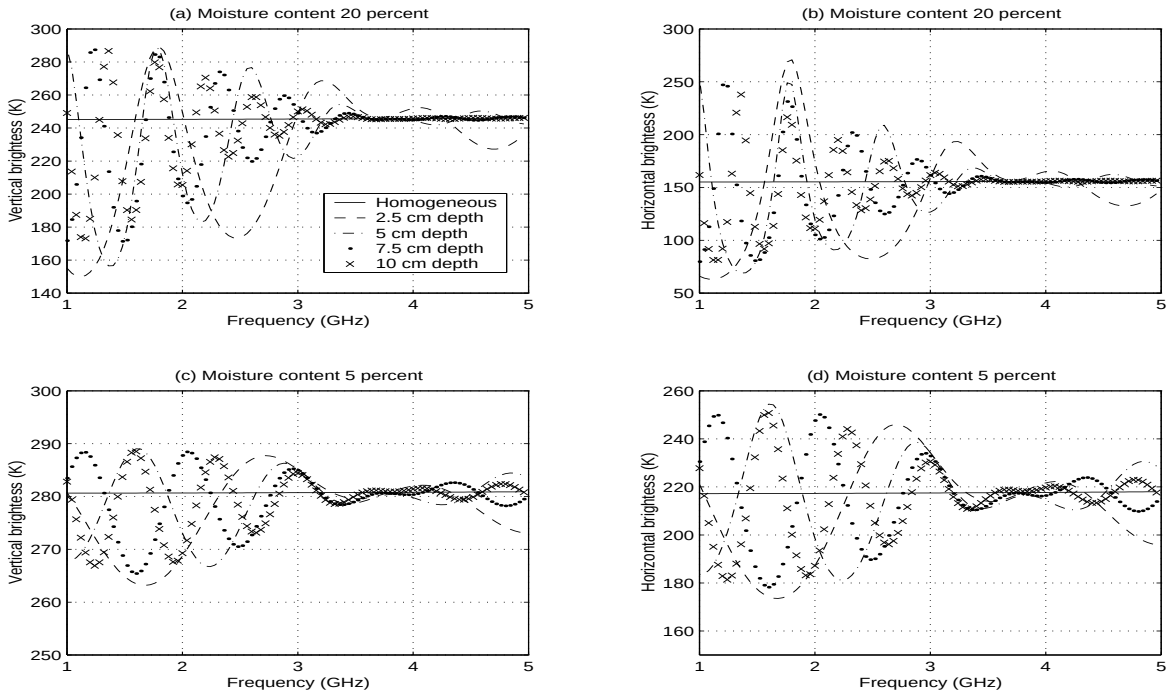


Figure 6. Layered medium brightness temperature versus frequency for observation at 50 degrees

Figure 7. Geometry of layered medium with a rough surface

layered medium geometry motivates further investigation. A model for including surface roughness effects as an additional clutter source is described in the next section.

3. SURFACE ROUGHNESS EFFECTS

Figure 7 illustrates a layered medium bounded by a rough surface $z = f(x, y)$, described statistically as a random process. Analytical models for emission from a homogeneous medium bounded by a rough surface (i.e., when $\epsilon_2 = \epsilon_1$) have been considered extensively,[14–17] again through the application of Kirchhoff's law for a medium at constant physical temperature. A small slope approximation (SSA) exists for this problem, which predicts average brightness temperature changes due to roughnesses with small slopes; the method is highly effective[16,17] and has been confirmed through comparison with results from numerical methods.[18] A theory for emission from a rough surface bounding a layered medium has also been developed[19] based on extension of the small perturbation method (SPM),[20–22] but in this case only a small height theory is obtained (i.e., the rough surface height must remain much smaller than an electromagnetic wavelength). Surface roughness clearly represents an additional clutter source for radiometric observations of buried objects, but brightness temperature variations are somewhat more difficult to characterize due to their strong dependence on the statistics of the surface. Current research efforts[23,24] are seeking to improve statistical descriptions of soil surface profiles, but at present only simple models are available for estimating surface roughness effects. Two such models are applied below, both of which model the soil surface as a Gaussian random process. The first model postulates an isotropic Gaussian function for the surface power spectral density, parametrized by the rms height, h, and correlation length, l. The second model postulates a power law form for the surface power spectral density, a_0/k^n, where a_0 and n are constants and k is the surface spatial frequency. The power law spectrum is truncated to exist only between spatial frequencies k_l and k_u to model surface variations on a limited set of length scales. Of these two models, the Gaussian model represents a process dominated by only a small range of length scales (around the correlation length), while the power law model is a more multi-scale or fractal type description. Both models are commonly applied in descriptions of soil surface scattering, although the multi-scale model is currently proposed as more realistic.

Figure 8 plots roughness induced changes in homogeneous medium brightnesses obtained from the SSA for a Gaussian roughness spectrum versus the rough surface correlation length in wavelengths. Surfaces in this plot are assumed to have a constant total rms slope ($2\frac{h}{l}$) of 0.1 as the correlation length is varied, but results at other rms slope values simply scale quadratically as long as the small slope assumption is satisfied. Curves for two medium permittivities are illustrated, corresponding approximately to sand with water contents 5 and 20 percent. Results demonstrate the moderate sensitivities to surface roughness obtained with radiometric measurements for small slope surfaces, particularly for near-nadir observation angles and large correlation length features. Brightness temperatures for nadir observation are observed to be completely insensitive to roughness in length scales much larger than the electromagnetic wavelength, as obtained with large correlation length Gaussian spectrum surfaces. Use of multiple frequency measurements is not observed necessarily to reduce rough surface clutter, since changes versus correlation



Figure 8. Roughness induced changes in flat surface brightness temperatures: Gaussian spectrum

length relative to wavelength are an implicit function of frequency, but for large length scale Gaussian spectrum surfaces use of nadir observations will prove advantageous in eliminating surface clutter effects. Figure 9 plots roughness induced changes in brightness temperatures for three observation angles versus frequency for $h = 5$ mm and $l = 5$ cm with soil permittivities (assumed constant versus frequency) of $12.12 + i0.43$ and $4.32 + i0.08$. Although the roughness clutter is appreciable, again the pattern observed versus frequency is not oscillatory, so that the oscillatory patterns associated with a subsurface layer could still potentially be extracted. Results with alternate values of h and l also show similar results.

Figure 10 plots roughness induced changes in homogeneous medium brightnesses versus observation angle obtained for power law spectrum surfaces with $n = 4$ and $a_0 = 10^{-3}$. Surfaces are truncated to include length scales between 2λ and 0.01λ only (i.e., $k_l = \frac{2\pi}{2\lambda}$, $k_u = \frac{2\pi}{0.01\lambda}$), where λ is the electromagnetic wavelength in free space; the results of Figure 10 are then independent of radiometer frequency. Again permittivities modeling sand moisture contents of approximately 5 and 20 percent are included. In this case, brightness temperatures remain sensitive to surface roughness at near-nadir observation angles, due to the multi-scale roughness which by definition includes roughness on length scales near the electromagnetic wavelength, so that the "large scale" assumption is not valid. Vertical polarization however is observed typically to have smaller sensitivity to surface roughness, and an angle exists (as has been described in studies of sea surface remote sensing[9]) at which sensitivity to surface roughness vanishes (the zero crossings in Figure 10, plot(a)). Thus, oblique incidence observations in vertical polarization seem better suited for reducing surface clutter effects with power law spectra. The exact angle at which surface clutter effects vanish is a function both of the surface permittivity as observed in Figure 10 and of the power law spectrum parameters a_0, n, k_l, and k_u. Thus, an improved understanding of the multiscale characteristics of soil surfaces will be necessary in order to determine the range of likely angles at which surface clutter sensitivity will be reduced. Figure 11 plots roughness induced brightness temperature changes versus frequency for a power law spectrum with $n = 4$, $a_0 = 10^{-3}$, $k_l = 2\pi/0.3$ rads/m and $k_h = 2\pi/0.003$ rads/m (i.e., a surface having variations on length scales ranging from 30 cm to 3 mm). Again note the non-oscillatory patterns of surface clutter contributions versus frequency.

As a final example, surface roughness effects in the presence of a subsurface layer are considered through the SPM. Figure 12 illustrates the differing levels of surface clutter that are obtained with and without a sub-surface layer

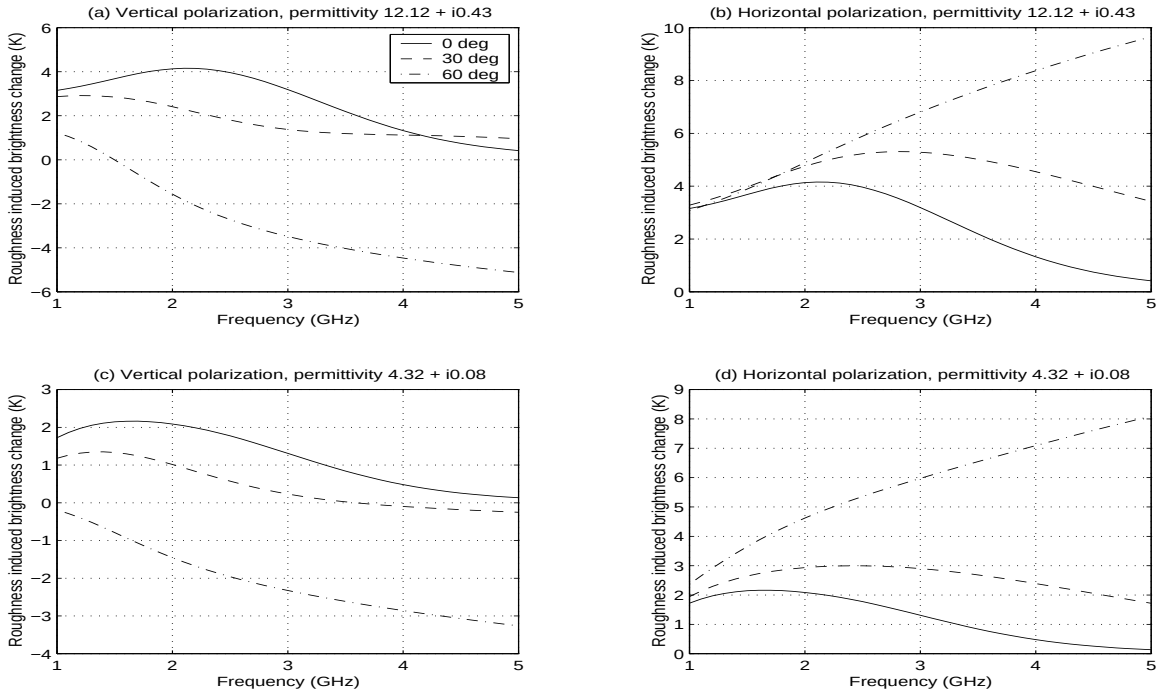


Figure 9. Roughness induced changes in flat surface brightness temperatures versus frequency for $h = 5$ mm, $l = 5$ cm

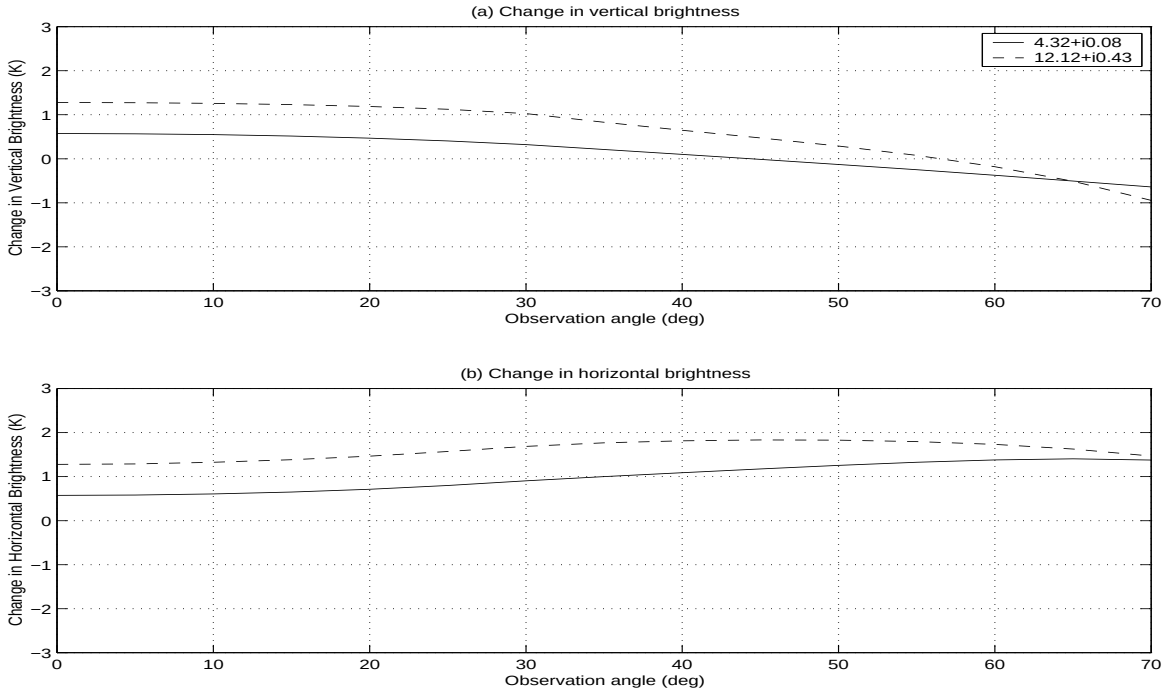


Figure 10. Roughness induced changes in brightness temperatures: Power law spectrum

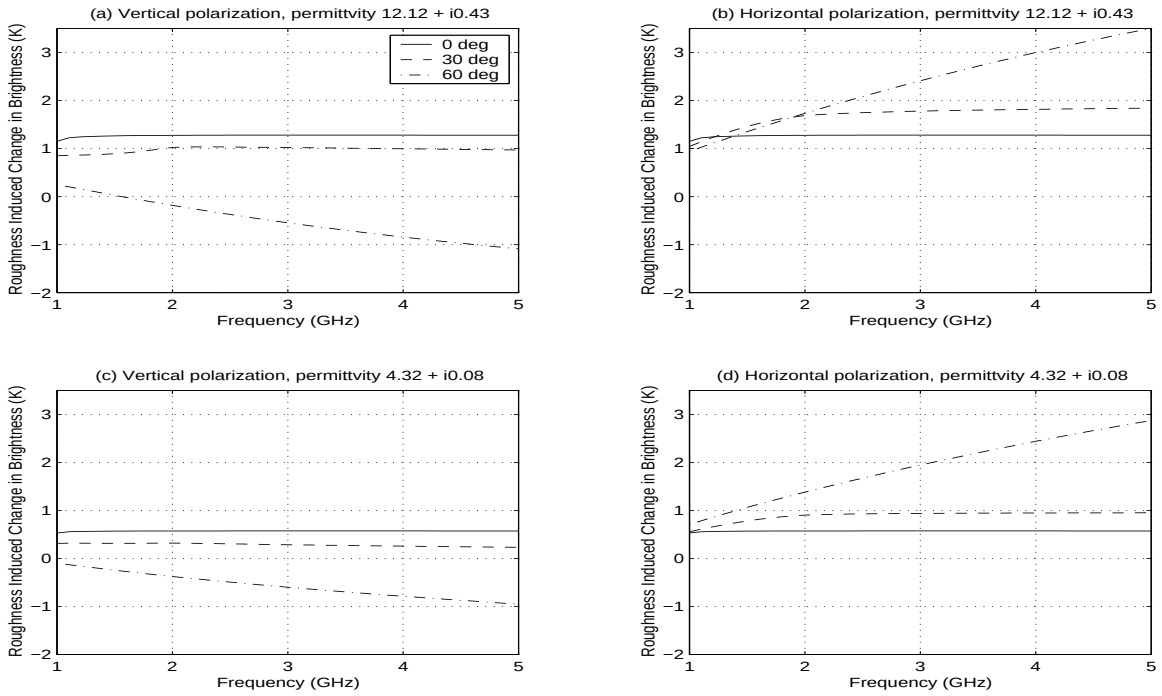


Figure 11. Roughness induced changes in brightness temperatures versus frequency: power law spectrum with $n = 4$, $a_0 = 10^{-3}$, $k_l = 2\pi/0.3$ rads/m, $k_h = 2\pi/0.003$ rads/m

versus observation angle for a Gaussian surface with $h = 0.05\lambda$ and $l = 0.33\lambda$ (note that brightnesses in the absence of surface clutter also vary depending on target depth and observation angle.) Results at both 1 and 5 GHz are plotted for soils with moisture content 20 percent. Although these results indicate that the presence of a subsurface layer can significantly modify surface clutter effects, the small height theory limits the current study to relatively small clutter effects only. Further conclusions regarding surface clutter changes in the presence of a subsurface layer will require an improved analytical theory (for example, a small slope theory) of emission from a layered medium.

4. CONCLUSIONS AND FUTURE WORK

A study of environmental and clutter effects on microwave radiometer systems for the detection of buried objects has shown that environmental clutter of $5-6$ K should not be unusual even for observation of a perfectly layered medium. The presence of a rough surface can further increase this clutter level depending on the surface statistics, but choices of observation angles and sensor polarizations exist to mitigate surface contributions if sufficient knowledge of surface statistics is available. Use of multiple frequency radiometric measurements is proposed to improve performance in the presence of clutter, with the oscillatory pattern versus frequency observed in the presence of a subsurface layer providing a clear means of distinguishing target and clutter contributions. Funds for the development of a multiple frequency radiometer system have recently been provided to The ElectroScience Laboratory under a Defense University Research Instrumentation Program (DURIP) grant, so research toward evaluation of multiple frequency radiometric measurements in practical environments will proceed as the sensor is developed. Further modeling efforts to include finite size target emission effects and to obtain a small slope theory of emission from a rough surface bounding a layered medium are also planned.

ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the Army Research Office (ARO) (OSD MURI on Humanitarian Demining) and by the Office of Naval Research (ONR). The findings, opinions and



Figure 12. Roughness induced changes in brightness temperatures in the presence of a sub-surface layer: Gaussian spectrum with $h = 0.05\lambda$, $l = 0.33\lambda$

recommendations expressed therein are those of the authors and are not necessarily those of Duke University, ARO, or ONR.

REFERENCES

1. L. Yujiri, B. Hauss, and M. Shoucri, "Passive millimeter wave sensors for detection of buried mines," *Detection and Remediation Technologies for Mines and Minelike Targets, Proceedings of SPIE* **vol. 2496**, pp. 2–6, 1995.
2. L. Yujiri, S. Fornaca, B. Hauss, M. Shoucri, and S. Talmadge, "Detection of metal and plastic mines using passive millimeter waves," *Detection and Remediation Technologies for Mines and Minelike Targets, Proceedings of SPIE* **vol. 2765**, pp. 330–336, 1996.
3. J. Groot, R. Dekker, and L. van Ewijk, "Landmine detection with an imaging 94 GHz radiometer," *Detection and Remediation Technologies for Mines and Minelike Targets, Proceedings of SPIE* **vol. 2765**, pp. 337–347, 1996.
4. B. Blume, A. Resnick, J. Foster, J. Albers, N. Witherspoon, and J. Holloway, "PMMW data collection results," *Detection and Remediation Technologies for Mines and Minelike Targets III, Proceedings of SPIE* **vol. 3392**, pp. 167–173, 1998.
5. L. Yuriji, B. Hauss, and M. Shoucri, "Microwave/millimeter wave radiometric detection of metal and plastic mines," *Detection and Remediation Technologies for Mines and Minelike Targets II, Proceedings of SPIE* **vol. 3079**, pp. 652–658, 1997.
6. R. Tan, R. Bender, and S. Stratton, "Synthetic aperture interferometric microwave radiometry for remote sensing of mines," *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proceedings of SPIE* **vol. 3710**, pp. 1003–1014, 1999.
7. G. D. Amici, B. Hauss, and L. Yujiri, "Detection of landmines via a passive microwave radiometer," *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proceedings of SPIE* **vol. 3710**, pp. 716–724, 1999.

8. T. J. Schmugge and B. J. Choudbury, "A comparison of radiative transfer models for predicting microwave emission from soils," *Radio Science* **vol. 16**, pp. 927–938, 1981.
9. F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, Artech House, Norwood, MA, 1986.
10. J. T. Johnson, J. Jenwatanavet, N. Wang, R. Caldecott, and J. Young, "Waveguide studies of soil modification techniques for enhanced mine detection with ground penetrating radar," *Detection and Remediation Technologies for Mines and Minelike Targets III, Proceedings of SPIE* **vol. 3392, part II**, pp. 716–724, 1998.
11. N. Skou, *Microwave Radiometer Systems: Design and Analysis*, Artech House, Norwood, MA, 1989.
12. I. R. Sendur and B. A. Baertlein, "Techniques for improving buried mine detection in thermal IR imagery," *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proceedings of SPIE* **vol. 3710**, 1999.
13. N. Skou, "Microwave radiometry for oil pollution monitoring, measurements, and systems," *IEEE Trans. Geosc. Rem. Sens.* **vol. GE-24**, pp. 360–367, 1986.
14. L. Tsang and R. W. Newton, "Microwave emissions from soils with rough surfaces," *J. Geophys. Res.* **vol. 87**, pp. 9017–9024, 1982.
15. S. H. Yueh, R. Kwok, F. K. Li, S. V. Nghiem, and W. J. Wilson, "Polarimetric passive remote sensing of ocean wind vectors," *Radio Science* **vol. 29**, pp. 799–814, 1994.
16. V. G. Irisov, "Small-slope expansion for thermal and reflected radiation from a rough surface," *Waves in Random Media* **vol. 7**, pp. 1–10, 1997.
17. J. T. Johnson and M. Zhang, "Theoretical study of the small slope approximation for ocean polarimetric thermal emission," *IEEE Trans. Geosc. Rem. Sens.* **vol. 37**, pp. 2305–2316, 1999.
18. J. T. Johnson, R. T. Shin, J. A. Kong, L. Tsang, and K. Pak, "A numerical study of ocean polarimetric thermal emission," *IEEE Trans. Geosc. Rem. Sens.* **vol. 37**, pp. 8–20, 1999.
19. J. T. Johnson, "Thermal emission from a layered medium bounded by a slightly rough interface," *submitted to IEEE Trans. Geosc. Rem. Sens.* , 1999.
20. S. O. Rice, "Reflection of electromagnetic waves from slightly rough surfaces," *Commun. Pure Appl. Math* **vol. 4**, pp. 361–378, 1951.
21. I. M. Fuks, "Radar contrast polarization dependence on subsurface sensing," *Proceedings of the 1998 International Geoscience and Remote Sensing Symposium (IGARSS'98)* **vol. 3**, pp. 1455–1459, 1998.
22. J. T. Johnson, "Third order small perturbation method for scattering from dielectric rough surfaces," *J. Opt. Soc. Am. A* **vol. 167**, pp. 2720–2736, 1999.
23. C. H. Huang and J. M. Bradford, "Applications of a laser scanner to quantify soil microtopography," *Soil Science Society of America Journal* **vol. 56**, pp. 14–21, 1992.
24. J. L. Salvati and J. T. Johnson, "Numerical and experimental studies of ground surface clutter with measured profiles," *Proceedings of the 1999 URSI General Assmeble: Toronto* , 1999.

Acoustic landmine detection: a 3D poroelastic model

Y. Q. Zeng[a] and Q. H. Liu[a]

[a]Department of Electrical and Computer Engineering,
Duke University, Durham, North Carolina 27708, U.S.A.

ABSTRACT

Acoustic waves can be a viable tool for the detection and identification of land mines, unexplored ordnance (UXO) and other buried objects. Design of acoustic instruments and interpretation and processing of acoustic measurements call for accurate numerical models to simulate acoustic wave propagation in a heterogeneous soil with buried objects. Compared with the traditional seismic exploration, high attenuation is unfortunately ubiquitous for shallow surface acoustic measurements because of the loose soil and the fluid in its pore space. To adequately model such acoustic attenuation, we propose a comprehensive multidimensional finite-difference time-domain model to simulate the acoustic wave interactions with land mines and soils based on the Biot theory for poroelastic media. For the truncation of the computational domain, we use the perfectly matched layer (PML). The method is validated by comparison with analytical solutions. Unlike the pure elastic wave model, this efficient PML-FDTD model for poroelastic media incorporates the interactions of waves and the fluid-saturated pore space. Several typical land mine detection measurements are simulated to illustrate the application.

Keywords: poroelastic, land mine detection, staggered-grid, elastic wave, PML, finite-difference

1. INTRODUCTION

Recently acoustic methods have received considerable attention in the characterization of shallow objects. It has been shown that acoustic waves can be a viable tool for the detection and identification of land mines and unexplored ordnance (UXO),[1] [2] (also see some experimental papers in this special issue). The critical issue of acoustic wave coupling into the soil has been addressed by using low-frequency loud-speakers, laser, off-set mechanical source, and electric spark sources.[2,?] These recent experimental efforts have shown promising results for acoustic landmine detection.

Design of acoustic instruments, interpretation and processing of acoustic measurements call for accurate numerical models to simulate acoustic wave propagation in a heterogeneous soil with buried objects. Conventional methods for acoustic modeling include the popular finite-difference time-domain (FDTD) solution of pure elastic media,[3] .[4] Unfortunately, the pure elastic model cannot adequately incorporate the attenuation mechanism in its governing equations, although some approximate models are possible.[5]

Compared with the traditional seismic exploration, high attenuation is unfortunately ubiquitous for shallow surface acoustic measurements because of the loose soil and the fluid in its pore space. To adequately model such acoustic attenuation, we propose a comprehensive model to simulate the acoustic wave interactions with land mines and soils based on the Biot theory[6]-[8] for poroelastic media. The finite-difference time-domain method is then used to solve the Biot equations.[8] Similar to the velocity-stress FD method,[9] a velocity-strain, finite-difference method is developed in a staggered grid for heterogeneous poroelastic media. In this method, Biot equations[8] are reformulated into first-order equations to arrive at a leap-frog system in a staggered grid both in time and space domains.

In order to simulate an unbounded medium, an absorbing boundary condition (ABC) must be implemented to truncate the computational domain in numerical algorithms. Several ABCs have been developed for numerical simulation of wave propagation. Cerjan *et al.*[10] introduce a simple damping taper to the boundaries that attenuates outgoing waves. Since this lossy layer is not perfectly matched to the interior region, however, it requires a substantial number of layers for the taper to be effective. Clayton and Engquist[11] (CE) use the paraxial approximation to the wave equation to make the boundary transparent to outgoing waves. The CE ABC can lead to instability when

Further author information: (Send correspondence to Y.Q.Zeng)
Y.Q.Zeng: E-mail: yz3@ee.duke.edu
Q.H.Liu: E-mail: qhliu@ee.duke.edu

the Poisson's ratio is greater than two.[12] Since Berenger[13] proposed the highly effective perfectly matched layer (PML) as an absorbing boundary condition for electromagnetic waves, it has been widely used for finite-difference and finite-element methods. Chew and Liu[14] first proposed the PML for elastic waves in solids, and proved the zero reflections from PML to the regular elastic medium. Hastings *et al.*[15] have independently implemented the PML ABC for two-dimensional problems by using potentials. The PML has also been extended to model acoustic waves and electromagnetic waves in lossy media[16] and to cylindrical and spherical coordinates.[17] Recently, PML has been applied to the second-order Biot's equations for 2-D fluid-saturated poroelastic media,[18] which requires a complicated convolution. In this paper, we apply the PML to 3-D using the first-order PDE system.

The difference between elastic model and poroelastic model is investigated by studying surface wave amplitude variation with offset (AVO) in three different types of soil: dry sand, fully water saturated sand and partly water saturated sand. The interaction of elastic wave with a plastic mine buried in dry sand is simulated. We also simulate two typical configurations for land mine detection, namely the excitation in soil,[1] and the excitation in air.[2] The results show that the wave responses are significantly affected by the existence of a mine-like object through surface waves. After processing, the target can be detected by using surface acoustic measurements. The 2-D and 3-D poroelastic codes provide a useful tool for design and detection issues in acoustic characterization of buried objects.

2. FORMULATION

2.1. The governing equations

The propagation of acoustic waves in fluid-saturated porous media is different from that in single phase elastic media. In addition to the regular P (compressional) waves and S (shear) waves in solid elastic media, a slow P wave resulting from the relative motion between the solid frame and pore fluid may be present in porous media. Thus energy dissipation in porous media is different from that in solid elastic media. Based on continuum mechanics and macroscopic constitutive relationship, Biot[6-8] developed a theory of wave motion in a poroelastic solid saturated with a viscous compressible fluid. Biot's theory was confirmed by Burridge and Keller[19] based on the dynamic equations governing the behavior of medium on a microscopic scale. Plona[20] also confirmed Biot's theory through experiments.

For a three-dimensional isotropic, heterogeneous and porous elastic medium, wave propagation is governed by Biot's equations[8]:

$$2\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) + \frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) = \frac{\partial^2}{\partial t^2}(\rho u_i + \rho_f w_i) \tag{1}$$

$$\frac{\partial}{\partial x_i}(\alpha M e - M\xi) = \frac{\partial^2}{\partial t^2}(\rho_f u_i + m w_i) + \frac{\eta}{\kappa}\frac{\partial w_i}{\partial t} \tag{2}$$

where $m = a\rho_f/\phi$ and

$$M = \frac{1}{\frac{\phi}{K_f} + \frac{(\alpha-\phi)}{K_s}}$$

$$\alpha = 1 - \frac{K_b}{K_s}$$

Let $\mathbf{v^s} = \frac{\partial \mathbf{u}}{\partial t}$ be the velocity of the solid particle, $\mathbf{v^f} = \frac{\partial \mathbf{w}}{\partial t}$ be the velocity of the pore fluid relative to the solid frame. Then the second-order equations (1) and (2) can rearranged as the first-order equations,

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial v_i^s}{\partial t} &= 2m\sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) + m\frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) \\ &\quad -\rho_f \frac{\partial}{\partial x_i}(\alpha M e - M\xi) + \rho_f \frac{\eta}{\kappa} v_i^f \end{aligned} \tag{3}$$

$$\begin{aligned}(m\rho - \rho_f^2)\frac{\partial v_i^f}{\partial t} &= \rho\frac{\partial}{\partial x_i}(\alpha M e - M\xi) - \rho\frac{\eta}{\kappa} v_i^f \\ &\quad -2\rho_f \sum_j \frac{\partial}{\partial x_j}(\mu e_{ij}) - \rho_f \frac{\partial}{\partial x_i}(\lambda_c e - \alpha M\xi) \end{aligned} \tag{4}$$

The time derivatives of strains e_{ij} and ξ can be expressed as

$$\frac{\partial e_{ij}}{\partial t} = [(\partial v_j^s/\partial x_i + \partial v_i^s/\partial x_j)/2] \tag{5}$$

$$\frac{\partial \xi}{\partial t} = -\nabla \cdot \mathbf{v^f} \tag{6}$$

In this work, an explicit second-order finite-difference scheme is applied to convert equations (3) to (6) into a leap-frog system for the strain field e_{ij}, ξ and velocity field $\mathbf{v^s}$ and $\mathbf{v^f}$. With proper absorbing boundary conditions, these equations can be solved numerically for the wave field in an unbounded medium.

Equations (3) to (6) predict the existence of three different waves in fluid-saturated poroelastic media: a shear wave and two compressional waves with a faster and a slower propagation velocities[6]-.[8]

2.2. Equations for the PML absorbing boundary condition

In this paper, the perfectly matched layer (PML) is used to truncate the unbounded medium to absorb all outgoing waves. This artificial absorptive medium is introduced in the regular medium by modifying equations (3) to (6) with complex coordinates.[14,17] In the frequency domain with $e^{-i\omega t}$ convention, a complex coordinate variable is chosen as

$$\tilde{x_j} = \int_0^{x_j} e_j({x_j}')d{x_j}' \tag{7}$$

$$e_j = a_j + i\frac{\omega_j}{\omega} \qquad (j = 1, 2, 3) \tag{8}$$

where $a_j \geq 1$ is a scaling factor, and $\omega_{x_j} \geq 0$ is an attenuation factor. The operator $\frac{\partial}{\partial x_j}$ can be expressed in terms of the regular coordinate,

$$\frac{\partial}{\partial \tilde{x_j}} = \frac{1}{e_j}\frac{\partial}{\partial x_j} \tag{9}$$

The PML formulation is to replace x_j in (4)-(6) by the corresponding complex coordinate $\tilde{x_j}$. In a regular non-PML region, $a_j = 1$ and $\omega_j = 0$. In order to simplify PML equations, the field variables are split as the following,

$$v_j^s = \sum_{k=1}^{3} v_j^{s(k)}$$

$$v_j^f = \sum_{k=1}^{3} v_j^{f(k)}$$

where $v_i^{s(k)}$ and $v_i^{f(k)}$ represent the split field variables containing space derivative $\frac{\partial}{\partial x_k}$ only .

The Biot's equations yield the time-domain PML equations (for component v_1),

$$\begin{aligned}(m\rho - \rho_f^2)(\frac{\partial v_1^{s(1)}}{\partial t} + \omega_1 v_1^{s(1)}) &= 2m\frac{\partial}{\partial x_1}(\mu e_{11}) + m\frac{\partial}{\partial x_1}(\lambda_c e - \alpha M\xi) \\ &\quad -\rho_f\frac{\partial}{\partial x_1}(\alpha M e - M\xi)\end{aligned} \tag{10}$$

$$(m\rho - \rho_f^2)(\frac{\partial v_1^{s(2)}}{\partial t} + \omega_2 v_1^{s(2)}) = 2m\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f\frac{\eta}{\kappa}(v_1^f + \omega_2\int_{-\infty}^{t} v_1^f d\tau) \tag{11}$$

$$(m\rho - \rho_f^2)(\frac{\partial v_1^{s(3)}}{\partial t} + \omega_3 v_1^{s(3)}) = 2m\frac{\partial}{\partial x_3}(\mu e_{13}) \tag{12}$$

$$\frac{\partial e_{11}}{\partial t} = \frac{\partial v_1}{\partial x_1} - \omega_1 e_{11} \tag{13}$$

$$\frac{\partial e_{12}^{(1)}}{\partial t} = \frac{1}{2}\frac{\partial v_2^s}{\partial x_1} - \omega_1 e_{12}^{(1)} \tag{14}$$

$$\frac{\partial e_{12}^{(2)}}{\partial t} = \frac{1}{2}\frac{\partial v_1^s}{\partial x_2} - \omega_2 e_{12}^{(2)} \tag{15}$$

$$\frac{\partial \xi^{(1)}}{\partial t} = -\frac{\partial v_1^f}{\partial x_1} - \omega_1 \xi^{(1)} \tag{16}$$

$$\frac{\partial \xi^{(2)}}{\partial t} = -\frac{\partial v_2^f}{\partial x_2} - \omega_2 \xi^{(2)} \tag{17}$$

$$\frac{\partial \xi^{(3)}}{\partial t} = -\frac{\partial v_3^f}{\partial x_3} - \omega_3 \xi^{(3)} \tag{18}$$

Similar equations can be obtained in the same way for other components v_2^s, v_3^s, v_1^f, v_2^f, v_3^f, and e_{13}, e_{23}. Within PML region, equations (3) and (4) are split into 18 equations for 3D. Equation (5) is split into 9 equations and equation (6) is split into 3 equations. So the total number of the equations is 30 for the PML region, compared to 13 for the regular interior region. Thus, the memory requirement within the PML region is about two and half times that required by a regular medium for three-dimensional problems. This extra memory requirement in PML region is offset by the effectiveness of PML in absorbing the outgoing waves.

2.3. Finite difference implementation

The governing equations for the PML absorbing boundary condition, such as equations (10)-(18), are first-order partial differential equations for particle velocity and strain. They can be solved with different numerical methods. For the evaluation of seismic and acoustic responses, the finite-difference method is widely used because of its flexibility and simplicity. Here, we use the explicit second-order finite-difference method with a staggered grid in both spatial and temporal domains.

To implement a 3-D finite-difference method to the PML equations, the material parameters and unknown field components are discretized on a regular 3-D grid at the intervals $\triangle x_1$, $\triangle x_2$ and $\triangle x_3$. The time domain is also discretized with time step $\triangle t$. For the FD implementation of Biot's equations on a staggered grid in Fig. 1, the velocity field components are located at the cell's face centers, while material parameters and normal strains are located at the center of the cell and shear strains are located at the six edge centers. The strain field is computed at $n\triangle t$ and velocity field is computed at $(n+\frac{1}{2})\triangle t$. This staggered grid is similar to that for elastic waves in a solid.[4,22]

With this discretization, a leap-frog time-stepping system can be obtained. In order to simplify the layout of the formulas, the governing equations with PML boundary condition can be generalized as the first-order differential equation. For examples, (11), (13) and (16) can be rewritten as

$$\frac{\partial v_1^{s(2)}}{\partial t} + c_0 v_1^{s(2)} = c_1 \int_{-\infty}^{t} v_1^f d\tau + c_2[2m\frac{\partial}{\partial x_2}(\mu e_{12}) + \rho_f \frac{\eta}{\kappa}(v_1^f)] \tag{19}$$

$$\frac{\partial e_{11}}{\partial t} + \omega_1 e_{11} = \frac{\partial v_2^s}{\partial x_1} \tag{20}$$

$$\frac{\partial \xi^1}{\partial t} + \omega_1 \xi^1 = -\frac{\partial {v_1}^f}{\partial x_1} \tag{21}$$

where c_0, c_1 and c_2 space-dependent coefficients. Then the time-stepping equations can be written as,

$$v_1^{s(2)}[j_1, j_2, j_3, (n+\frac{1}{2})] = f_1 v_1^{s(2)}[j_1, j_2, j_3, (n-\frac{1}{2})] + f_2 R_1 \tag{22}$$

$$e_{11}[j_1, j_2, j_3, (n+1)] = g_1 e_{11}[j_1, j_2, j_3, n] + g_2 R_2 \tag{23}$$

$$\xi^1[j_1, j_2, j_3, (n+1)] = g_1 \xi^1[j_1, j_2, j_3, n] - g_2 R_3 \tag{24}$$

$$f_1 = \frac{c_0/2 - 1/\triangle t}{1/\triangle t + c_0/2}$$

$$f_2 = \frac{1}{1/\triangle t + c_0/2}$$
$$g_1 = \frac{\omega_1/2 - 1/\triangle t}{1/\triangle t + \omega_1/2}$$
$$g_2 = \frac{1}{1/\triangle t + \omega_1/2}$$

where R_1, R_2 and R_3 are right-hand sides of equations (19),(20) and (21) respectively. It should be noted that the material parameters in the above equations must be properly averaged in order to arrive at a higher accuracy.[4] In order to save computer storage, the computational domain is divided into a PML region and an interior region. The absorption of outgoing waves is achieved by the PML region, which consists of several cells of PML materials with a quadratically tapered ω_j profile to increase the attenuation toward the outer boundary. In this paper, ω_j of the PML region is chosen as

$$\omega_j(j) = \frac{(M - 1/2 - j)^2}{(M - 1/2)^2}\omega_{j,max} \tag{25}$$

where $\omega_{j,max}$ is the value at the center of the cell at outermost boundary. At the outer boundary, the velocity components are forced to be zero. For convenience, $\omega_{j,max}$ can be expressed in term of dominant frequency and normalized coefficient. Then equation (25) becomes

$$\omega_j = \begin{cases} 2\pi a_0 f_0 (l_{x_j}/L_{PML})^2, & \text{inside PML,} \\ 0, & \text{outside PML} \end{cases} \tag{26}$$

where f_0 is the dominant frequency of the source, and L_{PML} is the thickness of the PML region, and l_{x_j} is the distance from the interface between the interior region and PML region.

3. NUMERICAL RESULTS

In the following numerical simulations, the source energy is partitioned linearly between the solid and the fluid phases with factors

$$W_f = \phi$$
$$W_s = (1 - \phi)$$
$$W_r = \phi|W_f - W_s|$$

where W_f is the weighting factor for the fluid motion, W_s is for the solid motion and W_r is for the relative motion between solid frame and pore fluid.

3.1. *A. Validation of Numerical Results*

An analytical solution for the particle velocity field in a homogeneous, fluid-saturated poroelastic medium subject to a point source in 3-D space or a line source in 2-D space can be obtained in a closed form via potential functions.[9] In a poroelastic medium with an ideal non-viscous fluid, a purely dilatational source will only excite P waves given by scalar potentials,

$$\psi_s(r,t) = \frac{\alpha s(t - \frac{r}{V_f}) + \beta s(t - \frac{r}{V_s})}{4\pi r}$$
$$\psi_f(r,t) = \frac{\alpha A_f s(t - \frac{r}{V_f}) + \beta A_s s(t - \frac{r}{V_s})}{4\pi r}$$

where r is the distance and $s(t)$ is the source time function, A_f and A_s represent the ratios between the solid and fluid motion for the fast P-wave and the slow P-wave, and the coefficients α and β are determined by the regularity conditions. V_f and V_s are the velocities of the fast P-wave and the slow P-wave, respectively. The displacement field in the solid and fluid is then written respectively as $\mathbf{u} = \nabla\psi_s$ and $\mathbf{U} = \nabla\psi_f$.

The validation of the numerical method can be done by comparing the numerical results with the above analytical solution. A homogeneous model of sand saturated with water is considered. A P-wave point source of the first derivative Gaussian time function with $f_0 = 40$Hz is located at (0,0,0). Then the solution at (30 m, 30 m, 5 m) is calculated. The numerical result and analytical solution in Fig. 2 have an excellent agreement.

3.2. *B. Comparison Between Elastic and Poroelastic Models*

In the conventional elastic model, a single-phase medium is considered. There is no energy dissipation in such an elastic medium. In poroelastic model, however, a multi-phase medium (usually solid frame and pore fluid) is considered. The in-phase motion between solid and fluid leads to the regular P wave and S wave, and the out-phase motion between solid and fluid leads to the slow P wave. Because the fluid is viscous, there is energy dissipation between the solid frame and the pore fluid. In Biot's theory, the ratio of viscosity and permeability ($b = \eta/\kappa$) determines this dissipation, as shown in equations (3) and (4). The difference between an elastic model and a poroelastic model is investigated here by studying the effect of b on AVO response in three different ground media: dry sand, fully water saturated sand and partly water saturated sand (70% air and 30% water). For a half-space problem, since the source and receivers are usually on the free surface and the surface wave is the dominant signal at low frequencies, only surface wave is investigated here. In this case, we model the source as a vertical dipole right on the ground surface and displaced horizontally away from the buried object, a scenario similar to the source in.[1]

Fig. 3, Fig. 4 and Fig. 5 show the effects of b on the surface wave AVO for dry sand, fully water saturated sand and partly water saturated sand models, respectively. For the dry sand, the difference between elastic and poroelastic models is obvious, but b has little effect on AVO. For the fully water saturated sand model, b has a significant effect on AVO. When b is very large, the AVO response is very close to that of elastic model. For partly water saturated sand model, the difference between elastic and poroelastic model is obvious and b also has a significant effect on AVO.

3.3. *C. Acoustic Land Mine Detection Systems*

This method is used to model a plastic antipersonnel mine that is difficult to detect with the more conventional electromagnetic induction sensors (*i.e.*, metal detectors). The geometry of the model of a buried mine is shown in Fig. 6. The source can be either on the surface or in the air. There are three arrays of receivers on the surface to pick up the signal of vertical velocity.

The parameters for the plastic mine are chosen as $\rho = 1200$ kg/m^3, S-wave velocity $v_{sh} = 1100$ m/s and P-wave velocity $v_p = 2700$ m/s. The material in the ground may be considered as a two-phase composite material consisting of granular solid and pore fluid. The nature of this composite varies with environment, geographic location, and with depth below the surface of ground. In this model, the soil is chosen as dry sand and the porosity is 0.35. The fast P wave velocity is 250 m/s and S wave velocity is 87 m/s.

3.3.1. *(a). Source on the Ground Surface*

To model the land-mine detection system in,[1] a half space with a dipole source on the ground surface is simulated. The source time function is the Blackman-Harris window function with the center frequency 450 Hz. Fig. 7 shows the four snapshots of the total field on the free ground surface at 2.3 ms, 4.0 ms, 6.0 ms and 7.8 ms. The primary wave on these snapshots are surface waves. Compared to the surface waves, the P wave is very weak. The obvious surface wave reflections by the mine are shown on the snapshots at 6.0 ms and 7.8 ms. The influence of the buried mine on the right-propagating surface wave can be seen clearly on the snapshots. There is a decrease on amplitude when the surface wave passes over the buried mine. Fig. 8(a) shows the waveforms of receiver array 1. The reflected surface waves are presented on the first 5 traces. Fig. 8(b) and 8(c) show the waveforms of receiver array 2 and 3 respectively. The decrease of the amplitude can be easily seen on the time traces of receiver 6 to 11.

3.3.2. *(b). Source in the Air*

Another land-mine detection prototype uses airborne acoustic waves to couple into the ground surface to detect minelike targets.[2] In our simulation, the model consists of two parts: air (pure fluid) and soil (poroelastic medium) with a plastic land mine. A monopole source is in the air 2 cm above the ground surface. The locations of the receivers are the same as those in the previous model. Fig. 9 shows the four snapshots of vertical velocity field on the ground surface at 2.3 ms, 4.0 ms, 6.0 ms and 7.8 ms. We can see that if a minelike target is present below the surface, the ground vibrational velocity will show distinct changes due to reflection and scattering of surface waves. Figs. 10(a), 10(b) and 10(c) show time traces of the scattered field at receiver arrays 1, 2 and 3 respectively. The signature of the buried land mine is clearly present in these figures. It is interesting to note that both direct arrival and the surface waves are scattered by the object into surface waves.

The above results are shown in time domain. However, the prototype in[2] actually displays the acoustic energy in the frequency domain. This display can also be easily done by our time-domain solution. Displayed in Fig. 11 is

the energy distribution of the scattered field for a monopole source directly above the mine, showing the presence of the buried object.

4. CONCLUSION

A particle velocity-strain, finite-difference method combined with the perfectly matched layer (PML) has been developed for the simulation of acoustic waves propagating in 3-D poroelastic media. The results show the scheme is stable even if the ratio of the fast P wave velocity to shear wave velocity is greater than 2.

The difference between elastic model and poroelastic model and moisture effect are investigated by studying surface wave amplitude variation with offset (AVO) in three different ground media: dry sand, fully water saturated sand and partly water saturated sand. The difference between these models is significant. For air saturated sand, the attenuation factor has little influence on AVO. For fully or partly water saturated sand, the attenuation factor greatly affects the AVO response.

This numerical method has been used to investigate the interaction of acoustic waves with a buried mine-like object. Two land-mine detection prototypes are simulated: a half space model with a dipole source on the free ground surface and a half space model with a monopole source in the air. The results show that the wave responses are substantially changed by the presence of the mine-like object, and surface waves play as a primary mechanism of detection at the low frequencies.

REFERENCES

1. W. R. Scott, Jr., C. T. Schröder and J. S. Martin, "An hybrid acousto/electromagnetic technique for locating land mines," in *Proc. International Geoscience and Remote Sensing Symposium,* Seattle, WA., pp. 216-18, 1998.
2. J. M. Sabatier and Ning Xiang, "Systematic investigation on acoustic-to-seismic responses of landmines buried in soil," *J. Acoust. Soc. Am.*, vol. 107, pp. 2896, 1999.
3. J. Virieux, "SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method," *Geophysics*, vol. 49, pp. 1933–1957, 1986.
4. Q. H. Liu, F. Daube, C. Randall, E. Schoen, H. Liu, and P. Lee, "A 3D finite difference simulation of sonic logging," *J. Acoust. Soc. Am.*, vol. 100, pp. 72–79, 1996.
5. Q. H. Liu, and C. Chang, "Compressional head waves in attenuative formations: forward modeling and inversion," *Geophysics*, vol.61, pp. 1908–1920, 1996.
6. M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. 1.Low-frequency range," *J. Acoust. Soc. Am.*, vol. 28, pp. 168–178, 1956.
7. M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. 1.High-frequency range," *J. Acoust. Soc. Am.*, vol. 28, pp. 179–191, 1956.
8. M. A. Biot, "Mechanics deformation and acoustic propagation in porous media," *J. Appl. Phys.*, vol. 33, pp. 1482–1498, 1962.
9. N. Dai, A. Vafidis and E. R. Kanasewich, "Wave propagation in heterogeneous, porous media: A velocity-stress, finite-difference method," *Geophysics*, vol. 60, pp. 327–340, 1995.
10. C. Cerjan, D. Kosloff, R. Kosloff, and M. Reshef, "A non-reflecting boundary condition for discrete acoustic and elastic wave equations," *Geophysics*, vol. 50, pp. 705–708, 1985.
11. R. Clayton and B. Engquist, "Absorbing boundary conditions for acoustic and elastic wave equations," *Bull. Seism. Soc. Am.*, vol. 67, pp. 1529-1540, 1977.
12. K.D. Mahrer, "An empirical study of instability and improvement of absorbing boundary conditions for elastic wave equation," *Geophysics*, vol. 51, pp. 1499–1501, 1986.
13. J. P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, vol. 114, pp. 185–200, 1994.
14. W. C. Chew and Q.H. Liu, "Perfectly matched layers for elastodynamics: A new absorbing boundary condition," *J. Comp. Acoust.* , vol. 4, pp. 72–79, 1996.
15. F. D. Hastings, J. B. Schneider, and S. L. Broschat, "Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation," *J. Acoust. Soc. Am.*, vol. 100, pp. 3061–3069, 1996.
16. Q. H. Liu and J. Tao, "The perfectly matched layer for acoustic waves in absorptive media," *J. Acoust. Soc. Am.*, vol. 102, pp. 2072–2082, 1997.

17. Q. H. Liu, "Perfectly matched layers for elastic waves in cylindrical and spherical coordinates," *J. Acoust. Soc. Am.*, vol. 105, pp. 2075–2084, 1999.
18. Y. Q. Zeng, J. He and Q. H. Liu, "The application of the perfectly matched layer in numerical modeling of wave propagation in poroelastic media," Research report, Duke University (1999).
19. R. Burridge and J. B. Keller, "Poroelasticity equations derived from microstructure," *J. Acoust. Soc. Am.*, vol. 70, pp. 1140–1146, 1981.
20. J. Plona, "Observation of the second bulk compressional wave in a porous medium at ultrasonic frequencies," *App. Phys. Lett.* vol. 36, pp. 259–261, 1980.
21. G. Kneib and C. Kerner, "Accurate and efficient seismic modeling in random media," *Geophysics*, vol. 58, pp. 576–588, 1993.
22. Q. H. Liu, "Some current trends in numerical methods for transient acoustic and elastic waves in multidimensional inhomogeneous media," *Current Topics in Acoustical Research*, vol. 2, pp. 31–42, 1998.

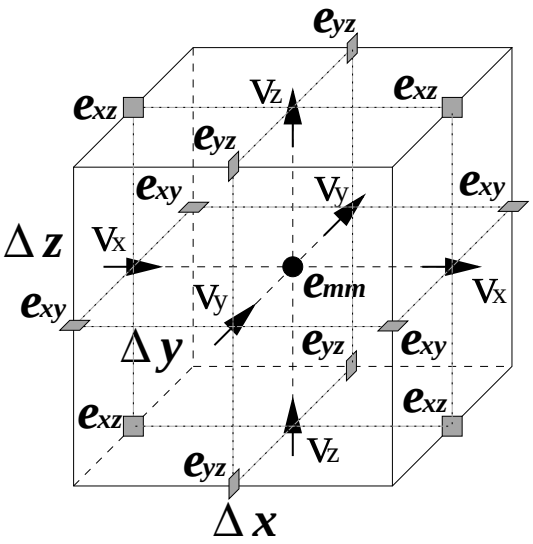


Figure 1. The relative locations of field components in a unit cell of staggered grid.

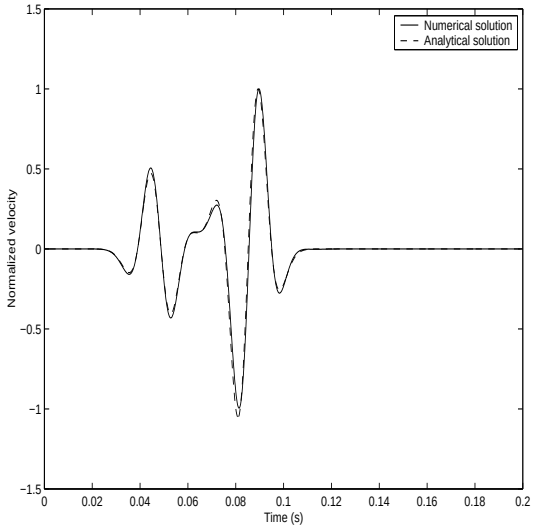


Figure 2. Comparison of analytical and numerical solutions for the vertical velocity component in a homogeneous solid.

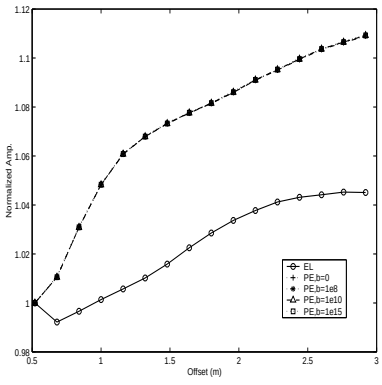


Figure 3. Comparison between elastic and poroelastic model for dry sand.

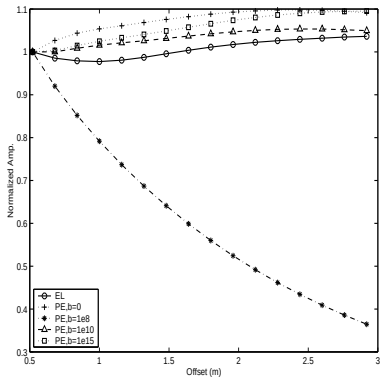


Figure 4. Comparison between elastic and poroelastic model for fully water saturated sand.

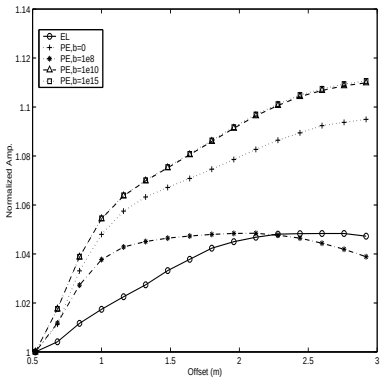


Figure 5. Comparison between elastic and poroelastic model for partly water saturated sand.

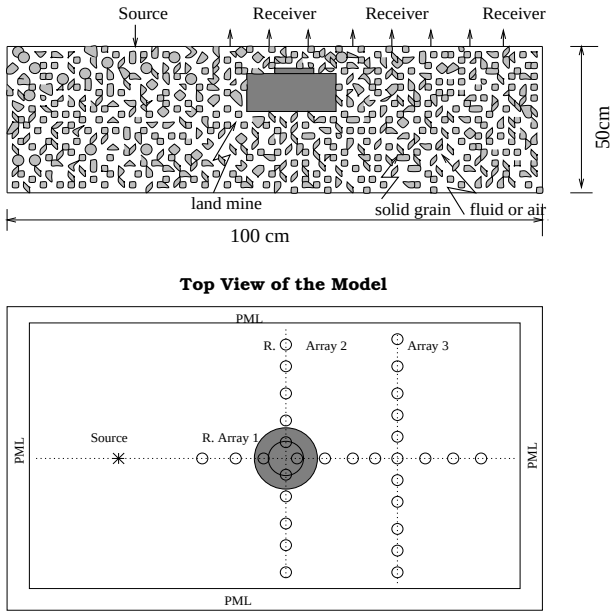


Figure 6. Geometry of a buried land mine model.

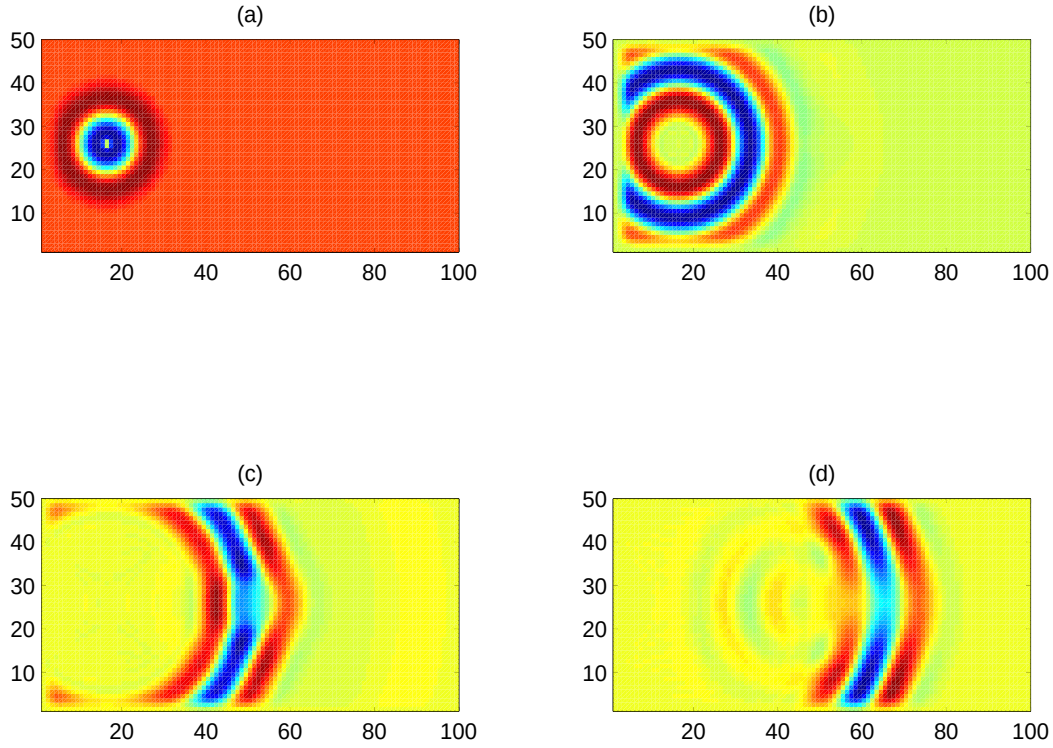


Figure 7. Snapshots at (a) 2.3 ms (b) 4.0 ms (c) 6.0 ms and (d) 7.8 ms. A dipole source is on the free surface.

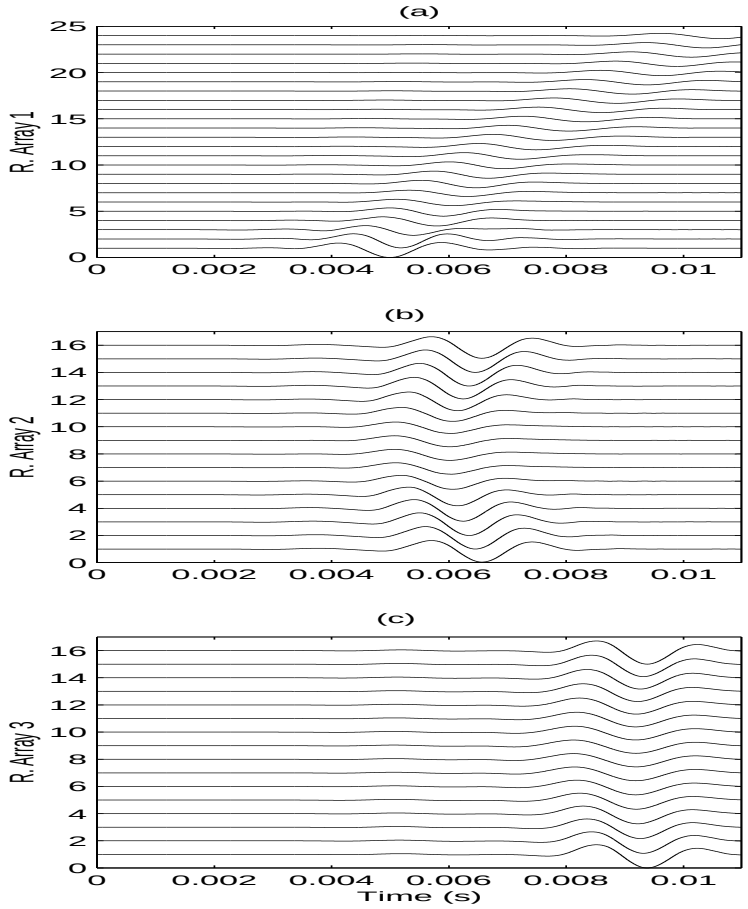


Figure 8. Waveforms of vertical velocity of solid frame at receiver (a) array 1 (b) array 2 (c) array 3; A dipole source is on the free surface.

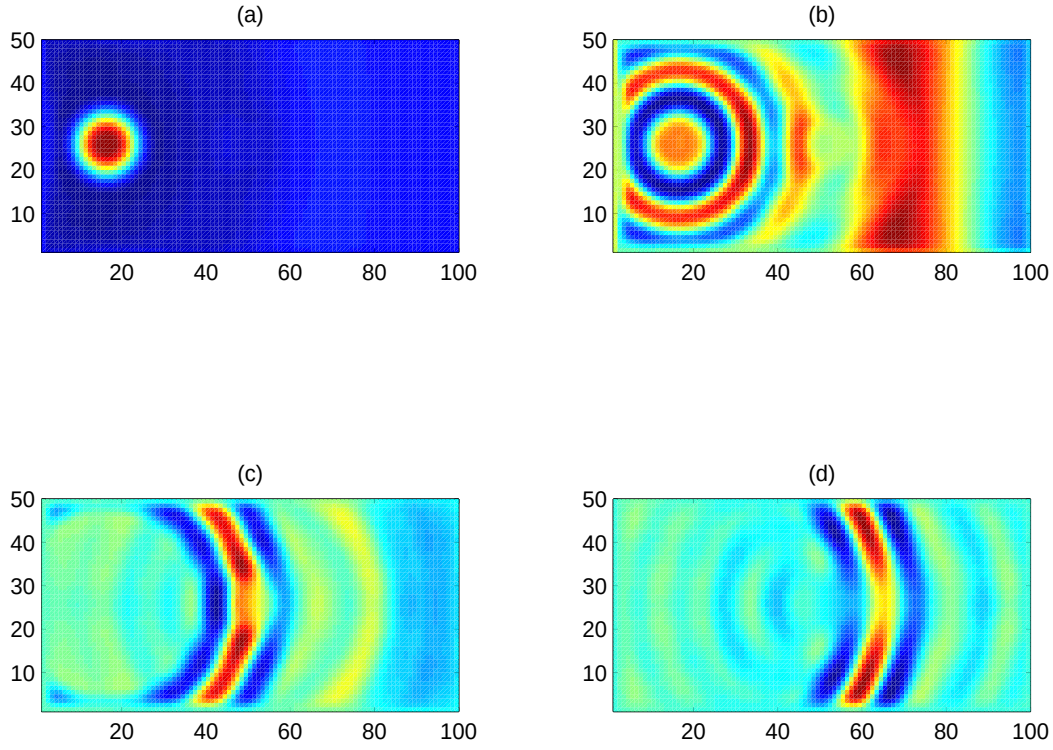


Figure 9. Snapshots at (a) 2.3 ms (b) 4.0 ms (c) 6.0 ms and (d) 7.8 ms. A monopole source is in the air.

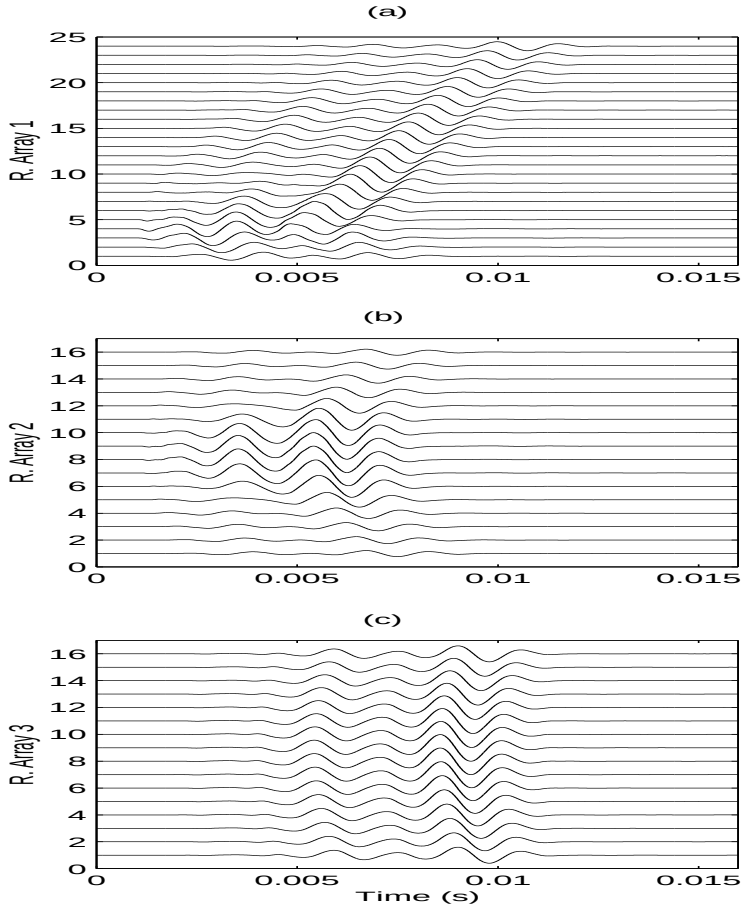


Figure 10. Waveforms of scattered field of solid frame at receiver (a) array 1 (b) array 2 (c) array 3; A monopole source is in the air.

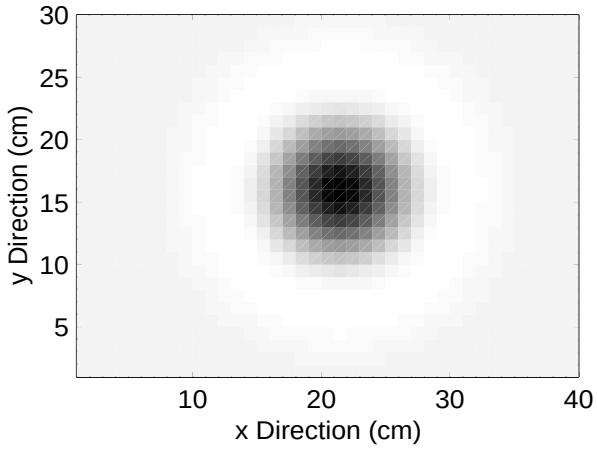


Figure 11. Energy distribution of the scattered field when the source is in the air above the mine.

Environmental and Target Influences on Microwave Radiometers for Landmine Detection

Joel T. Johnson, Baran U. Ungan, and David R. Wiggins

The Ohio State University ElectroScience Laboratory

1320 Kinnear Rd., Columbus, OH 43212-1191

johnson@ee.eng.ohio-state.edu

ABSTRACT

A simple layered medium model for microwave thermal emission from a buried object shows that multiple frequency emission measurements can potentially provide an effective means for target detection. Object detection is obtained from a search for oscillatory features in multiple frequency brightness temperatures, which occur due to interference effects between the surface and buried object interfaces. Previous studies have considered simple homogeneous temperature and water content models of the soil medium, and show that oscillatory features versus frequency are not obtained in the absence of a target even with medium temperature or soil moisture variations. However, the more realistic case of non-constant temperature and water content versus depth was not considered in previous studies; these effects can potentially modify interference phenomena. In addition, subsurface objects have typically been modeled as layers whose horizontal dimensions are infinite; models including the effects of finite target size are thus of interest.

In this paper, multiple frequency emission results from a layered medium with non-constant temperature and water content versus depth are examined in the target and no target cases. Thermal and water content profiles are obtained from models for water and thermal transport, so that reasonable variations with depth are included. Results are shown which indicate that the basic target detection methodology is not strongly influenced by these factors at microwave frequencies. Numerical models for computing emission from a finite size buried object are also presented, and results compared with those from the layered medium theories to quantify the influence of target size relative to the observed region on emission results. Again, results show that the basic detection methodology can remain valid even with finite target sizes.

Keywords: Thermal emission, microwave radiometry, landmine detection

1. INTRODUCTION

Microwave radiometers have been shown to be effective sensors for monitoring soil moisture and other geophysical data.[1,2] Models for geophysical medium brightness temperatures often consider horizontally stratified geometries, because exact evaluation of observed brightness temperatures is possible for this case in terms of a sum involving layered medium reflection quantities.[3] For cases involving media whose temperature is constant, the formulation reduces to "Kirchhoff's law" which relates emissivity to one minus the reflection coefficient of the layered medium. Previous studies of microwave radiometry for soil moisture remote sensing have considered both temperature and soil moisture variations with depth, and have shown that these environmental effects can have important influences on soil moisture retrievals.[4]

Recent studies have begun to consider the use of microwave radiometers for detecting shallow, sub-surface objects such as anti-personnel landmines.[5,6] Modeling studies[6] based on Kirchhoff's law and a three layer, horizontally stratified geometry (i.e. objects are infinite layers) show that significant brightness temperature contrasts can be obtained in the presence of a sub-surface object if sufficient dielectric contrast is available and if soil medium attenuation is not excessive. Previous results further demonstrated that use of multi-frequency brightness measurements could potentially provide detections even with low contrast or high attenuation due to the oscillatory behavior versus frequency observed in the presence of an object. Estimated environmental effects such as local surface temperature or soil moisture variations would not produce oscillatory frequency behavior (except in unusual circumstances[2]) so that detections would still be possible even in the presence of environmental "clutter".

However, previous buried object detection models have neglected variations in medium temperature (due to the use of Kirchhoff's law) and soil moisture (which was assumed constant) with depth. Since these factors can significantly impact soil moisture remote sensing, their effect on buried object detection requires consideration as well. In this paper, a multi-layer horizontally stratified emission model is coupled with a numerical solution of the heat equation and a model for water transport in the presence of a sub-surface object to estimate the importance of temperature and soil moisture variations. Results show that these factors can impact overall brightness temperatures, but that the concept of using oscillatory features in frequency swept data as an indicator of a sub-surface object remains valid.

Effects of finite target size were also neglected in previous work since sub-surface objects were modeled as horizontally infinite layers. Numerically calculated emission results from a finite-size buried object are thus also considered in this paper. In the case of a three dimensional finite size subsurface target, calculation of thermal emission from the object requires a numerical solution of the electromagnetic boundary value problem. If a constant temperature medium is assumed, brightness temperatures can be obtained from Kirchhoff's law by determining the total amount of power scattered into free space by the subsurface object when illuminated by a field incident from free space. The soil is modeled as a homogeneous half-space medium characterized by constant electromagnetic permittivity and physical temperature. Thermal emission from the target is numerically calculated using an iterative method of moments (including half-space Sommerfeld Green's functions) accelerated with the discrete dipole approximation.[7–9] Brightness temperatures and their variations with frequency are presented for differing target sizes and soil properties, and amplitudes of brightness temperature oscillations in frequency are shown to vary primarily according to the fraction of the observing antenna pattern occupied by the subsurface object. Finite size target solutions are also compared with results from the horizontally infinite layer model to determine the parameter space under which the layered model is applicable.

2. EFFECTS OF MOISTURE AND THERMAL PROFILES: LAYERED MODEL

Figure 1 plots the geometry considered for estimation of moisture and thermal profile effects on layered medium emission, and illustrates a 6.8 cm thick subsurface object (modeled as a horizontally infinite layer of constant relative permittivity $3 + i0.08$) located 4.25 cm below an air-soil boundary. Both permittivity $\epsilon(z)$ and temperature $T(z)$ are not constant in the soil background medium, and the temperature also varies in the object layer. Specification of the medium begins with a moisture transport model[10,11] for evaluating soil volumetric moisture contents in the presence or absence of a sub-surface object of dimensions $25.2 \times 25.2 \times 6.8$ cm. Water transport is computed on a grid $73.5 \times 73.5 \times 59.5$ cm in the presence of a periodic "rain" forcing function. Figure 2, plot (a) illustrates the resulting moisture profiles versus depth with (taken from directly underneath the object) and without (taken from one of the faces of the computational domain) the object; note the increased and decreased water contents above and below the object respectively. Moisture contents versus depth then directly determine the permittivity profile $\epsilon(z)$ through an empirical soil permittivity model[1] evaluated for sand with bulk density 1.12 g/cm^3. Obtained permittivity values ranged from $5.44 + i0.46$ to $12.74 + i2.15$; note the significant contrast with the subsurface object, although high attenuation is also obtained for this case.

Moisture contents are also used to determine the thermal conductivity and diffusivity parameters of the soil medium versus depth.[12] A one dimensional solution of the heat equation with or without the object layer[13] is then used to determine the temperature profile versus depth. Parameters for solar forcing, wind convection, and other effects are chosen according to standard models.[13] The heat equation is solved on a 100 layer grid up to depth 50 cm, beyond which the temperature is assumed constant. Figure 2, plot (b) illustrates typical temperature profiles obtained with and without the subsurface object at local time 15 : 42. In these simulations the subsurface object is modeled thermally as a near-insulator, allowing the relatively large temperature difference through the object layer to be maintained.

Finally, given the $\epsilon(z)$ and $T(z)$ profiles versus depth, brightness temperatures are computed from a multi-layer fluctuation-dissipation model.[3] This model includes all potential coherent emission effects, and again is evaluated on the 100 layer grid up to depth 50 cm. Results to be illustrated consider nadir observation, so that polarization effects can be neglected. Results are compared with computations which neglect either temperature variations (by assuming the entire medium is at the surface temperature) or permittivity variations (by assuming the soil medium is all of uniform permittivity) to clarify the influence of these factors.

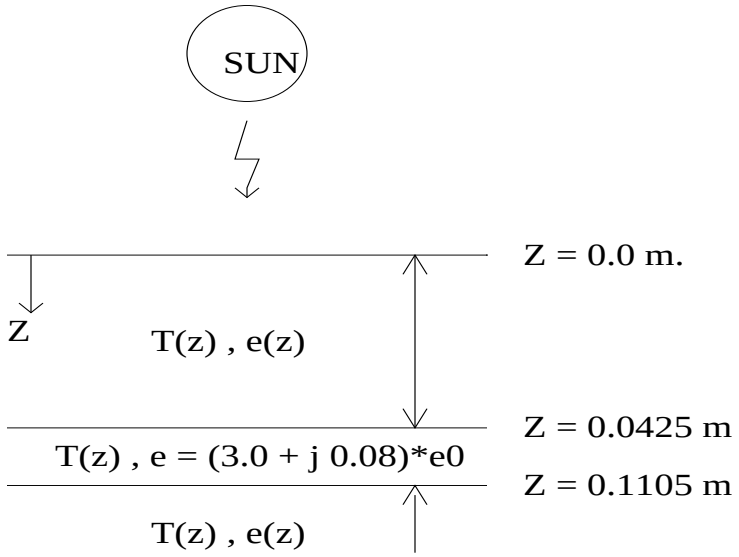


Figure 1. The stratified medium model

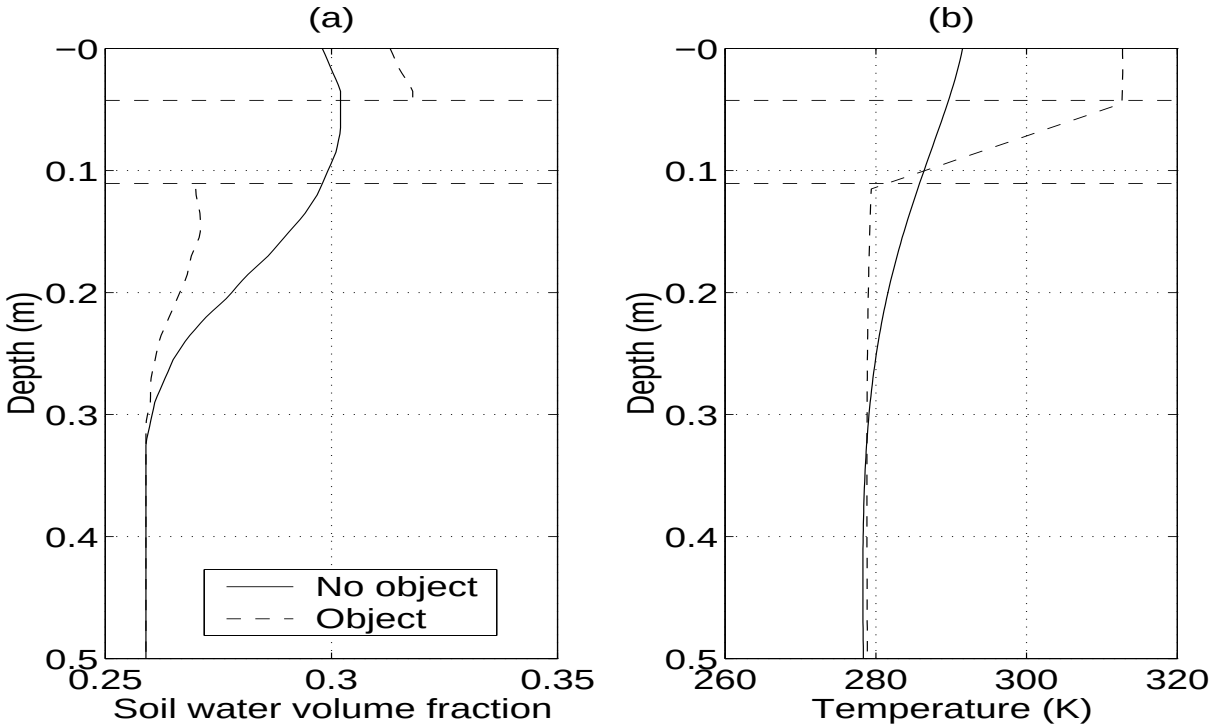


Figure 2. Moisture (a) and temperature (b) profiles with and without object

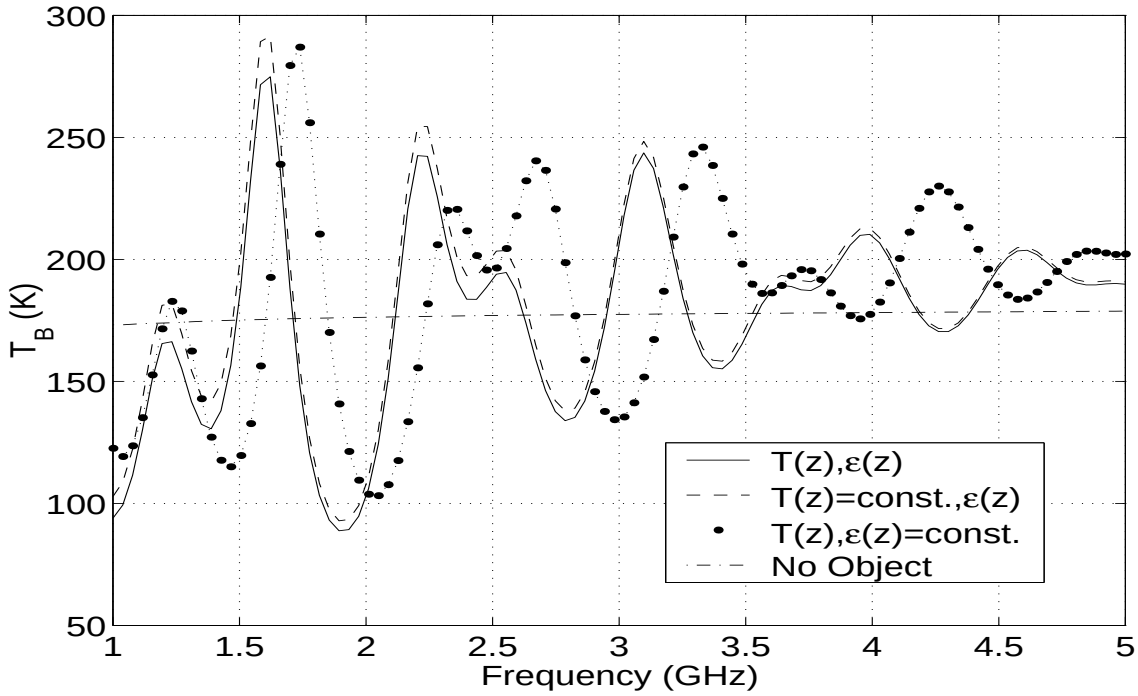


Figure 3. Brightness temperatures (T_B) vs. frequency for the high moisture case

Figure 3 compares brightness temperatures from the complete model with the object versus those assuming $T(z) = T_s = 313\ K$ or $\epsilon(z) = \epsilon$ (computed for 27% soil moisture). Results are plotted as a function of radiometer frequency from 1 to 5 GHz, and the oscillatory pattern versus frequency obtained in the presence of a subsurface object is observed. The complete model with no object is also included, and shows no strong oscillations versus frequency. For this relatively high attenuation case, the influence of temperature variations is observed to be small, since greater depths at which temperature changes are larger are not observed by the radiometer. Permittivity variations are found to have a larger effect, but the basic concept of searching for oscillatory features for object detection remains valid.

To illustrate an additional case with lower attenuation, the moisture profiles of Figure 2 are divided by 4 to obtain an average moisture content of 6.7%, and the thermal and emission models re-computed. Results are illustrated in Figure 4, and show smaller oscillations due to the reduced dielectric contrast and a greater influence of the temperature profile due to the reduced attenuation. Permittivity profile variations are less significant than in Figure 3 since the level of moisture and permittivity variations in the medium have been reduced. Overall, the basic features with and without the subsurface object again remain similar.

The results of this study confirm that temperature and moisture profiles can influence brightness temperatures, but that these factors are unlikely to produce the oscillatory features caused by a subsurface object.

3. FINITE SIZE TARGET EFFECTS

To study the influence of finite size target effects, a constant temperature medium is assumed so that brightness temperatures can be calculated using Kirchhoff's law:

$$T_B = T_s\,(1 - R) \tag{1}$$

where T_s is the physical temperature of the medium (taken as 290 K in the results of this section) and R is the total reflectivity of the medium. The total reflectivity R is equal to the total power scattered into the free space region when the subsurface target is illuminated by a plane wave field incident from the free space region, and can therefore be determined by solving the corresponding scattering problem. For a target modeled as an infinite layer, fields in the free space region consist only of the incident and reflected plane waves, with R determined by the amplitude of the reflected plane wave. However, in the presence of a finite size target, total fields in the free space region include

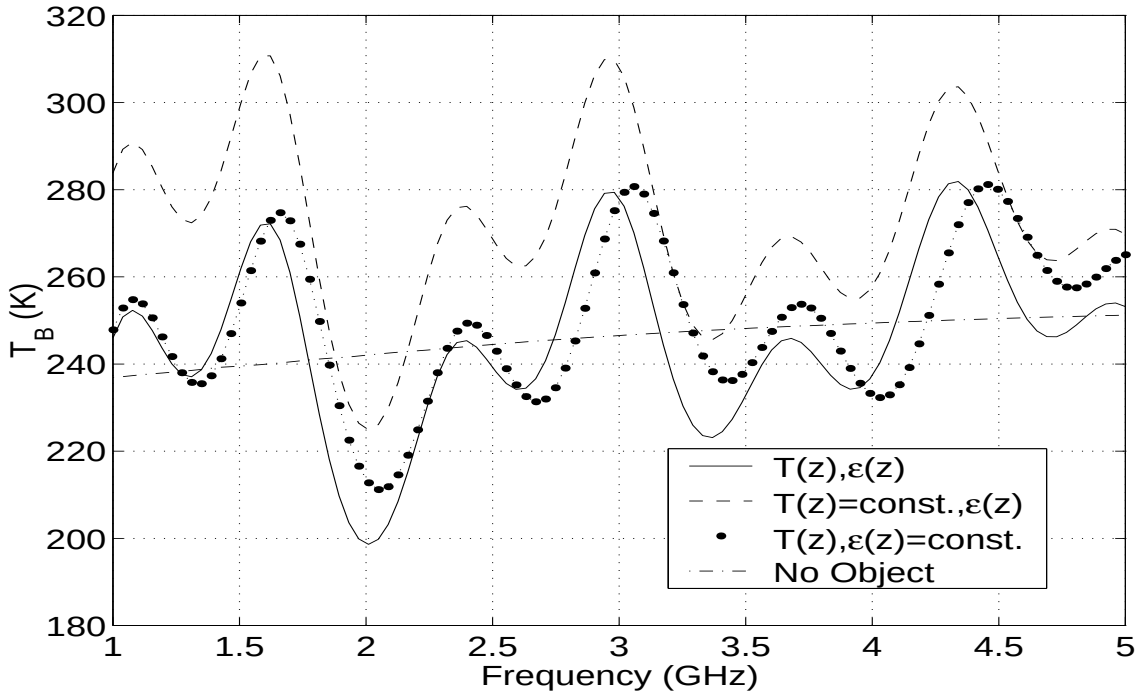


Figure 4. Brightness temperatures (T_B) vs. frequency for the reduced moisture case

both incident and reflected plane waves and spherical scattered wave components. The brightness temperature is then expressed as

$$T_B = T_s \left(1 - |\Gamma|^2 - \frac{P_{c0} + P_{s0}}{P_i}\right) \tag{2}$$

where Γ is the plane wave reflection coefficient of the half-space boundary, P_{s0} is the total scattered spherical wave power due to the sub-surface target and P_{c0} is a cross-power term due to the interaction of the reflected plane wave and scattered spherical waves. Both these terms are normalized by the total power incident on the half-space medium P_i. If the half-space medium and targets are assumed to be lossless, the same brightness temperature can also be calculated using the power radiated into the half-space medium as

$$T_B = T_s \left(|T|^2 \frac{\eta_0}{\eta_1} + \frac{P_{c1} + P_{s1}}{P_i}\right) \tag{3}$$

where T represents the plane wave transmission coefficient at the boundary, η_0 and η_1 are the characteristic impedances of free space and the half-space medium, respectively, and P_{s1} and P_{c1} are the scattered power and the cross-power terms, respectively, in the half-space medium. Comparison of the brightness temperatures obtained from equations (2) and (3) provides information on the level of power conservation obtained in the numerical solution; results will be plotted for brightnesses obtained from both methods.

For a plane wave incident field, the cross-power terms can be shown to reduce to

$$P_{c0} = \frac{2\pi}{\omega\mu} Im\{\hat{e_r} \cdot \bar{F_s}\Gamma\} \tag{4}$$

and

$$P_{c1} = \frac{2\pi}{\omega\mu} Im\{\hat{e_t} \cdot \bar{F_s}T\} \tag{5}$$

where ω is the radian frequency of the radiometer, μ is the permeability of free space, $\hat{e_r}$ and $\hat{e_t}$ are unit vectors in the directions of the reflected and the transmitted plane wave electric fields, respectively, and $\bar{F_s}$ is the scattered spherical wave amplitude in the reflected (for P_{c0}) or transmitted directions (for P_{c1}). The above equations can be used to derive an "optical theorem"[3] for an object in a lossless half-space.

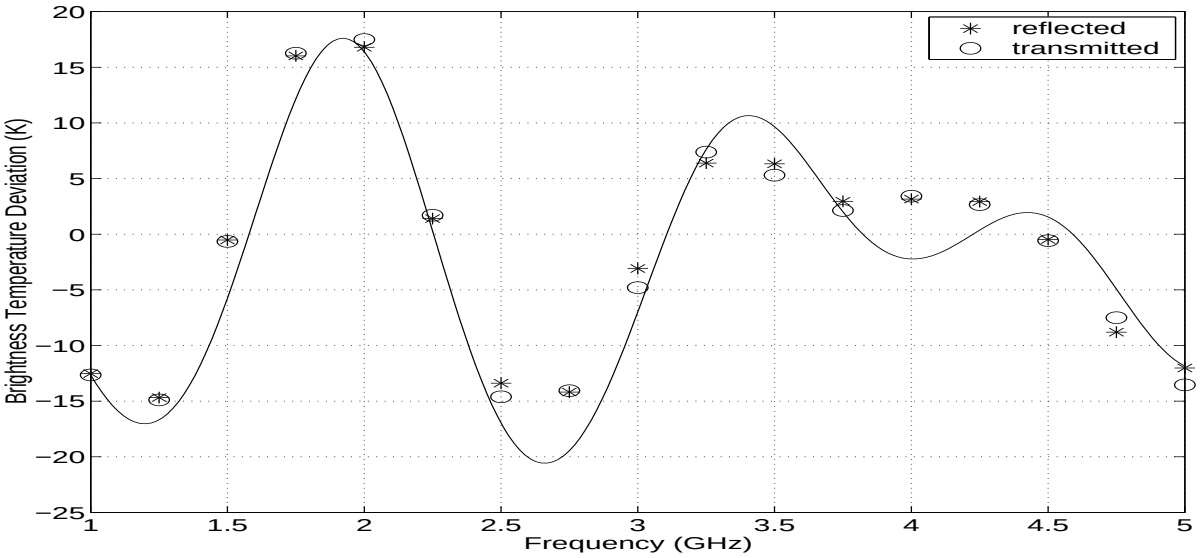


Figure 5. Brightness temperature deviation vs. frequency, numerical solution (discrete points) compared with layered model (smooth curve)

For the finite size target case, spherical wave scattered fields are numerically evaluated using an iterative method of moments (including half-space Sommerfeld Green's functions). Computational efficiency is improved through use of the discrete dipole approximation[7–9,14] so that electromagnetic coupling between all points on the object grid is computed in order $N \log N$, where N is the number of sampling points. In the brightness temperature calculations the cross-power term is computed using only specular scattered fields, whereas the scattered spherical wave power term requires a numerical integration of far-zone scattered powers over the upper or lower hemispheres. Tests varying the number of points in this integration were performed to insure accurate total scattered power computations.

To clarify the influence of a sub-surface target, results will be presented in terms of the brightness temperature change caused by the presence of the target. Since scattered power and the cross-power terms vanish when no target is present, the change in the brightness temperatures caused by the target is

$$\Delta T_B = -T_s \left(\frac{P_{c0} + P_{s0}}{P_i} \right) = T_s \left(\frac{P_{c1} + P_{s1}}{P_i} \right) \tag{6}$$

A final issue involves evaluation of the P_i term when plane wave observations are considered. Since a plane wave would illuminate an infinite portion of the half-space in the corresponding scattering problem, P_i approaches infinity and the effect of the target become negligible. In fact, a radiometer observing a sub-surface target would be sensitive to only a finite sized portion of the half-space, so that P_i remains finite and approximately equal to

$$P_i \approx \frac{1}{2\eta_0} A \tag{7}$$

for nadiral observations. Here A is the "spot-size" area of the observing antenna, which is assumed to have dimensions comparable to or larger than the electromagnetic wavelength so that the plane wave field model used in the numerical solution is valid. To simplify the results presented, brightness temperature deviations are re-written as

$$\Delta T_B = -T_s \left(\frac{P_{c0} + P_{s0}}{\frac{1}{2\eta_0} A_{obj}} \right) \left(\frac{A_{obj}}{A} \right) \tag{8}$$

$$= -T_s \left(\frac{P_{c0} + P_{s0}}{\frac{1}{2\eta_0} A_{obj}} \right) F \tag{9}$$

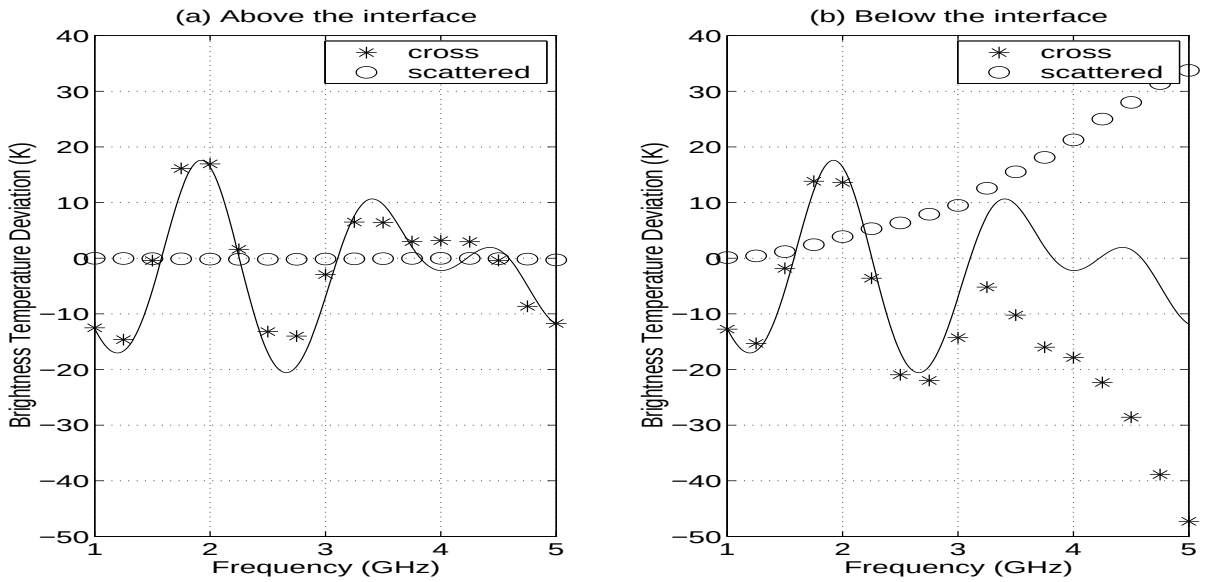


Figure 6. Contributions of scattered and cross- power terms to brightness temperature deviations using the fields (a) above and (b) below the ground. Numerical solution (discrete points) compared with total brightness deviation from layered model (smooth curve).

where A_{obj} is the cross-sectional area of the object and the term F in the final equation represents a "beam fill factor" which describes the fraction of the observing antenna pattern occupied by the sub-surface object. The following plots use $F = 1$ for simplicity.

Figure 5 illustrates ΔT_B as a function of frequency from the numerical solution (symbols) for a half-space relative permittivity of 4 and target relative permittivity of 3.15 (close to the values for dry soil and plastic, respectively.) The target is a cube with side length 2 cm centered 5 cm below the boundary, and was discretized into 16 by 16 by 16 points to insure accurate calculations over this range of frequencies. Numerical ΔT_B values are plotted from both the reflected and the transmitted powers, and are in good agreement, confirming power conservation in the numerical solution. Results from the layered medium model (i.e. with the target modeled as an infinite layer) are also included as the solid curve in Figure 5, and show good agreement with finite size target brightness deviations. Only slight effects due to finite target size are observed, although again the beam fill factor for a specific antenna would reduce finite target results proportionately. Overall, results show again that finite size target effects are not likely to modify the basic strategy of searching for oscillatory brightness features for detecting targets.

Figure 6 illustrates the individual contributions of the scattered spherical wave power and the cross-power terms to the ΔT_B results of Figure 5. The brightness temperature difference computed using the fields above the interface is presented in Figure 6 (a). For this case the scattered power term is negligible as compared to the cross-power term. Therefore ΔT_B can be obtained using only the cross-power term which depends on the field value at the specular reflection direction, hence increasing the computational efficiency. Unlike the case for the fields above the interface, the scattered power and the cross-power term both show important contributions with fields below the interface, as shown in Figure 6 (b). However the two power terms still add up to give an accurate ΔT_B. This increase in the scattered power below the ground is due to the fact that the object size is larger as compared to the wavelength in the medium.

Effects of varying the half-space relative permittivity are presented in Figure 7 for observations at frequency 2 GHz. The object size and permittivity are kept the same as in the previous cases, and the brightness temperature solution is computed using only the fields above the interface. Continuous curves in Figure 7 are results from the layered medium model with and without the target layer. Note that numerical results are closer to the no-target model for low permittivity values whereas they follow the three-layer model for higher permittivities. Again the

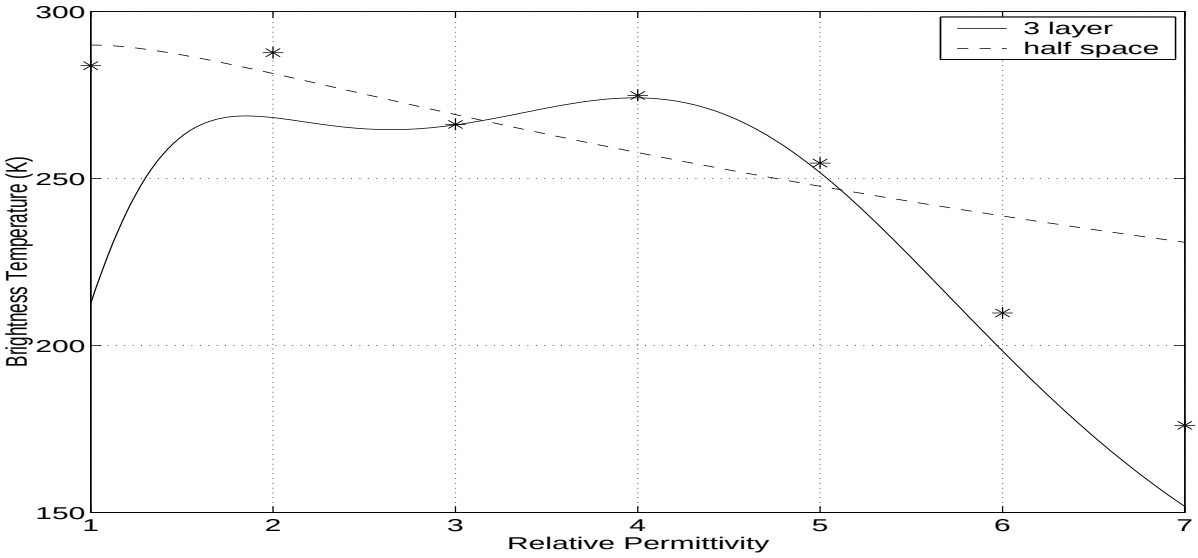


Figure 7. Brightness temperature vs. relative permittivity, numerical solution (discrete points) compared with analytical (smooth curves)

increasing size of the target relative to the half-space wavelength results in an increased target influence at higher permittivity values. Results were also calculated at a single frequency for varying target horizontal areas while keeping the thickness constant. The resulting brightness temperature deviations show only slight variations with target size so long as the beam fill factor is modeled as a constant.

4. CONCLUSIONS

Effects of soil moisture, medium temperature and target size on the detection of buried targets using microwave radiometry have been investigated. Thermal emission from buried objects has been shown to display an oscillatory behavior versus frequency which provide a means for detecting the target. Various moisture and temperature profiles and target sizes have been shown to affect the brightness temperature of the buried object, but the basic strategy of searching for oscillatory features versus frequency for target detection remains valid.

REFERENCES

1. F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, Artech House, Norwood, MA, 1986.
2. T. J. Schmugge, T. J. Jackson, P. E. O'Neill, and M. B. Parlange, "Observations of coherent emissions from soils," *Radio Sci.* **vol. 33**, pp. 267–272, 1998.
3. L. Tsang, J. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing*, John Wiley and Sons, New York, 1985.
4. Y.-A. Liou and A. W. England, "A land surface process/ radiobrightness model with coupled heat and moisture transport in soil," *IEEE Trans. Geosc. Rem. Sens.* **vol. 36**, pp. 273–286, 1998.
5. G. D. Amici, B. Hauss, and L. Yujiri, "Detection of landmines via a passive microwave radiometer," *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proceedings of SPIE* **vol. 3710**, pp. 716–724, 1999.
6. J. T. Johnson, "Theoretical study of microwave radiometry for buried object detection," *Detection and Remediation Technologies for Mines and Mine-Like Targets V, Proceedings of SPIE* , 2000.
7. J. T. Johnson and R. J. Burkholder, "Coupled canonical grid / discrete dipole approach for computing scattering from objects above or below a rough interface," *IEEE Trans. Geosc. and Rem. Sens.* **accepted**, 2001.

8. P. J. Flatau, "Improvements in the discrete-dipole approximation method of computing scattering and absorption," *Optics Letters* **vol. 22**, pp. 1205–1207, 1997.
9. B. T. Draine and P. J. Flatau, "Discrete-dipole approximation for scattering calculations," *J. Opt. Soc. Am. A* **vol. 11**, pp. 1491–1499, 1994.
10. J. Simunek, K. Huang, and M. van Genuchten, "The SWMS-3D code for simulating water flow and solute transport in three dimensional variably saturated media," *Research Report No. 139, U. S. Salinity Laboratory, USDA* , 1995.
11. B. A. Baertlein and I. K. Sendur, "The role of environmental factors and mine composition in thermal IR mine signatures," *Detection and Remediation Technologies for Mines and Mine-Like Targets VI, Proceedings of SPIE* , 2001.
12. P. Pregowski, W. Swiderski, R. T. Walczak, and K. Lamorski, "Buried mine and soil temperature prediction by numerical model," *Detection and Remediation Technologies for Mines and Mine-Like Targets V, Proceedings of SPIE* , 2000.
13. I. K. Sendur and B. A. Baertlein, "Techniques for improving buried mine detection in thermal IR imagery," *Detection and Remediation Technologies for Mines and Mine-Like Targets IV, Proceedings of SPIE* **vol. 3710**, pp. 1272–1283, 1999.
14. T. J. Cui and W. C. Chew, "Fast algorithm for electromagnetic scattering by buried 3-D objects of large size," *IEEE Trans. Geosc. Rem. Sens.* **vol. 37**, pp. 2597–2608, 1999.

Waveguide studies of soil modification techniques for enhanced mine detection with ground penetrating radar

Joel T. Johnson, Jatupum Jenwatanavet, Nan Wang, Ross Caldecott, and Jon Young

The Ohio State University ElectroScience Laboratory

1320 Kinnear Rd., Columbus, OH 43212-1191

ABSTRACT

Waveguide studies of the effectiveness of soil modification techniques for non-metallic mine detection with ground penetrating radar are described. Visibility improvements for a nylon target buried in sand are considered as varying amounts of water or liquid nitrogen are added to the sand. Experiments are performed in an S-band waveguide (2.6-3.8 GHz), and results show that increased water content improves shallow buried target visibility initially due to increased dielectric contrast between the target and background medium, but eventually obscures target responses due to increased loss. Initial tests of the addition of liquid nitrogen show reduced loss in the high moisture content case, so that targets can again be made visible. Analytical model results are also presented for the experimental configuration and found to be in good agreement with measured data, and further studies of soil modification effects with the analytical model are performed. A finite difference time domain (FDTD) electromagnetic model for more complicated geometries involving general target shapes and inhomogeneous water contents is also described.

Keywords: Soil modification, Ground penetrating radar, Landmine detection

1. INTRODUCTION

The detection and identification of non-metallic anti-personnel landmines remains a challenging problem for all current technologies.[1] Ground penetrating radar (GPR) systems are currently only of limited utility in this area because of the often low dielectric contrast between plastic mines and the surrounding soil. For example, the relative permittivity, ϵ_r, of most plastic materials is in the range of 2-4, while that of most dry soils is in the same range, making it difficult for a sensor which relies on scattering from dielectric contrasts to distinguish these two materials. It is well known however that the relative permittivity of soil can be dramatically modified through the external addition of certain chemical agents. A common example is the addition of water, which increases soil permittivity due to the highly polar nature of the water molecule. A plastic mine would thus appear as a large dielectric contrast to surrounding wet soil, improving the possibility of detection with a GPR system. However, the addition of water also significantly increases electromagnetic wave absorption in soil, so that more energy is lost as heat energy in the soil medium and does not return to the radar. There is thus a trade-off between these two effects - increased dielectric contrast but increased attenuation - that makes the exact effect of water addition difficult to completely characterize without experiments. The addition of liquid nitrogen to wet soils has been suggested as a method for reducing loss in the background medium, since electromagnetic loss in ice is much lower than that in liquid water. Coupled with the electromagnetic effects are the effects of fluid flow, which may enhance chemical content above a mine but decrease chemical content below, also potentially increasing mine detectability. Similar issues will arise with the addition of any chemical agents for soil modification.

To address these issues, experimental studies of target visibility improvements through the addition of water to dry sand have been performed at The Ohio State University ElectroScience Laboratory (ESL). The experimental configuration utilizes an S-band (2.6-3.8 GHz) waveguide geometry to reduce problem complexity while still providing insight into soil modification effects. Initial tests of target visibility improvements in wet sand through the addition of liquid nitrogen have also been performed. Results confirm that an optimal level of soil water content exists to maximize target visibility, and that liquid nitrogen can improve target detectability in wet soil. An analytical model for the experimental configuration has also been developed, and predictions are found to be in good agreement with measured data. Finally, a finite difference time domain (FDTD) model for the experimental configuration has been developed to allow studies of more general target cross sections and inhomogeneous water content effects.

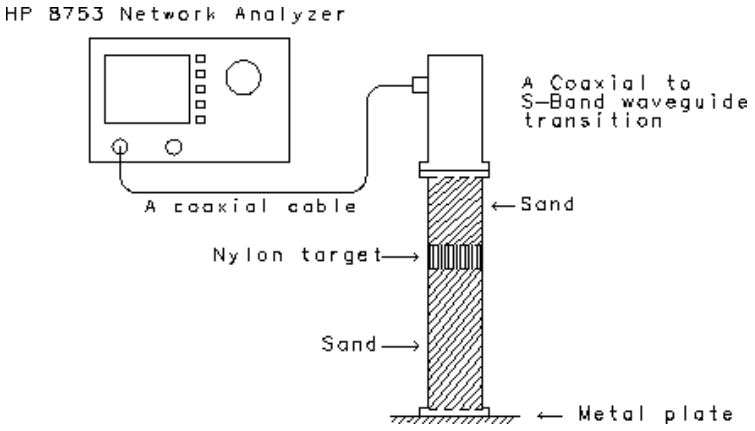


Figure 1. Experimental configuration

2. EXPERIMENTAL CONFIGURATION

To avoid the difficulties associated with an "open" electromagnetic experiment, in which scattering contributions can be obtained from many surrounding structures, a waveguide test range was developed for initial tests of the effectiveness of chemical addition. The "closed" geometry of a waveguide structure makes these experiments much more controlled than a GPR measurement, and enables the electromagnetic effects of a given chemical addition to be determined more directly than in a GPR measurement. In addition, the waveguide structure is more amenable to theoretical analysis.

S-Band waveguide (which operates from 2.6-3.8 GHz) was chosen for the test range, since this band coincides with a portion of the ESL GPR system frequency range of 2-6 GHz. The waveguide system consists of an HP 8753 network analyzer, a coaxial to S-Band waveguide transition and a 46 cm metal plate short circuited wave guide which acts as the sample holder. System calibration was performed using a three offset-short procedure (wideband load and open terminations were not available at S-band) and validated through reflection tests with known permittivity targets. Results from the measurement are obtained in terms of S_{11}, the reflected waveguide mode amplitude relative to the incident mode amplitude at the upper sand surface. Figure 1 illustrates the configuration of the waveguide test range.

Measurements were made in 0.003 GHz steps through the specified bandwidth, and time domain responses generated through a fast fourier transform after a 200 MHz raised cosine window was applied to the high and low frequency data. Although a waveguide inherently produces dispersive propagation, results show the effect to be tolerable, and reasonable time domain signals are still obtained which allow measurement data interpretation. Note that time domain signals for this configuration will show reflections from the top sand surface, nylon-sand interfaces, and the short circuit at the bottom of the guide if medium loss and dielectric constrasts are sufficiently low and high respectively. Time domain results will be presented only for the time interval before reflection from the short circuit appears, however, since this time interval more closely resembles a realistic GPR problem.

A surrogate "mine" target was created from a rectangular piece of nylon (relative permittivity $3.15 - j0.03$) which was cut to the same size as the waveguide cross section (7.2 cm by 3.4 cm) with a thickness of 2.64 cm. This target was buried at a depth of 7.9 cm when placed in the sand. The choice of this particular target was motivated by the simplicity of an analytical model for this geometry, with only a multilayer reflection coefficient[2] required. More general target shapes are also of interest, and would not completely fill the waveguide cross section so that water flow around the target would be possible, but analytical models become much more complicated.[3] These issues are considered with the FDTD model described in Sect. 5.

Sand samples (bulk density 1.55 g/cm^3) were prepared by first placing the sample in a 105 C oven overnight to insure a initially dry state. A specified volume of dry sand (measured by both mass and volume to insure that sand "packing" effects were avoided) was then mixed with a fixed volume of water to create the desired water content sand. The sample was then covered and again left overnight to insure a homogeneous mixture. Sand water content

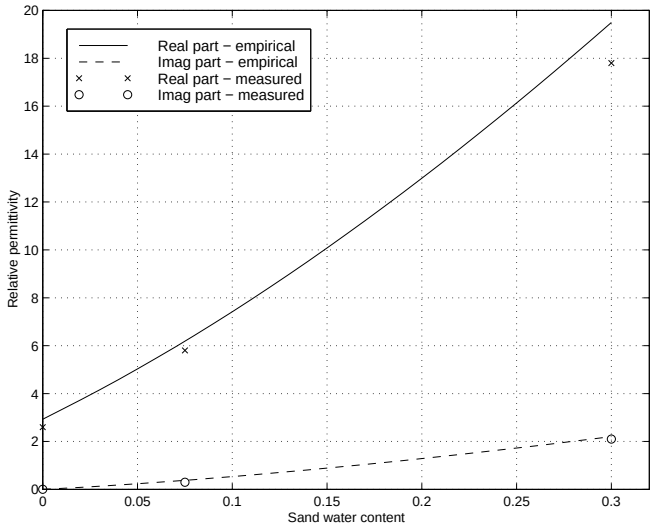


Figure 2. Comparison of retrieved and empirical model sand dielectric constant values versus water content at 3 GHz.

is described as the volumetric fraction of water in the sand, i.e. the ratio of the volume of water to the volume of dry sand that were mixed in preparing the sample. Water content fractions of 0%, 7.5%, and 30% are considered in this paper.

Initial tests of data repeatability showed problems which were corrected through more careful control of the experimental procedure. In particular, subtraction of the target-no target results was found to require careful cleaning of the waveguide interior to insure a repeatable result. Since "packing" effects are more significant with wet sands, a repeatable "packing" procedure was also required for the placement of wet sand in the waveguide. The procedure developed involved packing the sand after a specified mass had been added, and was found to provide repeatable masses of wet sand in the waveguide to within 5% with different identical water content fraction sand samples.

A smaller section of short circuited waveguide was used to measure sand dielectric constant (ϵ_r) for the three water content samples. In this procedure, a minimization algorithm was used with the multilayer reflection coefficient to determine the value of ϵ_r which most closely matched measured reflection data. A combined reflection-transmission measurement procedure,[4] well known to be more suited for permittivity retrieval than a reflection only measurement, is also being developed to improve this procedure. Figure 2 plots the retrieved values for real and imaginary parts of the dielectric constant at 3 GHz and compares them with an empirical model[5] for sand permittivity. Although some small discrepancies are observed, the overall good agreement between the empirical model and retrieved permittivity values further validates the measured data. The level of error observed is similar to that observed between the empirical model used and other empirical models for sand permittivity.[6]

3. EXPERIMENTAL RESULTS

Figure 3 illustrates measured values of S_{11} in the time domain using dry sand ($\epsilon_r = 2.60 - j0.005$). Plot (A) shows the case without the nylon target, plot (B) shows the case with the nylon target, and plot (C) illustrates the difference between plots (A) and (B), which should indicate the level to which the target appears. Also included are predictions from the multilayer reflection coefficient analytical model using the measured values of sand and nylon permittivity. A calculation using the waveguide group velocity in the sand medium at 3.2 GHz shows that the target response should be centered at 0.924 nS, as observed in plot (C). Clearly in this case, the low dielectric contrast between the target and background medium makes returns from the target very small and difficult to resolve with this system. The initial reflection off the top surface of the sand is very closely matched in plots (A) and (B), and subtracts very well in part (C), indicating that a repeatable sand sample is being obtained inside the waveguide. Experimental

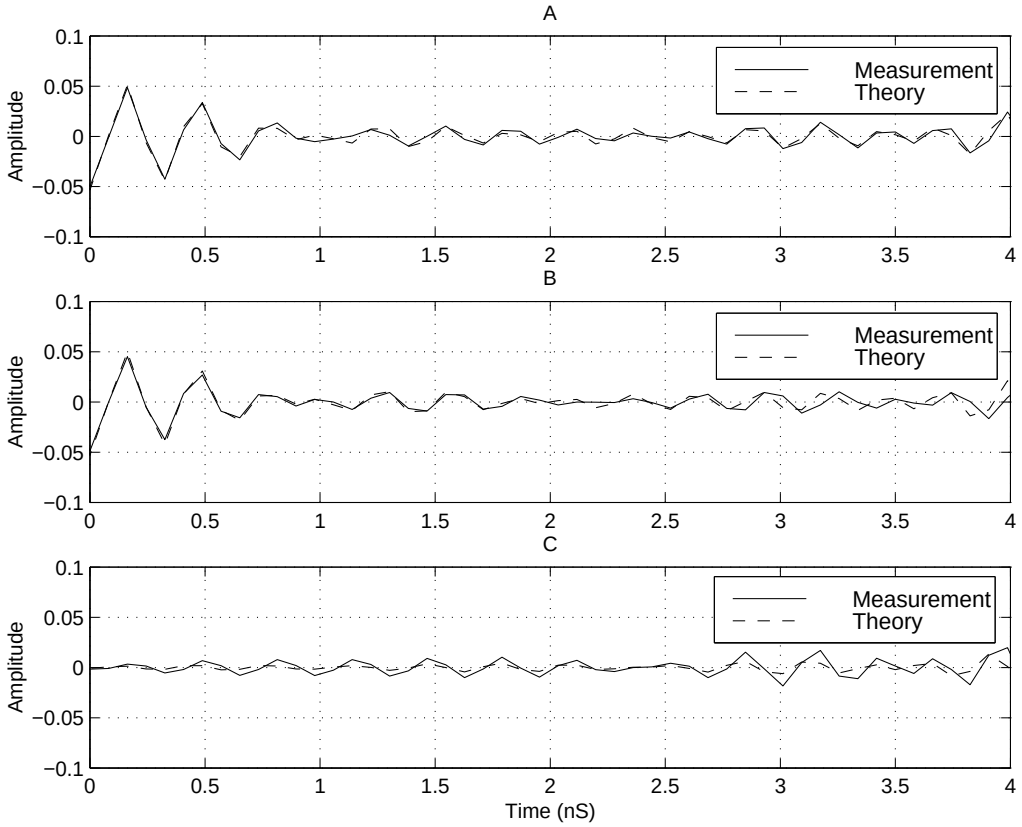


Figure 3. Time domain reflections from dry sand (A) with no target (B) with buried nylon target (C) subtraction of plots (A) and (B)

results also match the analytical theory model well, illustrating that calibration and sample errors have been reduced to a reasonable level in the experiment.

Figure 4 illustrates the same set of measurements using sand with a 7.5% water content ($\epsilon_r = 5.8 - j0.3$). In this case, plot (C) shows a much stronger return from the nylon target, given the increased dielectric contrast and relatively low loss in the background medium. The group velocity calculation shows that the target response should be centered at 1.3 nS, as observed in plot (C). Again, the reflection off the top sand surface subtracts well, and the results are in reasonable agreement with the analytical predictions.

Figure 5 illustrates the results using sand with a 30% water content ($\epsilon_r = 17.8 - j2.1$), where the target response should occur at 2.25 nS. In this case, the nylon target is again obscured in plot (C) due to the much higher loss in the background medium, so that target returns are strongly attenuated before returning to the top sand surface. Analytical predictions again match this data well.

Finally, Figure 6 shows the 30% water content case after an addition of liquid nitrogen to the top surface of the sand. Liquid nitrogen was added by tightly screwing a 7.2 cm piece of waveguide to the top of the sample holder, and then filling this piece of waveguide with liquid nitrogen and allowing the nitrogen to either evaporate or flow into the sand. Rapid evaporation of the liquid nitrogen and the tightly packed nature of the wet sand in the sample holder made obtaining nitrogen penetration into the wet sand difficult, so a substantial quantity (730 ml) of liquid nitrogen was required. This quantity was added by continuously pouring liquid nitrogen into the top waveguide in an effort to keep the top waveguide filled, and required approximately 5 minutes to add 730 ml. The results in Figure 6 (shown only for the with target case), clearly show that target visibility in very wet sands can be improved through the addition of liquid nitrogen. A comparison with Figure 5, plot (B) emphasizes this point; in fact, the response seems more similar to Figure 4, plot (B) in which the water content was 7.5%. Only this initial test of liquid

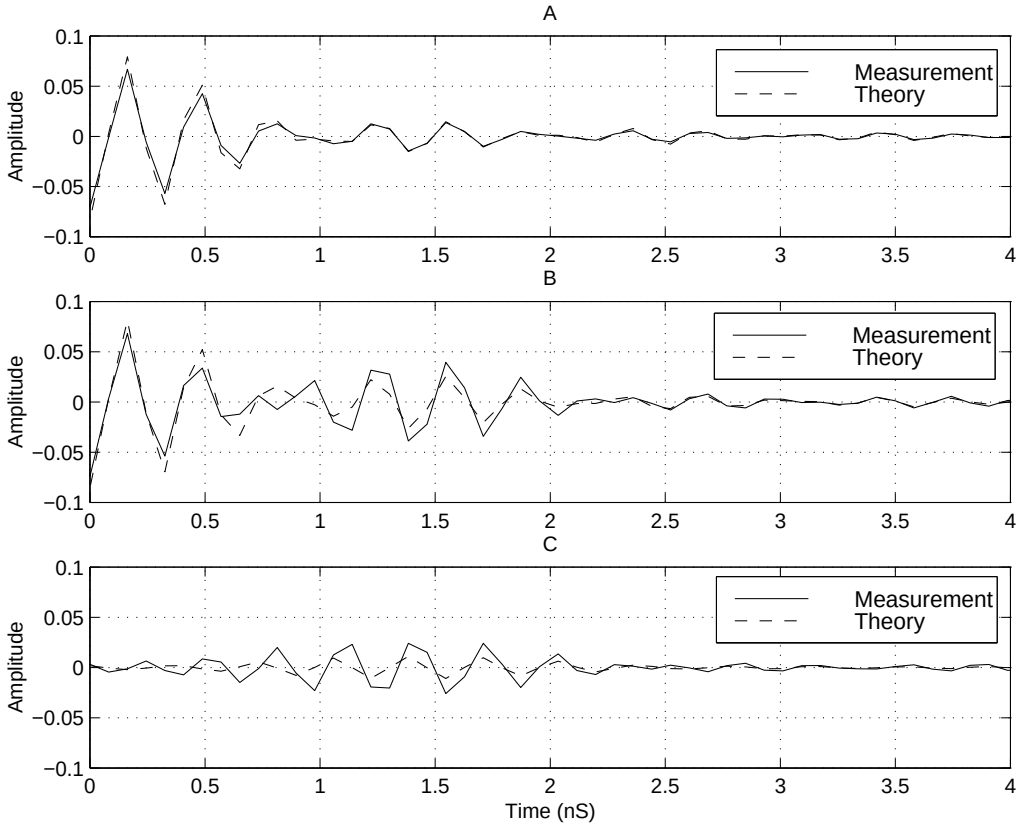


Figure 4. Time domain reflections from sand with 7.5% water content (A) with no target (B) with buried nylon target (C) subtraction of plots (A) and (B)

nitrogen addition has currently been performed, so further efforts will attempt to develop more efficient application procedures.

4. FURTHER ANALYTICAL STUDIES

The experiments of the previous section were performed with a fixed target depth and relative permittivity, and only for three values of soil water content. However, the success of the analytical model in matching the experimental data shows that results can be reliably predicted by the simple multilayer reflection model for the geometry considered, in combination with the empirical model[5] for sand dielectric constant as a function of water content. Use of the empirical model also allows soils types other than pure sand to be considered. A theoretical study of optimal soil water contents for varying target depths and soil types was performed to further investigate the effects of water addition on mine detection systems. The study was conducted for the same frequency range as the experiments and with the same nylon target, but with target depths ranging from 1 to 8 inches (0.0254 to 0.2032 m) in steps of 1 inch and with an infinitely long waveguide below the target (i.e. a matched load instead of the short circuit plate). Optimal soil water contents were determined by first calculating the frequency domain reflection coefficients with the analytical model for the target and no target cases, and then transforming the difference between the two into the time domain. The value of soil water content which produced the maximum amplitude of the time domain difference signal was taken as optimal. Soil water contents considered in the study varied from zero to 30% in steps of 1%.

Figure 7 illustrates optimal water contents obtained as a function of target depth for both pure sand and pure clay soils. The shape of the curves in both cases is similar, and indicates that increasing loss in the higher water content soils becomes more intolerable as target depth increases. The discrete values of target depths and soil moistures considered results in the somewhat non-smooth curve obtained, but the values observed can still serve

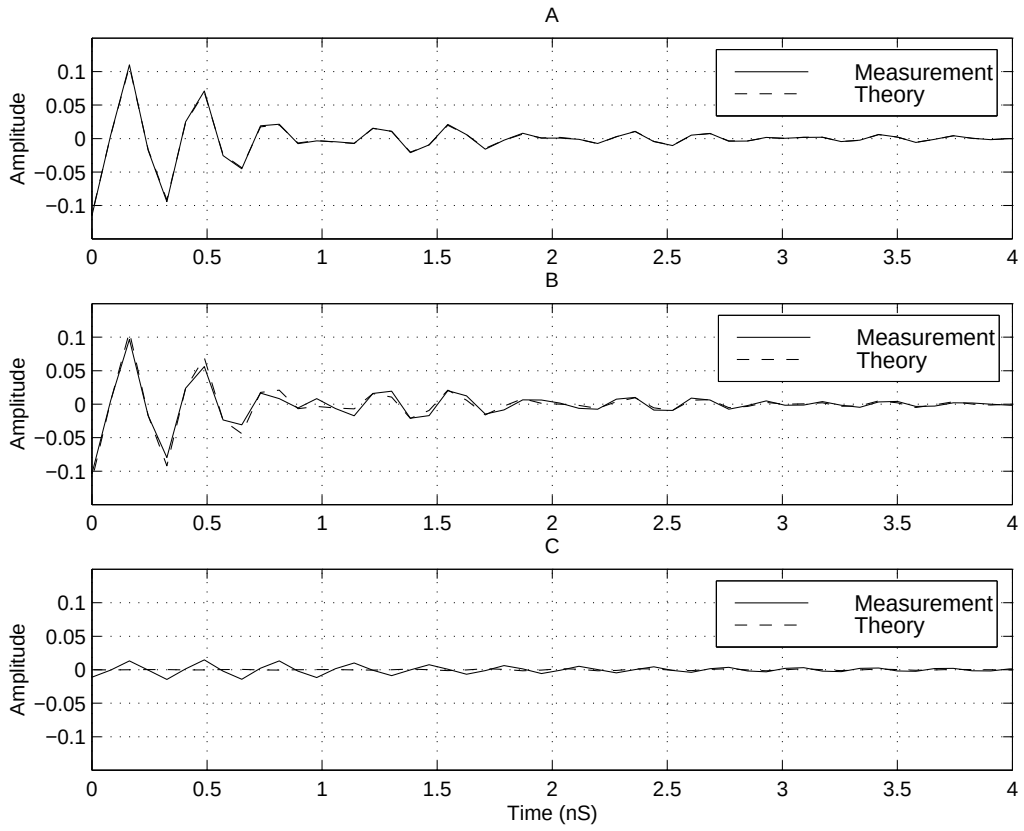


Figure 5. Time domain reflections from sand with 30% water content (A) with no target (B) with buried nylon target (C) subtraction of plots (A) and (B)

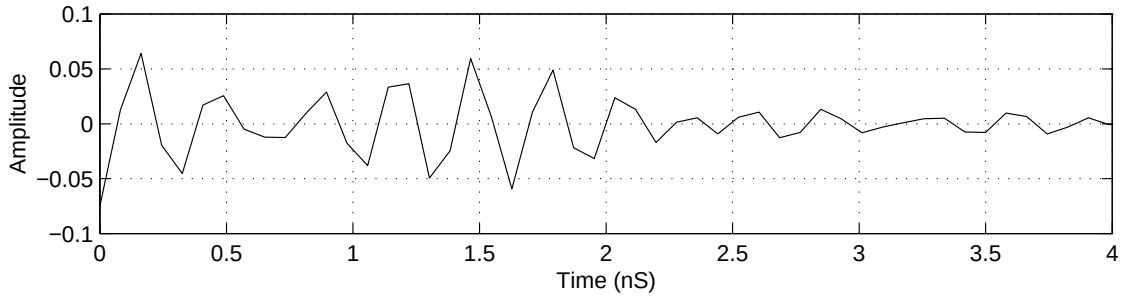


Figure 6. Same as Figure 5, plot (B) but after the addition of 730 ml of liquid nitrogen to the top sand surface

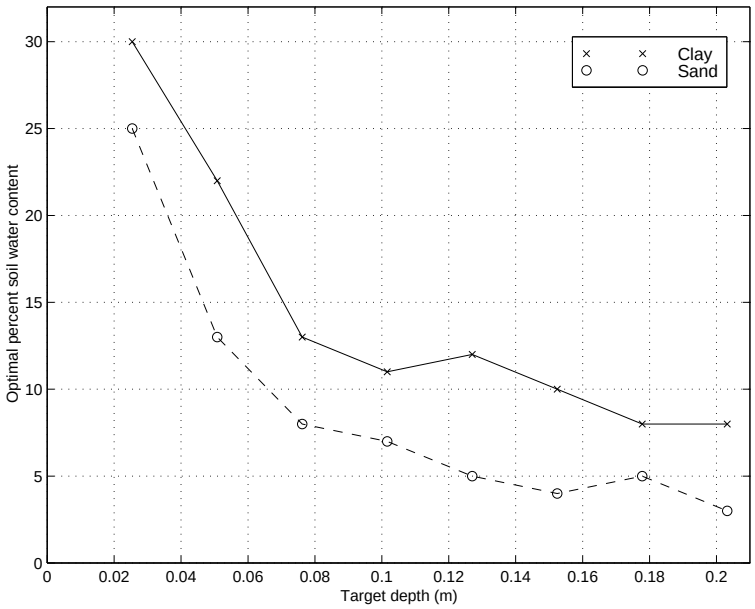


Figure 7. Optimal soil water contents obtained in analytical study versus target depth

as basic guidelines for further soil modification studies. Note also that the relatively flat nature of the curves for large target depths should not be expected to continue indefinitely, as loss even in the lower water content soils will become significant for very deep targets. The present study however is focused on anti-personnel mine efforts, so the range of target depths illustrated is reasonable.

5. FDTD MODEL AND RESULTS

The multilayer reflection coefficient analytical model is appealing in its simplicity and efficiency, but the geometry for which it applies (target completely filling waveguide cross section and homogeneous medium) is very simple and may not be representative of a realistic mine detection scenario. To address these issues, a numerical model based on the finite difference time domain[7] approach has been developed which can provide theoretical predictions for arbitrary cross section targets and inhomogeneous soil water contents. The FDTD approach discretizes Maxwell's equations in both the time and space domains, and solves for fields throughout the domain through a time stepping procedure. The perfectly matched layer (PML) absorbing boundary condition of Berenger[8] is used to terminate the computational domain at the top and bottom of the waveguide, and a point dipole source located above the top sand surface is used to excite fields inside the waveguide. Reflection coefficients are determined by completing runs with and without the sand surface, so that incident and total fields can be determined separately and subtracted to obtain the reflected field. Several tests of FDTD model accuracy were completed versus the analytical model and experimental results for the geometry of the previous sections, and showed the model to provide accurate calculations.

Figure 8 compares FDTD calculated time domain target-no notarget subtractions for the previous target and a smaller nylon target (thickness 0.0264 m but cross section reduced to half the waveguide dimensions) in sand with water content 7.5%. A total number of 12 x 24 x 240 cells were used in this calculation for a waveguide of dimensions 0.034 x 0.072 x 0.72 m, so that a worst case discretization of $\lambda/12$ was obtained in the entire band. A time step of 4.5 ps was used with 10,000 time steps, and the computation required approximately 7 hours on a 200 MHz Pentium Pro processor. The comparison of Figure 8 shows that the target response is not significantly reduced with the smaller target, and remains visible in this near optimal sand water content. Further studies with the FDTD model will continue, especially with inhomogeneous sand water contents, as experiments for these geometries are performed.

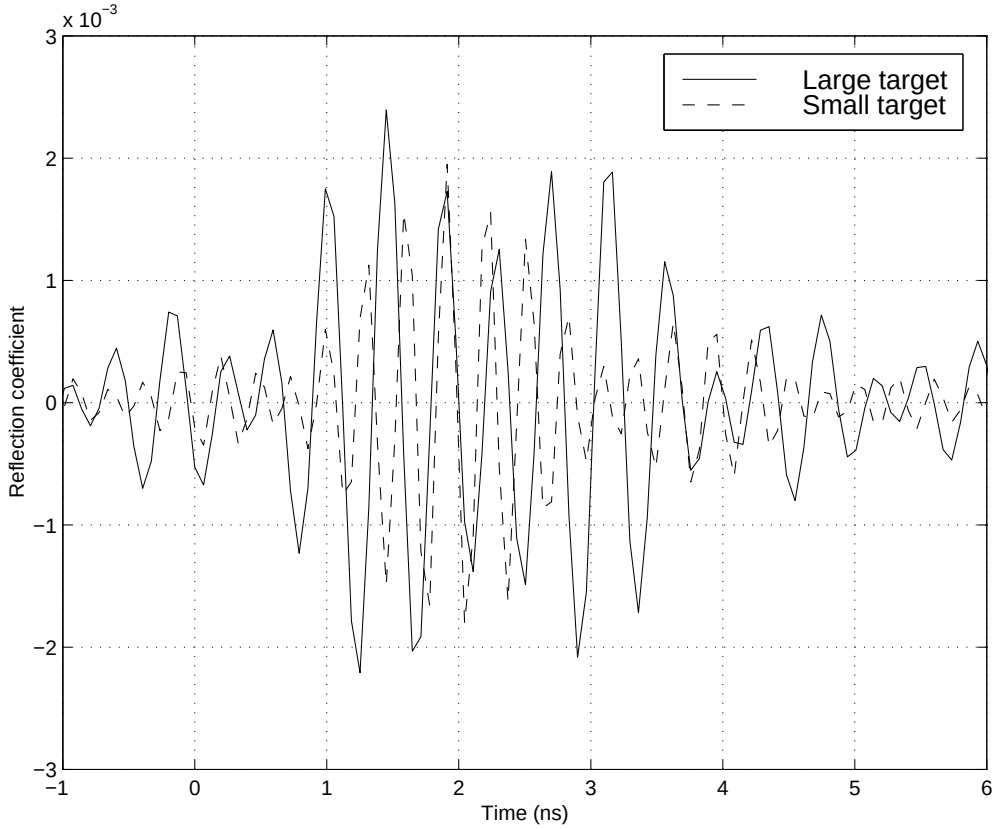


Figure 8. Comparison of FDTD calculated target-no target differences from sand with 7.5% water content for two target sizes

6. CONCLUSIONS AND FUTURE WORK

Waveguide studies of the effectiveness of water and liquid nitrogen in enhancing target detectability with ground penetrating radar systems have been described. Although the waveguide configuration does not completely capture all the details of a ground penetrating radar system, it provides a carefully controlled closed system for determination of the electromagnetic effects of various chemicals. Results confirm that low dielectric constant target visibility can be improved in dry soils through the addition of water, but that an optimal value of soil water content exists beyond which background medium loss begins to obscure the target. Initial experimental results show that the addition of liquid nitrogen can reduce loss in very wet soils and regain some target visibility. Analytical and numerical models for the waveguide configuration have been developed, and confirm experimental predictions. Further experimental studies involving varying target shapes and the effects of water flow are continuing. The necessity for improving soil/target dielectric contrasts to obtain better mine detection and discrimination with ground penetrating radar systems is clear, and results from these efforts could provide a very promising mine detection technology for future demining efforts.

ACKNOWLEDGMENTS

This work was supported by a grant from Duke University as part of the Multi-disciplinary university research initiative (MURI) on de-mining.

REFERENCES

1. A. Eisenberg, *War of the Mines: Cambodia and the Impoverishment of a Nation*, Pluto Press, London, 1994.
2. J. A. Kong, *Electromagnetic Wave Theory: 2nd edition*, Wiley, New York, 1990.
3. N. Marcuvitz, *Waveguide Handbook*, McGraw-Hill, New York, 1951.
4. Hewlett Packard Co., "Measuring the dielectric constant of soils with the HP 8510 network analyzer," *Product Note 8510-3* .
5. F. Ulaby, R. Moore, and A. Fung, *Microwave Remote Sensing: Active and Passive, vol. 3*, Addison-Wesley, Reading, 1981.
6. C. Matzler, "Microwave permittivity of dry sand," *IEEE Trans. Geosc. Rem. Sens.* **36**, pp. 317–319, 1998.
7. K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Ant. Prop.* **14**, pp. 302–307, 1966.
8. D. S. Katz, E. T. Thiele, and A. Taflove, "Validation and extension to three dimensions of the Berenger PML absorbing boundary condition for FDTD meshes," *IEEE Micro. and Guided Wave Letters* **4**, pp. 268–270, 1994.

Soil modification studies for enhanced mine detection with ground penetrating radar

Joel T. Johnson, Jatupum Jenwatanavet and Nan Wang

The Ohio State University ElectroScience Laboratory

1320 Kinnear Rd., Columbus, OH 43212-1191

ABSTRACT

The detection of non-metallic anti-personnel landmines with ground penetrating radar (GPR) is complicated by low dielectric contrasts with the surrounding background medium. Previous studies have shown that the addition of water can improve dielectric contrasts but also increases loss so that target detectability is not necessarily improved. Previous studies have also shown that the addition of liquid nitrogen to wet soils can reduce background medium loss and restore target visibility. In this paper, further waveguide studies of target detection through a controlled depth of nitrogen penetration are reported, and it is shown that scattering from known depth targets can be significantly enhanced if an optimal amount of nitrogen is added. The procedure can also be generalized to unknown depth targets if measurements are taken as gradually increasing amounts of liquid nitrogen are added. Both analytical models and waveguide experiments are presented to illustrate these ideas. Finally, initial tests of the soil modification techniques developed through waveguide experiments are reported with a dielectric rod GPR system; results indicate that these methods should be applicable to general GPR sensors.

Keywords: Soil modification, Ground penetrating radar, Landmine detection

1. INTRODUCTION

The detection and identification of non-metallic anti-personnel landmines remains a challenging problem for all current technologies.[1] Ground penetrating radar (GPR) systems are currently only of limited utility in this area because of the often low dielectric contrast between plastic mines and the surrounding soil. For example, the relative permittivity, ϵ_r, of most plastic materials is in the range of 2-4, while that of most dry soils is in the same range, making it difficult for a sensor which relies on scattering from dielectric contrasts to distinguish these two materials. Previous S-band waveguide studies[2] have shown that the addition of water to dry sand can increase target/background medium dielectric contrasts so that larger scattered fields can be obtained. However, attenuation in the background medium is also increased and reduces scattered fields. For homogeneous water content soils an optimal sand water content exists to provide maximum target visibility, but for deeper targets even this "maximum" response can be quite small due to attenuation. For the non-homogeneous water content soils created when water is poured onto a soil surface and allowed to penetrate over a moderate time period, previous waveguide experiments[3] have shown that target visibility is not improved because a substantial quantity of water is required to obtain penetration to target depths and results in extreme losses near the surface. Thus the addition of water alone may not be sufficient to enhance target detection in many environments.

To address problems with excessive loss in wet soils, the addition of liquid nitrogen to the soil surface has been proposed.[2,3] Due to the much smaller loss tangents in ice than in water,[4] target visibility can potentially be regained. However, the dielectric constant of frozen soil also becomes more similar to that of non-metallic targets, so again small dielectric contrasts become a problem in completely frozen soils. Experimental results[3] suggested that an "optimal" amount of liquid nitrogen should exist to enhance target visibility; this amount of nitrogen would be sufficient to freeze most of the soil above the target so that propagation loss is reduced, but would leave a small layer of wet soil around the target to retain dielectric contrast so that large scattered fields could be obtained. The importance of flow effects for water and liquid nitrogen have also been discussed[3]; improvements in target visibility are determined by subtracting measurements made with and without a surrogate mine target. Differences in the flow patterns of water and liquid nitrogen with and without a target can therefore also be measured, and careful control of these chemicals must be exercised to obtain repeatable measurements.

In this paper, further studies of soil modification through the addition of water and liquid nitrogen are presented. Results from a simplified analytical model for the waveguide configuration confirm the enhancement of known depth

target visibility through the addition of an optimal amount of liquid nitrogen. The analytical study also suggests a procedure for locating low-contrast targets at unknown depths through a gradual addition of increasing amounts of liquid nitrogen. Waveguide experiments confirm this concept, even with the complication of water and liquid nitrogen flow effects not considered in the analytical study. The resulting procedure allows unknown depth targets which have *exactly* the same dielectric constant as the original soil medium to be detected in principle. Finally, the soil modification techniques developed are applied in an initial measurement with the dielectric rod 2-6 GHz GPR system being developed at OSU,[5] and results suggest that the techniques proposed should be applicable to general GPR systems.

2. ANALYTICAL STUDY OF NITROGEN ADDITION

Figure 1 illustrates the soil modification procedure proposed to locate targets at a known depth. In plot (a) of Figure 1, a low contrast target (for example made of nylon which has $\epsilon_r \approx 3.15 - j0.003$ in the 2.6-3.8 GHz band considered for the S-band waveguide system) is buried at a known depth d in dry sand ($\epsilon_r \approx 2.6 - j0.005$). Scattered returns in plot (a) would be small due to the small dielectric contrast between the nylon and dry sand media. The second figure of plot (a) shows the experimental configuration after pouring a large amount of water onto the top sand surface. Scattered returns in this case would again be small due to an excessive attenuation in the very wet sand above the target. Figure 1 plots (b) through (d) now illustrate the configuration as increasing amounts of liquid nitrogen are added, so that the frozen sand region above the target extends to greater and greater depths. In plot (b) only a small quantity of nitrogen has been added, so that loss in the wet sand region above the target still causes excessive attenuation and the target remains obscured. In plot (c), an "optimal" amount of liquid nitrogen has been added, so that loss above the target is minimized, but a large dielectric contrast in the region surrounding the target remains to provide scattering. In plot (d), an excessive amount of nitrogen has been added and again the target is obscured due to the low contrast between the nylon and frozen sand media. If the target depth is assumed known, the optimal amount of nitrogen to be added can be determined and cataloged for future measurements of similar depth targets. However, it is also clear from Figure 1 that the procedure should be applicable to unknown depth targets as well, since non-optimal amounts of liquid nitrogen do not produce large target returns. Thus, a procedure in which measurements are taken as gradually increasing amounts of liquid nitrogen are added should be sufficient to improve visibility of unknown depth targets as well.

It is also clear from Figure 1 that scattering should be obtained not only from the target/background interfaces but also from the frozen/non-frozen sand interfaces. However, it is assumed that a no target result exists which can be subtracted to reduce the significance of these contributions. Such a result could be obtained in practice from performing measurements at a location where it is known a-priori that no target exists. Of course, the success of such a subtraction depends on the extent to which liquid nitrogen and water flow effects are similar in the target and no-target cases; experimental results presented later in the paper will confirm that this procedure is reasonable at least for the waveguide experiments performed.

An analytical model for reflected fields in the waveguide system has been developed based on multilayer reflection coefficients and applied to the waveguide configuration of previous experiments.[2] Unlike Figure 1, this model assumes that the target completely fills the waveguide aperture. Simulated target minus no-target S_{11} returns in the time domain based on the configurations of Figure 1, plots (a) through (d), are shown in Figure 2 plots (a) through (d) for a target at depth 12.7 cm. Permittivities of wet and frozen sand are assumed to be $17.8 - j2$ and $3.4 - j0.003$ respectively; these values were found in several tests to provide reasonable agreement with measured S_{11} returns. Note the no-target case has four configurations identical to those of Figure 1 except that the nylon target is removed. Figure 2, plot (a) confirms the small scattered returns obtained from the low contrast target in dry soil, and even smaller returns (not plotted) were obtained for the second figure of Figure 1, plot (a) after the addition of a large quantity of water. Figure 2, plot (b) illustrates the return for the configuration of Figure 1, plot (b); here the small amount of nitrogen added is insufficient to reduce excessive background loss and the target remains obscured. Figure 2, plot (c) shows the greatly enhanced target response obtained in the "optimal" configuration, in which 3.2 mm of wet sand remain above the target. Finally Figure 2, plot (d) shows much smaller responses as the entire background medium becomes frozen, again due to small dielectric contrast between nylon and frozen soil.

Figure 3 compares target minus no-target responses in the "optimal" liquid nitrogen case for targets buried at depths of 12.7 cm and 17.8 cm. The "optimal" case for the deeper target is again defined such that a 3.2 mm layer of wet sand remains above the target and therefore a differing total amount of liquid nitrogen has been added. The

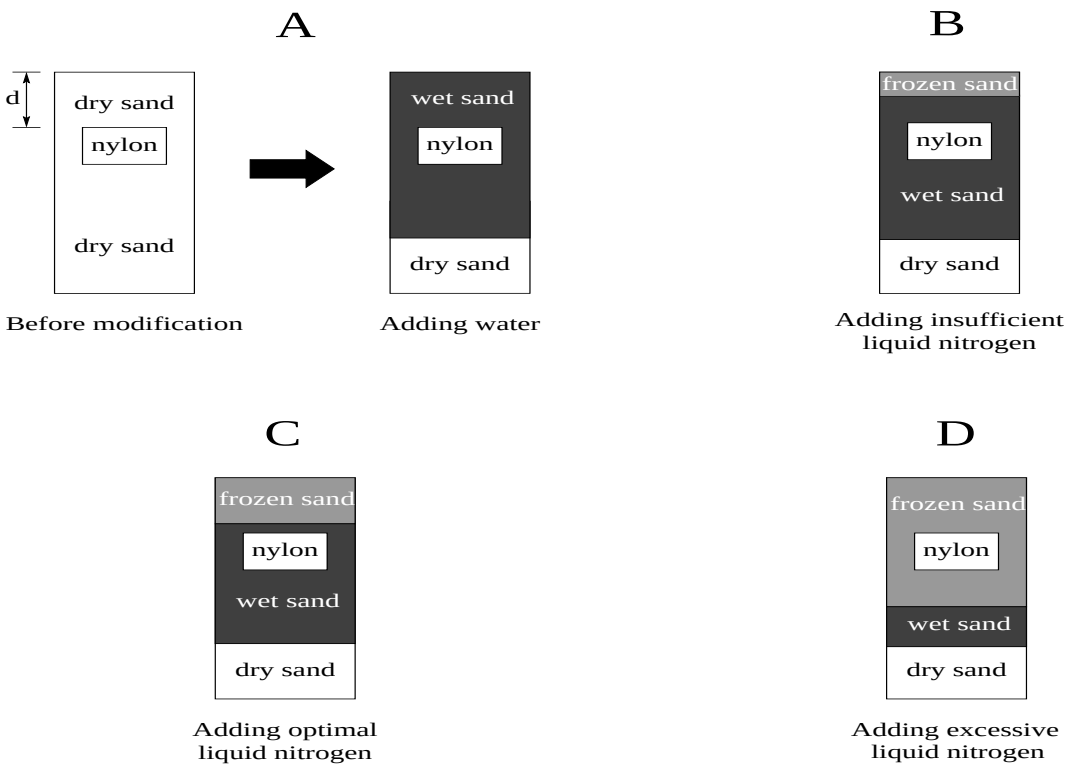


Figure 1. Soil modification procedure (a) Target in dry soil and after the addition of water (b) After a small amount of liquid nitrogen added (c) After an "optimal" amount of liquid nitrogen added (d) After an excessive amount of liquid nitrogen added

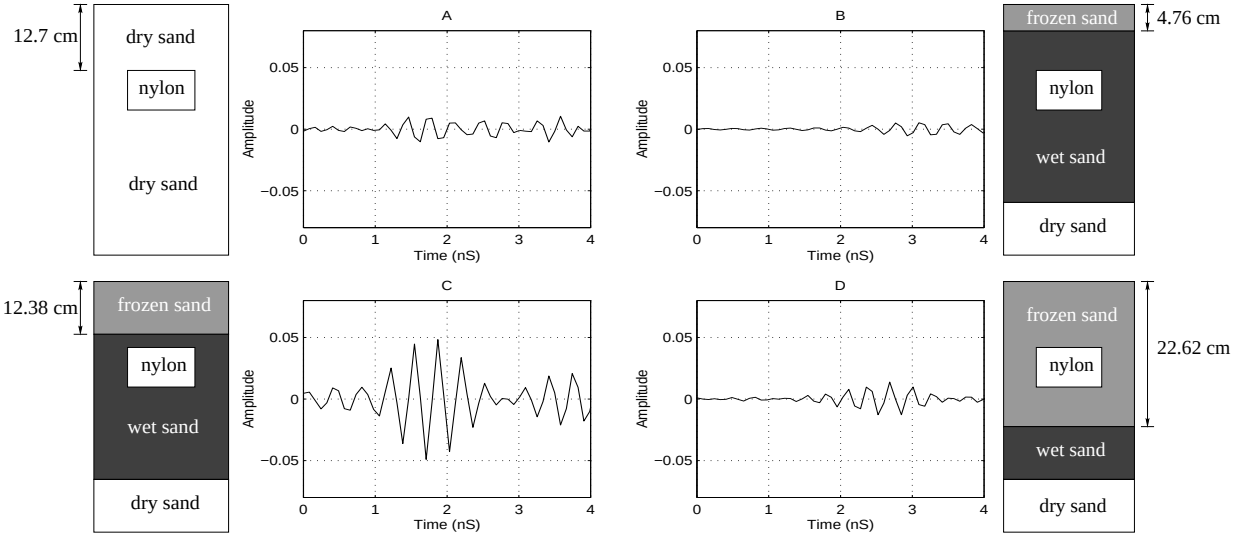


Figure 2. Time domain target minus no-target S_{11} for the configurations of Figure 1 and with a 12.7 cm depth target (a) Target in dry soil and after the addition of water (b) After a small amount of liquid nitrogen added (c) After an "optimal" amount of liquid nitrogen added (d) After an excessive amount of liquid nitrogen added

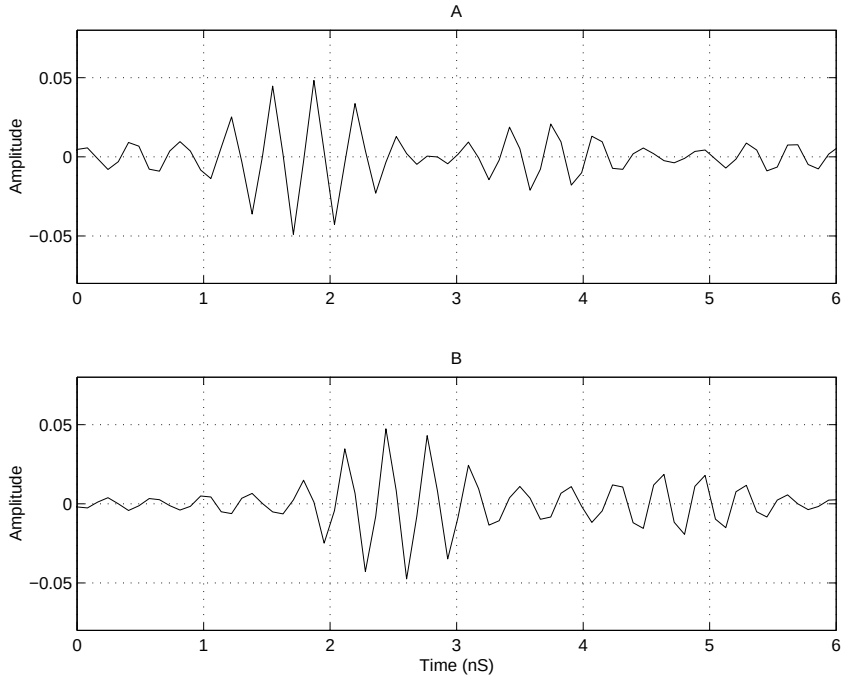


Figure 3. Time domain target minus no-target S_{11} with an "optimal" amount of liquid nitrogen added (a) 12.7 cm depth target (b) 17.8 cm depth target

similar level of target response obtained in these cases is as expected due to the very small losses in frozen sand. Also the time shift for the deeper target corresponds to the delay (0.68 nS) expected using the group velocity in a rectangular waveguide. Results for the deeper target using the "optimal" amount of nitrogen for the shallower target would provide very small scattering returns, so these results again suggest that measurements made as gradually increasing amounts of nitrogen are added should allow unknown depth, low contrast targets to be detected.

3. WAVEGUIDE MEASUREMENTS

To confirm the unknown depth target procedure suggested by the analytical models, measurements in an S-Band 2.6-3.8 GHz waveguide system[2] were performed. A 2.54 cm thick rectangular block nylon target was used which had cross sectional dimensions of one half those of the waveguide aperture, so that liquid flow effects around the target were possible. Note there are several complicating factors in the measurements not present in the analytical study which could invalidate the procedure; in particular, water and nitrogen flow effects around the target can produce differing background media in the target and no target cases so that reasonable subtractions would not be obtained. Also, obtaining repeatable flow patterns and depths of penetration of water and liquid nitrogen with identically prepared sand samples could also present problems. The procedures developed previously[2] were found adequate for adding water and liquid nitrogen, but several tests showed that considerable care was required in the timing of experiments to insure repeatable results. Timing is a more critical factor in waveguide experiments because the waveguide aperture must be opened to add nitrogen and re-sealed to perform measurements, incurring a time delay of several minutes.

Figure 4 plots measured target minus no-target S_{11} values in the time domain with the small nylon target buried at 12.7 cm. The target was initially buried in dry sand, and then 150 ml of water was added to the top sand surface. Plot (a) illustrates returns after 750 ml of liquid nitrogen was added, and the small returns demonstrate that this is too small an amount of nitrogen to obtain substantial penetration to target depths. Results after the addition of 1500 ml of nitrogen are shown in Plot (b), and the much larger target response demonstrates the enhanced target visibility predicted by the analytical model for the "optimal" configuration. A second curve is included in plot (b) also, which corresponds to the return obtained from a subtraction of two no target experiments and indicates the

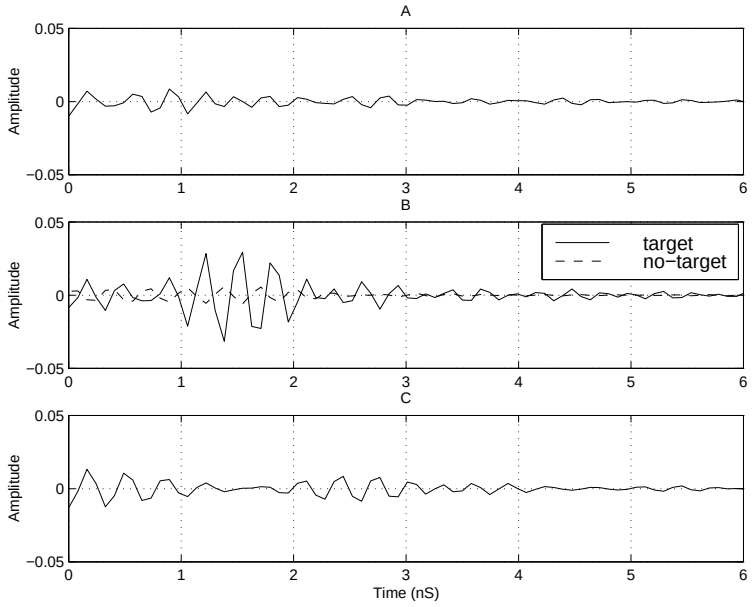


Figure 4. Measured time domain target minus no-target S_{11} with a 12.7 cm depth target after addition of 150 ml of water (A) with 750 ml of liquid nitrogen added (B) with 1500 ml of liquid nitrogen added (C) with 2250 ml of liquid nitrogen added.

level of measurement repeatability. Clearly the target response obtained is significantly larger than the experimental errors. The similarity between measured target minus no-target results in Figure 4, plot (b) and modeled results in Figure 3, plot (a) (which is also for a target at 12.7 cm depth) is quite striking. Target responses occur at similar time delays and have similar magnitudes, again confirming that the measured response indeed should be that of the target. Finally plot (c) illustrates results after the addition of 2250 ml of nitrogen, and confirms that excessive amounts of nitrogen addition again obscure target responses.

Results for a target buried at 17.8 cm depth are presented in Figure 5. In this case, 1250 ml of liquid nitrogen was added in plot (a), 2500 ml in plot (b), and 3250 ml in plot (c). The interpretation of these results is similar to that of Figure 4, and again the similarity between Figure 5, plot (b) and Figure 3, plot (b) is striking. Overall these experimental results confirm that even with the complications of chemical flow effects and measurement repeatability, an amount of liquid nitrogen exists which provides large target responses. Since this amount depends on the depth of the target, measurements taken as gradually increasing amounts of liquid nitrogen are added should reveal any obscured low contrast targets.

4. GPR MEASUREMENTS

Initial tests of the soil modification procedures developed through waveguide studies were also performed using the dielectric rod antenna GPR system currently under development at OSU.[5] Figure 6 depicts the configuration of this system, which operates from 2-6 GHz and uses a dielectric rod antenna located near the ground surface in order to obtain a very small illuminated area. The antenna is attached to a horizontal movement controller so that measurements can be taken as a function of position. The system has a basic calibration procedure involving small wire scatterers, but the near field nature of this system makes a complete calibration difficult, so only qualitative image responses are considered. Measurements were taken only in the presence of the target, so results are presented in terms of target responses only. The principle goal of these initial measurements was to qualitatively assess the performance of liquid nitrogen addition in restoring target responses previously obscured due to excessive loss in the background medium. The targets considered in these measurements were 7.62 cm diameter disks of 2.54 cm thickness buried at 6.35 cm depth; measurements were made both with metallic and nylon disks.

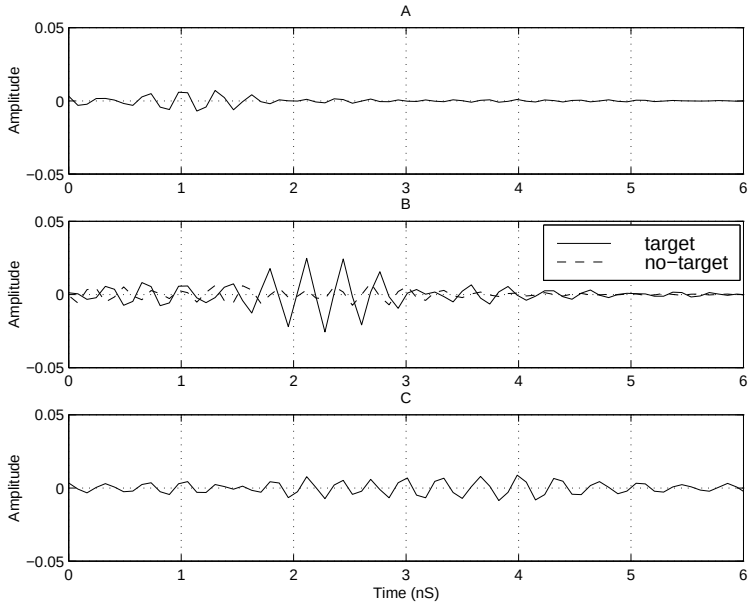


Figure 5. Measured time domain target minus no-target S_{11} with a 17.8 cm depth target after addition of 150 ml of water (A) with 1250 ml of liquid nitrogen added (B) with 2500 ml of liquid nitrogen added (C) with 3250 ml of liquid nitrogen added.

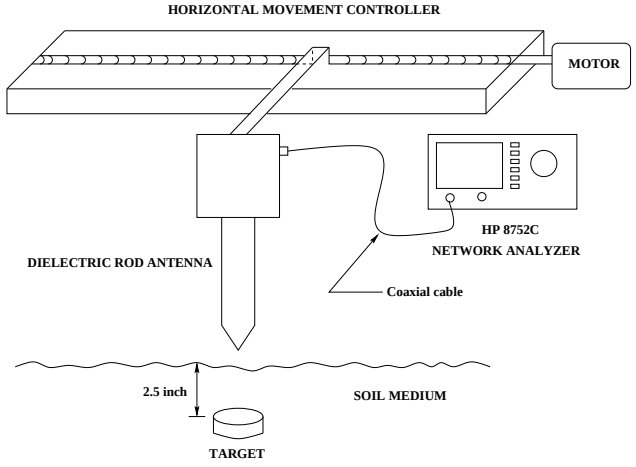


Figure 6. Experimental configuration for dielectric rod antenna GPR system

Again, several complicating issues arise when considering an "open" environment as with the GPR system compared to the "closed" environment of the waveguide measurements. First the soil considered is that of the outdoor ElectroScience Laboratory surrogate mine field (estimated complex permittivity $6 - j0.56$), which produces moderate attenuation even in the initial state before any water or nitrogen is added. The ground surface is not as easily controlled as in the waveguide system also, and clutter returns from the surface should be expected in the early time images which in practice require more advanced suppression methods[6] not considered in this study. Finally, the addition of liquid nitrogen is more difficult, since a confined area for the nitrogen to penetrate is not obtained as in the waveguide system. All of these issues will require more detailed studies to determine the most efficient methods for soil modification with GPR systems.

In this initial test, substantial quantities of water and liquid nitrogen were applied to insure that soil properties were modified. Measurements of the target were first made with the soil in its initial state, and then two gallons of water were poured over an area of approximately one half meter square which contained the target. Measurements in this "excessively wet" configuration were then taken, and finally approximately 15 liters of liquid nitrogen were applied over the same area and the final measurements made. Tests of the nitrogen penetration depth were made by digging up the soil after the measurements were completed, and it was estimated that penetration to approximately 8 cm was obtained with this quantity of nitrogen. Again these are substantial chemical amounts, but it is expected that further studies will develop more efficient application procedures that will reduce these quantities.

Measured results for the metallic and nylon targets are presented in Figures 7 and 8 respectively as images. These images are produced by plotting time domain measured fields as a function of antenna horizontal position. Results are calibrated so that time 0 nS corresponds to the top surface of the ground, and the target response should occur between antenna positions 60 and 70 at 1 to 1.5 nS time delay. Figure 7 and 8 both contain three plots: plot (a) is the image before any water or nitrogen has been added, plot (b) is the image after the addition of water, and plot (c) is image after the addition of water and liquid nitrogen. Figure 7 plots (b) and (c) clearly demonstrate the improved visibility of a metallic target in excessively wet soil obtained after the addition of liquid nitrogen. Even though the target is visible also in the un-modified soil, liquid nitrogen addition would be useful in situations where the un-modified soil had properties closer to the wet soil considered here. Results in Figure 8 are also interesting: clearly a much smaller response than the metallic target is obtained in plot (a) with the nylon target, and again the response is completely obscured in excessively wet soil. The enhanced response in part (c) again clearly shows the effectiveness of liquid nitrogen addition. The relative contributions of reduced loss in the background and increased dielectric contrast are difficult to assess from this single measurement, but these initial qualitative results clearly demonstrate that the soil modification techniques developed from the waveguide study should have applicability to general GPR systems.

5. CONCLUSIONS AND FUTURE WORK

A method for enhancing unknown depth, low contrast target visibility through a series of measurements taken as gradually increasing amounts of liquid nitrogen are added has been described. Both analytical model studies and measurements in a waveguide system demonstrate the success of the procedure, even with complicating factors such as liquid nitrogen and water flow effects. A target/no target subtraction is required, but no target results should be available if measurements are taken from a location where it is know a-priori that no target exists. Initial measurements with a dielectric rod antenna ground penetrating radar system also demonstrate that the addition of liquid nitrogen to excessively wet soils can reduce loss and enhance target visibility. Further tests regarding more efficient application procedures for liquid nitrogen in open environments and more detailed GPR system measurements are currently in progress. Note that portable "plants" exist for compressing liquid nitrogen from surrounding air[7,8] which could make these methods more feasible in remote environments.

Clearly the soil modification procedures described can produce improvements to any GPR system in two limiting cases: targets with strong scattering responses (i.e. metallic) obscured by excessive loss in the background medium, and targets whose dielectric constant is identical to that of the background medium. In addition, the soil modification techniques described can potentially produce advantages not only in target detection but also in target identification, since scattered target signatures should be enhanced under larger dielectric contrasts. These issues will also be considered in future studies.

ACKNOWLEDGMENTS

This project was supported by funds from Duke University under an award from the ARO (OSD MURI on Humanitarian Demining). The findings, opinions and recommendations expressed therein are those of the authors and are not necessarily those of Duke University or the ARO.

REFERENCES

1. A. Eisenberg, *War of the Mines: Cambodia and the Impoverishment of a Nation*, Pluto Press, London, 1994.
2. J. T. Johnson, J. Jenwatanavet, N. Wang, R. Caldecott, and J. Young, "Waveguide studies of soil modification techniques for enhanced mine detection with ground penetrating radar," *Detection and Remediation Technologies for Mines and Minelike Targets III, A. C. Dubey, J. F. Harvey, J. T. Broach, Editors, Proceedings of SPIE* **vol. 3392, part II**, pp. 716–724, 1998.
3. J. Jenwatanavet, J. T. Johnson, N. Wang, and R. Caldecott, "Development of soil modification techniques for enhanced detection/discrimination," *IGARSS '98 conference proceedings* **1**, pp. 210–212, 1998.
4. H. Frohlich, *Theory of dielectrics; dielectric constant and dielectric loss*, Oxford University Press, New York, 1958.
5. S. Nag, J. L. Peters, I. J. Gupta, and C.-C. Chen, "Ramp response for the detection of anti-personnel mines," *Detection and Remediation Technologies for Mines and Minelike Targets III, A. C. Dubey, J. F. Harvey, J. T. Broach, Editors, Proceedings of SPIE* [**3710-137**], 1999.
6. A. V. der Merwe, I. J. Gupta, and J. L. Peters, "A clutter reduction technique for gpr data from mine like targets," *Detection and Remediation Technologies for Mines and Minelike Targets III, A. C. Dubey, J. F. Harvey, J. T. Broach, Editors, Proceedings of SPIE* [**3710-106**], 1999.
7. Cryomech, "Liquid nitrogen plants," *http://www.cryomech.com/lnp.html* .
8. Teragon, "Liquid nitrogen controller," *http://pw2.netcom.com/ loulom/index.html* .

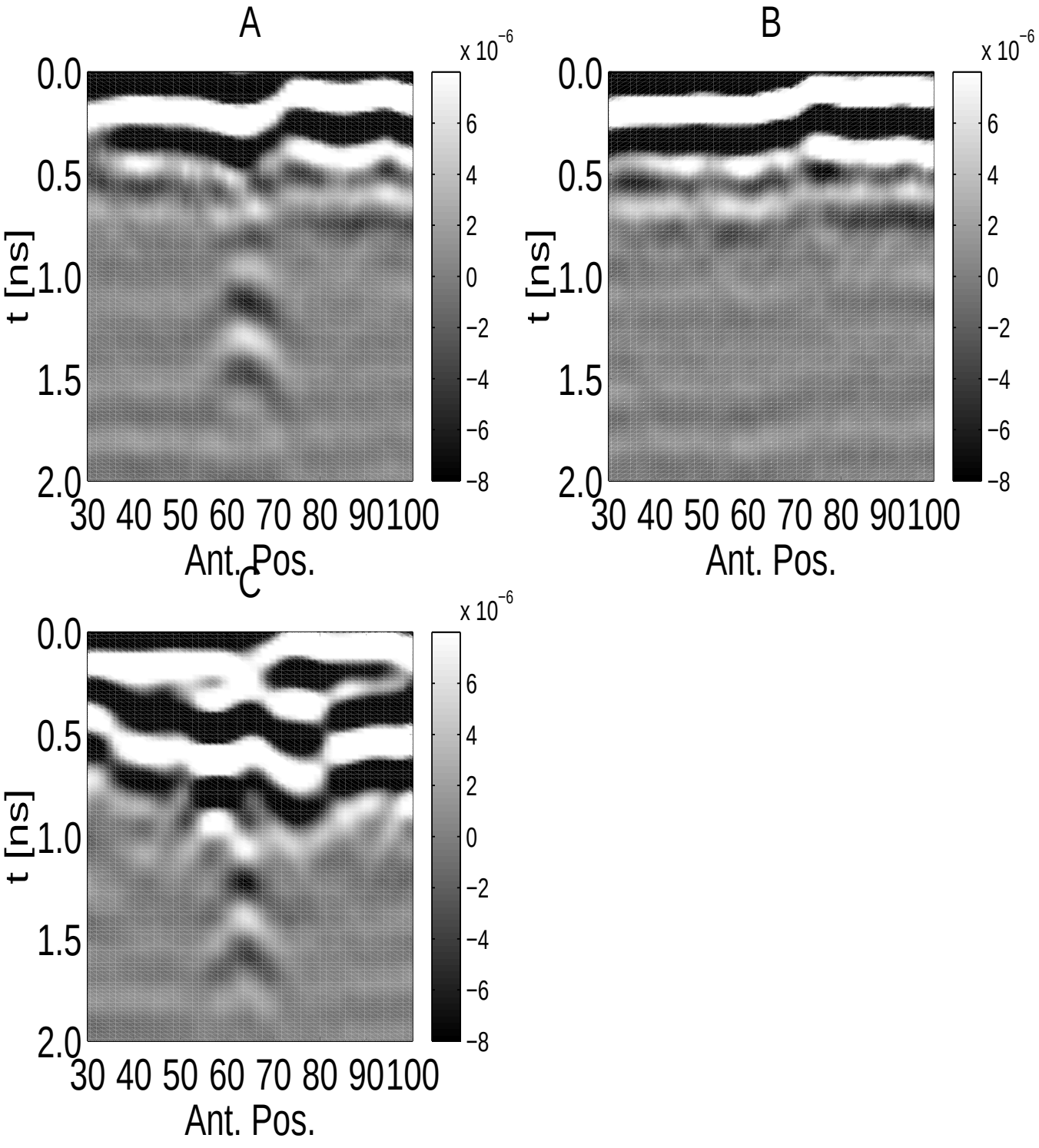


Figure 7. Image results of time domain measured fields for dielectric rod antenna with 6.35 cm depth metal target (a) Target in dry soil (b) After the addition of 2 gallons of water (c) After the addition of 15 liters of liquid nitrogen

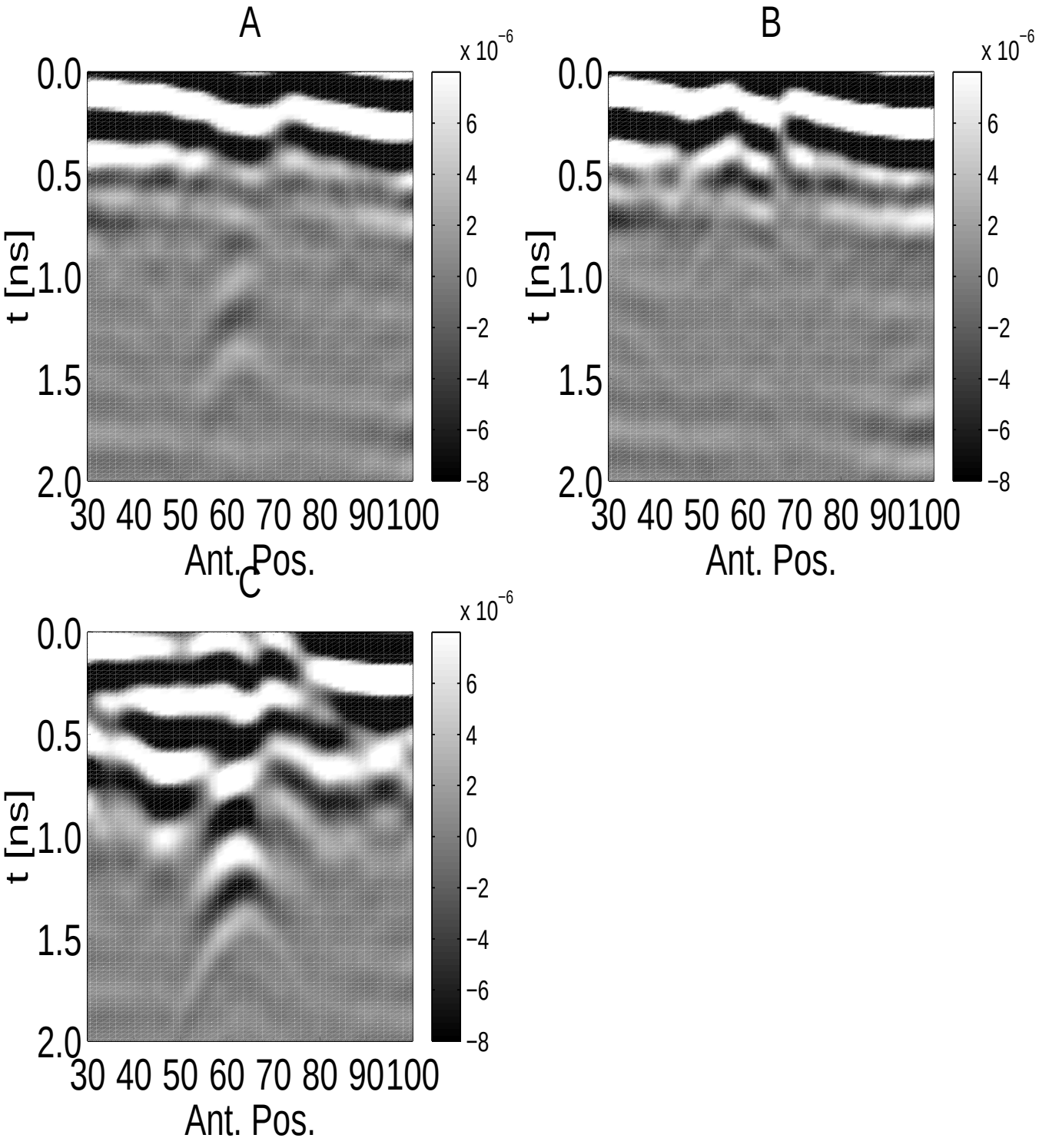


Figure 8. Image results of time domain measured fields for dielectric rod antenna with 6.35 cm depth nylon target (a) Target in dry soil (b) After the addition of 2 gallons of water (c) After the addition of 15 liters of liquid nitrogen

A Comparison of Optimal and Suboptimal Processors for Classification of Buried Metal Objects

Ping Gao, *Student Member, IEEE*, and Leslie Collins, *Member, IEEE*

***Abstract*—Classification of metal objects is important for landmine and unexploded ordnance applications. Previously, we have investigated optimal classification of landmine-like metal objects using wideband frequency-domain electromagnetic induction data [1]. Here, a suboptimal processor, which is computationally less burdensome than the optimal processor, is discussed. The data is first normalized, exploiting the fact that the level of the response changes significantly while the structure of the magnitude of the response changes only slightly as the target/sensor orientation changes for the class of objects considered. Results indicate that the suboptimal processor performance approaches that of the optimal classifier on normalized data. Thus, normalization mitigates the uncertainty resulting from the target/sensor orientation.**

***Index Terms*—Bayes procedures, data processing, electromagnetic induction, signal classification.**

I. Introduction

TIME-DOMAIN electromagnetic induction (EMI) sensors have been used extensively for landmine and unexploded ordnance (UXO) detection. In order to discriminate targets of interest from other pieces of metal, several modifications to traditional EMI sensors have been considered [2]–[6]. One promising approach is to operate the EMI sensor in the frequency-domain utilizing wideband excitation. The frequency dependence of the induced fields excited by buried conducting targets can then be exploited by a detector.

We previously considered classification of various metal objects using wideband frequency-domain EMI data [1]. The development of the optimal processor integrates a numerical model of wideband EMI responses [2] and a Bayesian decision-theoretic approach. It is of importance in this approach to accurately model the wideband EMI response and to consider uncertainties regarding the target/sensor orientation. We included appropriate treatment of such uncertainty, resulting in significant performance improvement over algorithms in which such information was ignored.

The drawback to the optimal processor is the computational burden. In some scenarios, such as UXO detection and classification where data processing can be performed offline, this computational complexity is not an issue. However, in humanitarian and military demining, real-time decision making is required, thus computational complexity becomes problematic.

Therefore, the goal of this work is to investigate suboptimal classifiers that continue to incorporate the physical nature of the wideband frequency-domain EMI signal, but are less computationally burdensome. It is shown on both simulated and experimental data that by normalizing the sensor output, a suboptimal processor operating on the *normalized* data can achieve nearly the same, or in some cases, better classification performance than that of the optimal processor operating on *unnormalized* data.

II. Approach

A model-based Bayesian decision-theoretic approach was investigated to discriminate man-made metal targets under conditions where the target/sensor orientation is unknown, since the exact sensor position (where measurements are obtained) relative to the underground objects is unknown in practice [1]. When the exact dimensions, constitutive parameters of the target, and the horizontal and vertical distance from the center of the sensor to that of the target are specified, the theoretical wideband EMI response can be calculated. In order to model the wideband EMI signature of these targets, a method of moment (MoM) analysis was used to predict the theoretical response (in-phase and quadrature voltages) from the targets at the frequencies of interest. To test the effectiveness of the approach, data was collected from a prototype wideband frequency-domain EMI sensor, the GEM-3 [7], developed by Geophex, Ltd. The discrimination algorithm was applied to both simulated data and experimental data measured using the sensor.

Five metal targets were used for the simulations: an aluminum cone, an aluminum barbell, an aluminum disk, a thick brass disk, and a thin brass disk. This target set was selected because of the similarities in shape of three of the targets to antitank mines and two of the targets to clutter. Twenty-one frequencies, ranging from 4 to 24 kHz in a 1-kHz spacing, were used for both simulations and measurements. It is assumed that the sensor is subject to a small amount of additive Gaussian noise, which was verified during the experimental data acquisition. Let H_i represent the hypothesis that the ith target is present, where $i = 1, \cdots, K$. The received data from the ith target can be expressed as $r_{ij} = A_{ij} + n_j$, where j corresponds to frequency, r_{ij} is the received data of the ith target from the sensor, A_{ij} is the model prediction for the ith target at the jth frequency when the target is at a known position, and n_j is Gaussian noise with zero mean and variance of $\sigma_{n_j}^2$. Let q_i represent the *a priori* probability that

Manuscript received March 4, 1999. This work was sponsored by the Army Research Office under Grant DAAH04-96-1-448 (Demining MURI). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Y. Hua.

The authors are with the Electrical and Computer Engineering Department, Duke University, Durham, NC 27708-0291 USA (e-mail: lcollins@ee.duke.edu).

Publisher Item Identifier S 1070-9908(99)05791-0.

1070–9908/99$10.00 © 1999 IEEE

hypothesis H_i is true, and assume that the cost of a correct decision is zero, and the cost of a wrong decision equals one. The optimal solution for this classification problem [8] is to decide that H_i is true if

$$\frac{p(H_i|r)}{p(H_k|r)} = \frac{q_i p(r|H_i)}{q_k p(r|H_k)} > 1 \quad (1)$$

is satisfied for any $k \neq i$, where $p(r|H_i)$ is the probability density or likelihood function of data r given $H_i, p(H_i|r)$ is the *a posteriori* distribution or discriminant function [9], and r is the in-phase and quadrature (assumed independent) data from the sensor at a known position. A uniform prior on q_i is assumed, therefore, $q_i = 1/K$. When data r is received, we decide in favor of hypothesis H_i where $q_i p(r|H_i) = \max_k \{q_k p(r|H_k)\}, k = 1, \cdots, K$. Thus, we decide in favor of a hypothesis that has the largest *a posteriori* probability among the K probability density functions. After simplifying, the alternative discriminant function [9] is

$$\log\, p'(r|H_i) = -(r - A_i)^T\, \Sigma^{-1} (r - A_i) \quad (2)$$

where A_i is the model prediction of the response of the ith target, and Σ is the covariance matrix of r.

The discriminant function in (2) is optimal only when a target is at a known position, when all the parameters are known, and when the sensor is subject to Gaussian noise. It can be implemented as a bank of "matched" filters. The noise is not identically distributed, and the signals are not of equal energy; thus, these two facts result in a formulation which is similar to, but not identical to, the traditional matched filter $r^T A_i$.

A more accurate assumption in practice is that the height and horizontal position between the center of the target and that of the sensor are uncertain. In this case, the matched filterbank is not the optimal solution. The optimal alternative discriminant is

$$p(r|H_i) = \iiint p(r|H_i, h, x, y) p(h) p(x, y)\, dh\, dx\, dy \quad (3)$$

where h represents the height of the sensor from the target; x, y represent the horizontal position of the center of the target relative to the sensor; and $p(h)$ and $p(x, y)$ are the *a priori* distributions of the position variables, which are assumed independent.

Four scenarios were considered in simulations:

1) target at a known, fixed position;
2) $h \sim N(20$ cm, 1.532 cm$^2)$, and $x, y = 0$;
3) x, y is uniformly distributed in a 20 $\times$ 20 cm square, and h is fixed;
4) both height and horizontal position are uncertain, following the same distributions as above.

The priors placed on the uncertain positional parameters are consistent with those observed in demining applications [10].

For each target, measurements were taken at 164 positions using the GEM-3. The choices of these locations follow a Gaussian distribution for height and a uniform distribution for horizontal position as in the simulations.

III. Alternative Processor Design

The wideband EMI responses for the same target at different heights and horizontal positions show a somewhat "parallel" structure as a function of frequency [11]. The overall level varies substantially as the target/sensor orientation changes, but the basic structure of the response as a function of frequency changes only slightly. This is because the transfer function of a nonferrous target can be expressed as $H(\omega) = \Sigma_n\, [a_n \omega [\omega - j\omega_n]$, where ω is radian frequency, ω_n is the nth resonant frequency, and a_n is the coefficient corresponding to each resonant frequency. When the target aspect ratios (length versus diameter) are approximately equal, and only the first mode dominates, the frequency response scales directly with variable orientations [11]. Based on this phenomenon, we hypothesized that if the wideband EMI signatures were normalized, performance could be evaluated using a processor that avoids calculating the integration over position uncertainty [as in (3)]. Then, the processor can be implemented as previously described for the "matched filter" case after normalization. Therefore, the outputs of the sensor are normalized so that the energy in the response is equal to unity. This operation essentially decreases the uncertainties of the data collected in an uncertain environment. The suboptimal processor takes the form of (2), which uses the mean signature after normalization over all uncertainties as A_i for the ith target.

The performance of a matched-filter-like processor cannot exceed the performance of an optimal classifier operating on the same data set. However, by using the suboptimal processor, computational savings are proportional to that required to evaluate the integral shown in (3), thus can be implemented in real time.

IV. Results

Table I illustrates the performance of the various processors operating on simulated data before and after normalization for the five metal objects. For the fixed position case, the matched-filter-like processor is optimal. The results indicate that when there is no uncertainty in the target/sensor orientation (fixed position), performing the normalization degrades the performance of the processor. However, when the target positions are uncertain, better performance can be achieved if the output is normalized. Also, the matched filter performance is quite close to the optimal classifier performance for the normalized data. This indicates that applying the sub-optimal processor to normalized data only slightly sacrifices performance, but computational time is reduced significantly. Interestingly, the optimal classifier operating on the normalized data outperforms the optimal classifier operating on the unnormalized data. This occurs because the normalization uncorrelates the signals to a small extent, which improves classification performance [12]. For instance, the average correlation coefficient decrease before and after normalization between the target 3 and 5 is 0.711×10^{-4}.

Table II illustrates the performance of various processors using measured data. Data was gathered from four out of five targets used in simulations. The suboptimal processor using normalized data also performs better in general than the processor using raw sensor output, and the optimal processor after normalization achieves better performance than the processor without normalization. This performance improvement validates the hypothesis that by normalizing the sensor data, the uncertainties associated with the sensor/target position can

TABLE I

PERFORMANCE OF THE VARIOUS PROCESSORS BEFORE AND AFTER NORMALIZATION USING SIMULATED DATA WHEN 1) TARGET POSITION IS KNOWN AND FIXED, 2) $h \sim N(20 \text{ cm}, 1.532 \text{ cm}^2)$ AND x, y IS FIXED, 3) HORIZONTAL POSITION IS UNIFORMLY DISTRIBUTED IN A 20 cm $\times$ 20 cm SQUARE AND HEIGHT IS FIXED, AND 4) BOTH HEIGHT AND HORIZONTAL POSITION ARE RANDOM, FOLLOWING THE DISTRIBUTIONS MENTIONED ABOVE. "NO NORM": PROCESSING OCCURRED WITHOUT NORMALIZATION. "NORM": PROCESSING FOLLOWED NORMALIZATION. "OPT": OPTIMAL CLASSIFIER. "MF": MATCHED-FILTER-LIKE PROCESSOR

			Probability of Correct Classification				
			Target 1	Target 2	Target 3	Target 4	Target 5
Fixed Position	No Norm.	MF/Opt.	1	1	1	1	1
	Norm.	MF/Opt.	1	.976	1	.976	1
Height Uncertain	No Norm.	MF	.832	.351	.518	.081	.620
		Opt.	.999	.857	.997	.830	1
	Norm.	MF	1	.973	1	.972	1
		Opt.	1	.977	1	.978	1
Horizontal Position Uncertain	No Norm.	MF	.804	.411	.573	.085	.655
		Opt.	1	.828	1	.810	.999
	Norm.	MF	1	.849	1	.754	1
		Opt.	1	.866	1	.810	1
Both Height & Horiz. Position Uncertain	No Norm.	MF	.758	.340	.506	.073	.492
		Opt.	.994	.701	.990	.700	.996
	Norm.	MF	1	.848	1	.759	1
		Opt.	1	.879	1	.823	1

TABLE II

PERFORMANCE OF THE SUBOPTIMAL AND OPTIMAL PROCESSORS BEFORE AND AFTER NORMALIZATION EVALUATED ON MEASURED DATA. "MF": MATCHED-FILTER-LIKE OR SUBOPTIMAL PROCESSOR. "NO NORM": PROCESSING OCCURRED WITHOUT NORMALIZATION. "OPT PROC": OPTIMAL PROCESSOR. "NORM": PROCESSING FOLLOWED NORMALIZATION

	Target 2	Target 3	Target 4	Target 5
Opt. Proc., No Norm.	0.91	0.77	0.69	0.81
Opt. Proc., Norm.	0	0.98	1	0.99
MF, No Norm.	0.23	0.24	0.03	0.14
MF, Norm.	1	0.99	0.89	0.83

be decreased. Also, the performance of a matched-filter-like processor operating on normalized data is very close to that of the optimal processor operating on normalized data. However, the computational load is reduced significantly.

V. DISCUSSION

Using wideband frequency-domain EMI data, metal objects of different dimensions and materials can be discriminated via classification algorithms carefully designed using signal detection theory. Since the uncertainty inherent in the sensor output is not only due to the additive noise but also the unknown target/sensor orientation, both facts should be considered during the development of the processor. The optimal classifier, which integrates over the uncertainty in the sensor/target orientation, can provide dramatic performance improvement over a matched-filter-like processor; however, it suffers from a computational burden and is not always practical. For the class of targets considered (target aspect ratio close to one), normalizing the wideband EMI responses, which exploits the physical nature of the signals, mitigates the uncertainties of the response due to the unknown target position. By performing the normalization, a suboptimal processor can provide significant performance improvement over suboptimal processors that operate on raw sensor output, and computational time can be dramatically decreased. Since for general targets the frequency-domain EMI response does not simply scale while the target/sensor orientation changes, we will investigate the relationship of the longitudinal and azimuthal modes of the response and integrate this information into the classifier design.

ACKNOWLEDGMENT

The authors would like to thank N. Geng, L. Carin, L. Nolte, S. Tantum, D. Keiswetter, I. J. Won, and R. Weaver for their generous help regarding this work.

REFERENCES

[1] P. Gao *et al.*, "Classification of landmine-like metal targets using wideband electromagnetic induction," submitted for publication.
[2] N. Geng *et al.*, "Wideband electromagnetic induction for metal-target identification: theory, measurement and signal processing," Tech. Rep., Duke Univ., Durham, NC, Sept. 1997.
[3] L. Collins, P. Gao, and L. Carin, "An improved Bayesian decision theoretic approach for land mine detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 811–819, Mar. 1998.
[4] C. E. Baum, Low frequency near-field magnetic scattering from highly, but not perfectly conducting bodies," Phillips Lab. Interaction Note 499, Nov. 1993.
[5] G. D. Sower and S. P. Cave, "Detection and identification of mines from natural magnetic and electromagnetic resonances," *Proc. SPIE*, 1995, FL.
[6] S. Vitebskiy and L. Carin, "Late-time resonant frequencies of buried bodies of revolution," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1575–1583, 1996.
[7] I. J. Won, D. A. Keiswetter, and D. R. Hansen, "GEM-3: A monostatic broadband electromagnetic induction sensor," *J. Environ. Eng. Geophys.*, vol. 2, pp. 53–64, Aug. 1997.
[8] J. C. Hancock and P. A. Wintz, *Signal Detection Theory*. New York: McGraw Hill, 1966, pp. 80–84.
[9] R. J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*. New York: Wiley, 1992, pp. 34–43.
[10] "Mine countermine operations," Army FM 20–32, May 1998.
[11] L. Carin, Wideband time- and frequency-domain EMI: phenomenology and signal processing," Tech. Rep., U.S. Air Force Philip's Lab. Project, Oct. 1998.
[12] J. G. Rasimas, "Signal detection and classification of targets using multiple aspect angles," M.S. thesis, Duke Univ., Durham, NC, 1998.

A Finite-Difference Model to Study the Elastic-Wave Interactions with Buried Land Mines

Christoph T. Schröder and Waymond R. Scott, Jr.

***Abstract*—A two-dimensional (2-D) finite-difference model for elastic waves in the ground has been developed. The model uses the equation of motion and the stress-strain relation, from which a first-order stress-velocity formulation is obtained. The resulting system of equations is discretized using centered finite-differences. A perfectly matched layer surrounds the discretized solution space and absorbs the outward traveling waves. The numerical model is validated by comparison to an analytical solution. The numerical model is used to study the interaction of elastic waves with a buried land mine. It is seen that the presence of an air-chamber within the mine gives rise to resonant oscillations that are clearly visible on the surface above the mine. The resonance is shown to be due to flexural waves being trapped within the thin layer between the surface of the ground and the air chamber of the mine. The numerical results are in good qualitative agreement with experimental observations.**

***Index Terms*—Acoustic, elastic wave, FDTD, finite-difference, land mine.**

I. Introduction

In a new technique, elastic and electromagnetic waves are used in a synergistic manner to detect buried land mines [1], [2]. Within this technique, elastic waves interact with a mine and cause the ground above the mine to vibrate. A radar detects these vibrations and, thus, the mine. To explore the mine–wave interactions, a numerical finite-difference model for elastic waves in the ground has been developed.

The equation of motion and the stress–strain relation, together with a constitutive relation, form a set of first-order partial differential equations that completely describe the elastic wave motion in a medium. Introducing finite differences, this set of equations can be discretized and adapted to the finite-difference time-domain modeling scheme, obtaining a second-order accurate stress-velocity formulation. Assuming that the field is known at one initial time t_0, this numerical scheme provides the field values at any later time $t > t_0$. The finite-difference model has been implemented in two dimensions. The solution space is discretized and a staggered finite-difference grid is introduced. The grid is surrounded by a perfectly matched layer, that absorbs the outward traveling waves. The numerical model has been validated by comparison to an analytical solution. The analytical solution for elastic waves in a homogeneous semi-infinite half-space is obtained in the form of an integral equation.

Manuscript received September 4, 1999; revised February 16, 2000. This work was supported in part by the OSD MURI program, U.S. Army Research Office, Washington, DC, Contract DAAH04-96-1-0448.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: christoph.schroeder@ee.gatech.edu).

Publisher Item Identifier S 0196-2892(00)05890-3.

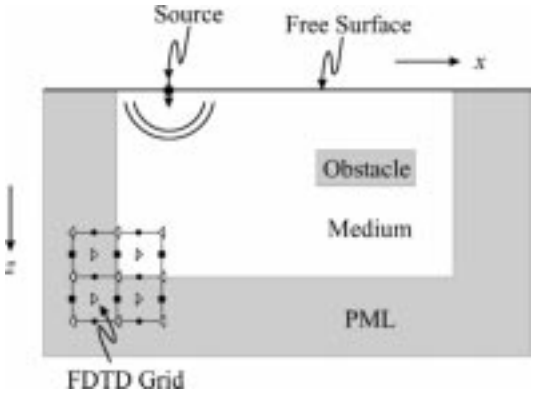


Fig. 1. Finite-difference model.

Excellent agreement is seen between the numerical model and the analytical solution.

By using the numerical model, the interaction of elastic waves with a buried antipersonnel mine is explored. Results for two simple models of a TS-50 antipersonnel mine are presented in this work: one containing an air-filled chamber and one without an air-filled chamber. The air-filled chamber is introduced to approximate the complex structure of a real mine, which contains explosives as well as chambers bearing the triggering mechanism etc. The results obtained with the numerical model are in remarkably good agreement with experimental observations even though the model is 2-D, whereas the experimental set-up is 3-D. In both experiments and numerical simulations, a strong resonance is observed at the mine. This resonance is shown to be due to flexural waves being trapped within the thin layer between the surface of the ground and the air chamber of the mine. In the experiments, the Rayleigh surface waves are seen to disperse while traveling along the surface. Assuming a depth-dependent shear wave speed, this effect can be also modeled numerically.

II. Two-Dimensional (2-D) Numerical Model

A. Finite-Difference Model

Fig. 1 shows the 2-D finite-difference model. The ground is modeled as an isotropic, lossless semi-infinite half-space. A normal point source is located on the free surface, exciting longitudinal (pressure) and transverse (shear) waves in the x-z-plane. The wave fields are invariant in the y direction. At $z = 0$, a free surface bounds the solution space. A perfectly matched layer (PML) terminates the solution space at the remaining edges, absorbing all outward traveling waves. The space is discretized using a staggered finite-difference grid.

0196–2892/00$10.00 © 2000 IEEE

A first-order velocity–stress formulation is used for the finite-difference model [3]–[5]. Here, the equation of motion and the stress–strain relation are discretized, leading to a system of first-order partial differential equations. Since only field components in the x-z-plane are excited, the only nonzero field components are the three unknown stress components τ_{xx}, τ_{xz}, τ_{zz}, and the two unknown particle velocities v_x and v_z. Due to the invariance in the y-direction, all derivatives with respect to y vanish. The wave motion is then completely described by a system of the following five partial differential equations:

$$\rho\frac{\partial v_x}{\partial t}=\frac{\partial \tau_{xx}}{\partial x}+\frac{\partial \tau_{xz}}{\partial z} \tag{1}$$

$$\rho\frac{\partial v_z}{\partial t}=\frac{\partial \tau_{xz}}{\partial x}+\frac{\partial \tau_{zz}}{\partial z} \tag{2}$$

$$\frac{\partial \tau_{xx}}{\partial t}=(\lambda+2\mu)\frac{\partial v_x}{\partial x}+\lambda\frac{\partial v_z}{\partial z} \tag{3}$$

$$\frac{\partial \tau_{zz}}{\partial t}=(\lambda+2\mu)\frac{\partial v_z}{\partial z}+\lambda\frac{\partial v_x}{\partial x} \tag{4}$$

$$\frac{\partial \tau_{xz}}{\partial t}=\mu\left(\frac{\partial v_x}{\partial z}+\frac{\partial v_z}{\partial x}\right) \tag{5}$$

where ρ is the material density, and λ and μ are Lame's constants. These equations are discretized using centered finite differences. By introducing the finite differences in space Δx and Δz and in time Δt, (1) and (3), for example, can be discretized as

$$\begin{aligned}\rho\frac{V_x^{k+0.5}|_{i,j-0.5}-V_x^{k-0.5}|_{i,j-0.5}}{\Delta t}&=\frac{T_{xx}^{k}|_{i+0.5,j-0.5}-T_{xx}^{k}|_{i-0.5,j-0.5}}{\Delta x}\\&\quad+\frac{T_{xz}^{k}|_{i,j}-T_{xz}^{k}|_{i,j-1}}{\Delta z}\end{aligned} \tag{6}$$

and

$$\begin{aligned}\frac{T_{xx}^{k+1}|_{i+0.5,j-0.5}-T_{xx}^{k}|_{i+0.5,j-0.5}}{\Delta t}&=(\lambda+2\mu)\frac{V_x^{k+0.5}|_{i+1,j-0.5}-V_x^{k+0.5}|_{i,j-0.5}}{\Delta x}\\&\quad+\lambda\frac{V_z^{k+0.5}|_{i+0.5,j}-V_z^{k+0.5}|_{i+0.5,j-1}}{\Delta z}.\end{aligned} \tag{7}$$

Here, the capital letters mark the numerical value of the correspondent field component at a discrete location in space and time. For example, $V_x^{k+0.5}|_{i,j-0.5}$ stands for the numerical value of the particle velocity v_x at the position $(x,z)=(i\Delta x,(j-0.5)\Delta z)$ at the time $t=(k+0.5)\Delta t$. Knowing $V_x^{k-0.5}$, T_{xx}^{k}, and T_{xz}^{k}, (6) can be solved for $V_x^{k+0.5}$, i.e. at the incremental time $t=(k+0.5)\Delta t$

$$\begin{aligned}V_x^{k+0.5}|_{i,j-0.5}&=V_x^{k-0.5}|_{i,j-0.5}\\&\quad+\frac{\Delta t}{\rho\Delta x}(T_{xx}^{k}|_{i+0.5,j-0.5}-T_{xx}^{k}|_{i-0.5,j-0.5})\\&\quad+\frac{\Delta t}{\rho\Delta z}(T_{xz}^{k}|_{i,j}-T_{xz}^{k}|_{i,j-1}).\end{aligned} \tag{8}$$

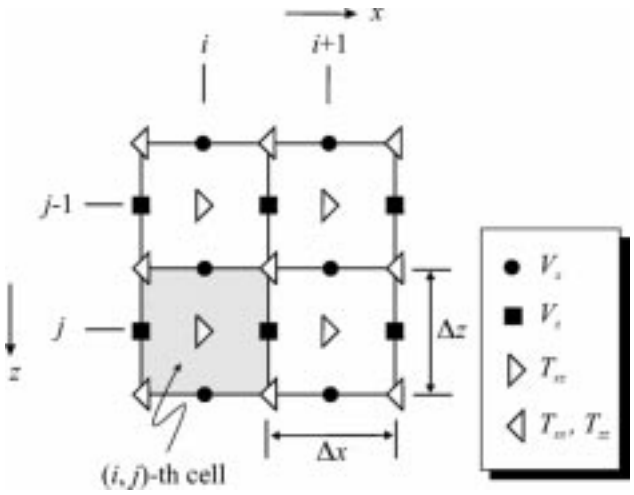


Fig. 2. Portion of the finite-difference grid. The field components are not known at the same locations but are offset by $\Delta x/2$ and $\Delta z/2$.

Similarly, T_{xx}^{k+1} is obtained from (7)

$$\begin{aligned}T_{xx}^{k+1}|_{i+0.5,j-0.5}&=T_{xx}^{k}|_{i+0.5,j-0.5}+(\lambda+2\mu)\frac{\Delta t}{\Delta x}\\&\quad\cdot(V_x^{k+0.5}|_{i+1,j-0.5}-V_x^{k+0.5}|_{i,j-0.5})\\&\quad+\lambda\frac{\Delta t}{\Delta z}(V_z^{k+0.5}|_{i+0.5,j}-V_z^{k+0.5}|_{i+0.5,j-1}).\end{aligned} \tag{9}$$

In the same manner, discretized equations can be obtained for all field components.

The discretization leads to the characteristic staggered finite-difference grid. Fig. 2 depicts the position of the field components in the finite-difference grid. In this grid, the velocity components and the stress components are not known at the same position in time and space but are offset by $\Delta t/2$ in time and by $\Delta x/2$ and $\Delta z/2$ in space. The offset in time leads to the so-called leapfrog scheme. In the leapfrog scheme, the field components are updated sequentially in time. The velocity components are calculated first, then the stress components from the velocity components, the velocity components again using the stress components, and so on. Thus, knowing the field components throughout the entire space at the time $t_0-0.5\Delta t$ and t_0, respectively, the field components can be determined for all later times $t>t_0$.

When implementing the finite-difference scheme, boundary conditions must be treated in a special manner. Three different kinds of boundaries arise: the source point, the internal boundaries (i.e., boundaries within the medium marked by a change in material properties), and the external boundaries (i.e., the grid edges).

A normal point source is implemented on the free surface. The normal stress component τ_{zz} is excited by a differentiated Gaussian pulse. An additive source is used. For each time step, the value of the excitation is added to the value calculated with the finite-difference scheme.

The conditions at internal boundaries (i.e., at the interfaces between different media) are usually satisfied implicitly. However, to ensure numerical stability, the material properties must be averaged for components on the boundary. While the material density ρ, appearing in the equation of motion, is averaged directly, the inverse of Lame's constants. For transitions between similar materials, the averaging may be omitted. However, it is necessary at an interface between media with greatly different material properties (for example, at an air–solid interface) in order to maintain stability [3].

Four external boundaries arise at the four outer grid edges. At its upper edge, the half-space is bounded by a free surface. Due to the continuity of normal stress, the normal stress components vanish at a free surface. In order to satisfy this condition, an extra row must be inserted into the finite-difference grid one step beyond the free-surface boundary. By using this extra row, all stress components are forced to zero on the free surface [3].

In order to model the semi-infinite half-space, all waves that are reaching the three remaining outer grid edges must be perfectly transmitted and absorbed. The boundary condition that does this most accurately is the PML boundary condition, first introduced by Berenger [6] and adapted to elastodynamics by Chew and Liu [7]. In continuous space, a nonphysical splitting of the wave fields allows the introduction of a lossy boundary layer that is perfectly matched to the solution space. It has been shown that an arbitrarily polarized wave incident on this PML is perfectly transmitted. The wave experiences the exact same phase velocity and wave impedance as in the solution space, while rapidly decaying along the axis normal to the PML/medium interface. However, in discrete space, the lossy layer will not be perfectly matched to the solution space, and slight reflections occur at the interface. To keep these reflections small, a tapered loss profile is chosen within the PML

$$\sigma(i) = \sigma_{\max} \cdot \left(\frac{i - i_{\mathrm{PML}}}{N_{\mathrm{PML}}} \right)^m \tag{10}$$

where $m \approx 2.1$ and $\sigma_{\max} = 0.1 N_{\mathrm{PML}}/\Delta t$; N_{PML} is the thickness of the PML in basis cells, and $i - i_{\mathrm{PML}}$ indicates the position within the PML. This loss profile (with a slightly different $\sigma_{\max}$) has been found to yield good performance in electromagnetic finite-difference modeling [8]. For this work, a thickness of ten cells has been found sufficient, yielding an attenuation of the incident waves of more than 70 dB.

B. Validation of Finite-Difference Results

An analytical solution for the particle displacement fields in a semi-infinite homogeneous half-space subjected to a normal harmonic line-load of finite width can be derived [9], [10]. Using a Fourier transform method, the particle displacement is obtained as an integral equation and can be computed using numerical integration [3].

The integral to be solved contains two poles and four branch cuts. Furthermore, the integration range is infinite and extends from $-\infty$ to ∞. While the branch cuts do not impose a problem during numerical integration, the poles cannot be integrated numerically. To account correctly for the pole contributions, the poles must be subtracted from the integral and integrated analytically. To accelerate the integration, the infinite integration range may be truncated. However, for this to be possible, the integrand has to converge sufficiently fast to zero. In order to accelerate the convergence of the integral, an asymptotic approximation for the tails of the integrals is derived. This asymptotic approximation is subtracted from the integrals and integrated analytically.

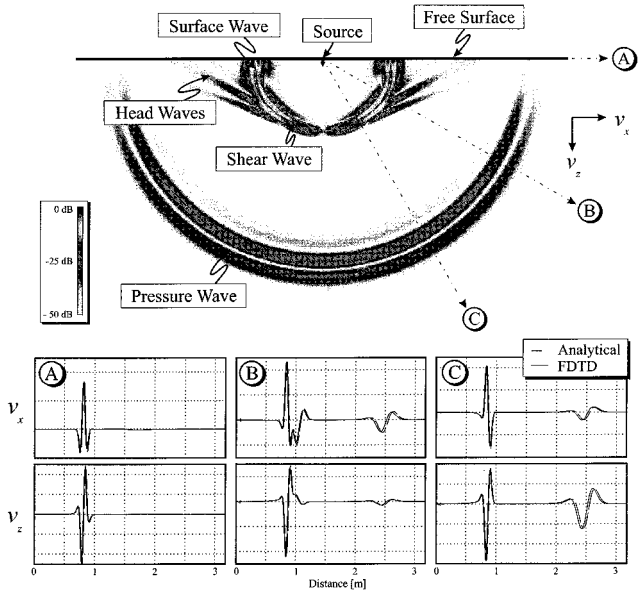


Fig. 3. Plot of the magnitude of the elastic waves in a half-space subjected to a normal line-load at $t = 10$ ms.

To obtain the particle displacement in the time-domain, the displacement field must be Fourier-transformed. Since the analytical solution is to be compared to the finite-difference result, where the particle velocity is determined, the particle displacement is differentiated with respect to time, and the particle velocity is obtained.

The integral is computed on a 600 × 600 grid with a grid spacing of 0.53 cm. Noting that the wave fields are symmetric to the vertical axis passing through the source, this yields an effective solution space of 6.37 m × 3.18 m. The finite-difference simulation is performed using a grid with 1239 × 620 nodes, incorporating a PML layer with a thickness of 20 cells. The finite-difference grid has the same dimensions and the same spacing as the grid used for the analytical result but does not exploit the symmetry and is consequently twice as big. The time step is chosen to be $\Delta t = 1.06 \cdot 10^{-5}$ s, and 1200 time steps are needed to calculate the wave fields up to $t = 10$ ms. The computations for the analytical solution took about 24 times longer than those for the numerical simulation.

In Fig. 3, the elastic waves due to a normal line-load on a free surface are shown. A differentiated Gaussian pulse with a center frequency of 400 Hz is launched at $t = 0$ s[1] from a source located on the free surface. In the upper half of Fig. 3, the magnitude of the particle velocity field at $t = 10$ ms is shown. A pressure, a shear, and a surface wave are seen to propagate. The pressure wave is the fastest of the three. The surface wave

[1]The center of the pulse is located at $t = 0$ s.

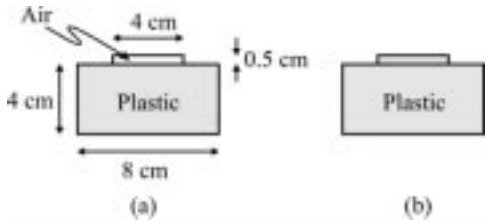


Fig. 4. Simple model of the TS-50 antipersonnel mine (a) with and (b) without an air-filled chamber.

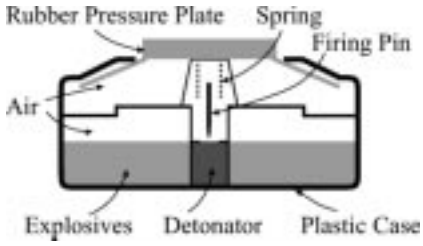


Fig. 5. Cross-sectional drawing of a TS-50 antipersonnel mine.

is propagating slightly slower than the shear wave and contains the most energy. A plane shear wave arises at the free surface, induced by the pressure wave. This wave is called a head wave. Head waves are downwardly directed shear waves generated by the passage of bulk waves along the free surface. Two more plane waves can be seen, which propagate at a steeper angle than the head wave induced by the pressure wave.

In Fig. 3, the particle velocity components v_x and v_z, according to the numerical and analytical solution, are plotted along the three section lines through the ground, denoted by A, B, and C. The agreement of numerical and analytical solutions is excellent. It can be seen that the finite-difference model predicts waves that travel slightly slower than those of the analytical solution. This is due to the well-known fact that the numerical phase velocity in the discrete finite-difference grid will come out to be slightly smaller than in continuous space [11].

The head waves obtained both with the finite-difference model and in the analytical solution have also been observed experimentally by the authors and others. In [12], a seismic field survey in a region with extremely high Poisson's ratio is described. In these measurements, head waves are detected that are very similar to the head waves observed here.

III. Interaction of Elastic Waves with a TS-50 Antipersonnel Mine

The interaction of elastic waves with a TS-50 antipersonnel mine, buried in sand, has been investigated. Two simple models for the TS-50 mine are used. In the first model [Fig. 4(a)], a small chamber filled with air is located on top of the mine's main chamber containing plastic explosives. The second model does not contain an air-filled chamber [Fig. 4(b)]. By including an air-filled chamber into the mine model, the effects of a real mine are approximated. A real mine is a complex mechanical structure with a flexible case, a trigger mechanism, air chambers, etc. Fig. 5 depicts a simplified cross sectional drawing of a real TS-50 antipersonnel mine. The elastic properties of the materials used for the numerical simulation are summarized in Table I. The parameters for the sand and for the plastic were measured experimentally by the authors, whereas the properties of the air are obtained from the literature.

A differentiated Gaussian pulse with a center frequency of 450 Hz is launched at $t = 0$ s from a source on the free surface. The left edge of the mine is located at a distance of 85 cm from the source. Its upper edge lies 1.5 cm beneath the surface of the ground. The space step within the finite-difference grid is set to $\Delta x = \Delta z = 0.5$ cm. The time step is chosen to be $\Delta t = \Delta x/(\sqrt{2}c_{p,\max}) = 1.309 \cdot 10^{-6}$ s and thus fulfills the

TABLE I
Parameters Used for Finite-Difference Simulation

Sand	Shear wave velocity	$c_{s,\text{sand}}$	87 m/s
	Pressure wave velocity	$c_{p,\text{sand}}$	250 m/s
	Material density	ρ_{sand}	1400 kg/m^3
Mine	Shear wave velocity	$c_{s,\text{mine}}$	1100 m/s
	Pressure wave velocity	$c_{p,\text{mine}}$	2700 m/s
	Material density	ρ_{mine}	1200 kg/m^3
Air	Shear wave velocity	$c_{s,\text{air}}$	0 m/s
	Pressure wave velocity	$c_{p,\text{air}}$	330 m/s
	Material density	ρ_{air}	1.3 kg/m^3

Courant condition. The elastic wave fields are computed on a grid containing 350 × 200 cells, including a PML with a thickness of ten cells and, consequently, yielding an effective solution space of 165 cm × 95 cm. The computation time was approximately 90 min for 22 000 time steps on a 450 MHz PC.

Fig. 6 shows the elastic wave fields on a cross section through the ground for (a) the model with the air-filled chamber and (b) the model without the air-filled chamber. The magnitude of the particle velocity is plotted on a logarithmic scale. The particle velocity field is shown at four different times: $t = 4$ ms, $t = 10$ ms, $t = 11$ ms, and $t = 16$ ms.

At $t = 4$ ms, a pressure wave (P), a shear wave (S), a Rayleigh surface wave (R), and head waves (H) are seen to propagate. The pressure wave, the fastest of the waves, just hits the mine. For the mine with the air-filled chamber, some energy is seen to be trapped between the surface and the mine, while for the mine without the air-filled chamber, no strong interaction occurs. At $t = 10$ ms and $t = 11$ ms, the surface wave (S) hits the mine and is partially transmitted (R) and partially reflected (rR). At $t = 16$ ms, the waves reflected from the mine are clearly seen [these are reflected Rayleigh waves (rR), a reflected shear wave (rS), and a reflected pressure wave (rP)]. For the mine without the air-filled chamber, no energy is seen to remain at the mine. However, for the mine with the air-filled chamber, energy is seen to be trapped above the mine and to be radiating waves, even though the incident wave has passed by.

In Fig. 7, waterfall graphs of the vertical particle velocity v_z at the surface are shown for (a) the mine with the air-filled chamber and (b) the mine without the air-filled chamber. Here, v_z is plotted as a function of time and vertically offset by the distance from the source. The slope of the traveling waves in the graph indicates the wave speed. Thus, by looking at their slope, the different waves can be distinguished. A Rayleigh surface wave (R) and a pressure wave (P) are seen to be incident

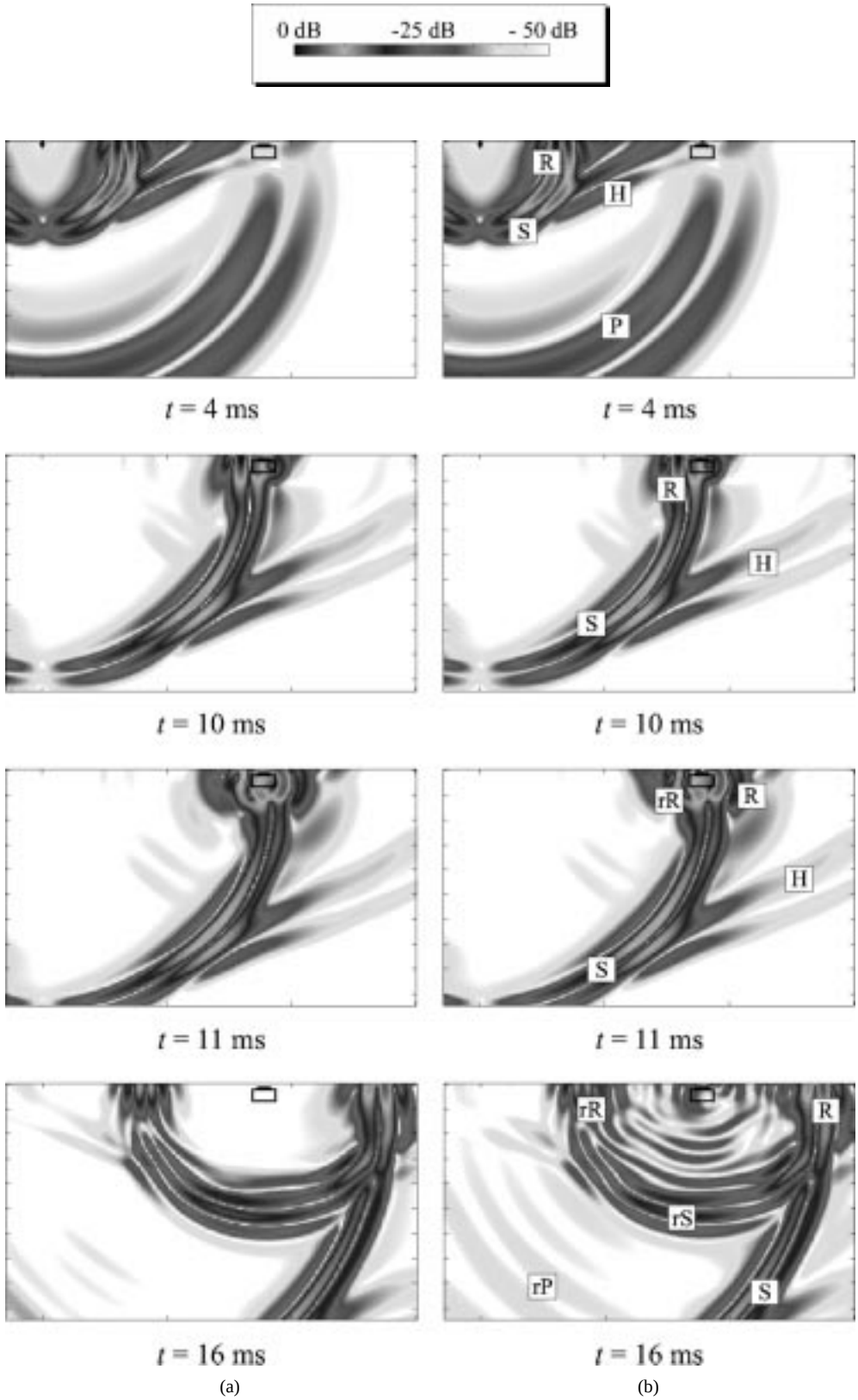
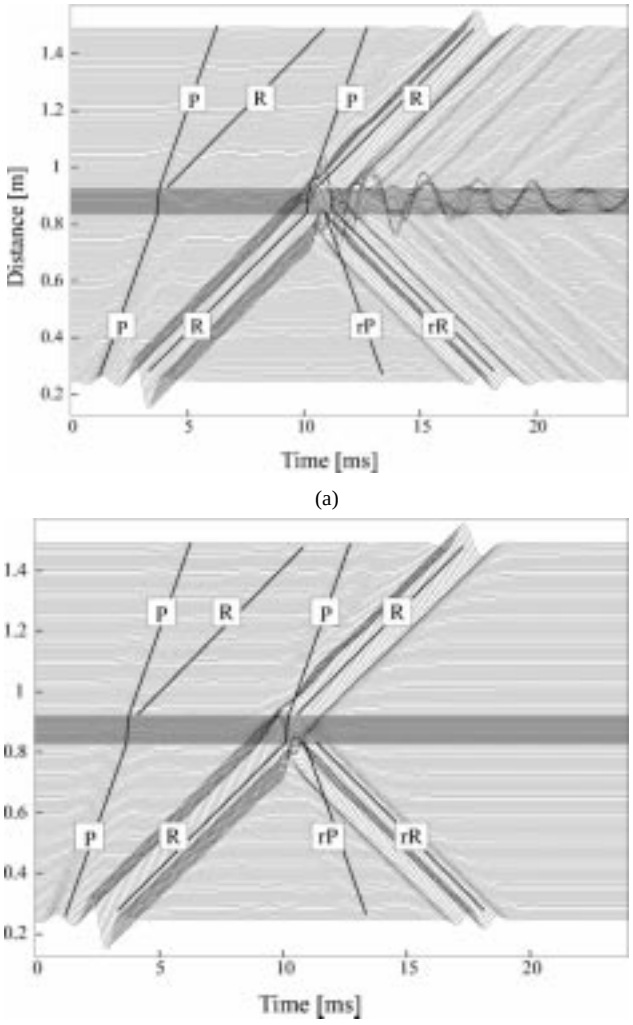


Fig. 6. Interaction of elastic waves with a buried antipersonnel mine and pseudo-color plots of the magnitude of the particle velocity on a cross section through the ground. (a) Mine with air-filled chamber. (b) Mine without air-filled chamber. R: Rayleigh wave; S: Shear wave; P: pressure wave; H: Head wave; rR: reflected Rayleigh wave; rS: reflected Shear wave; rP: reflected pressure wave. The arrow in the upper two plots denotes the source location.

onto the mine. The pressure wave is reflected and transmitted by the mine. It converts into a reflected pressure wave, a reflected surface wave, a transmitted pressure wave (P), and a transmitted surface wave (R). These waves are weak, due to the limited energy content of the pressure wave near the surface. The pressure wave is seen to travel faster over the mine than in the sand.



(a)

(b)

Fig. 7. Interaction of elastic waves with a buried antipersonnel mine and waterfall graph of the vertical particle velocity at the surface. (a) Mine with air-filled chamber and (b) mine without air-filled chamber. R: Rayleigh wave; P: pressure wave; rR: reflected Rayleigh wave; rP: reflected pressure wave. The gray box indicates the mine location.

This is due to the higher wave speed in the mine. The incident surface wave gives rise to a reflected pressure wave (rP), a reflected surface wave (rR), a transmitted pressure wave (P), and a transmitted surface wave (R). For the mine with the air-filled chamber [Fig. 7(a)], a strong resonance can be seen at the mine location. The resonance remains at the mine even after the incident surface wave has passed by, and causes the mine to radiate. This resonance is due to energy being trapped between the mine and the surface. The nature of this resonance will be explored further in Section III-B. For the mine without the air-filled chamber [Fig. 7(b)], no resonance occurs.

A. Comparison to Experimental Results

The results obtained with the numerical model are in fairly good agreement with experimental results [2]. In the experimental setup, an inert TS-50 antipersonnel mine (see Fig. 5) is buried in sand. A differentiated Gaussian pulse with a center frequency of 450 Hz is launched by an electrodynamic transducer placed on the surface of the ground, 85 cm away from

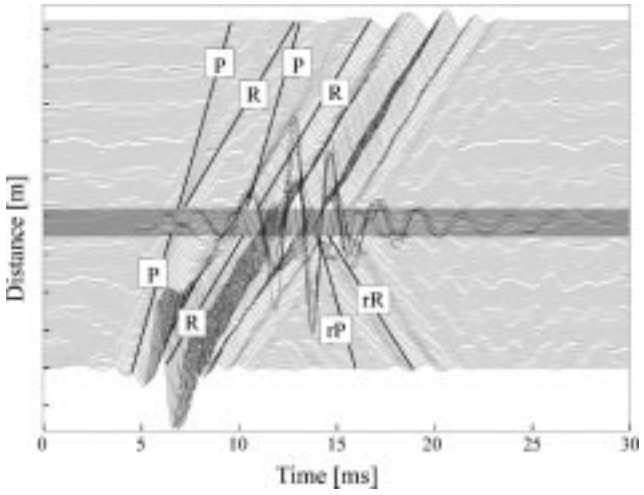


Fig. 8. Interaction of elastic waves with a buried antipersonnel mine (experimental results) and waterfall graph of the vertical particle displacement at the surface. R: Rayleigh wave; P: pressure wave; rR: reflected Rayleigh wave; rP: reflected pressure wave. The gray box indicates the mine location.

the mine. Above the surface, a radar that detects the vibrations of the surface is mounted. Fig. 8 shows a waterfall graph of the vertical particle displacement[2] on the surface as obtained in measurements with an inert TS-50 antipersonnel mine. The mine is buried 1.3 cm beneath the surface. The shear wave speed and the pressure wave speed were measured to be 87 m/s and 250 m/s, respectively. A pressure wave (P) and a Rayleigh wave (R) are seen to propagate and to interact with the buried mine. A resonance very similar to the one obtained in the numerical simulation of the mine with the air-filled chamber is observed. In the experimental model, the resonance appears stronger than in the numerical model, whereas in the numerical model, the reflections from the mine are much stronger. These differences are mainly due to the numerical model being 2-D, whereas the experimental model is 3-D. Furthermore, the model for the mine used in the numerical simulation is very simple and approximates the complex structure of a real land mine only very coarsely.

In the experiment, the Rayleigh surface wave is seen to be dispersed as it travels along the surface. This is believed to be due to the shear wave speed varying with depth. Due to increased cohesion of the sand, the sand becomes stiffer if subjected to increased pressure and thus, the shear wave speed increases with depth [13]. Empirically, the depth dependence of the sand has been estimated to be

$$c_s = [70 + 88 \cdot z^{0.3}] \text{ m/s} \tag{11}$$

where z is the depth in meters. Using this depth dependence within the numerical model, the Rayleigh wave disperses in a very similar manner as observed experimentally (Fig. 9).

B. Resonance

Both in the experiment and in the numerical simulation, a strong resonance is seen at the mine location. Though the nu-

[2]Note that experimentally, the particle displacement is measured, whereas the particle velocity is determined in the numerical model. Since the particle velocity is just the time derivative of the particle displacement, both wave fields bear the same characteristic behavior.

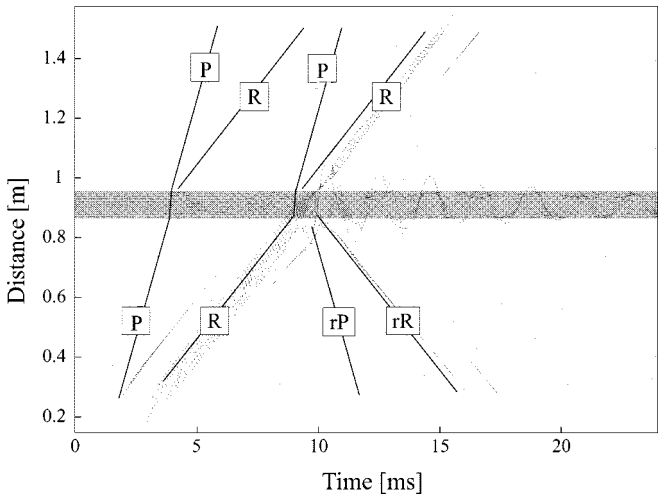


Fig. 9. Interaction of elastic waves with a buried antipersonnel mine (numerical simulation for a mine with air filled chamber) and waterfall graph of the vertical particle velocity at the surface (mine with air-filled chamber). The shear wave speed within the ground varies with depth and causes the waves to disperse. R: Rayleigh wave; P: pressure wave; rR: reflected Rayleigh wave; rP: reflected pressure wave. The gray box indicates the mine location.

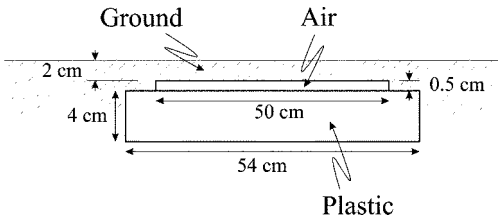


Fig. 10. Model to determine the nature of the resonance (elongated mine).

merical model for the mine is very simple, the shape of the resonance in the experiment strongly resembles the resonance obtained in the simulation. In order to explore the origin of the resonance and its nature, the mine is elongated within the numerical model, as shown in Fig. 10.

The resulting waves are shown in Fig. 11. The vertical particle velocity on the surface is depicted in a waterfall graph. As before, the waves can be easily distinguished by comparing their wave speeds. Pressure waves (P), Rayleigh waves (R), and head waves (H) are seen to propagate. They hit the mine and are partially reflected and partially transmitted. However, above the mine, a complex pattern of propagating waves arises. This can be explained as follows. First, the pressure wave hits the mine. The pressure wave couples its energy mostly into three different wave portions: reflected waves, waves that are transmitted through the plastic body underneath the air chamber, and waves that propagate within the thin layer between the surface and the air chamber. Because the plastic body lies underneath the air chamber, the wave portion transmitted through the plastic is not visible on the surface. This is evident in Fig. 11, where a transmitted pressure wave arises beyond the mine, apparently without being caused by any wave visible on the surface above the air chamber (the path the wave takes through the plastic is indicated by Pm). The Rayleigh wave carries much more energy than the pressure wave. When the Rayleigh wave hits the

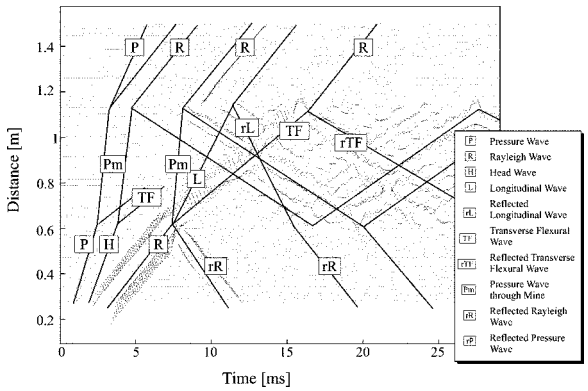


Fig. 11. Waves in the thin layer between the surface and the air chamber and waterfall graph of the vertical particle velocity.

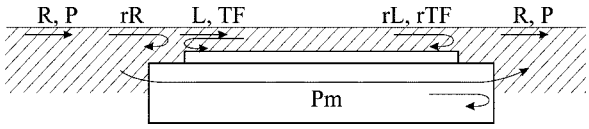


Fig. 12. Waves in the thin layer between the surface and the air chamber. The waves are almost totally reflected at the edge of the air chamber and are thus trapped within the thin layer above the air chamber.

mine, the same kinds of waves arise as for the incident pressure wave: reflected waves (pressure and Rayleigh waves, marked by rP and rR), waves transmitted through the plastic body (pM), and waves propagating within the thin layer between surface and air chamber. The waves through the plastic body behave as described for the pressure wave. However, the Rayleigh wave couples a significant amount of energy into the thin layer above the air chamber. Two different kinds of waves arise within the thin layer: a longitudinal wave and a transverse flexural wave. The longitudinal wave is the faster one of the two. It propagates within the thin layer between surface and air chamber (L), is reflected at the edge of the air chamber (rL), travels back within the layer and is reflected again. The transverse flexural wave also travels within the layer between surface and air chamber (TF), is reflected (rTF), travels back, and is again reflected. Note that the flexural waves are almost totally reflected at the edges of the air chamber, and the energy remains within the thin layer between surface and air chamber. Fig. 12 gives a schematic drawing of the principle behavior of the arising waves.

Using this model, the resonant behavior of the mine in the numerical model can be explained. When the waves hit the mine, they are partially reflected and partially transmitted. However, a large portion of the energy couples over into longitudinal and transverse (flexural) waves propagating within the thin layer between the surface and the mine's air chamber. Due to the short length of the mine, these waves form standing waves.

From the slope of the traveling waves in Fig. 11, the wave speeds of the longitudinal wave and the transverse flexural wave are determined to be $c_L \approx 160$ m/s and $c_{TF} \approx 50$ m/s, respectively. A simple theory for waves propagating in a thin plate predicts the wave speeds of the longitudinal wave to be [10]

$$c_L = \left(\frac{E}{\rho(1-\nu^2)} \right)^{1/2} \tag{12}$$

and the wave speed of the transverse flexural wave to be

$$c_{TF} = \sqrt{h\omega \left(\frac{E}{12\rho(1-\nu^2)} \right)^{1/2}} \tag{13}$$

where E and ν are Young's modulus and Poisson's ratio, respectively, ρ is the material density, and h is the thickness of the thin plate. For the longitudinal wave, this yields a wave velocity of $c_L \approx 163$ m/s. For a transverse flexural wave with a center frequency $f_c = 450$ Hz and for a plate thickness of $h = 2$ cm, a wave speed of $c_{TF} \approx 52$ m/s is predicted. This is in good agreement with the wave speeds obtained from the numerical simulation. Note that the wave speed of the transverse flexural wave is dependent on frequency, while the speed of the longitudinal wave is not. Thus, the transverse wave is strongly dispersive. The dispersion of the transverse flexural wave can be clearly seen in Fig. 11.

Note that the analysis of the resonant behavior of an actual mine is much more complex, and the argument outlined here gives only one possible cause for the occurring resonance. An actual mine is 3-D, has a flexible case that can support both flexural and longitudinal waves, and contains springs that can also give rise to resonances. The authors believe that the resonances observed in the experiments are due to flexural waves being trapped in both the case of the mine and the layer of soil above the mine. Experimentally, resonances are observed in mines flush with the surface of the ground. The authors are working to improve the model so that it more accurately predicts the mine–wave interactions.

IV. Conclusions

A finite-difference model has been developed and implemented in 2-D. An analytical solution was used to validate the finite-difference model. The finite-difference model was used to investigate the interaction of elastic waves with a buried antipersonnel mine. The TS-50 antipersonnel mine was approximated by two different models: one with an air-filled chamber and one without an air-filled chamber. Even though the numerical model was very simple, the results from the numerical model were found to be in good qualitative agreement with experimental results. Both in the experiment and in the numerical model containing the air-filled chamber, a strong resonanance at the mine location occured, whereas the model without the air-filled chamber did not show resonant oscillations. The nature of the resonance was investigated by slightly modifying the numerical model. If the air-chamber is elongated, flexural waves are seen to be excited and to propagate within the thin layer between the surface and air chamber. These waves are almost totally reflected at the edges of the air chamber and are thus trapped within the thin layer. In these experiments, the Rayleigh wave was seen to disperse as it travels along the surface. By assuming a depth-dependent shear wave speed, this effect can be reproduced within the numerical model.

References

[1] W. R. Scott, Jr., C. T. Schröder, and J. S. Martin, "An hybrid acousto/electromagnetic technique for locating land mines," in *Proc. Int. Geoscience and Remote Sensing Symp.*, Seattle, WA, 1998, pp. 216–218.

[2] W. R. Scott, Jr. and J. S. Martin, "Experimental investigation of the acousto-electromagnetic sensor for locating land mines," *Proc. SPIE*, vol. 3710, pp. 209–214, Apr. 1999.

[3] C. T. Schröder, "A Finite-Difference Model for Elastic Waves in the Ground," Diploma thesis, Univ. Braunschweig, Braunschweig, Germany, 1999.

[4] J. Virieux, "SH-wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophys.*, vol. 49, pp. 1933–1957, Nov. 1984.

[5] J. Virieux, "P-SV wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophys.*, vol. 51, pp. 889–901, Apr. 1986.

[6] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, vol. 114, pp. 185–200, 1994.

[7] W. C. Chew and Q. H. Liu, "Perfectly matched layer for elastodynamics; a new absorbing boundary condition," *J. Comput. Acoust.*, vol. 4, pp. 341–359, 1996.

[8] S. D. Gedney, "An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD-Lattices," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1630–1639, Dec. 1996.

[9] G. F. Miller and H. Pursey, "The field and radiation impedance of mechanical radiators on the free surface of a semi-infinite isotropic solid," *Proc. R. Soc. London*, vol. A 223, pp. 521–541, 1954.

[10] K. F. Graff, *Wave Motion in Elastic Solids*: Dover Publications, 1975.

[11] A. Tavlove, *Computational Electromagnetics: The Finite-Difference Time-Domain Method*. Norwood, MA: Artech House, 1995.

[12] M. Roth, R. Spitzer, and F. Nitsche, "Seismic survey across an environment with very high poisson's ratio," in *Proc. 12th Annu. Int. Meeting., Symp. Application of Geophysics to Engineering and Environmental Problems*, 1999, pp. 49–55.

[13] B. M. Das, *Principles of Soil Dynamics*. Boston, MA: PWS-Kent, 1993.



Christoph T. Schröeder [redacted] He received the M.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1997, and the Diplom Elekt. from the Technical University of Braunschweig, Braunschweig, Germany, in 1999. He is currently pursuing the Ph.D. degree in electrical engineering at the Georgia Institute of Technology.

His research interests include elastic wave propagation in solids, remote sensing, surface-acoustic wave devices, electromagnetics, and the finite-difference time-domain technique.



Waymond R. Scott, Jr. received the B.S., M.S., and Ph.D. degrees, all in electrical engineering, from the Georgia Institute of Technology, Atlanta, in 1980, 1982, and 1985, respectively.

He joined the Faculty of the School of Electrical Engineering, Georgia Institute of Technology, in 1986, where he is currently an Associate Professor of Electrical Engineering. His current research interests include the development of new techniques for measuring the electric and magnetic properties of materials, the development of materials with adjustable constitutive parameters, transient electromagnetic fields, ground penetrating radar, subsurface sensing, and numerical techniques such as the finite element and the finite-difference time-domain techniques.

Experimental Model for a Seismic Landmine Detection System

Waymond R. Scott, Jr., *Member, IEEE,* James S. Martin, and Gregg D. Larson

***Abstract*—A laboratory-scale experimental model has been developed and tested for a system that uses artificially generated high-frequency seismic waves in conjunction with a radar-based noncontact displacement sensor to detect buried landmines. The principle of operation of the system is to measure the transient displacement field very close to a mine location. In this way, the absorption and the geometrical spreading of the seismic waves have not reduced the effects of the mine. By using a seismic excitation, the system exploits the large difference between the elastic properties of a mine and the surrounding soil. This difference causes seismic wave interactions in the vicinity of a mine to be quite distinctive and provides a method for imaging mines and distinguishing them from typical buried clutter. Images of a variety of simulated and inert anti-tank and anti-personnel mines have been formed using this system. Burial scenarios involving natural clutter (rocks and sticks), light surface vegetation, localized burial effects, and multiple mines in close proximity have been studied. None of these scenarios appears to pose serious problems for detection performance.**

***Index Terms*—Acoustic, elastic waves, landmine, radar, seismic.**

I. Introduction

SEISMIC techniques show considerable promise for the reliable detection of all types of buried mines, even low-metal anti-personnel mines. The reason for this is that mines have mechanical properties that are significantly different from soils and typical forms of clutter. For example, the shear wave velocity is approximately 20 times higher in the explosive and the plastics used in typical mines than in the surrounding soil [1]. In addition, mines are complex mechanical structures with a flexible case, a trigger assembly, air pockets, etc. This complex structure gives rise to structural resonances, nonlinear interactions, and other phenomena that are atypical for both naturally occurring and man-made forms of clutter. This phenomenology can be used to distinguish a mine from clutter.

The range of burial depths typically associated with anti-personnel (AP) and anti-tank (AT) mines is from flush with the earth's surface to a depth of 10 to 20 cm. AT mines are typically larger and more deeply buried than AP mines. These characteristics make classical seismic techniques ill suited to the landmine detection problem. Seismic surveys with ground contacting geophones are intended to detect targets that are much larger and more deeply buried than landmines. These techniques usually involve the generation and detection of bulk waves (shear or compressional waves) in the earth. Surface guided (Rayleigh) waves are problematic for classical seismic measurements because their displacements decay exponentially away from the Earth's surface. They do not interrogate soil deeper than their wavelength, and they have relatively large surface displacements that can obscure other signals of interest. These features make the Rayleigh wave an excellent interrogation signal for the detection of land mines. Here, a signal with an appropriate frequency content will interrogate only the burial range of the targets of interest. Fig. 1 shows the wave fronts generated by an impulsive point source on an isotropic, homogenous elastic half space computed numerically [2]. The numerical computation was performed using a three-dimensional (3-D) finite-difference time-domain model. A free surface boundary condition is used to model the boundary between the air and the earth. A differentiated Gaussian pulse with a center frequency of 450 Hz is launched at $t = 0$ from a velocity source on the surface. The magnitude of the vertical component of velocity is plotted at $t = 10$ ms. The region plotted is a 3 m $\times$ 6 m cross-section through the ground. The incident signal, length scales, and time instant plotted have been selected so that the wavefronts have separated in space and correspond to parameters of the experimental system. The color map indicated in the figure is used throughout this paper, with the scales indicated in the figures. The Rayleigh wave and the bulk compressional and shear waves are indicated. The Rayleigh wave front can be seen to penetrate to a shallower depth and to be stronger near the free surface. This is the region in which mines are buried.

Previous attempts to detect buried mines by surface wave scattering have been confounded by the small size of mines and the high attenuation of most soils [3]. These force competing requirements for a classical, monostatic pulse–echo detection system: low frequencies are required to propagate measurable energy over the two-way path from the source to the mine location and back to the receiver while high frequencies are required for a mine to scatter waves of appreciable strength. This observation leads several authors to suggest that a noncontact sensor could be used to measure seismic motion at the mine [3], [4]. Such a sensor is at the heart of the system currently under investigation as it offers three distinct advantages over a remote-receiving array. First, it obviates the need for the scattered wave to propagate back to a receiver location in order to achieve detection, thus eliminating one half of the geometric spreading

Manuscript received June 8, 2000; revised October 30, 2000. This work was supported in part by the OSD MURI Program and by the U.S. Army Research Office under Contract DAAH04-96-1-0448.

W. R. Scott, Jr. is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: waymond.scott@ece.gatech.edu).

J. S. Martin and G. D. Larson are with the School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: james.martin@me.gatech.edu; gregg.larson@me.gatech.edu).

Publisher Item Identifier S 0196-2892(01)05107-5.

0196–2892/01$10.00 © 2001 IEEE

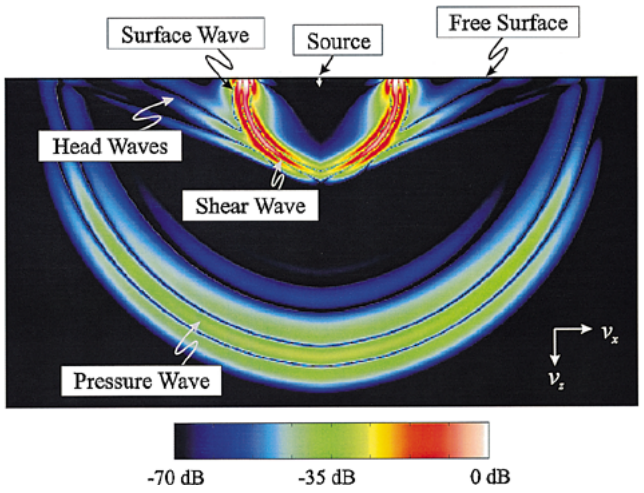


Fig. 1. Waves generated by a point source in an elastic half space.

and attenuation. Second, it removes ambiguities that might arise over the location of a target with a sparse remote array of receivers and/or a complicated propagation path. Third, it allows the measurement of nonpropagating motion above buried mines that may be detected only in the seismic near field of a mine.

All of the experiments reported in this paper have shown the existence of significant nonpropagating motion above buried mines. This is not surprising since nonpropagating wave-field components are a ubiquitous feature of near-field scattering and radiation problems. The most commonly cited examples of these are subsonic (evanescent) wavenumbers present on and near the surface of acoustically large objects. However, nonradiating components arise in extremely simple problems involving small sources also. One example of this would be the field produced by a radially oscillating sphere in an unbounded fluid. In addition to the propagating velocity component that is in phase with the acoustic pressure and has $1/r$ geometric spreading, there is a second, nonradiating, component that is in quadrature with the pressure and spreads as $1/r^2$. If the sphere under consideration is acoustically small ($ka \ll 1$), then the velocity close to the object is dominated by this nonradiating component [5].

Other investigators have used a laser-based local displacement sensor in conjunction with uniform, continuous aero-acoustic excitation of the soil surface for mine detection [6]. Although this method appears to be similar to the one presented here, the underlying physical principles are quite different for the two methods. In the aero-acoustic method, the seismic interrogation signal was comprised of compressional waves propagating into the porous surface soil layer. The primary detection cue was a change in the local input impedance. This change was attributed to the significant differences in the porosity of the soil and the mine. In the seismic method presented here, the primary detection mechanism is motion of the mine excited by its interaction with the Rayleigh wave. The spatial/temporal imaging scheme used here is tied to the nature of a transient incident signal which propagates in the measurement plane; this would not be possible for the downwardly propagating continuous excitation.

Both Rayleigh waves and air coupled pore compressional waves decay with depth into the soil. For a Rayleigh wave, penetration depth is related to the wavelength measured along the surface. However, the penetration depth of pore compressional waves is a function of soil parameters that are not directly accessible through surface displacement measurements. Thus, a Rayleigh wave incident signal offers the potential to exploit additional information concerning target depth.

II. Experimental Model

The mine detection system that has been modeled experimentally is depicted in Fig. 2. The system consists of a stationary seismic source and a moving radar sensor.

The source is an electrodynamic shaker that has been coupled to the ground by a narrow foot. This was designed by experimental iteration to preferentially generate Rayleigh waves. The foot, a thin rectangular aluminum bar approximately 25 cm long, was placed with the long dimension parallel to the y-axis as in Fig. 3. The Rayleigh wave generation appears to be dictated by the foot's perimeter and bulk wave generation by its surface area. Although this suggests that a knife edge geometry would be optimal, the foot must also carry a sufficient bias force (the weight of the shaker in this case) that contact with the surface is maintained. Without a sufficiently large contact area, the shaker foot will tend to bury itself.

The sensor, which measures the displacement of the Earth, is comprised of an 8 GHz CW radar. The radar illuminates an area of the Earth's surface, and the signal reflected from the surface is received and demodulated. The motion of the Earth changes the distance the electromagnetic waves travel, resulting in a phase modulation of the received signal that is proportional to the surface displacement. The surface displacement is obtained by demodulating the received signal. A homodyne system is used with in-phase and quadrature mixers to demodulate the received signal. Both the surface displacements due to the seismic waves and the variations in path length associated with the static contours of the ground are obtained separately from the outputs of the two mixers. The two biggest challenges to make the radar perform adequately for the mine detection system are 1) to make it sufficiently sensitive to be able to detect the small vibrations, and 2) to make the spot size (the area on the surface illuminated by the electromagnetic waves) sufficiently small to image the seismic wave field. The radar can measure displacements of 1 nm (10^{-9} m) as currently configured which is more than sufficient to measure the surface displacements due to the seismic waves that are on the order of 1 μm. To obtain this sensitivity, the radar was designed to minimize the effects of noise, such as the phase and amplitude noise of the source and electromagnetic interference from low-frequency magnetic fields. The spot size of the radar must be smaller than approximately one half of the shortest wavelength of the seismic waves. Currently, a small spot size is obtained by using an open-ended waveguide as the antenna for the radar. This antenna produces a spot size of about 2 cm in diameter when the open end of the antenna is placed within a few centimeters of the surface. A field operational mine detection system could contain many of these sensors in a planar

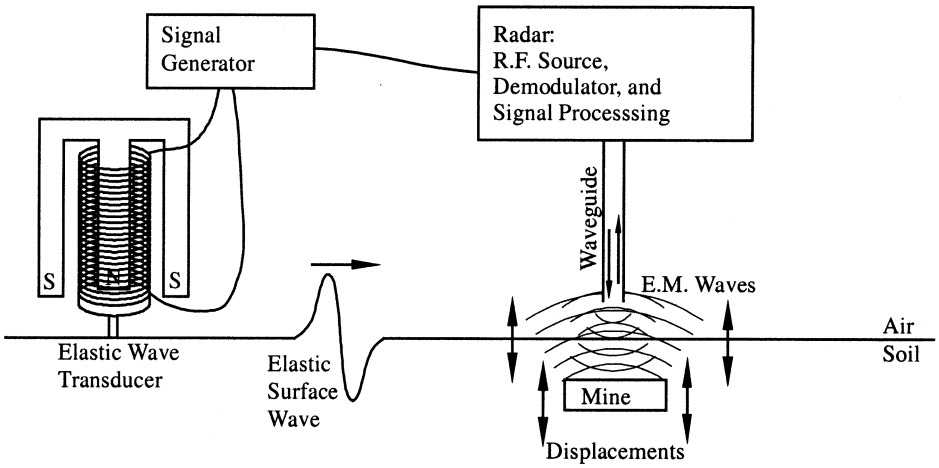


Fig. 2. Configuration of seismic mine detection system.

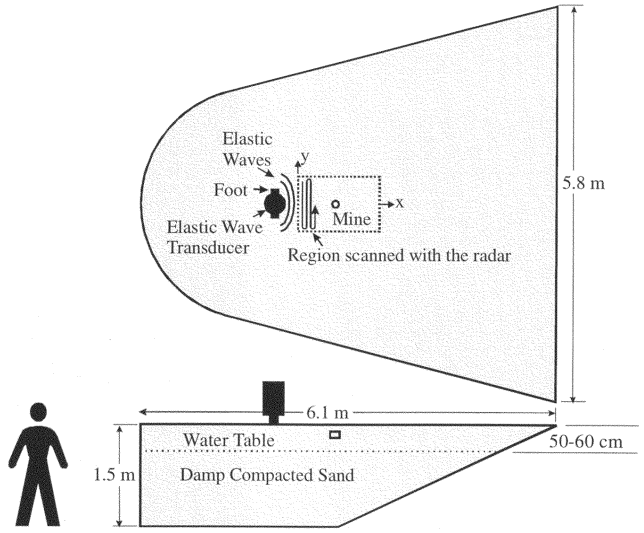


Fig. 3. Experimental soil tank facility.

array. For the laboratory system a single radar sensor is translated above the surface of the earth in order to synthesize this array.[1]

Laboratory testing has been conducted in a wedge-shaped tank, which is depicted in Fig. 3, filled with damp compacted sand to simulate soil. The seismic source is located as indicated in Fig. 3 and is bidirective toward the search area and the back wall. A water table is maintained 50 to 60 cm below the surface of the tank. Damp compacted sand was chosen as the soil surrogate, because it can be dug up, refilled, and repacked with fairly good repeatability. This repeatability is very important because it allows for comparisons between experiments and thus, easier interpretation of the experiments. The sand surface is periodically re-wet from above and is compacted prior to measurements. This maintains cohesion in the near surface layers of sand that have no bias force created by the weight of overlying material. Surface cohesion is needed to mimic naturally occurring soils. Natural cohesion of particles is due to the water content, weathering, and the presence of fine clays and minerals that cement the soil particles together. Field experiments and published data have verified the realism of compacted damp sand as a soil surrogate because of the comparable wave speeds [3], [6]–[9].

Simulated mines, inert mines, and clutter such as rocks and sticks can be buried within a 2 m $\times$ 2 m scannable region in the center of the tank. The typical scan region, 80 cm $\times$ 120 cm, is outlined in Fig. 3. The scan region is sufficiently far from the tank walls that wall reflections can be time-gated out of the data. The sensor can be scanned over this region with a three degree of freedom positioner fixed above the tank.

Experiments in the tank indicate the presence of two measurable propagating wave types: a slow large amplitude surface wave that propagates at 80 to 90 m/s and a smaller faster bulk wave which propagates at 190 to 250 m/s. The faster wave is consistent with previously reported bulk compressional waves in soil and sand [6], [7], and the slower wave matches reported measurements of high-frequency ($>$ 50 Hz) Rayleigh wave speeds [3], [8], [9]. Direct measurement of the compressional wave speed at the surface is difficult because these waves are refracted upwards, making the actual path between the source and receiver uncertain. Direct measurement of a bulk shear wave velocity is not possible at the free surface. This should be slightly faster than the Rayleigh wave speed and can be computed to be 90 to 100 m/s from the measured Rayleigh and compressional wave speeds.

Data taken in a one-dimensional (1-D) scan away from the source are depicted in Fig. 4. Here the time waveforms are depicted in a waterfall format or seismogram, where each trace has been offset vertically to represent the spatial separation of the measurement points. This creates a pseudo-surface effect. Individual modes of propagation can be seen as lines of delay connecting wave fronts. The lines on this figure represent subjective evaluations of the phase velocities for the two obvious propagation modes. Dispersion is apparent in the surface wave pulse. Absorption, vertical stratification, inhomogeneity, or nonlinear effects can cause this. Efforts are currently

[1]The synthetic array is formed on a seismic not an electromagnetic time scale by physically moving, rather than steering, the radar sensor. The seismic wavefronts represented in the data are, therefore, compilations of many distinct seismic events produced by identical source signals at different times (some several hours apart). This form of array synthesis is the reason for the inordinately long scan times associated with the experimental system.

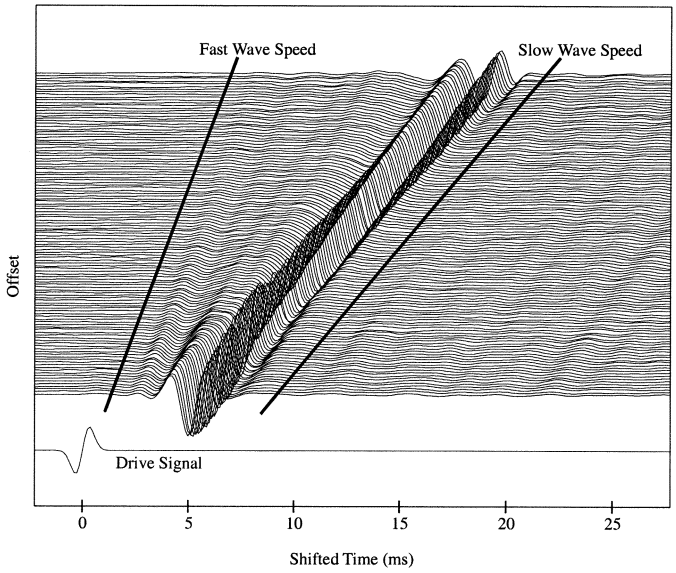


Fig. 4. Seismogram of wave propagation through soil tank measured in 1 cm increments on a 120 cm scan.

underway to model and measure each of these contributions. It is also possible that interference with simultaneous arrivals of refracted and reflected bulk waves may create the appearance of dispersion.

Similar data are depicted in Fig. 5 as a pseudo-color graph. Here, the amplitude of the surface displacement at a single instant in time is represented for all measurement locations within the scan region. Out of plane displacement magnitudes are shown on a color scale. The plot area represents the physical dimensions of the surface of the ground. The source is located to the left side of this image, and the incident waves propagate to the right. This convention is used for all the pseudo-color images presented in this paper. The wavefronts apparent in this figure are associated with the Rayleigh wave. At the time instant shown, the compressional wavefronts have propagated beyond the scan region. A sequence of these plots can be used to animate the wave propagation throughout the scan region and visualize interactions with buried objects.

A swept-frequency chirp is currently used to efficiently acquire the data. The sensor is moved to a measurement location, and the source is then driven with a 4-s chirp covering the frequency band of interest. The time record due to the chirp is used to reconstruct the transfer function of the drive signal to the measured displacement [4]. This is then used with a simulated drive signal of shorter duration, with similar bandwidth, to compute a corresponding transient displacement. For the present work, a differentiated Gaussian pulse with a center frequency of 450 Hz is used as the simulated drive signal. The drive signal is graphed in Fig. 4. The shorter pulse allows for the time separation of different wave types and propagation paths. Similar data have been taken using the shorter pulse as the initial drive signal. The number of record averages required to achieve an SNR comparable to the method using the chirp was time preclusive. However, the data acquired in this way were identical to the compressed waveforms.

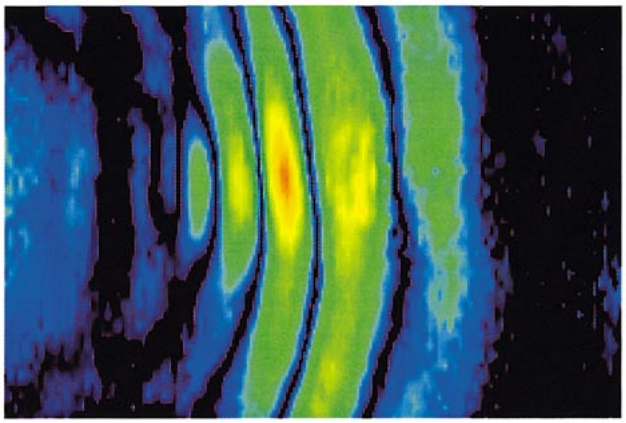


Fig. 5. Transient displacement of the sand surface on a 40 dB scale at an instant when a Rayleigh wavefront is midway through the 80 cm $\times$ 120 cm measurement region.

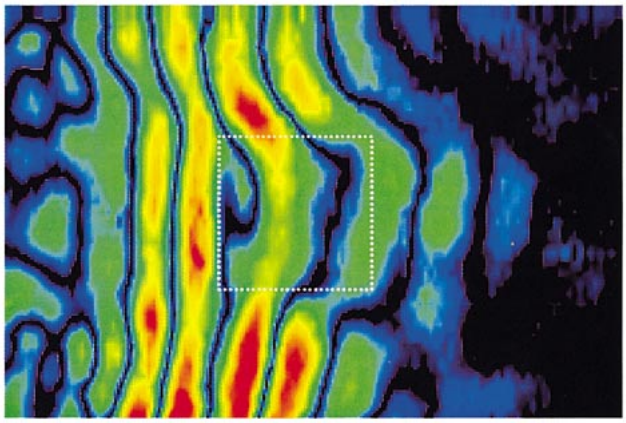


Fig. 6. Interaction of Rayleigh waves with a rigid mine simulant (white outline) buried 5 cm deep with the surface in an 80 cm $\times$ 120 cm scan region on a 40 dB scale.

The data acquisition scheme presumes a linear response of the sand and the mine. Sand, however, is a highly nonlinear media. Drive levels that are very low compared with the capabilities of the source are used to prevent nonlinear responses from producing artifacts in the processed data. Most of the observed nonlinearities occur at or near the source where displacements are largest. These nonlinearities significantly limit the amplitude of the waves that can be launched into the sand.

The experimental model has been configured to emphasize data integrity over scanning speed. Currently 9 h are required for a complete two-dimensional scan of 121 $\times$ 41 measurement points ($\sim$1 m^2). As this is clearly inappropriate for a field-operational mine detection system, several techniques are being considered to improve speed without sacrificing measured signal levels. Individually, each of these offers modest time savings, but taken in total, they should reduce scan times to a few seconds per square meter.

Physically arraying the displacement sensor will speed scanning by a factor of the number of elements in the array. If, in lieu of the chirp, an M-sequence of similar bandwidth is used, this should allow a 50% reduction in integration time [11]. Elimination of spatial oversampling with the current system will cut

measurement times by a factor of four to eight. Improved source design, reduced operating bandwidth, simultaneous processing of data, and an effective scheme for dealing with nonlinear artifacts will provide additional, interrelated improvements.

III. Rigid Mine Simulants

Some experiments on seismic mine detection were conducted using mine simulants that did not incorporate mine-like structures. A 30 cm $\times$ 30 cm $\times$ 7.5 cm acrylic block was used as an AT mine model representing only the size, shape, and density of an actual target. The primary cue for the identification of this simulant was found to be its combined stiffness and geometric regularity. The simulant responded in a nearly rigid way to the wave motion of the surrounding soil. The motion of the simulant was much like that of a ship at sea. Incident wave fronts appeared to break at the leading edge of the mine; waves were shed behind the mine prior to their arrival via the diffracted path around the mine, and the mine itself rocked about a point roughly 1/3 of the way along its length. Fig. 6 shows a pseudo-color graph of the displacement of the measurement surface above the acrylic mine buried 5 cm below the surface at an instant shortly after the Rayleigh wave has reached the mine location. The wavefronts to the right of the mine are seen to be curved, because of the accelerated arrival of the waves shed by the mine. The mine itself is moving less than the surrounding sand, and the edges of the mine are apparent as discontinuities in the surface displacement. Significant dispersion is also observed in a wave that propagates in the layer of soil over the mine. This layer forms a waveguide between the mine and the free surface. Although this sort of mine simulant has been imaged, it would be difficult to distinguish from buried clutter. The feature that makes this mine distinctive from rigid clutter such as rocks is its geometric regularity. Accurately outlining the shape of the simulant required wavelengths that would not penetrate to the full depth range expected for a mine of this size.[2]

IV. Resonant Mines

Although it was possible to image mine simulants without representative structural details, the dominant features observed in the signatures of inert AP and AT mines were found to be soil-loaded resonances of the mine case and trigger mechanism [12], [13]. These are excited by the passage of the Rayleigh wave and characterized by large displacements that persist after the passage of the incident pulse. Numerical models indicate that the resonance is due to flexural waves excited in the case/trigger of the mine and in the layer of soil above the mine [2]. The mine trigger is comprised of plastic that is much stiffer than the overlying sand. However, this plastic is thin in comparison to the soil layer. The stiffness and the mass of both the mine trigger and the overlaying layer of soil are believed to be important contributors to the resonance.

Although mines exhibiting resonances scatter a larger propagating wave field than similarly sized nonresonant objects, the most pronounced feature of the field scattered by these mines is its primarily local nature resulting from a dominantly reactive soil loading. For the mine types studied thus far, the localized resonant motion has been an excellent indicator of a mine's location and extent. Fig. 7 shows a seismogram for each of five different AP mine types compared to the no-mine case. Four of these can be seen in the photograph in Fig. 10. The waveforms have been windowed so that only information around the arrival time of the incident Rayleigh wave pulse is shown. To varying degrees, a resonance is apparent for each AP mine type. These mines were buried 0.6 cm below the surface.

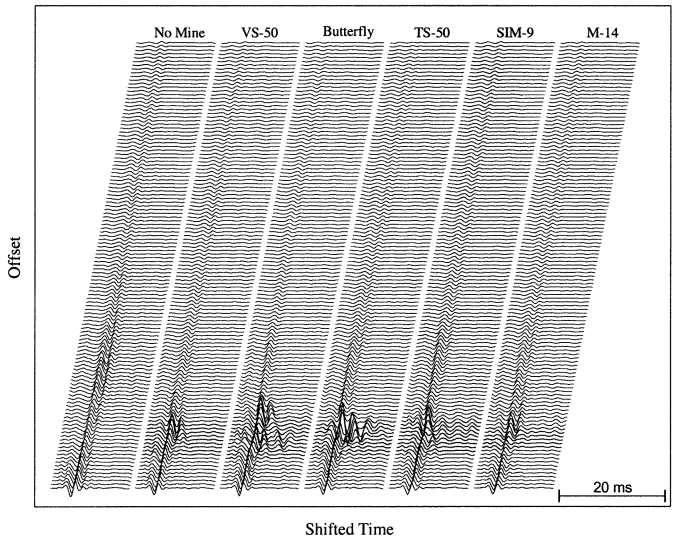


Fig. 7. Displacement of the surface of the sand above various AP mines, measured in 1 cm increments on a 120 cm scan. All AP mines were buried 0.6 cm deep.

2-D scans over buried inert mines show the mine resonance features more clearly. In Fig. 8, surface displacement magnitudes are represented over the scan region for three instants in time when an inert TS-50 mine is buried 1.3 cm deep in the scan region. At first (a), the incident wave has not reached the mine. In the second instant (b), the wave front has reached the mine and can be seen to produce large displacements in the soil layer above the mine. In the third instant (c), the incident signal has propagated beyond the mine. However, there is still ringing at the mine, and it is shedding energy by radiating small amplitude wave fronts.

Imaging of mines from the surface displacement measurements can be done in many ways. The current technique, which has been outlined in a previous paper [14], involves a multistep process. The 2-D scan data are filtered in the wavenumber domain to remove all components propagating away from the source, leaving the reflected waves and a portion of the nonpropagating waves. The remaining information is windowed in time around the arrival time of the incident signal and averaged to form a pixel in the final image. Fig. 9 shows an image formed in this way from the data that was used to generate the frames of Fig. 8.

V. Buried Clutter and Multiple Mines

Many objects, which are mine-sized or larger and have elastic properties quite different from soils, are buried at shallow depths

[2]This is a different problem from the competing operational frequency requirements encountered with the pulse-echo scheme previously discussed. Here, the competition between frequency, mine size, and depth pertains to target classification rather than detection.

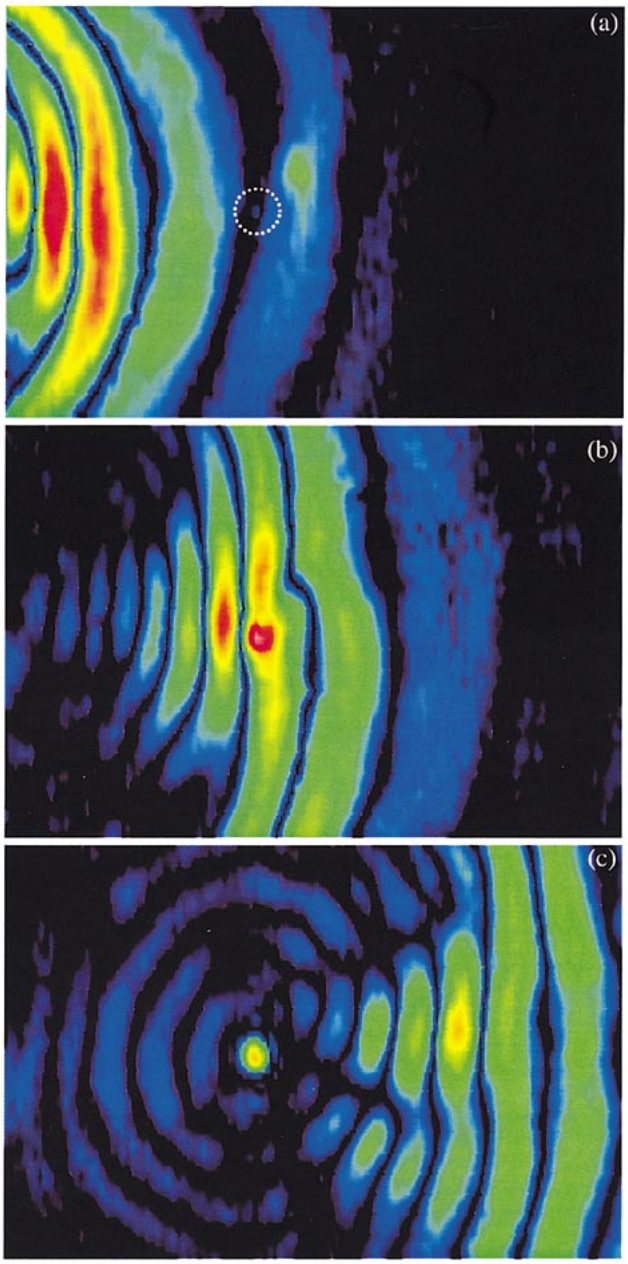


Fig. 8. Sand surface displacements in an 80 cm × 120 cm scan region at three instants on a 40 dB scale (a) before wavefronts reach a TS-50 AP mine (outlined) buried 1.3 cm deep at $t = 3.0$ ms, (b) during interaction of the waves with the mine at $t = 6.9$ ms, and (c) after the wavefronts have passed the mine at $t = 12.2$ ms.

in the ground. Rocks and tree roots are good examples of this sort of clutter. For a mine detection system to be effective, it must permit imaging that accurately depicts the size and location of a mine and distinguishes it from this type of clutter.

A common practice of mine warfare is to plant multiple AP mines in close proximity to AT mines. The AP mines thereby protect the AT mine from sappers who can more easily detect the larger object and remove it with little personal danger. This poses a unique detection problem in that it requires a system to operate with sensitivity appropriate to both mine types simultaneously. Also, the system must be capable of distinguishing

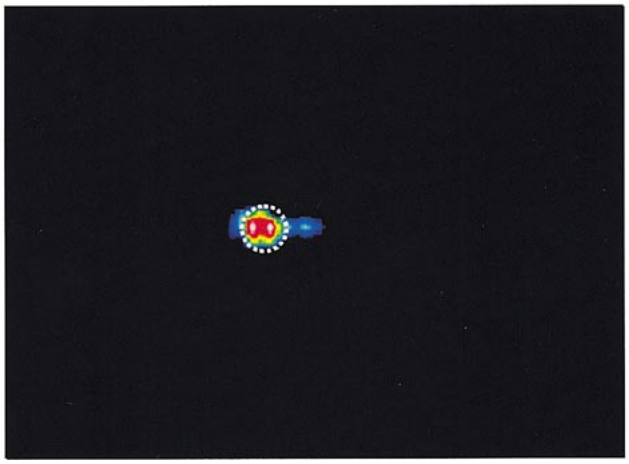


Fig. 9. Image formed of a single TS-50 AP mine (outlined) buried 1.3 cm deep in an 80 cm × 120 cm scan region on a 25 dB scale.

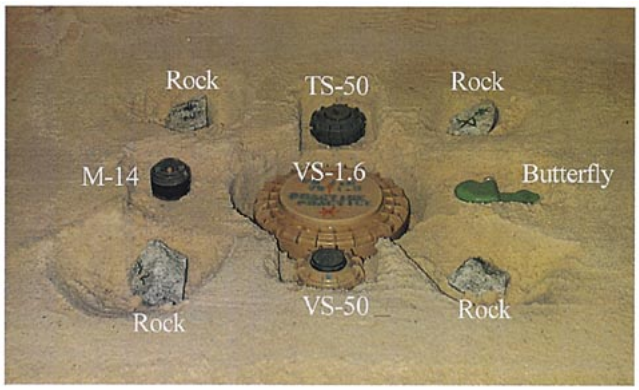


Fig. 10. VS-1.6 AT mine surrounded by TS-50, butterfly, VS-50, and M-14 AP mines and rocks. The burial depths for the mines were 4.5 cm for the VS-1.6; 2 cm for the TS-50, VS-50, and butterfly; and 0.5 cm for the M-14. The burial depths for the rocks were 3.5 cm, 1.5 cm, 2 cm, and 1 cm (clockwise, starting with the upper left rock).

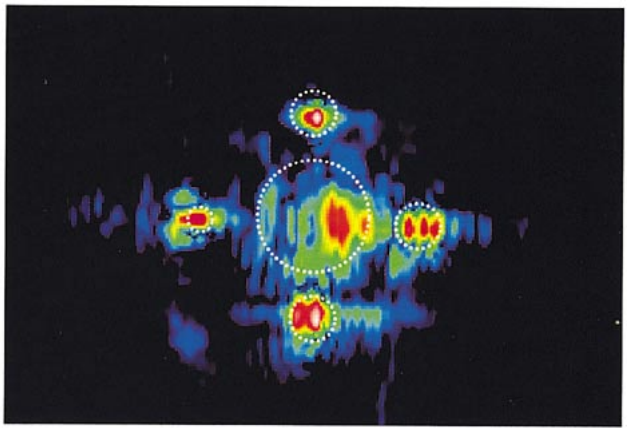


Fig. 11. Image of AT mine (outlined) and surrounding AP mines (outlined) in an 80 cm × 120 cm scan area on a 30 dB scale.

individual targets and rejecting ghost images formed by multiple scattering effects.

An experiment was performed to test the effects of buried clutter and multiple mines. In this experiment, an inert VS-1.6

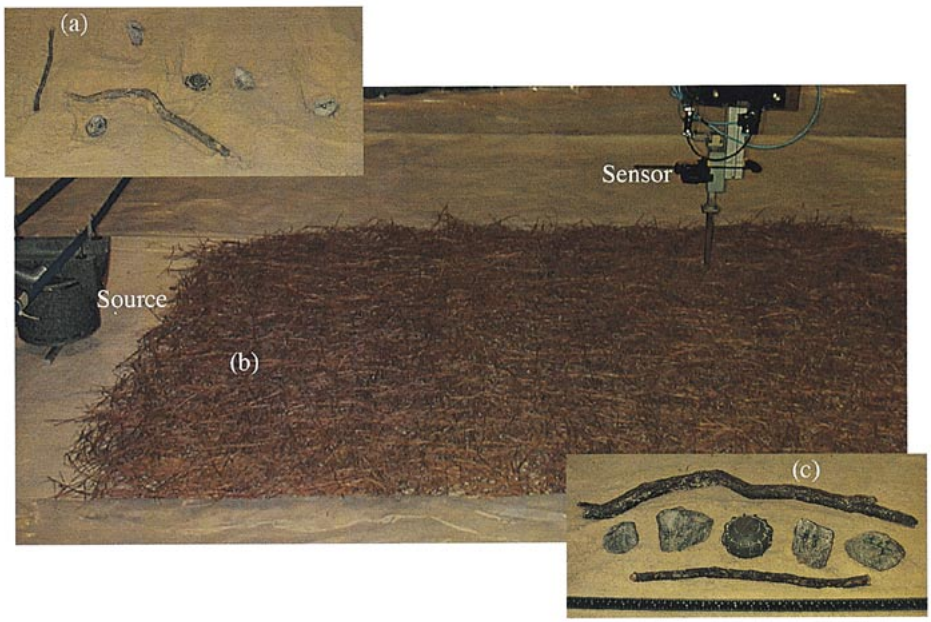


Fig. 12. Simulated minefield (a) relative locations of TS-50 AP mine and clutter objects (b) pine straw surface covering (c) relative scale of buried objects.

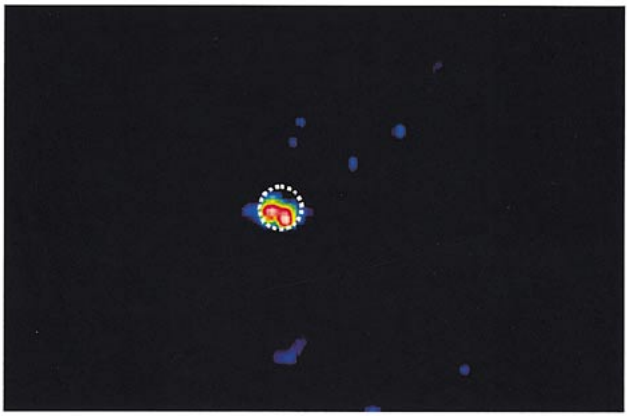


Fig. 13. Image formed of TS-50 AP mine (outlined) and clutter below pine straw surface cover in an 80 cm $\times$ 120 cm scan region on a 25 dB scale.

AT mine was surrounded by four different inert AP mines and four mine-sized rocks. The arrangement of this burial and relative scale of the objects can be seen in Fig. 10.

The image formed from the data taken over the multiple mine burial is depicted in Fig. 11. The number of mines present and their relative locations have been accurately imaged. The image of the AT mine is seen to be strongest at the back edge. This is due to the reflection at the back edge being stronger than that at the front. The effects of the rocks are much smaller than those of any of the mines. The largest rock is barely discernible on the 30-dB dynamic range used to generate the image. The reason for this is that the rocks do not exhibit resonances within the frequency range of the incident signal. This result has been modeled numerically with a 3-D finite-difference time-domain (FDTD) model [15]. The FDTD model incorporates measured material properties and approximations to the complex geometry of the rocks and mine. Results from the FDTD model are in good agreement with the experimental measurements.

VI. Ground Cover

In early testing of the seismic mine detection system, the surface of the soil surrogate used in the experiments was maintained level, smooth, and bare. This provided the radar sensor with a seismically modulated EM reflection from the surface that was nearly optimal (large and uncorrupted). In the field, the surface will be rough and lie under some type of ground cover. A sensor that sees through common surface cover such as grass or light vegetation is essential for practical system operation. Some types of ground cover, obviously, will be opaque to the radar's interrogation signal. These will require special consideration and possibly an alternate sensor design. Standing water is an example of such a problematic case. In general, vegetation is likely to be translucent to the radar sensor with a portion of the EM reflection originating in the vegetation and the remainder coming from the surface below. The signal that is reflected from the vegetation can corrupt the total return in two ways. First, it will reduce the relative contribution of the electromagnetic signal component reflected from the surface. This will reduce the level of the measured displacement signal. Second, any motion of the vegetation will produce an additional modulation of the carrier unrelated to the motion of the underlying surface. This will increase the effective noise floor.

To test the ability to penetrate surface cover, pine straw was selected both for the convenience of its application and because it is a commonly occurring ground covering. In the experiment, a 2.3 cm layer of pine straw was spread over the model's surface beneath which a TS-50 AP mine was buried 1.3 cm deep along with 4 mine sized rocks and two sticks at similar depths. The layout of this experiment can be seen in Fig. 12.

Fig. 13 shows the image formed from the pine straw covered surface. The location and extent of the TS-50 mine are apparent. There is less contrast in this image than for similar images formed in the absence of surface covering. An examination of the signals that contribute to the image reveals that the

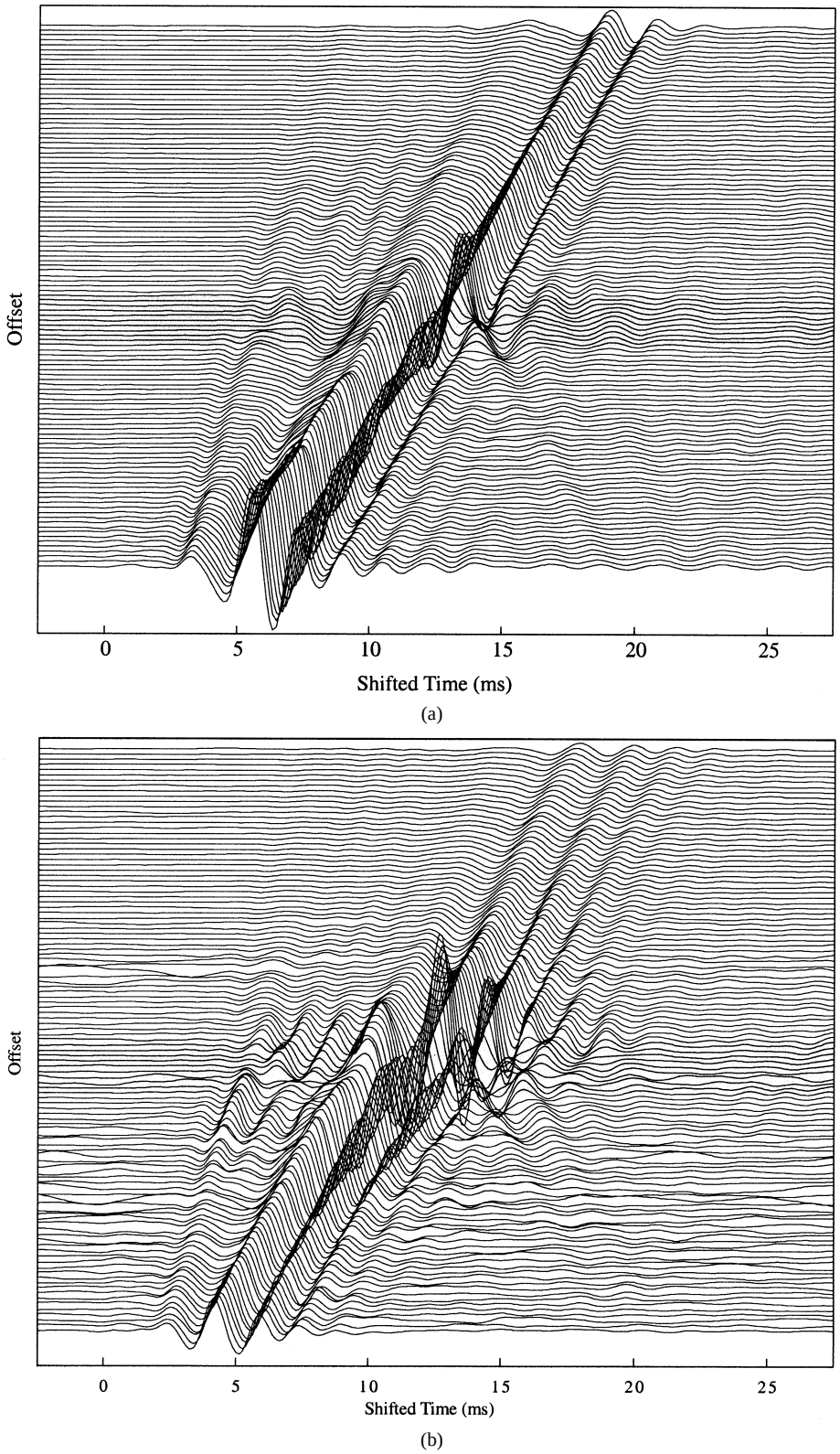


Fig. 14. Seismograms showing surface displacements measured in 1 cm increments on a 110 cm scan across a compliant AP mine simulant buried 4 cm deep (a) with the surrounding sand uniformly compacted and (b) with the sand locally disturbed around the mine.

dominant effect of the pine straw was to force an increase in the separation between the antenna and the ground's surface that resulted in a reduced signal level and a loss of some spatial resolution. The reduced signal level causes an increase in the noise floor that reduces the contrast in the image.

VII. Trenching Effects

It is well known that it is often easier to detect the recently disturbed earth around a buried mine than to detect the mine itself. This burial disturbance is called trenching. Observations

of trenching effects suggested an experiment to test the impact of recent burial on the detectability of an AP mine. When a mine has been recently buried, the soil around it is likely to be less compacted than in its greater surroundings. This is not surprising since an individual placing a mine is apt to be hesitant to compact the soil around it. Over time, compaction will occur due to weathering. The trenching effect is, therefore, a poor cue for reliable mine detection. A local lack of cohesion is not the primary detection cue under investigation with the current system. It was suspected that trenching could degrade system performance by decoupling the mine from the surrounding soil through which the seismic excitation is applied.

Fig. 14 shows seismograms generate 1-D scans over a resonant AP mine simulant buried 4 cm deep. The mine simulant was a hollow plastic case with thin walls approximately 9 cm in diameter and 2.2 cm in height. In the upper plot, the sand in the entire scan region was tilled and recompacted after the mine burial. In the lower plot, only the sand immediately around the mine was disturbed for the burial creating a local inhomogeneity. The resonance of the mine simulant is apparent in both images, but is much more pronounced in the presence of the local inhomogeneity. This surprising result is believed to be attributable to the hole focusing the incident wave to the location of the mine. The hole also reduces the radiation damping of the mine's resonance because of the poor match to the properties of its surroundings. These effects have been reproduced with numerical models that represent the filled hole as a cylinder of material having 20% lower wave speeds than the bulk of the medium. Both the experimental model and the numerical model predict that the trenching effect enhances the resonant response of the mine and thereby improves detection performance [16].

VIII. Depth Range of Detectability

There are two ranges of mine size and burial depth that are of concern for the mine detection problem. The first of these is for AP mines that are typically a few hundred milliliters in volume and buried such that the mine trigger is a few cm below the soil surface. AT mines are typically several liters in volume and have burial depths up to a few decimeters. Rayleigh waves penetrate the surface to a depth proportional to their wavelength. It is therefore likely that AT mine detection will dictate the low frequency requirements of the detection system. Since the primary mine detection cue that is exploited with the current system is a resonance of the mine, it is unlikely that the AT mine problem is a simple direct scaling of the AP mine imaging technique. It must, therefore, be specifically addressed. Currently, the operating band of the system (100–2000 Hz) is dictated by the design of the experimental model. Source response, ambient noise, and reverberation impose the low-frequency limit. Attenuation in the soil surrogate limits the operating bandwidth at high frequencies.

A VS-1.6 AT mine was examined as a function of burial depth using a 1-D scan. The results for this are depicted in the seismograms shown in Fig. 15. The seismograms shown in Fig. 15(a) have been windowed in the same manner as those in Fig. 7. In

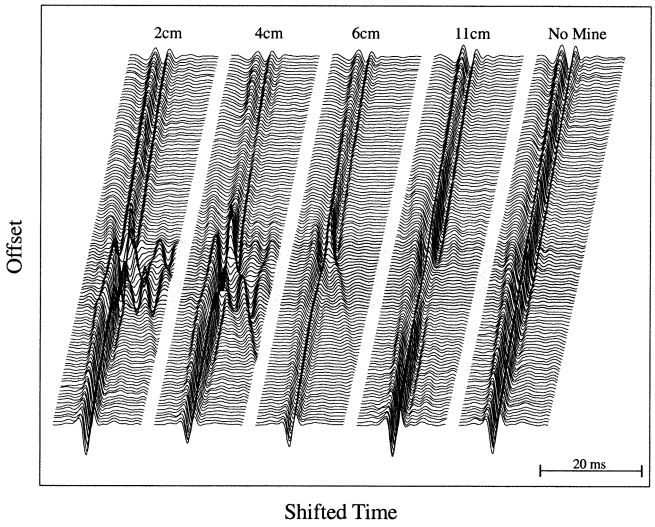


Fig. 15. Windowed seismograms of VS-1.6 AT mine at various depths, measured in 1 cm steps on a 120 cm scan. (a) Raw data and (b) forward propagating waves filtered out.

Fig. 15(b), the forward components of the wave field have been filtered out to emphasize the effects of the mine. The ripples seen in the no-mine graph in Fig. 15(b) are due to imperfect filtering of the incident wave. The root mean square (RMS) value of each trace shown in Fig. 15(b) would represent a pixel along the centerline if an image were formed.

Although the raw data covered the same band as in the previous seismograms, it was processed in a slightly different way. The measured transfer functions were convolved with a pulse shape that had a center frequency that was an octave lower (225 Hz as opposed to 450 Hz for the AP mine scans). The experimental data used to reconstruct either pulse response are identical. The difference is in the postprocessing. Many such processing techniques could be used simultaneously to interrogate for specific targets or burial depth ranges. The pulse with the lower center frequency was chosen for the AT mine in order to emphasize the effects of the spectral components which penetrated the soil to a sufficient depth to interact with the mine.

From data in Fig. 15 the presence of the mine is apparent up to a burial depth of about 11 cm. Above the mine, an amplification of the incident wave can be observed for 2, 4, and 6 cm burial depths. Waves that have been reflected from the mines are clearly seen propagating away for these burial depths. The strength of these waves for the 2 and 4 cm burial depths appears to be inversely related to the duration of the localized motion above the mine. This indicates that the radiation of Rayleigh waves constitutes a significant source of damping for the mine motion. There is also evidence of dispersion in the soil layer above the mine. The leading edge of the incident signal is clearly delayed as it passes above the mine. Also, the pulse shape and arrival time well beyond the mine are quite different than they are at the same location for the no-mine case.

IX. Conclusions

Detection of simulated and inert AP and AT mines using Rayleigh seismic waves and a radar-based noncontact displacement sensor has been demonstrated with a laboratory experimental model. Scenarios that mimic a variety of realistic field conditions have been modeled. The system has proven to be effective in these scenarios. Modeled conditions include the presence of natural surface covering, buried clutter, multiple mine burials, trenching, and deeply buried mines. Experimental system performance shows good agreement with published data and with analytical and numerical models of seismic propagation.

References

[1] D. L. Patel, "Handbook of land mines and military explosives for countermine exploitation," Tech. Rep. 2495, U.S. Army Belvoir Res. Devel. Center, Mar. 1992.

[2] C. T. Schroeder and W. R. Scott Jr., "A finite-difference model for elastic waves in the ground," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1505–1512, July 2000.

[3] "Feasibility of acoustic landmine detection," Tech. Rep. 7677, BBN Syst. Technol., Cambridge, MA, May 1992.

[4] C. Stewart, "Summary of mine detection research," Tech. Rep. 1636-TR, U.S. Army Eng. Res. Devel. Labs, Corps Eng., Belvoir, VA, vol. I, May 1960.

[5] M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction*: Acoust. Soc. Amer., 1993, pp. 31–34 and 60–64.

[6] J. M. Sabatier, H. E. Bass, L. N. Bolen, and K. Attenborough, "Acoustically induced seismic waves," *J. Acoust. Soc. Amer.*, vol. 80, pp. 646–649, Aug. 1986.

[7] R. Bachrach, A. Nur, and J. Dvorkin, "High-resolution shallow-seismic experiments in sand, Parts I and II," *Geophysics*, vol. 63, no. 4, pp. 1225–40, 1998.

[8] P. Smith, P. S. Wilson, F. W. Bacon, J. F. Manning, J. A. Behrens, and T. G. Muir, "Measurement and localization of interface wave reflections from a buried target," *J. Acoust. Soc. Amer.*, vol. 103, pp. 2333–2343, May 1998.

[9] M. Westebbe, J. F. Bohme, H. Krummel, and M. B. Matthews, "Model fitting and testing in near surface seismics using maximum likelihood in frequency domain," in *Conf. Record 32d Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1998.

[10] G. Godfrey, *Perturbation Signals for System Identification*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[11] A. Behboodian, W. R. Scott, Jr., and J. H. McClellan, "Relative performance of different perturbation signals for measuring the impulse response of linear time-invariant systems," *IEEE Trans. Instrum. Meas.*, to be published.

[12] W. R. Scott Jr., G. D. Larson, and J. S. Martin, "Simultaneous use of elastic and electromagnetic waves for the detection of buried land mines," in *Proc. SPIE: 2000 Annu. Int. Symp. Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, Apr. 2000.

[13] W. R. Scott Jr., C. Schroeder, and J. S. Martin, "A hybrid acoustic/electromagnetic technique for locating land mines," in *Proc. 1998 Int. Geoscience and Remote Sensing Symp.*, Seattle, WA, July 1998.

[14] A. Behboodian, W. R. Scott Jr., and J. H. McClellan, "Signal processing of elastic surface waves for localizing buried land mines," in *Proc. 33rd Assilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 1999.

[15] C. T. Schroeder and W. R. Scott Jr., "Three-dimensional finite-difference time-domain model for interaction of elastic waves with buried land mines," in *Proc. SPIE: 2000 Annu. Int. Symp. Aerospace/Defense Sensing, Simulation, and Controls*, Orlando, FL, Apr. 2000.

[16] W. R. Scott Jr., C. T. Schroeder, J. S. Martin, and G. D. Larson, "Investigation of a technique that uses both elastic and electromagnetic waves to detect buried land mines," in *AP2000 Millennium Conf. Antennas and Propagation*, Davos, Switzerland, Apr. 2000.



Waymond R. Scott, Jr. (M'00) [redacted] He received the B.E.E., M.S.E.E., and Ph.D. degrees from the Georgia Institute of Technology (Georgia Tech), Atlanta, in 1980, 1982, and 1985, respectively.

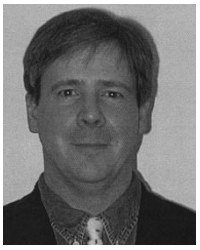
From 1979 to 1980, he was a Student Assistant and Graduate Research Assistant with the Georgia Tech Research Institute, and from 1980 to 1985, he was a Graduate Research Assistant with the School of Electrical Engineering at the Georgia Institute of Technology, where he is currently a Professor of Electrical and Computer Engineering. His research interests include methods for detecting buried objects using both electromagnetic and acoustic waves, measurement of the electromagnetic properties of materials, transient electromagnetic fields, and numerical methods including the finite element and the finite-difference time-domain techniques.



James S. Martin [redacted] He received the B.M.E. and M.S.M.E. degrees from the Georgia Institute of Technology (Georgia Tech), Atlanta, in 1989 and 1994, respectively, where he is currently pursuing the Ph.D. degree.

He has been working as a Research Engineer in the Woodruff School of Mechanical Engineering, Georgia Tech, since 1989. His research interests include acoustic transducer design, structural acoustics, acoustical oceanography, bioacoustics, and seismology. He is a registered Professional Engineer in the state of Georgia.

Mr. Martin is a member of the Acoustical Society of America.



Gregg D. Larson [redacted] He received the B.E. degree from Vanderbilt University, Nashville, TN, in 1988, and the M.S.M.E. and Ph.D. degrees from the Georgia Institute of Technology (Georgia Tech), Atlanta, in 1990 and 1996, respectively.

From 1990 to 1996, he was a Graduate Research Assistant and Graduate Teaching Assistant with the Woodruff School of Mechanical Engineering, Georgia Tech. Since 1996, he has been a Research Engineer II in the Woodruff School of Mechanical Engineering. His research interests include acoustic transduction, bioacoustics, active vibration control, smart materials, and seismic acoustics.

Dr. Larson is a member of the Acoustical Society of America and the American Society of Mechanical Engineers.

Elastic Waves Interacting with Buried Land Mines: a Study Using the FDTD Method

Christoph T. Schröder, *Student Member, IEEE*, Waymond R. Scott, Jr., *Member, IEEE*, and Gregg D. Larson

Abstract— **A three-dimensional finite-difference model for elastic waves in the ground has been developed and implemented. The model has been created to supplement the development of a sensor that uses elastic waves to detect buried land mines. The model is used to investigate the propagation characteristics of elastic waves in the ground and to explore the interaction of elastic waves with buried land mines. When elastic waves interact with a buried mine, a strong resonance occurs at the mine location. The resonance can be used to enhance the mine's signature and to distinguish the mine from clutter. Results presented in this paper explain the features of elastic wave propagation in the ground and show the interaction of elastic waves with both an anti-personnel mine and an anti-tank mine.**

Keywords— **land mine, elastic, FDTD, finite-difference**

I. INTRODUCTION

With over 120 million land mines deployed in over 60 countries of the world, causing an estimated 26,000 injuries and deaths each year, the problem of buried land mines has gained strong public attention in recent years. Various methods to detect buried land mines have been proposed and are being investigated. Recently, a sensor type has been developed that is mainly aimed at detecting land mines with a low-metal content, which are almost invisible to conventional ground-penetrating radar sensors or metal detectors. In this technique, buried land mines are located by synergistically using elastic and electromagnetic waves [1], [2]. Elastic waves propagating in the ground interact with a buried land mine and cause the mine and the surface above the mine to vibrate. An electromagnetic radar detects the vibrations and, thus, locates the mine. To better understand the interaction of the elastic waves with the buried land mines, a numerical model has been developed that simulates the elastic wave propagation in the ground. In this paper, the numerical model and its application are described.

The numerical model is based on the *finite-difference time-domain* (FDTD) method. Over the years, various finite-difference schemes for elastic modeling have been proposed and have frequently been employed [3] – [7]. The finite-difference model as presented here uses a first-order formulation, first introduced by Madariaga [8] and Virieux [9], [10]. It incorporates a perfectly matched layer absorbing boundary condition based on the formulation by Chew [11] and the well-known free-surface boundary condition [12] – [15]. For many years, the FDTD method has been applied to a wide range of different problems in the field of elasticity. For example, recently it has been used to model elastic waves in a bore hole environment [16], to simulate ultrasonic waves in human tissue or to compute elastic wave propagation in a porous medium [17], [18].

The finite-difference model that is presented in this paper has been implemented in two and three dimensions. The equation of motion and the stress-strain relation, together with a constitutive relation, form a set of first-order partial differential equations that completely describe the elastic wave motion in a solid medium. Introducing finite differences, this set of equations can be discretized and adapted to the FDTD modeling scheme. For the model, the solution space is discretized and a staggered finite-difference grid is introduced. The finite-difference model has been implemented in a fully parallel fashion.

The numerical model has been very helpful for explaining and understanding the experimental results. The numerical model bears several advantages over the experimental model. With the numerical model, it is possible to visualize the elastic wave motion within the ground, whereas with the experimental model only the wave fields on the surface can be observed. Furthermore, no noise is present in the numerical model, and the material parameters can be adjusted easily.

In Sec. II, the theoretical foundation of the model will be presented briefly. In Sec. III, results obtained with the model will be described. First, the waves that arise at the surface are explained. Then, the interaction of elastic waves with buried land mines is described as a function of burial depth for both an anti-personnel mine and an anti-tank mine.

II. THREE-DIMENSIONAL NUMERICAL MODEL

To investigate the interactions of the elastic waves with the buried mines, a three-dimensional finite-difference model has been developed. In this paper, the results of the model will be compared to experiments, which have been performed with mines buried in a large sand tank [1], [2]. In these experiments, elastic waves are excited by an electrodynamic transducer placed on the surface of the ground. The waves propagate within the ground and along its surface and interact with the buried land mines.

This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0448, by a grant from the US Office of Naval Research under contract N00014-99-1-0995, and by an equipment grant from the Intel Corporation.

C. T. Schröder and W. R. Scott, Jr. are with the School of Electrical and Computer Engineering and G. D. Larson is with the School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Ga 30332.

A. Finite-Difference Model

Fig. 1 (a) shows the three-dimensional finite-difference model. To reasonably simplify the model, the soil is assumed to be linear, isotropic and lossless. The surface of the ground is modeled as a free-surface, and a *Perfectly Matched Layer* (PML) terminates the solution space at the remaining outer grid faces.

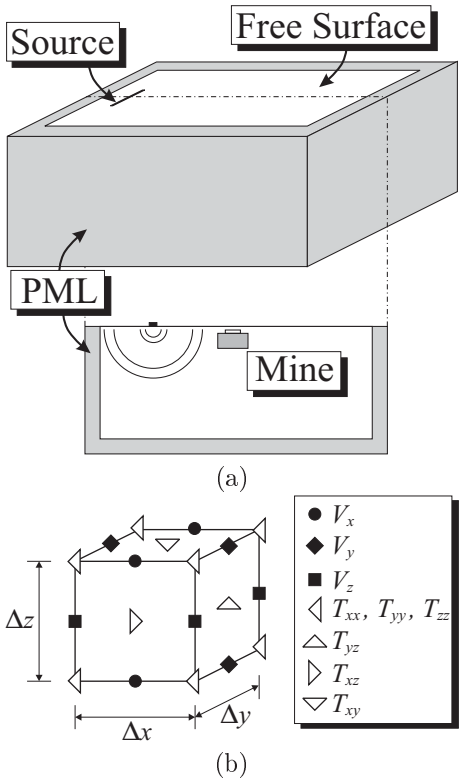


Fig. 1. Three-dimensional finite-difference model; (a) lay-out, (b) finite-difference cell.

The elastic wave motion in solids is described by a set of fundamental partial-differential equations: the equation of motion relating the particle velocity vector and the mechanic stress tensor, the strain-velocity relation and the elastic constitutive relation. Combining these equations, a first-order system of equations is obtained describing the elastic wave fields entirely in terms of the particle velocity and the mechanic stress. In three dimensions, three unknown velocity components and six unknown stress tensor components arise.

For the numerical finite-difference model, the derivatives of the partial-differential equations are approximated by central finite-differences. The finite-difference algorithm shall be explained by deriving the update equations for the x-component of the particle velocity, v_x, and for the longitudinal stress component, τ_{xx}. Update equations for the other field components can be derived in a similar manner.

The first-order system of equations consists of nine linear independent equations for the three unknown velocity vector components v_x, v_y and v_z, and the six stress tensor components τ_{xx}, τ_{yy}, τ_{zz}, τ_{yz}, τ_{xz} and τ_{xy}. The equations can be written out as

$$\rho\frac{\partial v_x}{\partial t} = \frac{\partial \tau_{xx}}{\partial x}+\frac{\partial \tau_{xy}}{\partial y}+\frac{\partial \tau_{xz}}{\partial z} \quad (1)$$

$$\rho\frac{\partial v_y}{\partial t} = \frac{\partial \tau_{xy}}{\partial x}+\frac{\partial \tau_{yy}}{\partial y}+\frac{\partial \tau_{yz}}{\partial z} \quad (2)$$

$$\rho\frac{\partial v_z}{\partial t} = \frac{\partial \tau_{xz}}{\partial x}+\frac{\partial \tau_{yz}}{\partial y}+\frac{\partial \tau_{zz}}{\partial z} \quad (3)$$

$$\frac{\partial \tau_{xx}}{\partial t} = (\lambda+2\mu)\frac{\partial v_x}{\partial x}+\lambda\frac{\partial v_y}{\partial y}+\lambda\frac{\partial v_z}{\partial z} \quad (4)$$

$$\frac{\partial \tau_{yy}}{\partial t} = \lambda\frac{\partial v_x}{\partial x}+(\lambda+2\mu)\frac{\partial v_y}{\partial y}+\lambda\frac{\partial v_z}{\partial z} \quad (5)$$

$$\frac{\partial \tau_{zz}}{\partial t} = \lambda\frac{\partial v_x}{\partial x}+\lambda\frac{\partial v_y}{\partial y}+(\lambda+2\mu)\frac{\partial v_z}{\partial z} \quad (6)$$

$$\frac{\partial \tau_{yz}}{\partial t} = \mu(\frac{\partial v_y}{\partial z}+\frac{\partial v_z}{\partial y}) \quad (7)$$

$$\frac{\partial \tau_{xz}}{\partial t} = \mu(\frac{\partial v_x}{\partial z}+\frac{\partial v_z}{\partial x}) \quad (8)$$

$$\frac{\partial \tau_{xy}}{\partial t} = \mu(\frac{\partial v_x}{\partial y}+\frac{\partial v_y}{\partial x}), \quad (9)$$

where λ and μ are *Lame's constants* and describe the stiffness of the isotropic solid, and ρ is the material density of the medium.

Taking Eq. (1) and Eq. (4), the update equations for the particle velocity v_x and the longitudinal stress τ_{xx} are derived. Introducing finite-differences in space and time, the partial-differential equations are discretized:

$$\rho\frac{V_x^{l+0.5}|_{i,j-0.5,k-0.5}-V_x^{l-0.5}|_{i,j-0.5,k-0.5}}{\Delta t} = \frac{T_{xx}^l|_{i+0.5,j-0.5,k-0.5}-T_{xx}^l|_{i-0.5,j-0.5,k-0.5}}{\Delta x} + \frac{T_{xy}^l|_{i,j,k-0.5}-T_{xy}^l|_{i,j-1,k-0.5}}{\Delta y} + \frac{T_{xz}^l|_{i,j-0.5,k}-T_{xz}^l|_{i,j-0.5,k-1}}{\Delta z}, \quad (10)$$

$$\frac{T_{xx}^{l+1}|_{i+0.5,j-0.5,k-0.5}-T_{xx}^l|_{i+0.5,j-0.5,k-0.5}}{\Delta t} = (\lambda+2\mu)\frac{V_x^{l+0.5}|_{i+1,j-0.5,k-0.5}-V_x^{l+0.5}|_{i,j-0.5,k-0.5}}{\Delta x} + \lambda\frac{V_y^{l+0.5}|_{i+0.5,j,k-0.5}-V_y^{l+0.5}|_{i+0.5,j-1,k-0.5}}{\Delta y} + \lambda\frac{V_z^{l+0.5}|_{i+0.5,j-0.5,k}-V_z^{l+0.5}|_{i+0.5,j-0.5,k-1}}{\Delta z}. \quad (11)$$

Here, the capital letters mark the numerical values of the field components at the discrete locations in space and time. For example, $V_x^{l+0.5}|_{i,j-0.5,k-0.5}$ stands for the numerical value of the particle velocity v_x at $(x,y,z) = (i\Delta x,(j-0.5)\Delta y,(k-0.5)\Delta z)$ at time $t=(l+0.5)\Delta t$. Knowing $V_x^{l-0.5}$, T_{xx}^l,T_{xy}^l and T_{xz}^l, Eq. (10) can be solved for $V_x^{l+0.5}$, i. e., for the particle velocity at the incremented time $t=(l+0.5)\Delta t$:

$$V_x^{l+0.5}|_{i,j-0.5,k-0.5} = V_x^{l-0.5}|_{i,j-0.5,k-0.5}+\frac{\Delta t}{\rho}\cdot$$

$$\cdot \left[\frac{T_{xx}^{l}|_{i+0.5,j-0.5,k-0.5} - T_{xx}^{l}|_{i-0.5,j-0.5,k-0.5}}{\Delta x} + \frac{T_{xy}^{l}|_{i,j,k-0.5} - T_{xy}^{l}|_{i,j-1,k-0.5}}{\Delta y} + \frac{T_{xz}^{l}|_{i,j-0.5,k} - T_{xz}^{l}|_{i,j-0.5,k-1}}{\Delta z} \right]. \quad (12)$$

Similarly, T_{xx}^{l+1} is obtained from Eq. (11):

$$T_{xx}^{l+1}|_{i+0.5,j-0.5,k-0.5} = T_{xx}^{l}|_{i+0.5,j-0.5,k-0.5} + \Delta t \cdot \left[(\lambda + 2\mu) \frac{V_x^{l+0.5}|_{i+1,j-0.5,k-0.5} - V_x^{l+0.5}|_{i,j-0.5,k-0.5}}{\Delta x} + \lambda \frac{V_y^{l+0.5}|_{i+0.5,j,k-0.5} - V_y^{l+0.5}|_{i+0.5,j-1,k-0.5}}{\Delta y} + \lambda \frac{V_z^{l+0.5}|_{i+0.5,j-0.5,k} - V_z^{l+0.5}|_{i+0.5,j-0.5,k-1}}{\Delta z} \right]. \quad (13)$$

In the same manner, discrete update equations can be obtained for all field components [10].

The discretization leads to the characteristic finite-difference grid. The discrete grid is distributed over the entire solution space, and the elastic field values are calculated only at the discrete grid nodes. The finite-difference grid can be thought of as being comprised of basis cells. One of these basis cells is shown in Fig. 1 (b). Note that, due to the central finite-difference approximation, the field components are not known at the same points in space and time. The grid is laid out such that each field component is surrounded by the field components it is dependent on.

When implementing the finite-difference scheme, boundary conditions have to be treated in a special manner. Three different kinds of boundaries arise: the source, internal boundaries (i. e., boundaries within the medium marked by a change in material properties), and external boundaries (i. e., the outer grid faces).

In the experimental model, an electrodynamic transducer placed on the surface launches the elastic waves. The transducer foot has the shape of a bar. In the numerical model, the transducer is emulated by exciting the particle velocity component normal to the surface, v_z, throughout an area equivalent to the area covered by the foot. The motion of the transducer foot has been measured using accelerometers and resembles closely the shape of a differentiated Gaussian pulse.

The conditions at internal boundaries, i. e., at the interfaces between different media, are usually satisfied implicitly. However, to ensure numerical stability, the material properties have to be averaged for components on the boundary. While the material density ρ, appearing in the equation of motion, is averaged directly, the inverse of Lame's constants, λ and μ from the stress-strain relation, must be averaged. A detailed study describing the necessity of the averaging can be found in [19].

Six external boundaries arise at the six outer grid faces. On its upper face, the half-space is bounded by a free surface. Due to the continuity of normal stress, the normal stress components vanish at a free surface. In order to satisfy this condition, an extra row must be inserted into the finite-difference grid one step beyond the free-surface boundary. In order to model the infinite half-space, all waves that are reaching the remaining outer grid edges must be perfectly transmitted and absorbed. The boundary condition that does this most accurately is the Perfectly Matched Layer boundary condition, first introduced by Berenger [20] and adapted to elastodynamics by Chew and Liu [11]. Here, a non-physical splitting of the wave fields allows the introduction of a lossy boundary layer that is perfectly matched to the solution space. In continuous space, it has been shown that an arbitrarily polarized wave incident on this PML is perfectly transmitted. In the PML, the wave experiences the exact same phase velocity and wave impedance as in the solution space, while rapidly decaying along the axis normal to the PML-medium interface. In discrete space, however, the PML will not be matched perfectly to the solution space. To keep reflections at the interface small, a tapered loss profile is chosen for the PML. For all results presented in this paper, a thickness of ten cells for the PML has been assumed.

B. Parallelization

The finite-difference model has been implemented in a fully parallel fashion. When implemented on a parallel computer, the three-dimensional finite-difference grid is divided into several sub-grids, and each sub-grid is assigned to one processor of the parallel machine. The processors compute the wave fields on their sub-grids and share only the field values on the common interfaces with their neighbors. The finite-difference model has been implemented to run both on a Cray T3E supercomputer with 544 processors (located at the ERDC Massively Shared Resource Center in Vicksburg, Miss.) and on a PC computer cluster located at the Georgia Institute of Technology. The PC cluster contains 50 Pentium III processors and 6 GBytes of memory.

III. INTERACTION OF ELASTIC WAVES WITH BURIED LAND MINES

In laboratory experiments, inert land mines are buried in a large sand-filled tank. Elastic waves are excited, and the displacement of the surface above the mine is measured using a radar system. To analyze the experimental results, the experiments are duplicated with the numerical model. In this section, results obtained with the numerical model are presented. First, the numerical model is used to explore the propagation of the elastic waves along the surface of the ground. Second, the interaction of elastic waves with buried anti-personnel and anti-tank mines is computed numerically and the results are compared to experimental observations.

For the numerical model, the soil is assumed to be linear, isotropic and lossless. In that case, the elastic properties of the soil can be completely described by three independent parameters: the material density, the shear wave speed, and the pressure wave speed. In order to model the soil and its behavior accurately within the numerical model, these

parameters have been measured as a function of depth in the sand tank in which the experiments are conducted.

The wave speeds in the ground have been estimated by burying a vertical array of accelerometers. By using the accelerometers and displacing the transducer, the arrival times of the wave fronts as a function of depth and distance are determined. From the arrival times, the wave speeds of the different waves in the ground can be estimated.

The material density, the pressure wave speed and the shear wave speed as a function of depth are shown in Fig. 2. All three quantities have been found to be depth dependent. The material density has a value of roughly 1420 kg/m^3 at the surface and increases to about 1670 kg/m^3 at a depth of half a meter. The variation in the material density is due to changes in the water content within the soil. At the surface, the sand is fairly dry. Within the ground, the water content increases steadily, until the water table is reached at a depth of about half a meter. The pressure wave speed and the shear wave speed both also vary with depth. At the surface, a layer with fairly slow wave speeds exists. Beyond this layer, the pressure wave speed and the shear wave speed increase rapidly. At a depth of half a meter, the pressure wave speed and the shear wave speed have values of 210 m/s and 130 m/s, respectively.[1] The depth variation of the wave speeds is influenced by several mechanisms. The cohesion between the grains of the sand, for example, is affected by the water content, because the water binds the grains of the sand. The cohesion is also altered by the pressure within the ground, which, due to gravity, increases with depth. The increase in the cohesion will cause an increase in the wave speeds. On the other hand, the increase in the material density with depth would, if the stiffness was constant, reduce the wave speeds (see also [21]). It must be noted that, although the measured depth profiles are fairly good approximations of the actual depth profiles in the soil, they are only approximations and the actual depth profile might be more complex. The material properties of all other materials used in the following are summarized in Table I.

TABLE I
PARAMETERS USED FOR FINITE-DIFFERENCE SIMULATION.

	Shear Wave Speed c_s [m/s]	Pressure Wave Speed c_p [m/s]	Material Density ρ [kg/m^3]
Plastic	1100	2700	1200
Air	0	330	1.3
Rubber	500	800	1000

For the results that are presented in this paper, the spacing between the nodes of the discrete finite-difference grid is chosen to be $\Delta x = \Delta y = \Delta z = 0.5$ cm; the time step is $\Delta t = 1.06$ μs and, thus, fulfills the Courant condition (the necessary condition for stability of the finite-difference algorithm). The excitation has the shape of a differentiated Gaussian pulse with a center frequency of 450 Hz and resembles closely the shape of the excitation in the experiment.

[1]The material properties are measured only to a depth of half a meter and are assumed to be constant deeper in the ground. This is a vague assumption and probably not true within the water table, but is justifiable, because, within the frequency range of interest, the variation of the material properties deeper in the ground will not have a significant impact onto the wave propagation at and close to the surface.

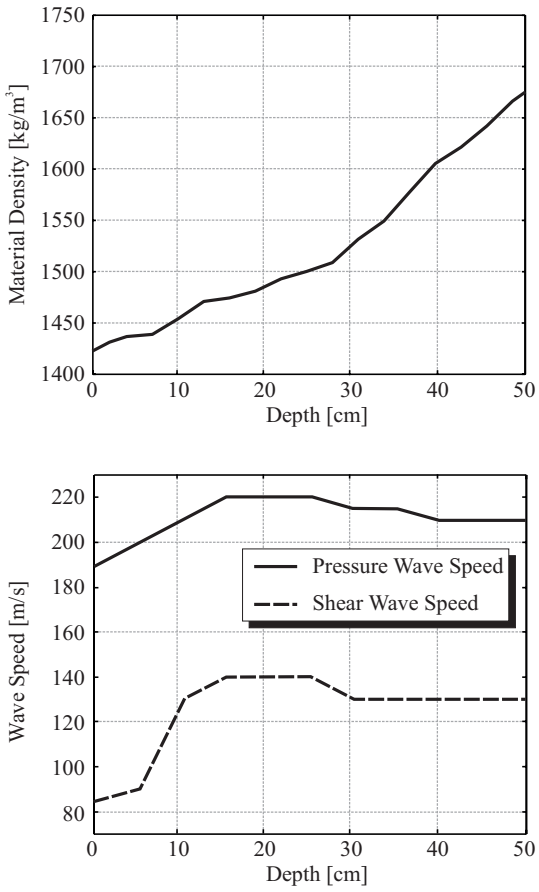


Fig. 2. Variation of the material properties with depth.

A. Surface Waves

Using the numerical model, the elastic wave fields in the ground are computed. The simulation space spans a surface area of 1.67 m by 0.8 m, with a depth of 0.6 m. Fig. 3 shows waterfall graphs of the vertical particle displacement at the surface as measured experimentally and as determined numerically. In these graphs, the displacement at a number of points along a line on the surface is plotted as a function of time and offset by the distance from the source. The source is located at $x = 0$ m. Due to the fact that distance is plotted vs. time, the slopes of the traveling waves in the graph indicate the wave speeds. Thus, by determining the slope, the different wave types can be distinguished.

Both in the experiment and in the numerical model, three types of waves are seen to propagate, indicated by LS, S and R in Fig. 3. The fastest wave of the three (LS) decays quickly. Its speed is determined to be approximately 180 m/s. The second-fastest wave (S) has a speed of roughly 120 m/s and is fairly weak on the surface.

The slowest wave (R), on the other hand, barely decays, indicating that its energy is confined to the surface. This wave corresponds to the Rayleigh surface wave. From the graph, the average speed of the Rayleigh wave is determined to be approximately 80 m/s. Clearly, the Rayleigh wave is dispersed as it propagates, an effect caused by the soil being inhomogeneous. The agreement of experiment and numerical simulation is fairly good. The authors believe that the differences are mostly due to inaccuracies in the depth profile of the material properties that is used for the numerical model and to non-linear effects occurring in the sand. The non-linearities dampen the high-frequency components and, in this way, cause the pulse to broaden and to change in shape. By decreasing the amplitude of the excitation in the experiment, the agreement of experiment and numerical simulation can be improved. However, the signal-to-noise ratio of the radar output in the experiment will also decrease and, thus, the noise content of the data will increase. Thus far, the non-linear effects are not incorporated into the numerical model.

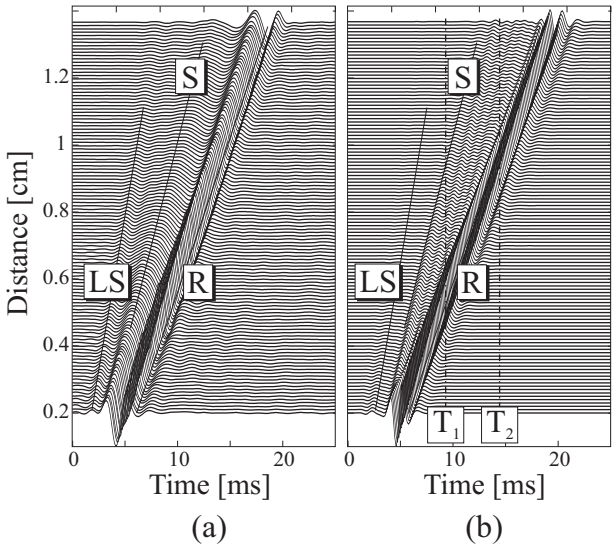


Fig. 3. Waterfall graphs of the vertical particle displacement on the surface according to (a) experiment and (b) numerical simulation.

The origin of the waves that are visible in Fig. 3 becomes clear, if the particle displacement beneath the surface is analyzed. Fig. 4 shows the vertical particle displacement on the surface and on a cross section through the ground as computed with the numerical model. Pseudo-color plots are used to represent the wave fields, with a logarithmic scale ranging from red (0 dB) to blue (-50 dB). The wave fields are shown at two instants in time, T_1 and T_2, corresponding to the vertical lines in Fig. 3 (b). The source is located on the surface, to the left of the plot. The top edge of the cross section coincides with the surface. At T_1, a Rayleigh surface wave (R), a shear wave (S) and a pressure wave (P) are seen to propagate. On the cross section, an additional surface wave (LS) is observed. This surface wave travels along the surface with a wave speed faster than the shear wave, but slower than the pressure wave. Because

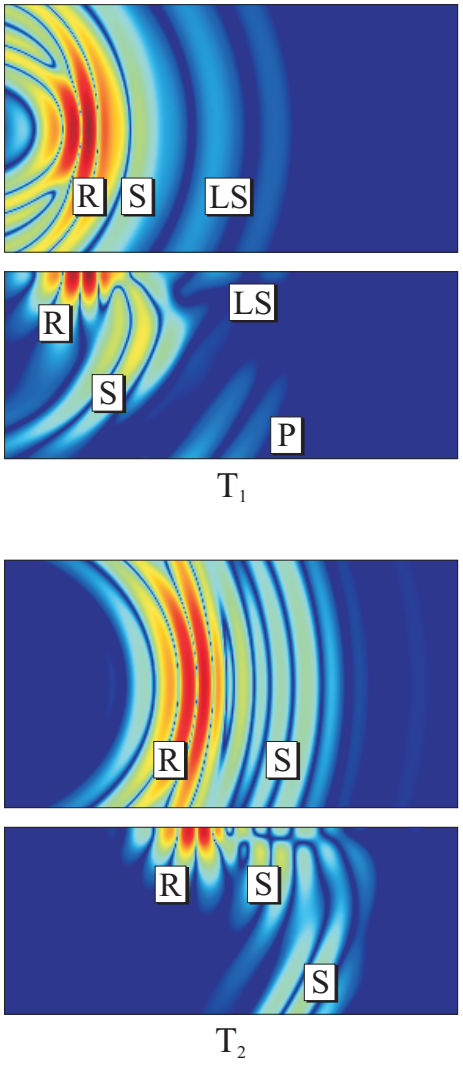


Fig. 4. Normal particle displacement on the surface (top) and on a cross section through the ground (bottom). T_1 and T_2 correspond to the vertical lines in Fig. 3 (b).

it is faster than the shear wave, the surface wave couples into a plane shear wave and, therefore,"leaks" energy into the ground. Due to these coupling effects, the surface wave decays rapidly in its propagation direction. A wave of this type is commonly called a *leaky surface wave*. On the surface plane, the leaky surface wave is clearly visible ahead of the other waves, whereas the pressure wave is not recognizable on the 50 dB scale used. This suggests that the fast wave that has been observed in Fig. 3 in both experiment and numerical simulation corresponds to the leaky surface wave. At a later time, at T_2, the waves have propagated farther. The leaky surface wave has mainly decayed, and the pressure wave has left the range of the plot. Due to the rapid increase of the shear wave speed beyond the surface layer, a portion of the shear wave is refracted back towards the surface. A guided shear wave arises, and it is this guided shear wave that is the second-fastest wave on the surface in Fig. 3. The guided shear wave is faster than the Rayleigh surface wave, but propagates at a slower speed than the shear wave within the ground.

In Fig. 5, the particle motion due to the waves on the surface is analyzed. The particle displacement is plotted at one point on the surface (60 cm from the source) as a function of time. Two hodograms are shown beneath the graph. In these hodograms, the vertical displacement is plotted versus the horizontal displacement and, thus, the actual particle motion is traced. The hodogram on the left describes the leaky surface wave, the hodogram on the right corresponds to the Rayleigh surface wave. When looking at the direction of the particle motion, it is evident that the leaky surface wave describes a prograde (clockwise) motion, whereas the Rayleigh surface wave causes a retrograde (counterclockwise) motion. The prograde particle motion is caused by the vertical displacement being ahead in phase of the horizontal displacement. The retrograde motion, on the other hand, is due to the vertical displacement lagging behind in phase. The guided shear wave, which arrives just ahead of the Rayleigh wave, has a predominantly vertical particle motion.

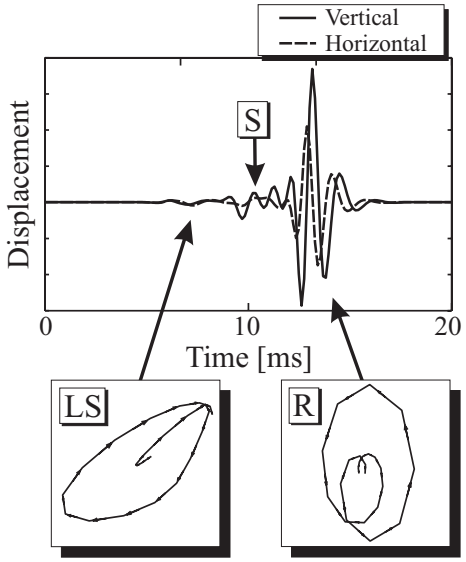


Fig. 5. Top: particle motion on the surface at a distance of 60 cm from the source as a function of time. Bottom: vertical vs. horizontal displacement for the leaky surface wave and the Rayleigh surface wave.

The leaky surface wave described above has also been observed by various other authors. For example, Roth *et al.* noticed a rapidly decaying seismic surface wave in an environment with a very high Poisson ratio that had a phase velocity larger than the Rayleigh wave, but smaller than the pressure wave [22]. Smith *et al.* identified prograde and retrograde surface wave modes in a geologic study conducted on the shore of the gulf of Mexico [23]. And Glass and Maradudin found a leaky surface wave to exist in the flat-surface limit of a corrugated crystal surface [24]. A detailed theoretical derivation of the leaky surface wave can be found in [25].

B. Interaction with Buried Land Mines

To investigate the interaction of elastic waves with a buried anti-personnel mine, a very simple model is used to approximate the mine's structure within the numerical model. Fig. 6 shows the simple mine model, together with a simplified cross sectional drawing and a photograph of a TS-50 anti-personnel mine. The simple mine model consists of a large chamber containing plastic explosives, and a small air-filled chamber. In the cubic finite-difference grid, the cylindrical shape of the mine is approximated by cubes, leading to the stair-case form evident in Fig. 6. The air-filled chamber is inserted into the model to approximate the structure of the TS-50 mine. A real land mine has a fairly complex structure, containing explosives, a trigger mechanism and several chambers.

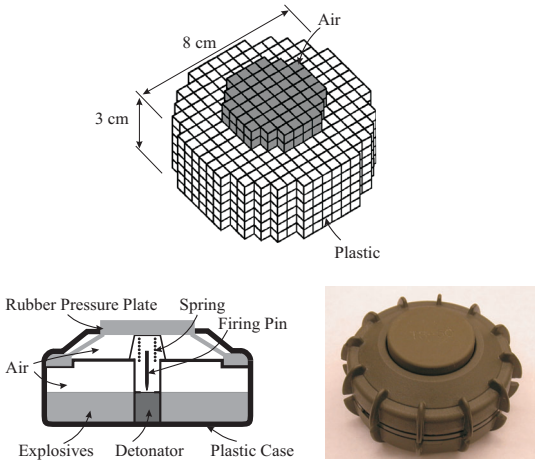


Fig. 6. Simple model for the TS-50 anti-personnel mine; cross-sectional drawing and photograph of a real TS-50 mine.

With the numerical model, the wave fields are computed throughout a solution space with a surface area of 1.67 m by 0.8 m and a depth of 0.4 m. Both in experiments and numerical simulations, the mine is placed at a distance of 97 cm from the source. The mine-wave interaction is investigated as a function of burial depth.

Fig. 7 shows pseudo-color plots of the vertical particle displacement on the surface plane (top) and on a cross section through the ground (bottom). The color scale ranges from red (0 dB) to blue (-50 dB). The wave fields are shown at four instants in time. The mine is placed 2 cm beneath the surface. At T_1, the guided shear wave (S) just hits the mine. On the cross section, the shear wave (S) and the Rayleigh wave (R) are seen. The leaky surface wave is faintly visible on the surface plane (LS). The pressure wave has already left the range of the plot. At T_2, the Rayleigh wave hits the mine. Scattered cylindrical wave fronts appear, which are induced by the incident guided shear wave. At T_3, scattered waves from both the guided shear wave and the Rayleigh wave are visible. The leading wave fronts are reflections from the incident guided shear wave. The strong waves that follow are reflections from the Rayleigh wave. At T_4, resonant oscillations are visible at the location of the buried mine. In cylindrical wave fronts, energy is radiated out. It appears that the resonant oscillations are confined to the thin soil layer above the mine. Due to these oscillations, the mine location can be clearly

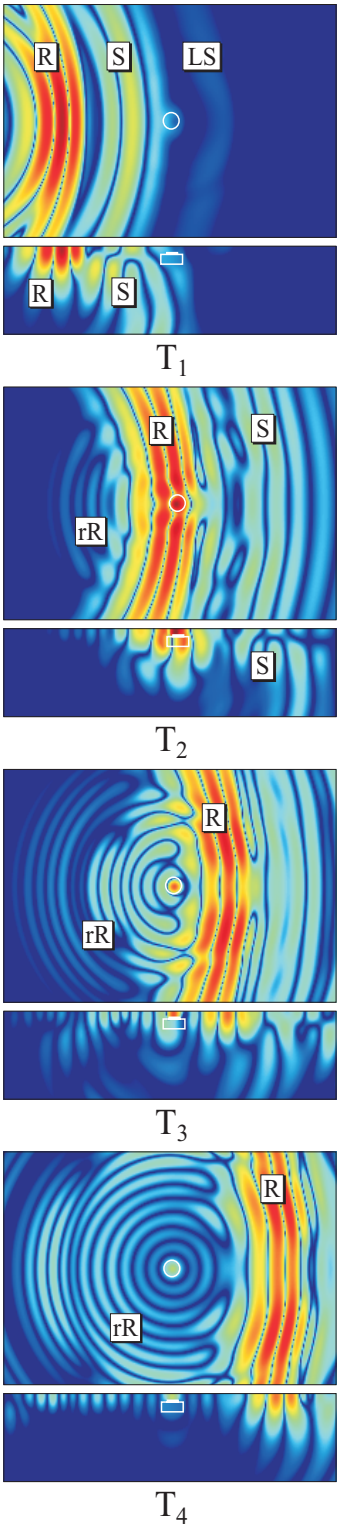


Fig. 7. Interaction of elastic waves with a buried anti-personnel mine; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instants in time.

identified on the surface.

To compare numerical to experimental results, the vertical surface displacement is plotted in a set of waterfall graphs. Fig. 8 shows these waterfall graphs for (a) experimental and (b) numerical results. Mines at four different depths are considered, at 1 cm, 2 cm, 3 cm, and 6 cm (from left to right). The vertical displacement is plotted as a function of time and vertically offset by the distance from the source. The mine, indicated by the gray shading, is located at a distance of 97 cm from the source. For the mine at 1 cm, strong oscillations are visible at the mine location. In the experiment, the amplitude of the oscillations appears larger, but the oscillations decay faster than in the numerical simulation. For the mines at 2 and 3 cm, the oscillations are weaker, but still clearly visible. For the mine at 6 cm, the oscillations are very weak and cannot be seen on this graph. However, by applying signal processing techniques the effects of the mine can be enhanced, making it easier to locate the buried mine [2].

The agreement between the numerical and experimental results is surprisingly good considering the uncertainties of the material parameters for the ground in the experimental model and the simplifications made in the numerical model. For the numerical model, it can be shown that the oscillations at the mine location are caused by the incident waves coupling into flexural waves in the thin soil layer above the mine [26]. These flexural waves are confined to the thin layer and form a standing wave pattern, giving rise to the resonant oscillations. As the burial depth of the mine is increased, the soil layer becomes thicker, causing the flexural waves to be dampened, because more energy is radiated out. The resonance in the experimental results has a more complex shape. In the experimental model, the resonance is strongly influenced by the resonant behavior of the actual buried mine. A TS-50 mine has several chambers, it has a flexible case that can support both flexural and longitudinal waves, and it contains springs that can also give rise to resonances. Furthermore, the soil above the mine might act in a strongly non-linear way, thus altering the shape of the resonance.

To compare the interaction of the elastic waves with a buried mine to, for example, the interaction with a buried rock, a solid body is inserted into the numerical model. The body has a spherical shape and a diameter of 8 cm. Its properties are assumed to be the same as the ones of plastic (see Table I). The results are shown in Fig. 9, for the same depths as previously for the buried mine. When the incident waves interact with the solid sphere, reflections occur which have about the same strength and shape as the reflections that have been observed for the buried anti-personnel mine. However, no continuous oscillations are visible at the location of the sphere, thus clearly distinguishing the solid body from the buried mine. Very similar results have also been observed experimentally.

Next, the interaction with a large anti-tank mine is to be analyzed. Fig. 10 shows the simple model for the anti-tank mine used in the numerical simulation as well as a photograph of a VS 1.6 mine. The mine has a diameter of about 23 cm and, thus, is substantially larger than the anti-personnel mine. The simple mine model contains three parts: a large plastic body, an air-filled chamber, and a rubber plate simulating the pressure plate of the mine.

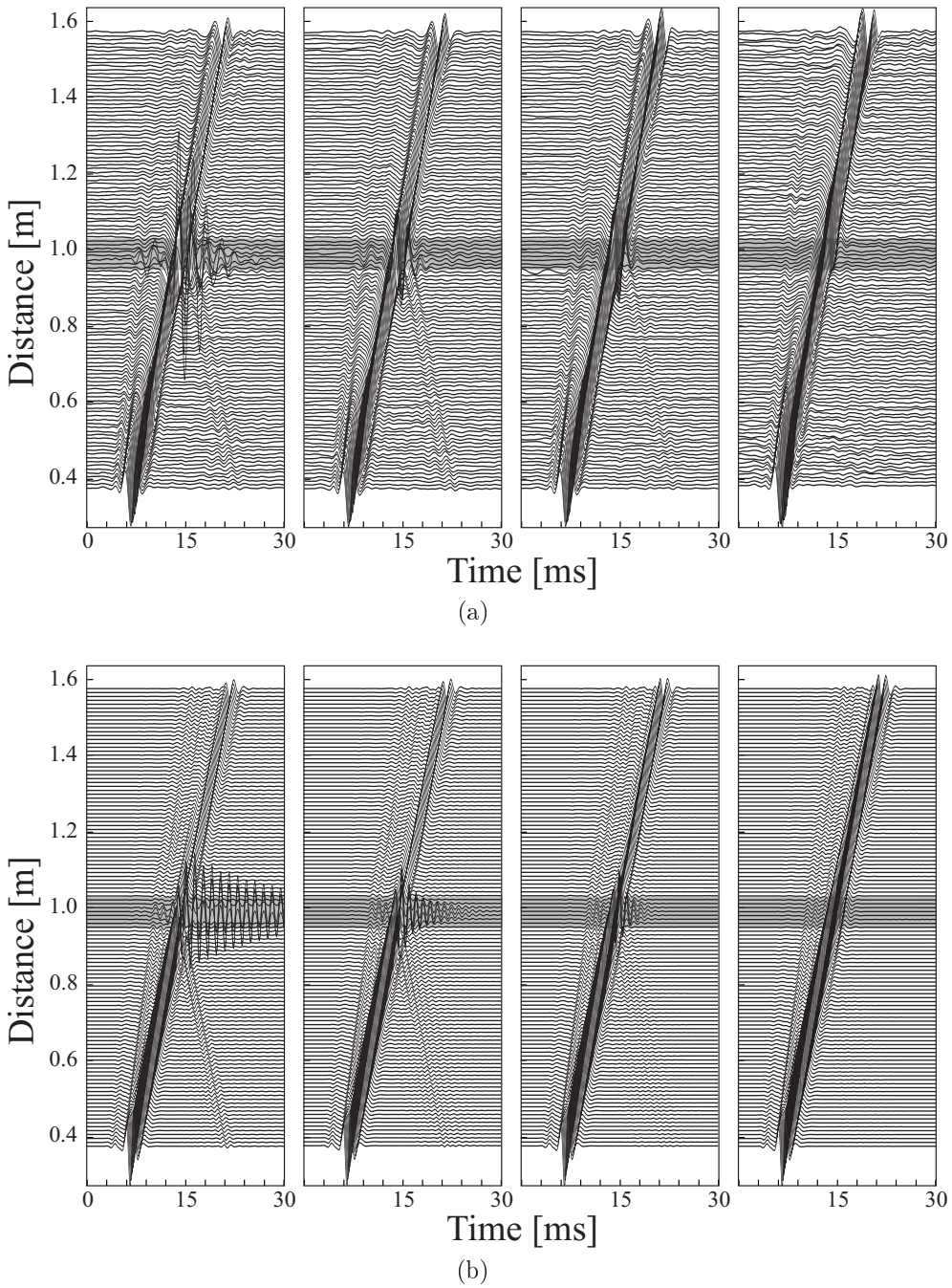


Fig. 8. (a) Experimental results for a buried TS-50 anti-personnel mine and (b) numerical results for a simple anti-personnel mine model; waterfall graphs of the vertical particle displacement on the surface. Land mine buried at 1 cm, 2 cm, 3 cm, and 6 cm (from left to right).

Fig. 11 shows the results according to experimental measurements and numerical simulations. The center of the mine is placed at a distance of 77 cm from the source. The location of the mine is indicated by the gray shading. The mines are buried at four depths: at 2 cm, 4 cm, 6 cm and 10 cm (from left to right). For the mine at 2 cm and 4 cm, the interaction with the elastic waves is strong. At the mine location, a complex pattern of waves traveling back and forth above the mine is visible. This can be explained by using the argument given before to interpret the oscillations observed for an anti-personnel mine. The incident wave partially couples into flexural waves in the thin soil layer above the mine. For the anti-personnel mine, these flexural waves form a standing wave pattern. For the larger anti-tank mine, however, the flexural waves propagate in the thin layer, are reflected at the edge of the mine, propagate back, are again reflected at the other edge of the mine, and so forth. Due to the large size of the anti-tank mine, the waves do not form standing waves, but are propagating back and forth above the mine. For the anti-tank mine at

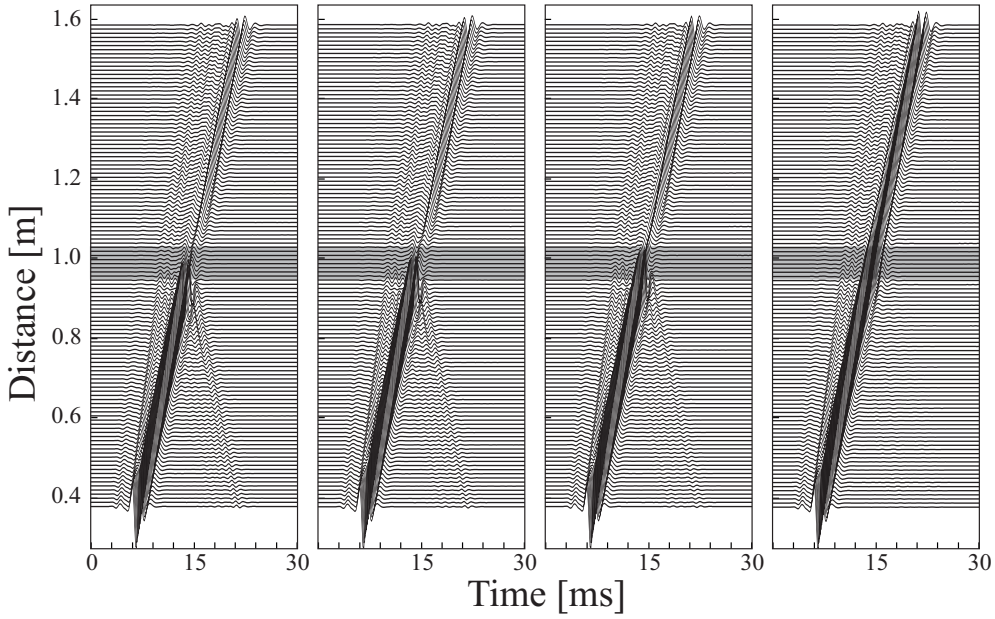


Fig. 9. Numerical results for a buried rock; waterfall graphs of the vertical particle displacement on the surface. Rock buried at 1 cm, 2 cm, 3 cm, and 6 cm (from left to right).

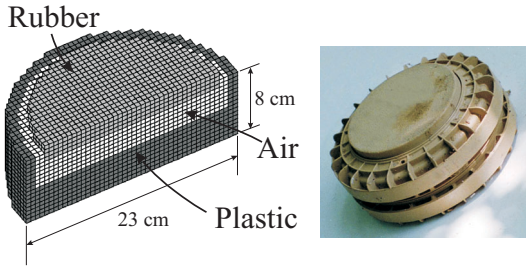


Fig. 10. Simple model for the VS 1.6 anti-tank mine; schematic picture of the model and photograph of a VS 1.6 mine.

2 cm, the flexural waves are almost completely reflected at the edges of the mine and are confined to the soil layer. For the mine at 4 cm, more energy leaks out of the thin layer and is radiated away. And for the mine at 6 cm, the waves are reflected only very weakly at the mine's edges. For the mine at 10 cm, barely any interaction of the incident waves with the buried mine is visible. The results of experimental measurements and numerical simulations are in fairly good agreement. For the numerical results, the waves above the mine seem to travel faster. This may be due to the depth profile for the shear wave speed being especially inaccurate close to the surface, and to the sand behaving in a strongly non-linear way at the surface. However, although the mine model is very simple, the same qualitative behavior for the interaction of the elastic waves with the buried mine is observed.

IV. CONCLUSIONS

A three-dimensional finite-difference model for elastic waves in the ground has been developed and implemented in a fully parallel fashion. The model has been used to investigate the interaction of elastic waves with buried land mines. The results that are presented in this paper describe the propagation of elastic waves along the surface of the ground and the interaction of elastic waves with buried land mines. All numerical results are compared to experimental results. During the study of elastic surface wave propagation, a leaky surface wave has been observed at the surface. Contrary to the Rayleigh surface wave, the leaky surface wave describes a prograde particle motion at the surface and decays fairly quickly. When elastic waves interact with a buried mine, oscillations occur that clearly distinguish the mine from its surroundings. For an anti-personnel mine, a resonance is visible at the mine location that, for the numerical model, can be attributed to flexural waves forming a standing wave pattern in the soil layer above the mine. For an anti-tank mine, rather than forming standing waves, the flexural waves propagate back and forth in the soil layer above the mine. The numerical model and the experimental model qualitatively exhibit the same behavior.

V. ACKNOWLEDGEMENTS

The authors would like to thank Prof. Glenn S. Smith, Prof. Peter H. Rogers and Mr. James S. Martin for their advice and consideration throughout this research and for their help in obtaining support for the project.

REFERENCES

[1] W. R. Scott, Jr., G. D. Larson, and J. S. Martin, "Simultaneous use of elastic and electromagnetic waves for the detection of buried land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets V, Proc. SPIE*, Apr. 2000, vol. 4038, pp. 667–678.

[2] W. R. Scott, Jr., J. S. Martin, and G. D. Larson, "Experimental model for a seismic landmine detection system," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 39, no. 6, pp. 1155–1164, Jun. 2001.

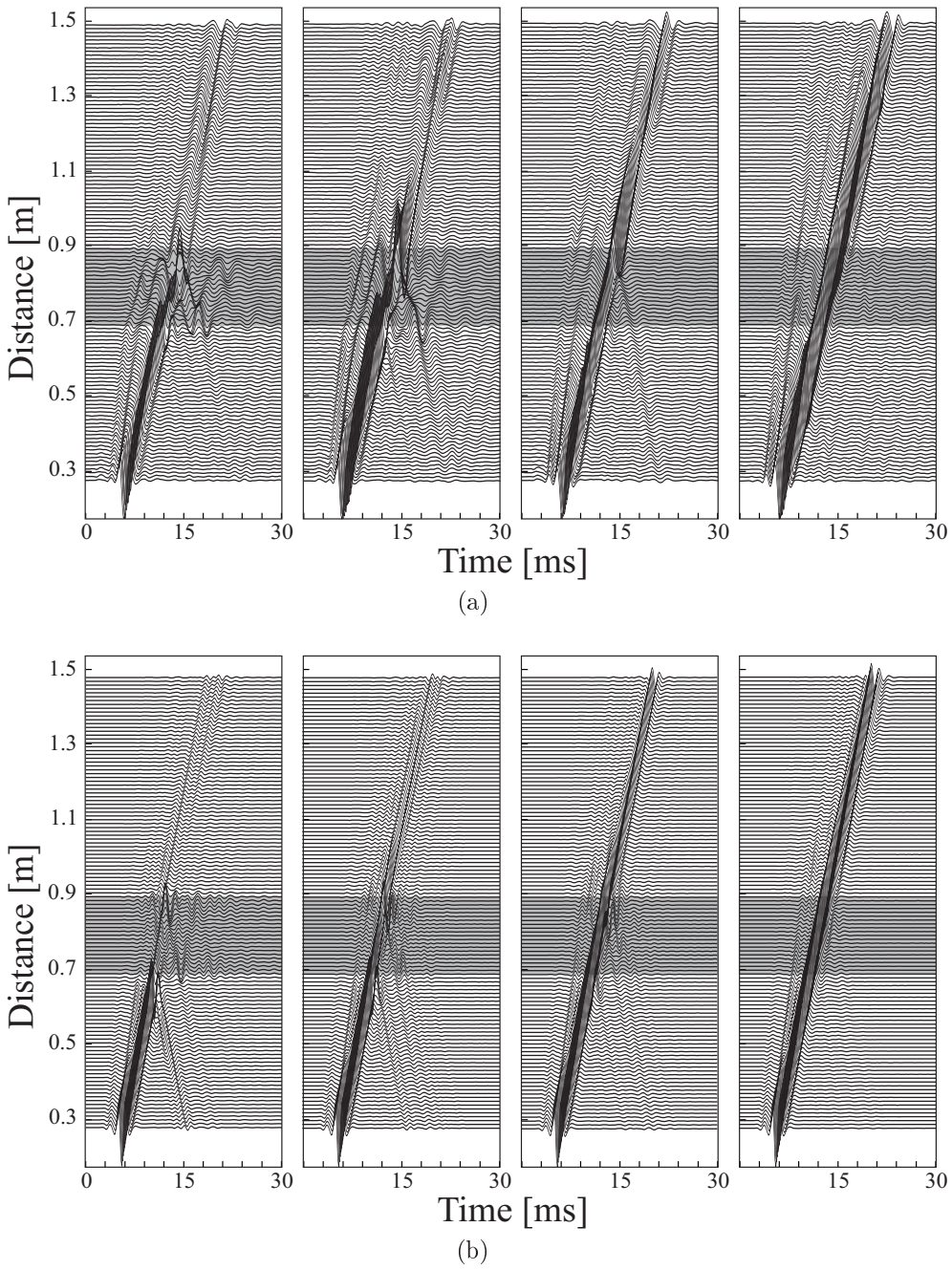


Fig. 11. (a) Experimental results for a buried VS 1.6 anti-tank mine and (b) numerical results for a simple anti-tank mine model; waterfall graphs of the vertical particle displacement on the surface. Land mine buried at 2 cm, 4 cm, 6 cm, and 10 cm (from left to right).

[3] Z. Alterman and F. C. Karal, Jr., "Propagation of elastic waves in layered media by finite-difference methods," *Bulletin Seismological Society of America*, vol. 58, no. 1, pp. 367–398, Feb. 1968.

[4] M. Ottaviani, "Elastic wave propagation in two evenly-welded quarter-spaces," *Bulletin Seismological Society of America*, vol. 61, no. 2, pp. 1119–1152, 1971.

[5] K. R. Kelly, R. W. Ward, S. Treitel, and R. M. Alford, "Synthetic seismograms: a finite-difference approach," *Geophysics*, vol. 41, no. 1, pp. 2–27, Feb. 1976.

[6] B. Kummer and A. Behle, "Second-order finite-difference modeling of SH-wave propagation in laterally inhomogeneous media," *Bulletin Seismological Society of America*, vol. 72, no. 3, pp. 793–808, Jun. 1982.

[7] S. H. Emerman and R. A. Stephen, "Comment on 'Absorbing boundary conditions for acoustic and elastic wave equations'," *Bulletin Seismological Society of America*, vol. 73, no. 2, pp. 661–665, Apr. 1983.

[8] R. Madariaga, "Dynamics of an expanding circular fault," *Bulletin Seismological Society of America*, vol. 66, pp. 639–666, 1976.

[9] J. Virieux, "SH-wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophysics*, vol. 49, no. 11, pp. 1933–1957, Nov. 1984.

[10] J. Virieux, "P-SV wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophysics*, vol. 51, no. 4, pp. 889–901, Apr. 1986.

[11] W. C. Chew and Q. H. Liu, "Perfectly matched layer for elas-

todynamics; a new absorbing boundary condition," *J. Comput. Acoustics*, vol. 4, pp. 341–359, 1996.

[12] A. Ilan and D. Loewenthal, "Instability of finite difference schemes due to boundary conditions in elastic media," *Geoph. Prosp.*, vol. 24, no. 3, pp. 431–453, Sep. 1976.

[13] K. J. Marfurt, "Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations," *Geophysics*, vol. 49, no. 5, pp. 533–549, May 1984.

[14] S. H. Emerman, W. Schmidt, and R. A. Stephen, "An implicit finite-difference formulation of the elastic wave equation," *Geophysics*, vol. 47, no. 11, pp. 1521–1526, Nov. 1982.

[15] J. E. Vidale and R. W. Clayton, "A stable free-surface boundary condition for two-dimensional elastic finite-difference wave simulation," *Geophysics*, vol. 51, no. 12, pp. 2247–2249, Dec. 1986.

[16] B. K. Sinha, Q.-H. Liu, and S. Kostek, "Acoustic waves in pressurized boreholes: a finite-difference formulation," *J. Geophysical Research*, vol. 101, no. B11, pp. 25,173–25,180, Nov. 1996.

[17] N. Dai, A. Vaifidis, and E. R. Kanasewich, "Wave propagation in heterogeneous, porous media: a velocity-stress, finite-difference method," *Geophysics*, vol. 60, no. 2, pp. 327–340, Apr. 1995.

[18] Y. Q. Zeng and Q.-H. Liu, "Acoustic detection of buried objects in 3-D fluid saturated porous media: Numerical modeling," *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, no. 6, pp. 1165–1173, Jun. 2001.

[19] C. T. Schröder and W. R. Scott, Jr., "On the stability of the FDTD algorithm for elastic media at a material interface," *submitted to the IEEE Trans. on Geoscience and Remote Sensing.*

[20] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Physics*, vol. 114, pp. 185–200, 1994.

[21] R. Bachrach and A. Nur, "High-resolution shallow seismic experiments in sand," *Geophysics*, vol. 63, no. 4, pp. 1225–1240, Jul. 1998.

[22] M. Roth, R. Spitzer, and F. Nitsche, "Seismic survey across an environment with very high Poisson's ratio," in *Proc.* 12^{th} *Annu. Int. Meeting, Symp. Application of Geophysics to Engineering and Environmental Problems*, 1999, pp. 49–55.

[23] E. Smith, P. S. Wilson, F. W. Bacon, J. F. Manning, J. A. Behrens, and T. G. Muir, "Measurements and localization of interface wave reflections from a buried target," *J. Acoust. Soc. Am.*, vol. 103, no. 5, pp. 2333–2343, May 1998.

[24] N. E. Glass and A. A. Maradudin, "Leaky surface-elastic waves on both flat and strongly corrugated surfaces for isotropic, nondissipative media," *J. Appl. Phys.*, vol. 54, no. 2, pp. 796–805, Feb. 1983.

[25] C. T. Schröder and W. R. Scott, Jr., "On the complex conjugate roots of the rayleigh equation: the leaky surface wave," *submitted to the J. Acoust. Soc. Am.*

[26] C. T. Schröder and W. R. Scott, Jr., "A finite-difference model to study the elastic-wave interactions with buried land mines," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 38, pp. 1505–1512, Jul. 2000.

On the Stability of the FDTD Algorithm for Elastic Media at a Material Interface

Christoph T. Schröder, *Student Member, IEEE* and Waymond R. Scott, Jr., *Member, IEEE*

Abstract

In this paper, the stability behavior of the first-order finite-difference time-domain algorithm for elastodynamics at the interface between two different materials is investigated. A necessary condition for stability is established, which, dependent on the material properties of the two media, might be more restrictive than the well-known Courant condition. It is shown that this more restrictive stability condition can be avoided if the material properties are averaged on the boundary.

Keywords

FDTD, finite-difference, elastic, stability

I. Introduction

For many years now, the finite-difference time-domain (FDTD) method has been applied to a wide range of different problems in various fields. Advances in computer technology have made it possible to handle increasingly more complex problems, thus drastically enhancing the versatility and potentiality of the FDTD method. In elastodynamics, the FDTD was introduced in the late 1960's, and due to its accurate results and its algorithmic simplicity it fast gained popularity [1] – [3]. Throughout the years, various elastic finite-difference schemes have been proposed, and the theoretical foundations of the algorithm have been a topic of continuous research ever since.

A significant question that must be considered when applying a numerical algorithm is the question of stability of the algorithm. The stability of difference equations in general and the finite-difference algorithm in particular has been investigated in much detail [4] – [14]. Various aspects of the stability of the FDTD algorithm and its different schemes have been explored, and the mathematical tools to analyze the algorithmic stability are well-established. In this paper, the stability of the first-order finite-difference formulation at an interface between two different materials is investigated.

This work is supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0448 and by a grant from the US Office of Naval Research under contract N00014-99-1-0995.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Ga 30332.

The first-order elastic finite-difference formulation was introduced by Madariaga [15] and Virieux [16], [17]. Contrary to the second-order scheme, which is based on the elastic wave equation, the first-order formulation uses a set of first-order partial differential equations, consisting of the *equation of motion*, the *strain-velocity* relation and the *elastic constitutive relation*. A staggered numerical grid is introduced in which the field components are not known at the same points in space and time. Within the first-order FDTD algorithm, internal boundaries, i. e., boundaries between different materials, are satisfied implicitly. The finite-difference algorithm is usually stable, if the Courant condition is satisfied. However, if media with greatly different material properties are considered, the algorithm turns out to be unstable at the interface and a more restrictive condition for stability arises. In this paper, this stability behavior is analyzed in terms of the difference in material properties between the two adjacent materials. A longitudinal wave normally incident onto a material interface is considered. The matrix method is employed, and a necessary condition for stability is established. It is shown that, if the material parameters are averaged properly on the boundary, the Courant condition poses a necessary condition for stability, and no further restrictions on the stability condition due to the presence of the material interface arise.

To establish the stability condition, some well-known mathematical theorems and procedures are utilized in this paper. The matrix method in conjunction with Gershgorin's Circle theorem has frequently been employed and is well-documented (see [4], [8], [10]). For example, Ilan and Loewenthal used a fairly similar procedure as presented in this paper to analyze the stability of the second-order finite-difference formulation incorporating a free-surface boundary condition [9]. This paper, although similar in its mathematical foundations, explores a quite different aspect by addressing the stability of the first-order scheme at a material interface.

This paper is structured as follows. First, a theoretical analysis is provided which indicates that a more restrictive stability condition arises at a material interface. Second, it is shown that this more restrictive condition can be avoided by averaging the material parameters on the interface. Third, the theoretical results are verified by numerical results. Finally, in an appendix, the averaging procedure is derived and explained.

II. Stability Analysis: Theory

The fundamental condition for stability of the finite-difference algorithm, i. e., the condition that relates the size of the time increment to the spacing of the discrete nodes in the FDTD grid and that must be fulfilled for the finite-difference algorithm to be stable, is the *Courant condition* [18]. The Courant condition states that the physical wave speed of a propagating wave must not exceed the

velocity by which information can travel in the discrete grid. Mathematically, for a space step of size Δl, the Courant condition in an n-dimensional grid is

$$\frac{\Delta t}{\Delta l} c_{max} \leq \frac{1}{\sqrt{n}}. \tag{1}$$

Here, Δt denotes the time step, and c_{max} is the maximum wave speed occurring in the solution space. If an infinite homogeneous medium is considered, the Courant condition poses a necessary as well as a sufficient condition for stability [4], [6], [8], [10].

In this section, a necessary condition for stability for the finite-difference algorithm in the presence of an interface between two differing materials is derived. It is shown that this condition might, dependent on the ratio of the two materials, pose a more restrictive condition on stability than the Courant condition. The analysis presented here is performed for a longitudinal wave being normally incident onto a material interface. Results for a transverse wave, however, can be obtained in an analogous way.

A. 1-D Longitudinal Wave Incident onto a Material Interface

The stability behavior of a 1-D finite-difference grid for a longitudinal wave normally incident onto a material interface is to be investigated. The propagation direction coincides with the x-direction, and the only non-zero velocity component is v_x. Figure 1 shows a portion of the discrete first-order finite-difference grid. The material interface is located between T_{xx} at $i - 0.5$ and v_x at i. The first-order finite-difference system of equations can be combined to obtain a second-order formulation in terms of solely the particle velocity. By combining the discrete first-order equations rather than discretizing the second-order wave equation, the stability behavior of the first-order finite-difference formulation is preserved. In the second-order formulation, the stress T_{xx} is eliminated, but the Lamé constants, which are associated with the stress components, are still located half a step in between the velocity components. The second-order finite-difference update equation for the longitudinal particle velocity v_x becomes

$$\begin{aligned} v_x^{k+1}(n) \;=\; & v_x^k(n)\left[2 - \frac{\Delta t^2}{\Delta x^2}\frac{\lambda(n-0.5) + 2\mu(n-0.5) + \lambda(n+0.5) + 2\mu(n+0.5)}{\rho(n)}\right] \\ & + v_x^k(n+1)\left[\frac{\Delta t^2}{\Delta x^2}\frac{\lambda(n+0.5) + 2\mu(n+0.5)}{\rho(n)}\right] \\ & + v_x^k(n-1)\left[\frac{\Delta t^2}{\Delta x^2}\frac{\lambda(n-0.5) + 2\mu(n-0.5)}{\rho(n)}\right] \\ & - v_x^{k-1}(n), \end{aligned} \tag{2}$$

where $n = 1 \ldots N$ are the nodes of the 1-D grid and k represents the discrete time step. Again, note that the Lamé constants are placed at the half steps in between the velocity nodes.

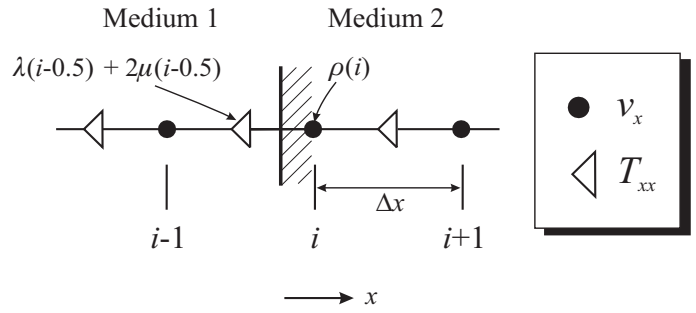


Fig. 1. 1-D finite-difference grid for a longitudinal wave normally incident onto a material boundary.

For a finite-sized problem, the finite-difference scheme can be formulated as a matrix equation [8]. Considering a one-dimensional grid with nodes $n = 1 \ldots N$, and assuming $v_x(0) = v_x(N+1) = 0$ on the outer grid edges at all times, the finite-difference scheme can be written as

$$\mathbf{v_x}^{k+1} = \mathbf{A} \cdot \mathbf{v_x}^{k} - \mathbf{I} \cdot \mathbf{v_x}^{k-1}, \tag{3}$$

where

$$\mathbf{v_x}^{k} = \begin{bmatrix} v_x^k(1) \\ \vdots \\ v_x^k(N) \end{bmatrix}$$

is a vector containing the velocity values at the N nodes at time k, and $\mathbf{I}$ is the $N \times N$ identity matrix. $\mathbf{A}$ is a tri-diagonal matrix:

$$\mathbf{A} = \begin{bmatrix} \ddots & \ddots & & & & & \\ \ddots & 2 - 2\frac{\Delta t^2}{\Delta x^2}c_1^2 & \frac{\Delta t^2}{\Delta x^2}c_1^2 & & & 0 & \\ & \frac{\Delta t^2}{\Delta x^2}c_1^2 & 2 - 2\frac{\Delta t^2}{\Delta x^2}c_1^2 & \frac{\Delta t^2}{\Delta x^2}c_1^2 & & & \\ & & \frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1+2\mu_1}{\rho_2} & 2 - \frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1+2\mu_1}{\rho_2}\right) & \frac{\Delta t^2}{\Delta x^2}c_2^2 & & \\ & & & \frac{\Delta t^2}{\Delta x^2}c_2^2 & 2 - 2\frac{\Delta t^2}{\Delta x^2}c_2^2 & \frac{\Delta t^2}{\Delta x^2}c_2^2 & \\ & 0 & & & \frac{\Delta t^2}{\Delta x^2}c_2^2 & 2 - 2\frac{\Delta t^2}{\Delta x^2}c_2^2 & \ddots \\ & & & & & \ddots & \ddots \end{bmatrix}, \tag{4}$$

where

$$c_1 = \sqrt{\frac{\lambda_1 + 2\mu_1}{\rho_1}}, \quad c_2 = \sqrt{\frac{\lambda_2 + 2\mu_2}{\rho_2}} \tag{5}$$

describe the longitudinal wave speeds in Medium 1 and Medium 2, respectively. In this matrix, rows $n = 1 \ldots i-1$ describe Medium 1, the i-th row is the material interface, and the rows $n = i+1 \ldots N$

represent Medium 2. Eq. (3) can be written in a more compact form:

$$\begin{bmatrix} \mathbf{v_x}^{k+1} \\ \mathbf{v_x}^{k} \end{bmatrix} = \mathbf{B} \cdot \begin{bmatrix} \mathbf{v_x}^{k} \\ \mathbf{v_x}^{k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{v_x}^{k} \\ \mathbf{v_x}^{k-1} \end{bmatrix}, \tag{6}$$

where each submatrix of $\mathbf{B}$ is $N \times N$ and, consequently, $\mathbf{B}$ is $2N \times 2N$.

For the finite-difference algorithm to be stable, the magnitude of all eigenvalues of $\mathbf{B}$ must be smaller than or equal to one [4], [8], [10]. However, the eigenvalues cannot be easily determined in closed form and, thus, they are estimated and a necessary condition for stability is established.

The analysis of the eigenvalues is made considerably easier considering the structure of the matrix $\mathbf{B}$. $\mathbf{B}$ consists of four blocks of size $N \times N$: $\mathbf{A}$, $\mathbf{I}$, $-\mathbf{I}$ and $\mathbf{0}$. If a matrix can be divided into $M \times M$ square sub-blocks of equal size $N \times N$, and the sub-blocks have a common set of N linearly independent eigenvectors, then the eigenvalues of the entire matrix are given by the eigenvalues of the matrices

$$\begin{bmatrix} \lambda_{11}^p & \cdots & \lambda_{1M}^p \\ \vdots & & \vdots \\ \\ \lambda_{M1}^p & \cdots & \lambda_{MM}^p \end{bmatrix}, \qquad p = 1 \ldots N \tag{7}$$

where p indicates the p-th eigenvector to each sub-block [8].

Since any vector is an eigenvector of $\mathbf{I}$ and $\mathbf{0}$, the four sub-blocks of $\mathbf{B}$ indeed have a common set of eigenvectors: all eigenvectors of $\mathbf{A}$ are also eigenvectors to the identity matrix as well as the zero matrix. The identity matrix has the N-fold eigenvalue one and the zero-matrix has the N-fold eigenvalue zero. Thus, letting λ_{A} indicate any eigenvalue of $\mathbf{A}$, the eigenvalues of $\mathbf{B}$ are given by the eigenvalues of the matrix

$$\begin{bmatrix} \lambda_{\mathrm{A}} & -1 \\ 1 & 0 \end{bmatrix}, \tag{8}$$

and are determined from the quadratic equation

$$1 - \lambda(\lambda_{\mathrm{A}} - \lambda) = 0. \tag{9}$$

The eigenvalues are then

$$\lambda = \frac{\lambda_{\mathrm{A}}}{2} \pm \sqrt{\frac{\lambda_{\mathrm{A}}^2}{4} - 1}. \tag{10}$$

Note here that all eigenvalues of $\mathbf{A}$ are real.[1]

[1]A tridiagonal matrix with either all its off-diagonal elements positive or all its off-diagonal elements negative is diagonizable and has only real eigenvalues [8].

The magnitude of all eigenvalues of $\mathbf{B}$ must be smaller than or equal to one: $|\lambda| \leq 1$. For this to be true, it can be shown from Eq. (10) that the magnitude of all eigenvalues of $\mathbf{A}$ must be smaller than or equal to two:

$$|\lambda_{\mathrm{A}}| \leq 2. \tag{11}$$

It is then sufficient to find or estimate the eigenvalues of $\mathbf{A}$ and conclude the stability condition from these.

The eigenvalues of a matrix can most easily be approximated using *Gershgorin's Circle Theorem* [4], [8], [9], [19]. Gershgorin's Circle theorem states that the eigenvalues of a matrix lie within circles in the complex plane whose centers are the elements of the matrix's main diagonal and whose radii are equal to the sum of the magnitude of the off-diagonal row elements:

$$|\lambda_n - a_{nn}| \leq \sum_{\substack{m \\ n \neq m}} |a_{nm}|. \tag{12}$$

From Gershgorin's Theorem it is seen that the eigenvalues of $\mathbf{A}$ lie within three circles:

$$\left|\lambda_{\mathrm{A}} - \left(2 - 2\frac{\Delta t^2}{\Delta x^2} c_1^2\right)\right| \leq 2\frac{\Delta t^2}{\Delta x^2} c_1^2; \tag{13}$$

$$\left|\lambda_{\mathrm{A}} - \left(2 - \frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1 + 2\mu_1}{\rho_2}\right)\right)\right| \leq \frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1 + 2\mu_1}{\rho_2}\right); \tag{14}$$

$$\left|\lambda_{\mathrm{A}} - \left(2 - 2\frac{\Delta t^2}{\Delta x^2} c_2^2\right)\right| \leq 2\frac{\Delta t^2}{\Delta x^2} c_2^2. \tag{15}$$

Eq. (13) and Eq. (15) yield the Courant condition for Medium 1 and 2, respectively. Because all eigenvalues of $\mathbf{A}$ are real, λ_{A} according to Eq. (13) will lie in the range

$$2 - 4\frac{\Delta t^2}{\Delta x^2} c_1^2 \leq \lambda_{\mathrm{A}} \leq 2. \tag{16}$$

If $|\lambda_{\mathrm{A}}| \leq 2$ is to be satisfied,

$$\frac{\Delta t^2}{\Delta x^2} c_1^2 \leq 1, \tag{17}$$

which is the Courant condition for Medium 1. Similarly, from Eq. (15), the Courant condition for Medium 2 is obtained as

$$\frac{\Delta t^2}{\Delta x^2} c_2^2 \leq 1. \tag{18}$$

Eq. (14) describes a necessary condition for stability at the node between Medium 1 and 2. For the boundary node, the eigenvalues lie in the range

$$2 - 2\frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1 + 2\mu_1}{\rho_2}\right) \leq \lambda_{\mathrm{A}} \leq 2. \tag{19}$$

This yields the stability criterion for the boundary node:

$$\frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1 + 2\mu_1}{\rho_2}\right) \leq 2. \tag{20}$$

The most restrictive of the three conditions (Eq. (17), (18) and (20)) poses the stability condition for the entire system. Note that the overall stability condition is a *necessary* condition. Numerical experiments, however, indicate that the condition is also sufficient for stability.

Eq. (20) can be written as

$$\frac{\Delta t^2}{\Delta x^2}c_2^2\left(1 + \frac{\lambda_1 + 2\mu_1}{\lambda_2 + 2\mu_2}\right) = \frac{\Delta t^2}{\Delta x^2}c_1^2\frac{\rho_1}{\rho_2}\left(1 + \frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1}\right) \leq 2. \tag{21}$$

For $c_1 > c_2$, the stability condition at the boundary will be most restrictive, if

$$\frac{\rho_1}{\rho_2}\left(1 + \frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1}\right) > 2. \tag{22}$$

On the other hand, for $c_2 > c_1$, the boundary stability criterion is most restrictive, if

$$\left(1 + \frac{\lambda_1 + 2\mu_1}{\lambda_2 + 2\mu_2}\right) > 2. \tag{23}$$

Expressed differently, the material interface does not impose an additional constraint on the stability of the finite-difference scheme, if

$$\sqrt{\frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1}} \geq 1 \qquad \text{or} \tag{24}$$

$$\sqrt{\frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1}} \leq \sqrt{2\frac{\rho_2}{\rho_1} - 1}. \tag{25}$$

This stability bound is plotted in Fig. 2 as a function of $\sqrt{(\lambda_2 + 2\mu_2)/(\lambda_1 + 2\mu_1)}$ and $\sqrt{\rho_2/\rho_1}$. The radial lines in Fig. 2 correspond to lines of constant velocity ratios. In the outer region, indicated by *Courant Region* in Fig. 2, the Courant condition is the limiting condition for stability of the finite-difference algorithm. In the region called *Boundary Region*, the stability condition at the boundary is more restrictive than the Courant condition and poses the decisive condition for stability of the finite-difference algorithm.

The same results as above are also obtained when applying the *von-Neumann* (or *Fourier-Series*) method [8], [10]. The von-Neumann method checks for the local stability at a single node in the finite-difference grid in terms of its surrounding nodes. The stability conditions obtained with the von-Neumann method for Medium 1, Medium 2 and the boundary are equivalent to Eqs. (17), (18) and (20), respectively.

Note that, if the boundary in Fig. 1 is placed in between the velocity at node i and the stress at node $i + 0.5$, the bounds on stability will be exactly the same as depicted in Fig. 2, but the ratios

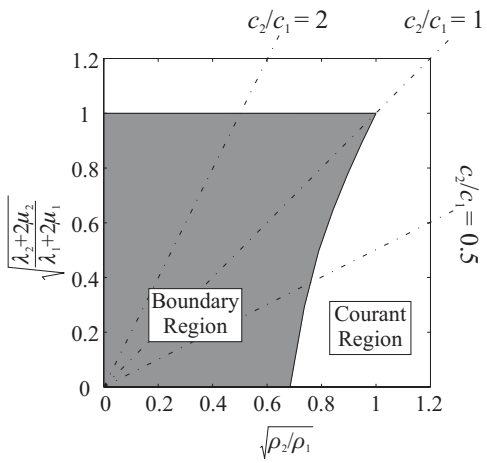


Fig. 2. Stability bounds due to the presence of a material interface.

on the axes are inverted: $\sqrt{\rho_1/\rho_2}$ on the horizontal axis and $\sqrt{(\lambda_1+2\mu_1)/(\lambda_2+2\mu_2)}$ on the vertical axis. Thus, the stability condition at the boundary node now becomes the restrictive condition for stability of the finite-difference algorithm, if the density and the stiffness in Medium 1 are *smaller* than in Medium 2. The analysis of a 1-D shear wave incident onto a material interface is analogous. In fact, by substituting μ_2/μ_1 for $(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)$ in Eq. (24) and Eq. (25), the bounds for the shear wave case are obtained.

The above description is not very intuitive but rather describes the derivation of a mathematical bound on the stability of the finite-difference algorithm at a boundary. To obtain a more intuitive picture of the actual reason why instabilities occur at the boundary, one can look at the boundary row of $\mathbf{A}$ in Eq. (4). It is seen that in the boundary row a mixed term appears which is composed of the stiffness in Medium 1 and the density in Medium 2: $(\lambda_1+2\mu_1)/\rho_2$. Note that this mixed term has the units of a squared wave speed. Thus, by comparing the value of the mixed term to the wave speeds in the general media, a reason for the instabilities can be suggested. If the longitudinal stiffness in Medium 1 is large or the density in Medium 2 is small (which, for example, is true at the interface between a solid and air), the mixed term might be quite large and, thus, the square root of its value might exceed the wave speeds of Medium 1 and Medium 2. In that case, the stability condition at the boundary can become more restrictive than the Courant condition in the general medium, and neither the Courant condition for Medium 1 nor the one for Medium 2 is a sufficient condition for stability.

B. Averaging the Material Parameters on the Interface

To ensure that the Courant condition is a sufficient condition for stability, the material parameters are averaged for the field components on the boundary. The averaging procedure is briefly described in the Appendix. For the 1-D case discussed above, the velocity component v_x is placed on the boundary

(see Fig. 1) and the material density is averaged according to Eq. (37). In the matrix equation (Eq. (4)), only the boundary row of $\mathbf{A}$ will change, and $\mathbf{A}$ for the averaged case is obtained by replacing ρ_2 by the averaged density, $(\rho_1 + \rho_2)/2$, throughout the boundary row.

By applying Gershgorin's Circle Theorem to the interface row of the averaged matrix, the eigenvalue associated with this row is determined to lie in the range

$$2 - 2\frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1 + 2\mu_1 + \lambda_2 + 2\mu_2}{(\rho_1 + \rho_2)/2} \leq \lambda_{\mathrm{A}_{avg}} \leq 2. \tag{26}$$

As shown previously, the magnitude of the eigenvalues of $\mathbf{A}$ must be smaller than two, and the stability condition is obtained as

$$\frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1 + 2\mu_1 + \lambda_2 + 2\mu_2}{\rho_1 + \rho_2} \leq 1. \tag{27}$$

This can be rewritten using the wave speeds as

$$\frac{\Delta t^2}{\Delta x^2}\frac{c_1^2\rho_1 + c_2^2\rho_2}{\rho_1 + \rho_2} = \frac{\Delta t^2}{\Delta x^2}c_1^2\frac{\rho_1 + \rho_2(c_2/c_1)^2}{\rho_1 + \rho_2} = \frac{\Delta t^2}{\Delta x^2}c_2^2\frac{\rho_1(c_1/c_2)^2 + \rho_2}{\rho_1 + \rho_2} \leq 1. \tag{28}$$

When having a closer look at this equation, it becomes clear that the stability condition for the averaged boundary is always less restrictive than the Courant condition in the general medium, and the finite-difference scheme will be stable as long as the Courant condition is satisfied. If $c_1 > c_2$, the stability condition for the entire system is

$$\frac{\Delta t^2}{\Delta x^2}c_1^2 \leq 1, \tag{29}$$

which is the Courant condition for Medium 1. From the second term in Eq. (28), it can be seen that the stability condition for the averaged boundary layer is always less restrictive than the Courant condition for Medium 1, and thus it does not pose any further constraints on the stability of the system. Similarly, if $c_2 > c_1$, the stability condition for the system is

$$\frac{\Delta t^2}{\Delta x^2}c_2^2 \leq 1, \tag{30}$$

which is the Courant condition for Medium 2. From the third term in Eq. (28), it is clear that again the stability criterion for the averaged boundary layer is always satisfied, if the Courant condition for Medium 2 is fulfilled.

C. Numerical eigenvalues of A

Rather than deriving a necessary stability condition by estimating the eigenvalues of $\mathbf{A}$ using Gershgorin's Circle theorem, the stability condition can also be determined accurately by computing the eigenvalues numerically. For this, the matrix $\mathbf{A}$ describing the same problem as previously (i. e., a 1-D

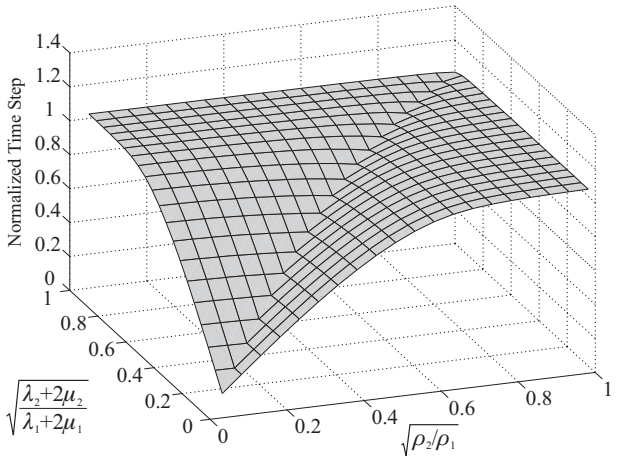
longitudinal wave and assuming the velocity to be zero on the boundaries) is created and its eigenvalues are determined. A root finder seeks the value of the time step, Δt, for which the magnitude of one or several eigenvalues just exceeds two and, thus, determines the maximum size of the time step that must not be exceeded for the finite-difference algorithm to be stable. The time step is determined for a 1-D grid with 10 nodes.

Fig. 3 shows the necessary stability condition as determined numerically. The stability condition is plotted as a function of the square root of the ratios of the material density and the longitudinal stiffness in the two media. The stability condition is normalized to the value of the time step calculated from the Courant condition in the general medium (Eq. (1)). Due to the normalization, a value of one corresponds to the value predicted by the Courant condition. Values smaller than one represent a condition for stability more restrictive than the Courant condition. In Fig. 3 (a) the material properties are not averaged and in Fig. 3 (b) the material properties are averaged for the velocity component on the interface.

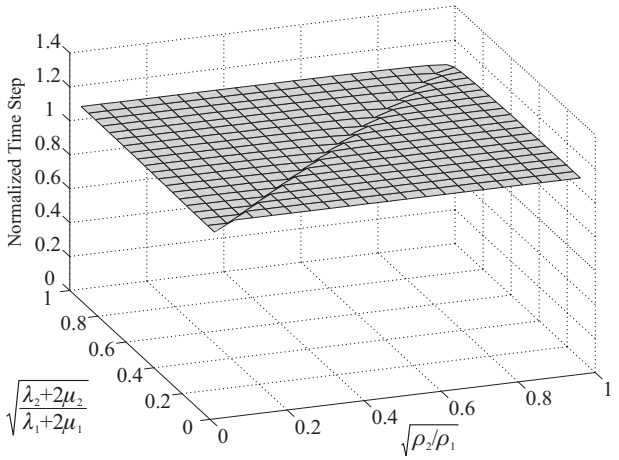
The stability condition derived from the numerically determined eigenvalues verifies the results presented earlier. If the material parameters are not averaged on the material interface, the interface will impose a constraint on the stability condition. If the density ratio $\sqrt{\rho_2/\rho_1}$ and the stiffness ratio $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)}$ fall below some critical value (≈ 1), the time step must be chosen smaller than dictated by the Courant condition in order to achieve a stable behavior. When $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)}$ approach zero, the necessary stability condition also approaches zero, and, thus, the required time step becomes infinitely small. For $\sqrt{\rho_2/\rho_1} > 1$ or $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)} > 1$, the stability condition at the boundary is less restrictive than the Courant condition. For the averaged case, the boundary does not pose an additional stability constraint, and the Courant condition is the limiting condition for stability independent of the ratios for the material density and Lamé's constants. Similar results have also been obtained for the 2-D first-order formulation.

III. Conclusions

The stability of the first-order FDTD algorithm at a material interface has been analyzed. It has been theoretically shown and verified numerically that the presence of a material interface can pose restrictions on the stability condition of the FDTD algorithm. By averaging the material properties on the interface, the restrictions can be avoided, and the well-known Courant condition will be the limiting condition for stability. A simple averaging procedure is presented in the Appendix. Note that the kind of instabilities described in this paper are not limited to only the first-order FDTD algorithm for elastodynamics. Due to the similarity of the governing equations, the same effects will also be



(a)



(b)

Fig. 3. Normalized stability condition;(a) material properties not averaged on the boundary, (b) material properties averaged on the boundary.

observed, for example, for the first-order electromagnetic FDTD scheme.

References

[1] Z. Alterman and F. C. Karal, Jr., "Propagation of elastic waves in layered media by finite-difference methods," *Bulletin Seismological Society of America*, vol. 58, no. 1, pp. 367–398, Feb. 1968.

[2] M. Ottaviani, "Elastic wave propagation in two evenly-welded quarter-spaces," *Bulletin Seismological Society of America*, vol. 61, no. 2, pp. 1119–1152, 1971.

[3] K. R. Kelly, R. W. Ward, S. Treitel, and R. M. Alford, "Synthetic seismograms: a finite-difference approach," *Geophysics*, vol. 41, no. 1, pp. 2–27, Feb. 1976.

[4] R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial Value Problems*, John Wiley and Sons, New York, 2nd edition, 1967.

[5] M. L. Buchanan, "A necessary and sufficient condition for stability of difference schemes for second-order initial value problems," *J. Soc. Indust. Appl. Math.*, vol. 11, no. 2, pp. 474–501, Jun. 1963.

[6] P. Fox, "The solution of hyperbolic partial differential equations by difference methods," in *Mathematical Methods for Digital Computers*. A. Ralston and H.S. Wilf, Editors, Ch. 16, Wiley, New York, 1964.

[7] P. D. Lax and B. Wendroff, "Difference schemes for hyperbolic equations with high order of accuracy," *Comm. Pure Applied Math.*, vol. 17, pp. 381–398, 1964.

[8] G. D. Smith, *Numerical Solution of Partial Differential Equations; Finite Difference Methods*, Clarendon Press, 2nd edition, 1978.

[9] A. Ilan and D. Loewenthal, "Instability of finite difference schemes due to boundary conditions in elastic media," *Geoph. Prosp.*, vol. 24, no. 3, pp. 431–453, Sep. 1976.

[10] A. R. Mitchell and D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations*, John Wiley and Sons, New York, 1980.

[11] B. Kummer and A. Behle, "Second-order finite-difference modeling of SH-wave propagation in laterally inhomogeneous media," *Bulletin Seismological Society of America*, vol. 72, no. 3, pp. 793–808, Jun. 1982.

[12] K. J. Marfurt, "Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations," *Geophysics*, vol. 49, no. 5, pp. 533–549, May 1984.

[13] A. Sei, "A family of numerical schemes for the computation of elastic waves," *SIAM J. Scientific Computations*, vol. 16, no. 4, pp. 898–916, Jul. 1995.

[14] J. W. Thomas, *Numerical Partial Differential Equations; Finite Difference Methods*, Springer, 1995.

[15] R. Madariaga, "Dynamics of an expanding circular fault," *Bulletin Seismological Society of America*, vol. 66, pp. 639–666, 1976.

[16] J. Virieux, "SH-wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophysics*, vol. 49, no. 11, pp. 1933–1957, Nov. 1984.

[17] J. Virieux, "P-SV wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophysics*, vol. 51, no. 4, pp. 889–901, Apr. 1986.

[18] A. Tavlove, *Computational Electromagnetics: The Finite-Difference Time-Domain Method*, Artech House, Boston, 1995.

[19] D. S. Watkins, *Fundamentals of Matrix Computations*, John Wiley and Sons, New York, 1991.

[20] K. R. Umashankar, A. Tavlove, and B. Beker, "Calculation and experimental validation of induced currents on coupled wires in an arbitrary shaped cavity," *IEEE Trans. Antennas Propagat.*, vol. 35, no. 11, pp. 1248–1257, Nov. 1987.

Appendix

I. Averaging

To ensure the stability of the finite-difference scheme at a material interface, the material parameters are averaged for the field components on the interface. Figure 4 shows a part of a 2-D finite-difference grid including two different media. The boundary between the two media is placed such that the normal particle velocity components are located on the interface. Thus, at a vertical boundary, the velocity component v_x lies on the boundary, whereas v_z is placed on a horizontal boundary. The only stress component that is placed on the boundary is the shear stress T_{xz}. A simple linear averaging procedure is presented which is derived in a straightforward manner from the governing equations in integral form. The derivation is well-known (see, for example, [18], [20]) and is given here for the sake of completeness.

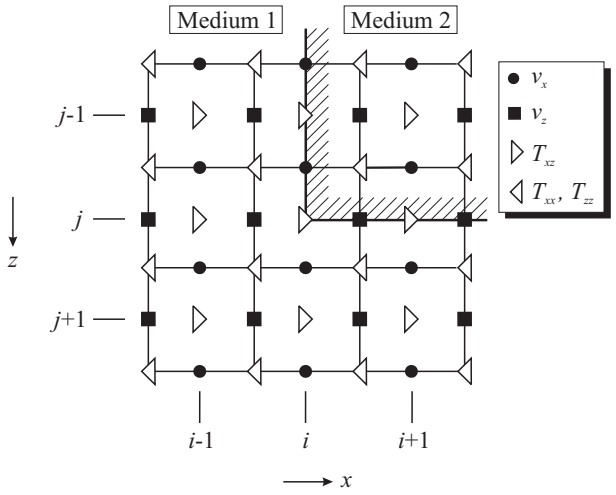


Fig. 4. Boundary between two media in a 2-D finite-difference grid; the normal particle velocities and the shear stress are located on the interface.

A. Material Density

The procedure for averaging the material density on the boundary is derived from the equation of motion. The derivation of the averaging is analogous for both the horizontal and vertical boundary and, thus, the averaging procedure is presented only for the horizontal particle velocity on a vertical interface. The equation of motion,

$$\nabla \cdot \mathbf{T} = \rho \frac{\partial \mathbf{v}}{\partial t}, \tag{31}$$

is expanded for the 2-D Shear-Vertical (S-V) case in the x-z plane:

$$\frac{\partial}{\partial x} T_{xx} + \frac{\partial}{\partial z} T_{xz} = \rho \frac{\partial v_x}{\partial t} \tag{32}$$

$$\frac{\partial}{\partial x} T_{xz} + \frac{\partial}{\partial z} T_{zz} = \rho \frac{\partial v_z}{\partial t}. \tag{33}$$

The only non-zero velocity components are v_x and v_z. Considering the horizontal particle velocity v_x, the procedure for averaging the material density on the vertical boundary is derived by integrating Eq. (32) over a box enclosing a portion of the boundary as depicted in Fig. 5:

$$\iint_{Box} \left[\frac{\partial}{\partial x} T_{xx} + \frac{\partial}{\partial z} T_{xz} \right] dx\, dz = \oint \left(T_{xx} \hat{\mathbf{x}} + T_{xz} \hat{\mathbf{z}} \right) \hat{\mathbf{n}}\, ds = \iint_{Box} \rho \frac{\partial v_x}{\partial t}\, dx\, dz, \tag{34}$$

where the density ρ is a function of position, and Gauss' integral theorem has been applied to convert the surface integral on the left hand side into a contour integral; $\hat{\mathbf{n}}$ is the normal vector to the contour of the area, and $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ are normal vectors in the x- and z-direction, respectively.

Using the discrete finite-difference formulation, the integration is carried out over a box of area $\Delta A = \Delta x \cdot \Delta z$, equivalent to the size of one cell of the finite-difference grid (see Fig. 5). Assuming the

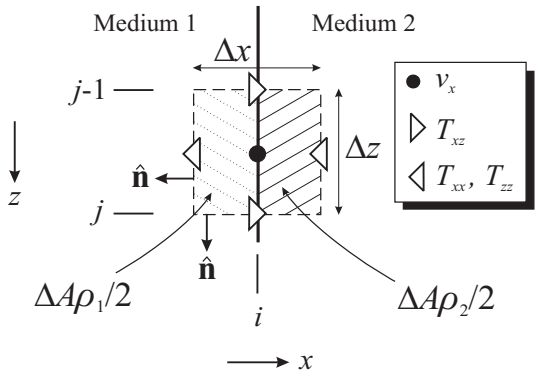


Fig. 5. Averaging of the material density ρ; the normal stress components are integrated along the contour.

boundary between Medium 1 and Medium 2 divides the area of integration in half, the integral comes out to be

$$\frac{\rho_1+\rho_2}{2}\Delta A\cdot\frac{v_x^{k+0.5}(i,j-0.5)-v_x^{k-0.5}(i,j-0.5)}{\Delta t}= \\ \Big(T_{xx}^k(i+0.5,j-0.5)-T_{xx}^k(i-0.5,j-0.5)\Big)\cdot\Delta z+\Big(T_{xz}^k(i,j)-T_{xz}^k(i,j-1)\Big)\cdot\Delta x. \tag{35}$$

After dividing by ΔA, the finite-difference approximation to Eq. (32) is obtained:

$$\frac{\rho_1+\rho_2}{2}\cdot\frac{v_x^{k+0.5}(i,j-0.5)-v_x^{k-0.5}(i,j-0.5)}{\Delta t}= \\ \frac{T_{xx}^k(i+0.5,j-0.5)-T_{xx}^k(i-0.5,j-0.5)}{\Delta x}+\frac{T_{xz}^k(i,j)-T_{xz}^k(i,j-1)}{\Delta z}. \tag{36}$$

Thus, for the velocity components on the boundary, the material densities of the two media are averaged linearly:

$$\rho_{avg}=(\rho_1+\rho_2)/2. \tag{37}$$

The same result is obtained for the vertical particle velocity on a horizontal interface.

B. Lamé's Constants

The procedure for averaging Lamé's constants for the stress components on the boundary is derived by using the elastic constitutive relation together with the strain-velocity relation. The elastic constitutive relation combines the mechanic strain and the stress and is usually expressed as

$$\mathbf{T}=\mathbf{c}\cdot\mathbf{S}, \tag{38}$$

where the stress $\mathbf{T}$ and the strain $\mathbf{S}$ are 3×3 tensors and the stiffness matrix $\mathbf{c}$ is a four-dimensional tensor of size $3\times3\times3\times3$. For an isotropic and lossless medium, the entries of the stiffness matrix are completely described by two independent constants, Lamé's constants λ and μ.

In the 2-D finite-difference gird (Fig. 4), the shear stress T_{xz} is placed on the boundary and, consequently, Lamé's constants are averaged for only T_{xz}. After combining the strain-velocity relation and the constitutive relation, an equation relating v_x and v_z to T_{xz} arises, from which, in the finite-difference formulation, the shear stress is determined:

$$\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} = \frac{1}{\mu}\frac{\partial T_{xz}}{\partial t}. \tag{39}$$

The averaging of μ is achieved by integrating over a box enclosing parts of the boundary, and Eq. (39) becomes

$$\iint_{Box} \left[\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right] dx\, dz = \iint_{Box} \frac{1}{\mu}\frac{\partial T_{xz}}{\partial t}\, dx\, dz. \tag{40}$$

Figure 6 shows the integration contour around T_{xz}. Applying Stoke's integral theorem, the left hand side of Eq. (40) can be written as

$$\iint_{Box} \left[\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}\right] dx\, dz = \iint_{Box} \nabla \times (v_x \cdot \hat{\mathbf{x}} - v_z \cdot \hat{\mathbf{z}})\, d\vec{A} = \oint (v_x \cdot \hat{\mathbf{x}} - v_z \cdot \hat{\mathbf{z}})\, d\vec{s}. \tag{41}$$

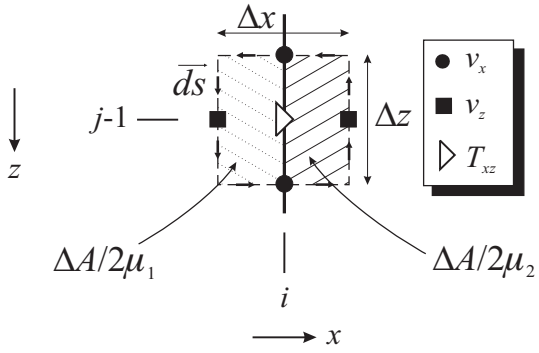


Fig. 6. Averaging of the shear stiffness μ; the tangential velocity components are integrated along the contour.

Using the discrete finite-difference formulation and Eq. (41), Eq. (40) becomes

$$\left(v_x^{k+0.5}(i, j-0.5) - v_x^{k+0.5}(i, j-1.5)\right) \cdot \Delta x + \left(v_z^{k+0.5}(i+0.5, j-1) - v_z^{k+0.5}(i-0.5, j-1)\right) \cdot \Delta z = \frac{\Delta A}{2}\left(\frac{1}{\mu_1} + \frac{1}{\mu_2}\right) \frac{T_{xz}^{k+1}(i, j-1) - T_{xz}^{k}(i, j-1)}{\Delta t}. \tag{42}$$

Thus, the inverse of the shear stiffness μ is averaged on the boundary between two media as

$$\frac{1}{\mu_{avg}} = \frac{1}{2}\left(\frac{1}{\mu_1} + \frac{1}{\mu_2}\right). \tag{43}$$

The same is obtained for the field components on a horizontal boundary. For the shear stress on a corner of the boundary (see $T_{xz}(i, j)$ in Fig. 4), the integration yields

$$\frac{1}{\mu_{avg}} = \frac{1}{4}\left(\frac{3}{\mu_1} + \frac{1}{\mu_2}\right). \tag{44}$$

A Theoretical Performance Analysis and Simulation of Time-Domain EMI Sensor Data for Land Mine Detection

Ping Gao, *Student Member, IEEE,* and Leslie M. Collins, *Member, IEEE*

***Abstract*—In this paper, the physical phenomenology of electromagnetic induction (EMI) sensors is reviewed for application to land mine detection and remediation. The response from time-domain EMI sensors is modeled as an exponential damping as a function of time, characterized by the initial magnitude and decay rate. Currently deployed EMI sensors that are used for the land mine detection process the recorded signal in a variety of ways in order to provide an audio output for the operator to judge whether or not the signal is from a mine. Sensors may sample the decay curve, sum it, or calculate its energy. Based on exponential decay model and the assumption that the sensor response is subject to additive white Gaussian noise, the performance of these, as well as optimal, detectors are derived and compared. Theoretical performance predictions derived using simplifying assumptions are shown to agree closely with simulated performance. It will also be shown that the generalized likelihood ratio test (GLRT) is equivalent to the likelihood ratio test (LRT) for multichannel time-domain EMI sensor data under the additive white Gaussian noise assumption and specific assumptions regarding the statistics of the decay rates of targets and clutter.**

I. Introduction

In order to evaluate the performance of a detection system, there are two fundamental quantities required: the probability of detection (P_d) and the probability of false alarm (P_{fa}). The receiver operating characteristic (ROC), which is a plot of P_d versus P_{fa}, is the standard method of illustrating the overall performance of a detector. With the motivation of obtaining quantitative measurements of the performance bounds, and given the fact that experimental minefield data necessarily suffers from a paucity of emplaced targets [1], sensor responses were modeled, and performance was evaluated theoretically and through simulation for a variety of detection algorithms using time-domain electromagnetic induction (EMI) sensor data. The goal of this study was to theoretically compare the performance of algorithms currently implemented in EMI sensors to that of algorithms derived using signal detection theory.

In references [2]–[4], Bayesian decision theoretic approaches to land mine detection using EMI sensor data obtained experimentally in the Defense Advanced Research Projects Agency (DARPA) Backgrounds Clutter Data Collection Experiment [5] were investigated. Two types of time-domain pulsed EMI systems were considered: sensors that either sample or integrate the time-domain information to provide a single data point (single-channel systems) and those that provide a sampled version of the time-domain waveform (multichannel sensors). The standard detection strategy for single-channel time-domain sensor data is simply a threshold test on the raw data recorded from the sensor. Extensions of this approach to multichannel EMI data are a threshold test on the energy present at the location under test, which is called an energy detector, and a threshold test on the integral (sum) of the entire time domain signal, called an integral or summed-data detector. In our previous work, we applied signal detection theory to the DARPA experimental data and incorporated the underlying physics of the sensor [6]–[10] as well as the statistical properties of the responses due to target and clutter/noise to generate the likelihood ratio test (LRT) at each surveyed location. The probability density functions (PDF's) of the responses from mines and clutter/noise were used to formulate the likelihood ratio. However, theoretical performance analyses of these detectors were not performed and the statistical significance of the conclusions which could be drawn based on this preliminary work were limited by the number of targets emplaced in the test sites. In this paper, rigorous theoretical analyses of the performance of the detectors mentioned above will be presented.

When using detection theory to form an LRT, it is sometimes difficult to determine an explicit form for the likelihood ratio since the data is often dependent on one or more unknown parameters. To obtain PDF's of the response, r, from target and clutter ($p(r|H_1)$ and $p(r|H_0)$ respectively), one must integrate over these parameters (details shown in Section III), which can be computationally expensive. (Here, H_1 is the hypothesis that a target (a mine), is present, and H_0 is the hypothesis that a target is not present, i.e. clutter or noise is present.) To avoid computational complexity, the generalized LRT (GLRT) (see Section III) provides a simplified, yet sub-optimal version of the LRT. In some cases, the GLRT achieves the same performance as the LRT, however, this is not generally the case [11]. In [2], [3], it has been shown that for single-channel EMI data, the optimal processor at each surveyed location is equivalent to a threshold test performed on the raw data under some assumptions on the statistics of the underlying process. In this paper, we show that for multichannel EMI data, the GLRT performed at a single surveyed location is the optimal detector under a similar set of assumptions. We also show that under the necessary assumptions, the form of the GLRT is a filter matched to the difference between the mean responses from targets and clutter.

Manuscript received March 12, 1999; revised December 2, 1999. This work was supported by Army Research Office Grant DAAH04-96-1-0448 (Demining MURI).

The authors are with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708-0291.

Publisher Item Identifier S 0196-2892(00)06221-5.

0196–2892/00$10.00 © 2000 IEEE

II. Sensor Phenomenology

An EMI system is essentially a metal detector. It records the induced electromagnetic field due to an incident electromagnetic field which impinges on underground objects, clutter, etc. An EMI system can detect mines that have metal content, as well as other unexploded ordnance (UXO) or anthropic clutter which contain metal. In order to detect such targets, the EMI system normally operates at low frequencies (< 1 MHz), at which the conductivity- and permeability-dependent skin depth of the materials varies significantly [6], [7]. Furthermore, at these frequencies the displacement current is weak enough to be overlooked [12]. Hence, the response of the pulsed EMI system, $\boldsymbol{r}$, at each location surveyed with the sensor can be modeled as a superposition of weighted resonances

$$r(t) = \sum_{i=1}^{N} A_n e^{j\omega_n t} \tag{1}$$

where ω_n is the nth natural resonant frequency of the object, and A_n is the initial magnitude of the response corresponding to that natural resonant frequency. In practice, the real part of ω_n is very small, and thus can be ignored [12]. The late time field, which is the field recorded by EMI sensors, is usually dominated by the lowest mode. Therefore, only one exponential damping coefficient need be considered. The response can thus be modeled (approximately) as

$$r(t) = Ae^{-\alpha t} \tag{2}$$

where A is the initial magnitude of the response, and α is the lowest natural resonant frequency. A is strongly dependent on the excitation level, as well as the depth and the orientation of the underground objects. The resonant frequency α is a function of conductivity and permeability. Both theoretical and experimental data demonstrate that α is unique to each metal type [6], [8], [12]. Consequently, it can be used to identify land mines. On average, metal mines have a lower natural resonant frequency than clutter, i.e. the decay rate of the target signature is slower.

In this paper, we consider data from two classes of EMI sensors: sensors that sample or perform some sort of integration on the time-domain information to provide a single data point (standard metal detector) and those that provide a sampled portion of the time-domain waveform. An example of the first system, the Geonics EM61, consists of a single-channel pulsed induction system with a 0.5 m transmitter coil positioned approximately 0.3 m above the ground. Data is received in both the transmitter coil and a secondary receiver coil that is located 0.4 m above the transmitter. The system operates at a center frequency of 75 Hz; the received signal is integrated from 0.18 to 0.87 ms after each transmitting pulse. The result, a single scalar value, is stored for later processing. Therefore, its response can be expressed as

$$r = \sum_{i=0}^{N} Ae^{-\alpha(t_0 + i\Delta t)} \tag{3}$$

where t_0 is the initial time for integration and the integration ends at $t_0 + N\Delta t$. Another example of a single channel system is the standard army device, the PSS-12, which is manufactured by Schiebel. The output of this sensor is a single sample of the time-domain waveform. An example of a multichannel time-domain pulsed EMI sensor is the Geonics EM61-3D sensor. It is a three-component, time-domain induction system having a 1 m square transmitter coil and three orthogonal 0.5 m receiver coils positioned approximately 0.3 m above the ground. The system operates at a center frequency of 7.5 Hz. Sensor output is measured and recorded at 20 geometrically spaced time gates, spanning a time range from 320 μs to 30 ms following the incident pulse [5]. Thus, the received signal of Geonics EM61-3D sensor, $\mathbf{r(t)}$, can be expressed as

$$\mathbf{r(t)} = Ae^{-\alpha \mathbf{t}} \tag{4}$$

where $\mathbf{t}$ is a 20×1 vector whose elements are the sampling times.

In this work, we have normalized the initial value, A, and have thus concentrated on the information conveyed by the resonant frequency α. Note that since the actual sensor output is subject to noise, only approximate normalization for real data is possible. Because the SNR at the early time response is high (since the response is essentially an exponential decay), the estimate of A is reasonably accurate.

III. Performance Analysis

In our previous work [2]–[4], a set of signal processing algorithms were applied to the EMI data collected in conjunction with the DARPA Background Clutter Data Collection Experiment [5]. However, no theoretical performance analyses and simulations on synthetic data were performed to validate the conclusions from the experimental study. Such analyses are necessary since so few targets were emplaced in the DARPA experiment. In this section, theoretical analyses of the performance for the LRT, the GLRT, the integral/sum, and the energy detectors and single-point detector are investigated. Additionally, when we consider a single time sample of the time-domain EMI response, the sample time that can provide the best performance using this type of data is derived. These results are then verified by simulation (Section IV).

We have assumed that the sensor is subject to independent and identically distributed (IID) Gaussian noise at each time sample. Therefore, the conditional density of $\mathbf{r}\,(r_1, r_2, \cdots, r_N)$ is jointly Gaussian. For the Gaussian noise problem, the performance of the detector is completely characterized by the quantity d'^2 [11], [13], where d' is defined as the distance between the means of the two hypotheses when the variance is normalized to one

$$d'^2 \equiv \frac{(E(l|H_1) - E(l|H_0))^2}{\mathrm{Var}(l|H_0)} \tag{5}$$

where l is the output of the processor, which is a function of received signal $\mathbf{r}$, and $E(\cdots)$ and $\mathrm{Var}(\cdots)$ represent the mean and variance. Hence, based on the assumptions stated previously, the d'^2 value of each detector is derived under a variety of assumptions regarding the amount of information which is known. In Section IV, we compare the ROC for the theoretical calculation of d'^2 with the simulation results.

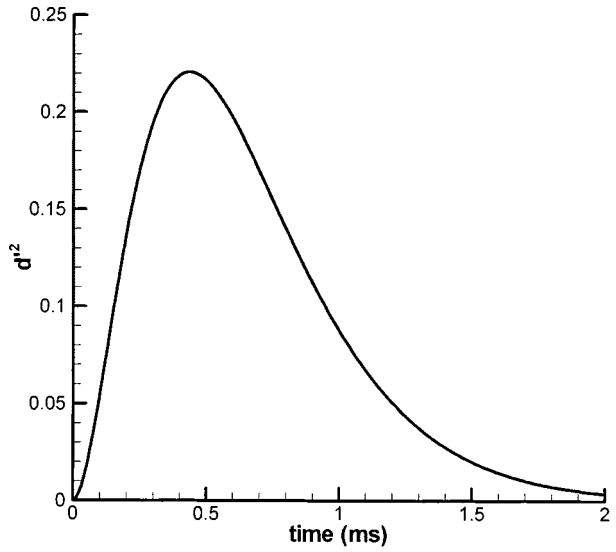


Fig. 1. d'^2 as a function of sample time t_s when parameters are fixed and known values. $\alpha_1 = 2.1$, $\alpha_0 = 2.5$, and $\sigma = 0.14$.

A. Fixed Parameter Case

As described previously, the output from the EMI sensor depends on whether or not a mine is present at the location under test. It is a function of the noise variance σ^2 and the decay rate α_1 for H_1, and α_0 for H_0. There are also other factors, such as temperature, humidity, environmental noise, etc., which affect the response. In this study, we have not considered these additional parameters. If the parameters α_1, α_0, and σ^2 are known values or can be estimated (by MLE, MMSE, etc.) the d'^2 corresponding to single time sample data, the LRT, an integral, and an energy detector on multichannel time-domain EMI data can be calculated as follows.

1) Single Time Sample of Time Domain (TD) EMI Data: If a time-domain EMI sensor samples the induced response at a single point in time, the performance of this sensor depends on the sample time used by the sensor. Thus, in order to achieve the best performance under the physical constraints imposed by the sensor, the time at which a sample is taken after the incident pulse vanishes becomes a key issue.

It is assumed that the sensor is subject to additive white Gaussian noise, i.e. $\mathcal{N}(0, \sigma^2)$. From detection theory, we know that the optimal detector operating on this type of data is simply a threshold test. Therefore

$$d'^2 = \frac{(e^{-\alpha_1 t_s} - e^{-\alpha_0 t_s})^2}{\sigma^2} \tag{6}$$

where t_s is the operating sample time of the sensor. Fig. 1 plots d'^2 as a function of sample time t_s. To determine the sample time that maximizes d', the derivative of d' with respect to t_s is taken and set to zero. The sample time maximizing d' is

$$t_{\max} = \frac{\ln \alpha_0 - \ln \alpha_1}{\alpha_0 - \alpha_1} \tag{7}$$

where ln(-) is the natural logarithm. Note that the best sample time is a function of the decay rates of the target and clutter, so it is object dependent.

2) LRT (Matched Filter) on Multichannel TD EMI Data: The LRT is defined as $\lambda(r) \equiv (p(r|H_1)/p(r|H_0))$ [11], [13], [14], where r is the sensor output, H_1 is the target present hypothesis, and H_0 is the no-target hypothesis. If the decay rates for targets and clutter α_1 and α_0 are known constants, from the definition of the LR, the LRT reduces to a matched filter $y = \sum_i r_i(e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i})$, where in this case, the data is matched to the difference of the two main signals. This is a signal known exactly (SKE) case, and the LR is distributed as a Gaussian random variable, i.e. under H_1, $y \sim \mathcal{N}(\mu_1, \sigma_1^2)$, and under H_0, $y \sim \mathcal{N}(\mu_0, \sigma_0^2)$, where

$$\mu_1 = \sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i}) e^{-\alpha_1 t_i},$$
$$\sigma_i^2 = \sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i})^2 \sigma^2,$$
$$\mu_0 = \sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i}) e^{-\alpha_0 t_i}$$

and $\sigma_0^2 = \sigma_1^2$. Thus

$$d'^2 = \frac{(\mu_1 - \mu_0)^2}{\sigma_0^2} = \frac{\sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i})^2}{\sigma^2} \tag{8}$$

As expected, d' is proportional to the energy of the difference signal and inversely proportional to the noise variance.

3) Integral/Summed-Data Detector on Multichannel TD EMI Data: An integral detector, also called a summed-data detector, integrates (sums) the sampled time sequence, then compares the sum to a threshold to make a decision as to whether a mine is present or not. The output of an integral detector is $x = \Sigma\, r_i$ and thus under H_1, $x \sim \mathcal{N}(\sum\, e^{-\alpha_1 t_i}, N\sigma^2)$ and under H_0, $x \sim \mathcal{N}(\sum\, e^{-\alpha_0 t_i}, N\sigma^2)$. Hence

$$d'^2 = \frac{\left[\sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i})\right]^2}{N\sigma^2} \tag{9}$$

4) Energy Detector on Multichannel TD EMI Data: The energy detector compares the energy ($E = \Sigma\, r_i^2$) to a threshold to make a decision. Based on the assumptions made before, under H_1

$$E = \sum_i (e^{-\alpha_1 t_i} + n_i)^2$$
$$= \sum_i e^{-2\alpha_1 t_i} + 2 \sum_i e^{-\alpha_1 t_i} n_i + \sum_i n_i^2. \tag{10}$$

The first term of the right-hand side of (10) is a constant, the second term is a Gaussian distributed random variable with mean of zero and variance of $4\sigma^2 \sum e^{-\alpha_i t_i}$, and the third term is distributed as $\sigma^2 \chi_N^2$ [16], where χ_N^2 is a chi-square distribution with N degrees of freedom. Based on the central limit theorem, the third term can be approximated by a Gaussian random

variable with a mean of $N\sigma^2$ and a variance of $2N\sigma^4$. The correlation between the second and the third terms in (10) is

$$E\left[\sum_i (e^{-\alpha_1 t_i} n_i)\left(\sum n_i^2\right)\right] = E\left[\left(\sum_i e^{-\alpha_1 t_i} n_i\right)(n_1^2+n_2^2+\cdots+n_N^2)\right] = \sum_i E[(e^{-\alpha_1 t_i} n_i)(n_1^2+n_2^2+\cdots+n_N^2)]. \quad (11)$$

Since n_i and n_j are independent, if $i \neq j$, $E[n_i n_j^2] = E[n_i]E[n_j^2] = 0$, and $E[n_i^3] = 0$ (since the mean of the odd power of a zero-mean Gaussian random variable is zero), then

$$E\left[\sum_i e^{-\alpha_1 t_i} n_i\right] E\left[\left(\sum_i n_i^2\right)\right] = \left(\sum_i e^{-\alpha_1 t_i} E[n_i]\right) E\left[\sum_i n_i^2\right] = 0. \quad (12)$$

$$E\left[\left(\sum_i e^{-\alpha_1 t_i} n_i\right)\left(\sum n_i^2\right)\right] = 0. \quad (13)$$

Therefore, the correlation between the second and the third terms of (10) is zero, implying that they are independent. since they are (approximately) Gaussian distributed [15]. Thus, the energy under H_1 can be approximated by a Gaussian distribution with a mean of $\sum e^{-2\alpha_1 t_i} + N\sigma^2$ and variance $4\sigma^2 \sum e^{-2\alpha_1 t_i} + 2N\sigma^4$. Similarly, under H_0. the energy also follows a Gaussian distribution with a mean of $\sum e^{-2\alpha_0 t_i} + N\sigma^2$ and variance $4\sigma^2 \sum e^{-2\alpha_0 t_i} + 2N\sigma^4$. Therefore

$$d'^2 = \frac{(\mu_1-\mu_0)^2}{(\sigma_1^2+\sigma_0^2)/2} = \frac{\left[\sum_i (e^{-2\alpha_1 t_i} - e^{-2\alpha_0 t_i})\right]^2}{2\sigma^2\left[\sum_i e^{-2\alpha_1 t_i} + e^{-2\alpha_0 t_i}\right] + 2N\sigma^4}. \quad (14)$$

B. Random Parameter Case

In practice, the decay rates of both targets and clutter are not known constants, but vary within some range. Without compromising generality and for simplicity of further calculation, a reasonable assumption on the distribution of α_1 and α_0 is Gaussian. Furthermore, α_1 and α_0 are independent of additive sensor noise, since they are intrinsic properties of the objects being considered. Practically, the decay rate cannot be negative, thus, the distributions of α_1 and α_0 can be considered as truncated Gaussian distributions

$$p(\alpha_1) = \left(\Phi\left(\frac{\alpha_1^u-\mu_{\alpha_1}}{\sigma_{\alpha_1}}\right) - \Phi\left(\frac{\alpha_1^l-\mu_{\alpha_1}}{\sigma_{\alpha_1}}\right)\right)^{-1} \cdot \frac{1}{\sqrt{2\pi\sigma_{\alpha_1}^2}}\, e^{-(\alpha_1-\mu_{\alpha_1})^2/2\sigma_{\alpha_1}^2}$$

and

$$p(\alpha_0) = \left(\Phi\left(\frac{\alpha_0^u-\mu_{\alpha_0}}{\sigma_{\alpha_0}}\right) - \Phi\left(\frac{\alpha_0^l-\mu_{\alpha_0}}{\sigma_{\alpha_0}}\right)\right)^{-1} \cdot \frac{1}{\sqrt{2\pi\sigma_{\alpha_0}^2}}\, e^{-(\alpha_0-\mu_{\alpha_0})^2/2\sigma_{\alpha_0}^2}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt.$$

$\mu_{\alpha_1}, \alpha_{\alpha_1}^2$ and $\mu_{\alpha_0}, \sigma_{\alpha_0}^2$ are the mean and the variance of α_1 and α_0, respectively, and α_1^u, α_1^l, and α_0^u, α_0^l are the left and right truncated points of α_1 and α_0, respectively. $\alpha_1^l = \alpha_0^l = 0$ and $\alpha_1^u \leq \alpha_0^u$ based on the underlying physics illustrated in Section II. The limits of α_1^u, α_0^u are infinity. Based on the field data collected in the DARPA experiment [5], $\Phi(\alpha_1^u - \mu_{\alpha_1}/\sigma_{\alpha_1}) - \Phi(\alpha_1^i - \mu_{\alpha_1}/\sigma_{\alpha_1})$, and $\Phi(\alpha_0^u - \mu_{\alpha_0}/\sigma_{\alpha_0}) - \Phi(\alpha_0^l - \mu_{\alpha_0}/\sigma_{\alpha_0})$ are very close to 1.

The LR test with uncertain parameters is defined as [11], [13], [14]

$$\lambda(r) = \frac{\int p(r|H_1,\theta_1)p(\theta_1|H_1)\, d\theta_1}{\int p(r|H_0,\theta_0)p(\theta_0|H_0)\, d\theta_0} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \quad (15)$$

where r is the measured, or available, data (which can be a scalar or a vector), H_1 and H_0 are the hypotheses of target present and target not present, respectively, $p(r|H_1,\theta_1)$ and $p(r|H_0,\theta_0)$ are the PDF's describing the statistical nature of the response, r, given hypotheses (H_1 or H_0), and parameters (θ_1 or θ_0). θ_1 and θ_0 are unknown parameter sets associated with each hypothesis which follow the distributions $p(\theta_1|H_1)$ and $p(\theta_0|H_0)$. In this problem, θ_1 and θ_0 correspond to the decay rate under the two hypotheses, respectively. When the LR is greater than a threshold, γ, H_1 is chosen, i.e. the target is present. Otherwise, H_0 is chosen, i.e. no target is present at the test location.

For a more general distribution of α_1 and α_0, no quantity such as d'^2 can easily be derived to characterize the performance. However, an ROC can be obtained through simulation using synthetic data. The noise variance σ^2 is not considered to be a random parameter.

1) GLRT/Matched Filter: The GLRT simplifies the calculation of a test statistic. Instead of integrating over the θ_1 and θ_0 parameters, as shown in (15), it is defined as [11]

$$\Omega(r) \equiv \frac{p(r|H_1,\hat{\theta}_1)}{p(r|H_0,\hat{\theta}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \quad (16)$$

where r is defined as before, and $\hat{\theta}_1$ and $\hat{\theta}_0$ are the maximum likelihood estimates (MLE) of the parameters θ_1 and θ_0. Based on the assumptions made before, the generalized LR is

$$\Omega = \frac{(2\pi\sigma^2)^{-N/2}\exp\left[\frac{1}{2\sigma^2}\sum_i (r_i - e^{-\hat{\mu}_{\alpha_1}t_i})^2\right]}{(2\pi\sigma^2)^{-N/2}\exp\left[\frac{1}{2\sigma^2}\sum_i (r_i - e^{-\hat{\mu}_{\alpha_0}t_i})^2\right]} = \exp\left[\frac{1}{2\sigma^2}\sum_i [2r_i(e^{-\hat{\mu}_{\alpha_0}t_i} - e^{-\hat{\mu}_{\alpha_1}t_i}) + e^{-2\hat{\mu}_{\alpha_1}t_i} + e^{-2\hat{\mu}_{\alpha_0}t_i}]\right] \quad (17)$$

where $\hat{\mu}_{\alpha_1}$ and $\hat{\mu}_{\alpha_0}$ are estimates of μ_{α_1} and μ_{α_0}. ML estimates were used in this study. By taking the logarithm and incorporating the constant into the threshold, the log-GLR is

$$y = \sum_i r_i(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i}). \quad (18)$$

This is essentially a matched filter, which is matched to the difference between the target response and the clutter response at the estimated decay rates.

Thus

$$y|H_1 = \sum_i (e^{-\alpha_1 t_i} + n_i)(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i}) = \sum_i e^{-\alpha_1 t_i}(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i}) + \sum_i n_i(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i}). \quad (19)$$

When the variance of α_1 is relatively small compared to the variance of the noise (this often is the case, as has been verified experimentally [3]), $e^{-\alpha_1 t_i}$ can be approximated by a straight line through the mean value of α_1 with a slope of the derivative of $e^{-\alpha_1 t_i}$ at $\hat{\mu}_{\alpha_1}$, i.e.

$$e^{-\alpha_1 t_i} \approx -t_i e^{-\hat{\mu}_{\alpha_1}t_i}\alpha_1 + e^{-\hat{\mu}_{\alpha_1}t_i} + t_1 e^{-\hat{\mu}_{\alpha_1}t_i}\hat{\mu}_{\alpha_1}. \quad (20)$$

Thus, by substituting (20) into (19), y is seen to be a linear combination of Gaussian-distributed random variables. Therefore, $y \sim \mathcal{N}(\mu_{1y}, \text{var}(y_1))$, where

$$\mu_{1y} = \sum_i e^{-\hat{\mu}_{\alpha_1}t_i}(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})$$
$$\text{var}(y_1) = \sum_i [(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})t_i e^{-\hat{\mu}_{\alpha_1}t_i}]^2\sigma^2_{\alpha_1} + \sum_i (e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})^2\sigma^2. \quad (21)$$

Similarly, $y|H_0 \sim \mathcal{N}(\mu_{0y}, \text{var}(y_0))$, where

$$\mu_{0y} = \sum_i e^{-\hat{\mu}_{\alpha_0}t_i}(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})$$
$$\text{var}(y_0) = \sum_i [(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})t_i e^{-\hat{\mu}_{\alpha_0}t_i}]^2\sigma^2_{\alpha_0} + \sum_i (e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})^2\sigma^2. \quad (22)$$

Therefore, d'^2 is as shown in (23), at the bottom of the page.

It is shown in Section IV that the LRT reduces to the GLRT for EMI data. Therefore, no additional analysis is required to determine the performance of the LRT.

2) Integral/Summed-Data Detector: Letting $x = \Sigma\, r_i$ correspond to an integral detector output, and using the approximation expressed in (20), then under H_1, $x = \sum_i e^{-\alpha_1 t_i} + n_i$, so

$$x|H_1 \sim N\left(\sum e^{-\hat{\mu}_{\alpha_1}t_i}, \left(\sum t_i e^{-\hat{\mu}_{\alpha_1}t_i}\right)^2\sigma^2_{\alpha_1} + N\sigma^2\right). \quad (24)$$

Similarly,

$$x|H_0 \sim N\left(\sum e^{-\hat{\mu}_{\alpha_0}t_i}, \left(\sum t_i e^{-\hat{\mu}_{\alpha_0}t_i}\right)^2\sigma^2_{\alpha_0} + N\sigma^2\right). \quad (25)$$

Hence, the result for d'^2 is (26), as shown at the bottom of the next page.

3) Energy Detector: The output of an energy detector on EMI data is either $x = \sum_i (e^{-\alpha_1 t_i} + n_i)^2$ under H_1, or $\sum_i (e^{-\alpha_0 t_i} + n_i)^2$ under H_0. Since both α_1 and α_0 are random variables, the distribution of the energy cannot be accurately approximated by a Gaussian random variable. Therefore, it is not valid to calculate d', since the definition of d' is based on Gaussian-distributed data. Even though d' is not applicable, we can explain the fact that an energy detector exhibits better performance than an integral detector. Since the noise is IID at each sample time, the noise variance at each sample of the sensor is

$$d'^2 = \frac{(\mu_1 - \mu_0)^2}{(\sigma_1^2 + \sigma_0^2)/2} = \frac{\left[\sum (e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})\right]^2}{\frac{1}{2}\left(\sum_i t_i^2(e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})^2(e^{-2\hat{\mu}_{\alpha_1}t_i}\sigma^2_{\alpha_1} - e^{-2\hat{\mu}_{\alpha_0}t_i}\sigma^2_{\alpha_0})\right) + \sum_i (e^{-\hat{\mu}_{\alpha_1}t_i} - e^{-\hat{\mu}_{\alpha_0}t_i})^2\sigma^2}. \quad (23)$$

the same. For lower-level signals, corresponding to later samples, the SNR is lower than that for higher-level signals. Because the operation of calculating the energy puts more weight on higher values and less weight on lower signals. When the value of the signal is greater than 1, lower SNR time samples contribute less, and consequently, the noise affects the results of the energy detector less than the integral detector, which assigns equal weights to each time sample.

4) Single Time Sample of TD EMI Data: If the decay rate of target or clutter is not a known constant, d' is more complicated. Borrowing the result of Section IV, the LRT and GLRT is equivalent for TD EMI data under some specific assumptions. Hence, using the approximation in (20), it can be shown that

$$d'^2 = \frac{(e^{-\hat{\mu}_{\alpha_1} t_s} - e^{-\hat{\mu}_{\alpha_0} t_s})^2}{\sigma^2 + \dfrac{t^2}{2}\left[e^{-2\hat{\mu}_{\alpha_1} t_s}\sigma^2_{\alpha_1} + e^{-2\hat{\mu}_{\alpha_0} t_s}\sigma^2_{\alpha_0}\right]}. \tag{27}$$

It is difficult to obtain an analytical expression of the sample time that maximizes d', since it involves transcendental function. However, it is easy to find the maximum numerically. If $\sigma^2_{\alpha_1} \ll 1$, $\sigma^2_{\alpha_0} \ll 1$, then

$$d'^2 \approx \frac{(e^{-\hat{\mu}_{\alpha_1} t_s} - e^{-\hat{\mu}_{\alpha_0} t_s})^2}{\sigma^2}$$

so

$$t_{\max} \approx \frac{\ln\,\hat{\mu}_{\alpha_0} - \ln\,\hat{\mu}_{\alpha_1}}{\hat{\mu}_{\alpha_0} - \hat{\mu}_{\alpha_1}}.$$

IV. Equivalence of the GLRT and the LRT for Multi-Channel EMI Data

In many applications, the GLRT is often used instead of the LRT to reduce computational complexity at the cost of sacrificing performance. However, it is not in general an optimal processor [11]. In the following analyses, we show that under some reasonable assumptions made for the statistics governing the land mine detection problem using EMI sensor data, the GLRT and the LRT provide the same performance. In this problem, the GLRT (shown in Section III that it is essentially a matched filter, $z = \Sigma_{i=1}^{N} r_i(e^{-\hat{\mu}_{\alpha_1} t_i} - e^{\hat{\mu}_{\alpha_0} t_i})$) always provides the same performance as that of the LRT as long as the statistics of the decay rates of targets and clutter follow certain assumptions.

1) Proposition: Assume the multichannel EMI sensor response is modeled as $\mathbf{r} = e^{-\alpha_n \mathbf{t}}$, where $n = 1$ or 0, representing H_1 and H_0, respectively, $\mathbf{t}$ is an N by one vector of the sampling times, α_1 and α_0 are truncated Gaussian-distributed random variables with means μ_{α_1} and μ_{α_0}, and variances $\sigma^2_{\alpha_1}$ and $\sigma^2_{\alpha_0}$, and left truncated points of α_1^l, α_0^l, and right truncated points of α_1^u, α_0^u, respectively. Furthermore, it is assumed that $\alpha_0^l \geq \alpha_1^l \geq 0$ and $\alpha_0^l < \alpha_1^u \leq \alpha_0^u$, based on the underlying physics illustrated in Section II. The limits of α_1^u, α_0^u are infinity. The means and variances of α_1 and α_0 satisfy $\mu_{\alpha_1} < \mu_{\alpha_0}$, $(\mu_{\alpha_0}/\mu_{\alpha_1}) \geq (\sigma^2_{\alpha_0}/\sigma^2_{\alpha_1}) \geq 1$. Also, assume the sensor is subject to IID Gaussian noise at each sample time. The GLRT on $\mathbf{r}$ is in the form of $z = \Sigma_{i=1}^{N} r_i(e^{-\hat{\mu}_{\alpha_1} t_i} - e^{-\hat{\mu}_{\alpha_0} t_i})$, and the LRT is of the form shown in (15). Then, the GLRT and LRT are equivalent, thus providing the same performance.

Proof: If it can be shown that $d\lambda/dz \geq 0$, which means the LR λ, is a nondecreasing function of the output of a matched filter z, and thus, a monotonic function of z, the LRT, and the GLRT are equivalent. Because λ is not an explicit function of z, the chain rule is utilized to prove this relationship

$$\frac{d\lambda}{dz} = \sum_{i=1}^{N} \frac{\partial \lambda}{\partial r_i}\,\frac{1}{\partial z/\partial r_i} \tag{28}$$

where r_i is the received signal from the sensor at the ith sample time at one location, and N is the number of times at which the sensor samples the response. Under H_1, the output of multichannel time-domain EMI sensor is $r_i = e^{-\alpha_1 t_i} + n_i$ $(i = 1, 2, \cdots, N)$, and under H_0, the sensor output is $r_i = e^{-\alpha_0 t_i} + n_i$, where r_i represents the output from the sensor at the sample time t_i, n_i is IID white Gaussian noise with zero mean and variance of σ^2. Since the noise terms are IID, the covariance matrix for $\boldsymbol{n}$ is the identity matrix scaled by σ^2.

The LR is then (29), shown at the bottom of the next page, where

$$F(\mathbf{r}, \alpha_1) = K_1 \exp\left[-\sum_{i=1}^{N} (r_i - e^{-\alpha_1 t_i})^2/2\sigma^2\right] \cdot \exp[-(\alpha_1 - \mu_{\alpha_1})^2/2\sigma^2_{\alpha_1}] \tag{30}$$

and

$$G(\mathbf{r}, \alpha_0) = K_0\ \exp\left[-\sum_{i=1}^{N} (r_i - e^{-\alpha_0 t_i})^2/2\sigma^2\right] \cdot \exp[-(\alpha_0 - \mu_{\alpha_0})^2/2\sigma^2_{\alpha_0}] \tag{31}$$

and

$$K_1^{-1} = \left(\Phi\left(\frac{\alpha_1^u - \mu_{\alpha_1}}{\sigma_{\alpha_1}}\right) - \Phi\left(\frac{\alpha_1^l - \mu_{\alpha_1}}{\sigma_{\alpha_1}}\right)\right) \cdot (2\pi)^{(N+1)/2}\sigma^N \sigma_{\alpha_1}$$

$$K_0^{-1} = \left(\Phi\left(\frac{\alpha_0^u - \mu_{\alpha_0}}{\sigma_{\alpha_0}}\right) - \Phi\left(\frac{\alpha_0^l - \mu_{\alpha_0}}{\sigma_{\alpha_0}}\right)\right) \cdot (2\pi)^{(N+1)/2}\sigma^N \sigma_{\alpha_0}$$

thus, $K_0 > 0$ and $K_1 > 0$. Taking the partial derivative of λ with respect to r_i yields (32), shown at the bottom of

$$d'^2 = \frac{\left(\sum_i (e^{-\hat{\mu}_{\alpha_1} t_i} - e^{-\hat{\mu}_{\alpha_0} t_i})\right)^2}{\frac{1}{2}\left[\left(\sum_i t_i e^{-\hat{\mu}_{\alpha_1} t_i}\right)^2 \sigma^2_{\alpha_1} + \left(\sum_i t_i e^{-\hat{\mu}_{\alpha_0} t_i}\right)^2 \sigma^2_{\alpha_0}\right] + N\sigma^2}. \tag{26}$$

the page, where the $(\cdots)'$ notation denotes a derivative with respect to r_i. We now write $\partial\lambda/\partial r_i = h(r_i)/w(\mathbf{r})$. Since the denominator $w(\mathbf{r})$ is always positive, if $h(r_i)$ is nonnegative, $\partial\lambda/\partial r_i$ is nonnegative as well. Based on a theorem presented in [15], the derivative of the integral is equal to the integral of the derivative since the integrand is differentiable with respect to r_i, i.e.

$$\left(\int_{\alpha_1^l}^{\alpha_1^u} F(\mathbf{r},\alpha_1)\,d\alpha_1\right)'_{r_i} = \int_{\alpha_1^l}^{\alpha_1^u} (F(\mathbf{r},\alpha_1))'_{r_i}\,d\alpha_1. \quad (33)$$

Therefore, the numerator is

$$\begin{aligned} h(r_i) &= \int_{\alpha_1^l}^{\alpha_1^u} \frac{\partial F(\mathbf{r},\alpha_1)}{\partial r_i}\,d\alpha_1 \int_{\alpha_0^l}^{\alpha_0^u} G(\mathbf{r},\alpha_0)\,d\alpha_0 \\ &\quad - \int_{\alpha_0^l}^{\alpha_0^u} \frac{\partial G(\mathbf{r},\alpha_0)}{\partial r_i}\,d\alpha_0 \int_{\alpha_1^l}^{\alpha_1^u} F(\mathbf{r},\alpha_1)\,d\alpha_1 \\ &= \int_{\alpha_0^l}^{\alpha_0^u}\int_{\alpha_1^l}^{\alpha_1^u} \left(\frac{\partial F(\mathbf{r},\alpha_1)}{\partial r_i} G(\mathbf{r},\alpha_0) \right. \\ &\quad \left. - \frac{\partial G(\mathbf{r},\alpha_0)}{\partial r_i} F(\mathbf{r},\alpha_1)\right) d\alpha_1\,d\alpha_0. \end{aligned} \quad (34)$$

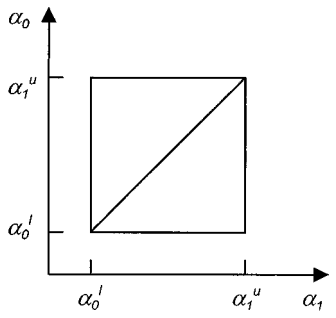


Fig. 2. Integration area for (38). α_1 and α_0 are integration variables. The square is divided into upper and lower triangular regions.

TABLE I
NOISE VARIANCE σ^2 REQUIRED TO ACHIEVE $P_d = 0.7$ AND $P_{fa} = 0.1$ FOR FIXED AND RANDOM DECAY RATES

	Fixed Parameter	Random Parameter
Matched Filter	17.5	8.6
Energy Detector	16.3	7.3
Integral detector	10.8	4.9
Threshold Test on Single Time Sample	7.9	3.7

$$\begin{aligned} \lambda &= \frac{\displaystyle\int_{-\infty}^{\infty} p(\mathbf{r}|H_1,\alpha_1)p(\alpha_1|H_1)\,d\alpha_1}{\displaystyle\int_{-\infty}^{\infty} p(\mathbf{r}|H_0,\alpha_0)p(\alpha_0|H_0)\,d\alpha_0} \\ &= \frac{\displaystyle\int_{\alpha_1^l}^{\alpha_1^u} (2\pi\sigma^2)^{-N/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}[(r_i - e^{-\alpha_1 t_i})^2]\right]\left(\Phi\left(\frac{\alpha_1^u-\mu_{\alpha_1}}{\sigma_{\alpha_1}}\right)-\Phi\left(\frac{\alpha_1^l-\mu_{\alpha_1}}{\sigma_{\alpha_1}}\right)\right)^{-1}}{\cdot(2\pi\sigma_{\alpha_1}^2)^{-1/2}\exp\left[-\frac{1}{2\sigma_{\alpha_1}^2}(\alpha_1-\mu_{\alpha_1})^2\right]\,d\alpha_1} \\ &\quad \frac{\displaystyle\int_{\alpha_0^l}^{\alpha_0^u} (2\pi\sigma^2)^{-N/2}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}[(r_i - e^{-\alpha_0 t_i})^2]\right]\left(\Phi\left(\frac{\alpha_0^u-\mu_{\alpha_0}}{\sigma_{\alpha_0}}\right)-\Phi\left(\frac{\alpha_0^l-\mu_{\alpha_0}}{\sigma_{\alpha_0}}\right)\right)^{-1}}{\cdot(2\pi\sigma_{\alpha_0}^2)^{-1/2}\exp\left[-\frac{1}{2\sigma_{\alpha_0}^2}(\alpha_0-\mu_{\alpha_0})^2\right]\,d\alpha_0} \\ &= \frac{\displaystyle\int_{\alpha_1^l}^{\alpha_1^u} F(\mathbf{r},\alpha_1)\,d\alpha_1}{\displaystyle\int_{\alpha_0^l}^{\alpha_0^u} G(\mathbf{r},\alpha_0)\,d\alpha_0}. \end{aligned} \quad (29)$$

$$\frac{\partial\lambda}{\partial r_i} = \frac{\displaystyle\left(\int_{\alpha_1^l}^{\alpha_1^u} F(\mathbf{r},\alpha_1)\,d\alpha_1\right)'_{r_i}\int_{\alpha_0^l}^{\alpha_0^u} G(\mathbf{r},\alpha_0)\,d\alpha_0 - \int_{\alpha_1^l}^{\alpha_1^u} F(\mathbf{r},\alpha_1)\,d\alpha_1\left(\int_{\alpha_0^l}^{\alpha_0^u} G(\mathbf{r},\alpha_0)\,d\alpha_0\right)'_{r_1}}{\displaystyle\left[\int_{\alpha_0^l}^{\alpha_0^u} G(\mathbf{r},\alpha_0)\,d\alpha_0\right]^2}. \quad (32)$$

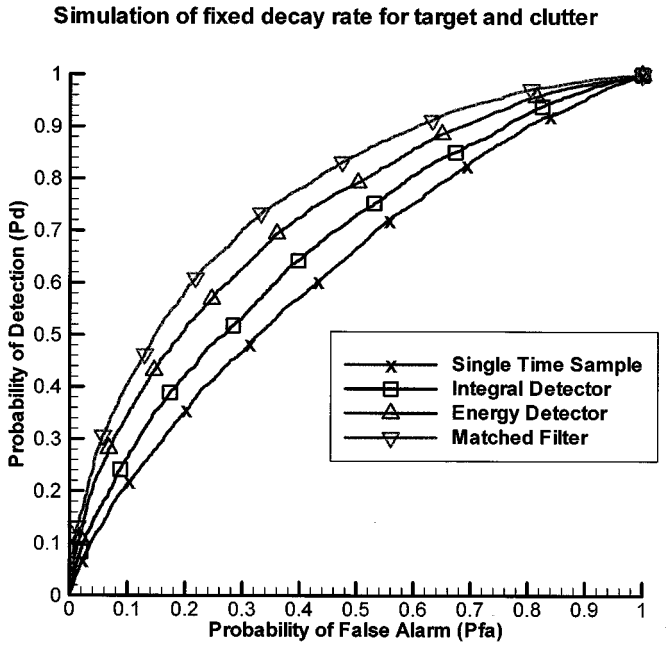


Fig. 3. ROC of simulations of various detection algorithms for a multichannel time-domain EMI sensor at a fixed decay rate for both target and clutter. The decay rate of the target is $\alpha_1 = 2.1$, the decay rate for clutter is $\alpha_0 = 2.5$, and the initial magnitude is 220 for both target and clutter. Gradient: matched filter; delta: energy detector; square: integral detector; cross: threshold test on single time sample.

Taking the derivative of $F(\mathbf{r}, \alpha_1)$ with respect to r_i

$$\frac{\partial F(\mathbf{r},\alpha_1)}{\partial r_i} = K_1\sigma^2(e^{-\alpha_1 t_i} - r_i) \cdot \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j - e^{-\alpha_1 t_j})^2\right]\right] \cdot \exp\left[-\frac{(\alpha_1 - \mu_{\alpha_1})^2}{2\sigma_{\alpha_1}^2}\right]. \quad (35)$$

Similarly

$$\frac{\partial G(\mathbf{r},\alpha_0)}{\partial r_i} = K_0\sigma^2(e^{-\alpha_0 t_i} - r_i) \cdot \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j - e^{-\alpha_0 t_j})^2\right]\right] \cdot \exp\left[-\frac{(\alpha_0 - \mu_{\alpha_0})^2}{2\sigma_{\alpha_0}^2}\right]. \quad (36)$$

After combining constants, $h(r_i)$ can be expressed as

$$h(r_i) = K_0K_1\sigma^2 \int_{\alpha_0^l}^{\alpha_0^u}\int_{\alpha_1^l}^{\alpha_1^u} \cdot \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j - e^{-\alpha_1 t_j})^2\right]\right] \cdot \exp\left[-\frac{1}{2\sigma^2}(r_j - e^{-\alpha_0 t_j})^2\right] \cdot \exp\left[-\frac{(\alpha_1 - \mu_{\alpha_1})^2}{2\sigma_{\alpha_1}^2}\right]\exp\left[-\frac{(\alpha_0 - \mu_{\alpha_0})^2}{2\sigma_{\alpha_0}^2}\right] \times (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i})\, d\alpha_1\, d\alpha_0. \quad (37)$$

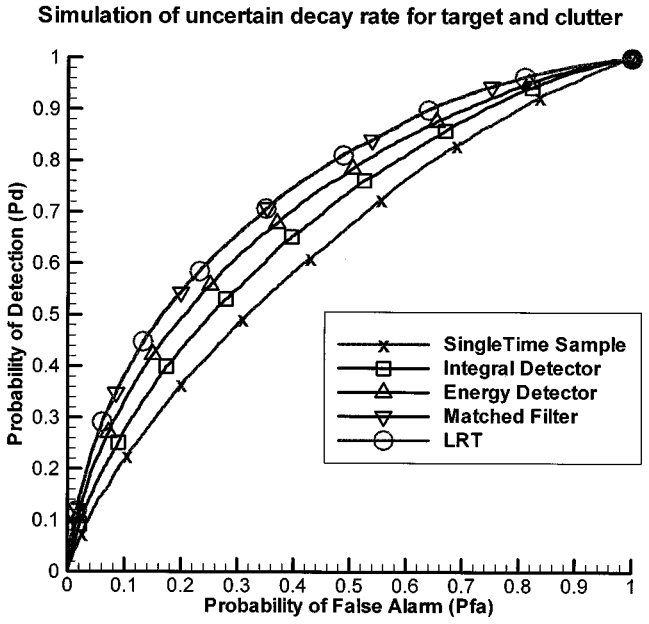


Fig. 4. ROC of simulations of various detection algorithms for multichannel time-domain EMI sensor data with uncertain decay rate for both target and clutter. Distribution of decay rate α_1 is $N(2.1, 0.2^2)$ and for α_0 is $N(2.5, 0.2^2)$. The initial magnitude is 220 for both target and clutter. Circle: the LRT; gradient: matched filter, delta: energy detector; square: integral detector; cross: threshold test on single time sample.

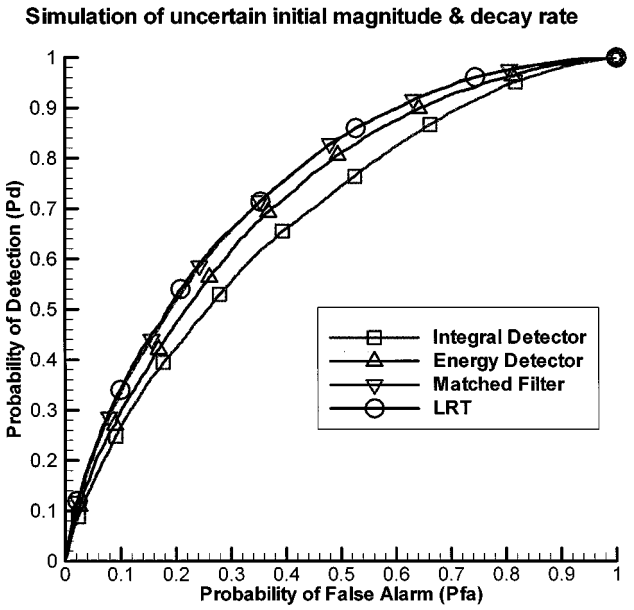


Fig. 5. ROC of simulations of various detection algorithms for multichannel time-domain EMI sensor data with uncertain decay rate and uncertain initial magnitude of target and clutter. Distribution of decay rate α_1 is $N(2.1, 0.2^2)$ and for α_0 is $N(2.5, 0.2^2)$. The distribution of the initial magnitude for target A, is $N(220, 20^2)$ and for clutter B is distributed as $N(210, 40^2)$. Circle: LRT; gradient line: matched filter, delta: energy detector; square: integral detector.

Because exponential functions are always positive, the first four terms of the integrand of the integral in (37) are always positive. The constant term associated with the integral is also positive. Obviously, if the decay rate of the target α_1 is always slower than that of clutter α_0, the fifth term in (37), $(e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i})$, is also positive, so $h(r_i) \geq 0$. Although in practice, α_1

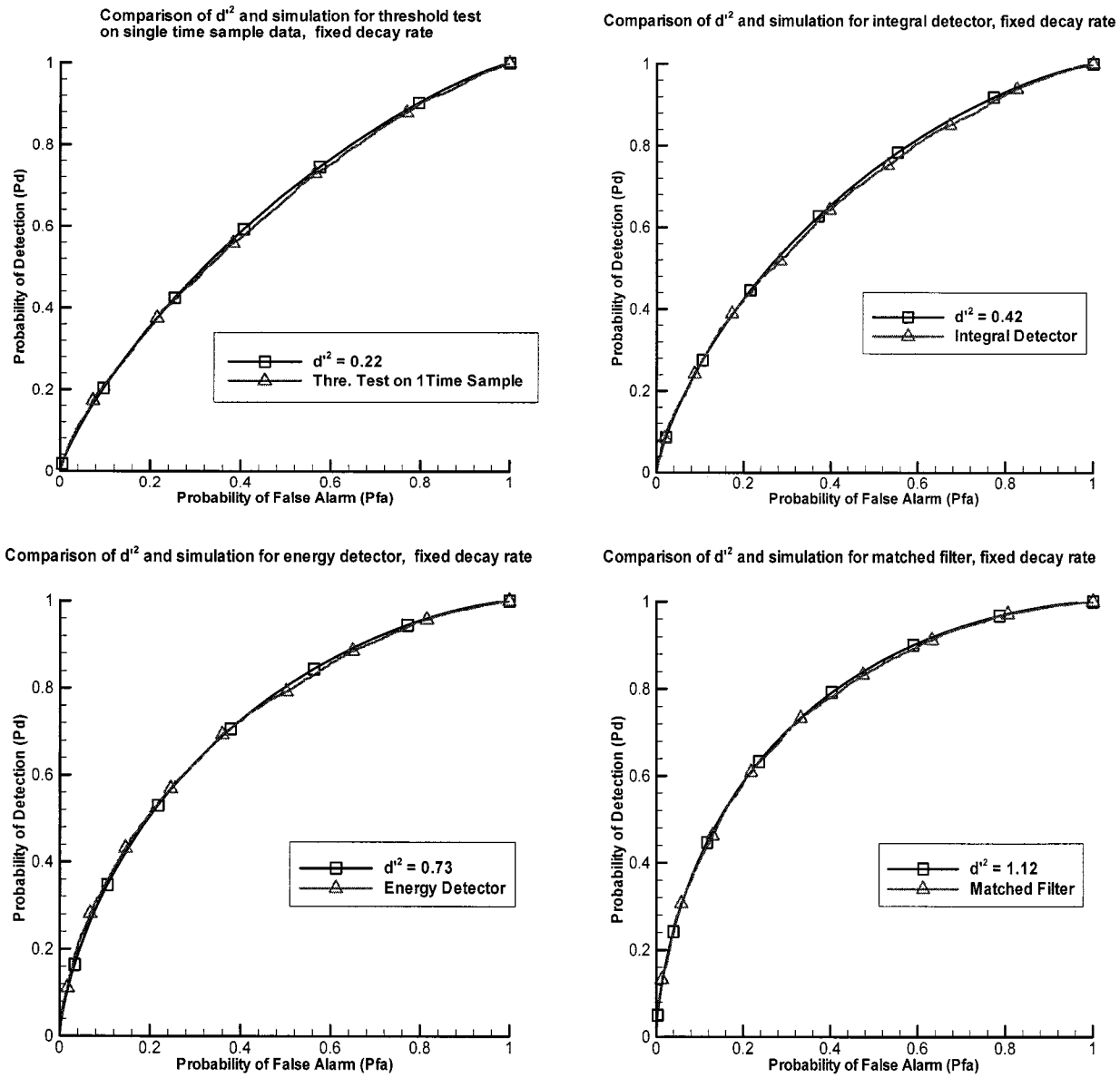


Fig. 6. Comparison of ROC's of the theoretical d' value and simulations of various detectors for the fixed decay rate case. (Top left) Threshold test for single time sample time-domain EMI sensor data. (Top right) Integral detector. (Bottom left) Energy detector. (Bottom right) Matched filter for TD EMI sensor data. Delta: simulation; square: theoretical d'.

TABLE II
THE THEORETICAL VALUES OF d'^2 FOR VARIOUS PROCESSORS AND KNOWN/RANDOM PARAMETER ENVIRONMENTS

	Fixed Parameter	Random Parameter
Matched Filter	1.12	0.87
Energy Detector	0.73	≈0.72
Integral Detector	0.42	0.38
Threshold Test on Single Time Sample	0.22	≈0.22

is not always smaller than α_0, we have assumed the relationships of $\mu_{\alpha_1} < \mu_{\alpha_0}$, $\alpha_1^l \le \alpha_0^l < \alpha_1^u \le \alpha_0^u$. The integration area can be divided into four sub-areas, and each sub-area can be integrated separately. Let $K = K_0 K_1 \sigma^2$. Thus, $K > 0$

$$h(r_i) = K \int_{\alpha_0^l}^{\alpha_1^u} \left\{ \int_{\alpha_1^l}^{\alpha_0^l} (\cdot)\, d\alpha_1 + \int_{\alpha_0^l}^{\alpha_1^u} (\cdot)\, d\alpha_1 \right\} d\alpha_0 + K \int_{\alpha_1^u}^{\alpha_0^u} \left\{ \int_{\alpha_1^l}^{\alpha_0^l} (\cdot)\, d\alpha_1 + \int_{\alpha_0^l}^{\alpha_1^u} (\cdot)\, d\alpha_1 \right\} d\alpha_0. \tag{38}$$

It is easily shown that for the first, third, and fourth integrals in (38), the integration is over an area in which α_1 is always less than α_0. Thus, integration in these three sub-areas contributes positive or nonnegative value to $h(r_i)$. In the second integral in (38), α_1 and α_0 are integrated over an identical region. Let $h_2(r_i)$ represent the second integral in (38) after neglecting the

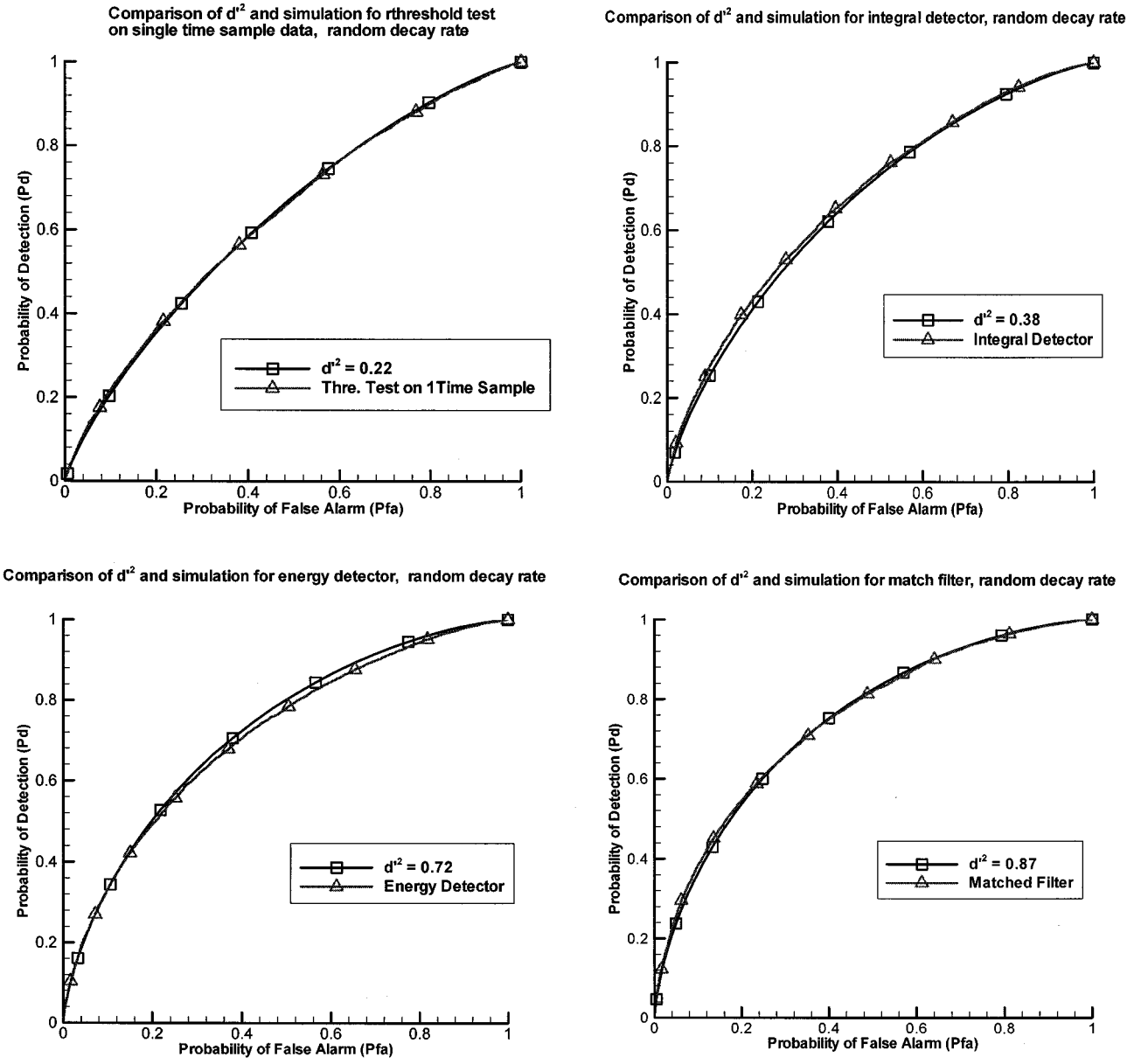


Fig. 7. Comparison of ROC of the theoretical d'^2 value and simulation of various detectors for the random decay rate case. (Top left) Threshold test for single time sample time-domain EMI sensor data. (Top right) Integral detector. (Bottom left) Energy detector. (Bottom right) GLRT/matched filter for time-domain EMI sensor data. Delta: simulation; square: theoretical d'.

positive constant K. The integration area can be further divided into upper triangular and lower triangular portions (Fig. 2 illustrates how the integration areas are divided.) Therefore, $h_2(r_i)$ can be written as

$$
\begin{aligned}
&h_2(r_i)\\
&\quad=\int_{\alpha_0^l}^{\alpha_1^u}\int_{\alpha_0^l}^{\alpha_0}(e^{-\alpha_1 t_i}-e^{-\alpha_0 t_i})\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_1 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_0 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{(\alpha_1-\mu_{\alpha_1})^2}{2\sigma_{\alpha_1}^2}\right]\exp\left[-\frac{(\alpha_0-\mu_{\alpha_0})^2}{2\sigma_{\alpha_0}^2}\right]\,d\alpha_1\,d\alpha_0\\
&\quad+\int_{\alpha_0^l}^{\alpha_1^u}\int_{\alpha_0^l}^{\alpha_1}(e^{-\alpha_1 t_i}-e^{-\alpha_0 t_i})\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_1 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_0 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{(\alpha_1-\mu_{\alpha_1})^2}{2\sigma_{\alpha_1}^2}\right]\exp\left[-\frac{(\alpha_0-\mu_{\alpha_0})^2}{2\sigma_{\alpha_0}^2}\right]\,d\alpha_0\,d\alpha_1.
\end{aligned}
\tag{39}
$$

Then, keeping the first integral in (39) the same and exchanging variables α_1 and α_0 via a change of variables in the second integral in (39), and changing the integration limits correspondingly, $h_2(r_i)$ becomes

$$\begin{aligned}
&h_2(r_i)\\
&\quad=\int_{\alpha_0^l}^{\alpha_1^u}\int_{\alpha_0^l}^{\alpha_0}(e^{-\alpha_1 t_i}-e^{-\alpha_0 t_i})\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_1 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_0 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{(\alpha_1-\mu_{\alpha_1})^2}{2\sigma_{\alpha_1}^2}\right]\exp\left[-\frac{(\alpha_0-\mu_{\alpha_0})^2}{2\sigma_{\alpha_0}^2}\right]d\alpha_1 d\alpha_0\\
&\qquad+\int_{\alpha_0^l}^{\alpha_1^u}\int_{\alpha_0^l}^{\alpha_0}(e^{-\alpha_0 t_i}-e^{-\alpha_1 t_i})\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_1 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_0 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{(\alpha_0-\mu_{\alpha_1})^2}{2\sigma_{\alpha_1}^2}\right]\exp\left[-\frac{(\alpha_1-\mu_{\alpha_0})^2}{2\sigma_{\alpha_0}^2}\right]\,d\alpha_1\,d\alpha_0\\
&\qquad\int_{\alpha_0^l}^{\alpha_1^u}\int_{\alpha_0^l}^{\alpha_0}(e^{-\alpha_1 t_i}-e^{-\alpha_0 t_i})\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_1 t_j})^2\right]\right]\\
&\qquad\cdot\exp\left[-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{N}(r_j-e^{-\alpha_0 t_j})^2\right]\right]\\
&\qquad\cdot\left[\exp\left[-\frac{(\alpha_1-\mu_{\alpha_1})^2}{2\sigma_{\alpha_1}^2}\right]\exp\left[-\frac{(\alpha_0-\mu_{\alpha_0})^2}{2\sigma_{\alpha_0}^2}\right]\right.\\
&\qquad-\exp\left[-\frac{(\alpha_0-\mu_{\alpha_1})^2}{2\sigma_{\alpha_1}^2}\right]\\
&\qquad\left.\cdot\exp\left[-\frac{(\alpha_1-\mu_{\alpha_0})^2}{2\sigma_{\alpha_0}^2}\right]\right]\,d\alpha_1\,d\alpha_0. \qquad (40)
\end{aligned}$$

In (40), the integration is in the upper triangular portion, where $\alpha_1 < \alpha_0$, so $e^{-\alpha_1 t_i}-e^{-\alpha_0 t_i} > 0$. The second and third terms of the integrand of (40) are exponential functions. Therefore, they are always positive, and the fourth term, represented by $v(\alpha_1,\alpha_0)$, can be simplified to

$$\begin{aligned}
v(\alpha_1,\alpha_0)&=\exp\left[-\frac{\mu_{\alpha_1}^2}{\sigma_{\alpha_1}^2}-\frac{\mu_{\alpha_0}^2}{\sigma_{\alpha_0}^2}\right]\\
&\quad\cdot\left[\exp\left[-\frac{\alpha_1^2-2\mu_{\alpha_1}\alpha_1}{2\sigma_{\alpha_1}^2}-\frac{\alpha^2-2\mu_{\alpha_0}\alpha_0}{2\sigma_{\alpha_0}^2}\right]\right.\\
&\quad\left.-\exp\left[-\frac{\alpha_0^2-2\mu_{\alpha_1}\alpha_0}{2\sigma_{\alpha_1}^2}-\frac{\alpha_1^2-2\mu_{\alpha_0}\alpha_1}{2\sigma_{\alpha_0}^2}\right]\right]\\
&=\exp\left[-\frac{\mu_{\alpha_1}^2}{\sigma_{\alpha_1}^2}-\frac{\mu_{\alpha_0}^2}{\sigma_{\alpha_0}^2}\right]\\
&\quad\cdot\left[\exp\left[-\frac{\alpha_1^2\sigma_{\alpha_0}^2+\alpha_0^2\sigma_{\alpha_1}^2}{2\sigma_{\alpha_1}^2\sigma_{\alpha_0}^2}\right]\right.\\
&\quad\exp\left[\frac{\mu_{\alpha_0}\alpha_0\sigma_{\alpha_1}^2+\mu_{\alpha_1}\alpha_1\sigma_{\alpha_0}^2}{\sigma_{\alpha_1}^2\sigma_{\alpha_0}^2}\right]\\
&\quad-\exp\left[-\frac{\alpha_1^2\sigma_{\alpha_1}^2+\alpha_0^2\sigma_{\alpha_0}^2}{2\sigma_{\alpha_1}^2\sigma_{\alpha_0}^2}\right]\\
&\quad\left.\cdot\exp\left[\frac{\mu_{\alpha_1}\alpha_0\sigma_{\alpha_0}^2+\mu_{\alpha_0}\alpha_1\sigma_{\alpha_1}^2}{\sigma_{\alpha_1}^2\sigma_{\alpha_0}^2}\right]\right]. \qquad (41)
\end{aligned}$$

If it is assumed $(\mu_{\alpha_0}/\mu_{\alpha_1}) \geq (\sigma_{\alpha_0}^2/\sigma_{\alpha_1}^2) \geq 1$, then $\mu_{\alpha_0}\sigma_{\alpha_1}^2 \geq \mu_{\alpha_1}\sigma_{\alpha_0}^2$. Multiplying a positive value on each side of the inequality ($\alpha_1 \leq \alpha_0$) in the upper triangular integration area for $v(\alpha_1,\alpha_0)$, gives $(\alpha_0-\alpha_1)\mu_{\alpha_0}\sigma_{\alpha_1}^2 \geq (\alpha_0-\alpha_1)\mu_{\alpha_1}\sigma_{\alpha_0}^2$. This is equivalent to $\mu_{\alpha_0}\alpha_0\sigma_{\alpha_1}^2+\mu_{\alpha_1}\alpha_1\sigma_{\alpha_0}^2 \geq \mu_{\alpha_1}\alpha_0\sigma_{\alpha_0}^2+\mu_{\alpha_0}\alpha_1\sigma_{\alpha_1}^2$. Thus

$$\begin{aligned}
&\exp\left[\frac{\mu_{\alpha_0}\alpha_0\sigma_{\alpha_1}^2+\mu_{\alpha_1}\alpha_1\sigma_{\alpha_0}^2}{\sigma_{\alpha_1}^2\sigma_{\alpha_0}^2}\right]\\
&\quad\geq\exp\left[\frac{\mu_{\alpha_1}\alpha_0\sigma_{\alpha_0}^2+\mu_{\alpha_0}\alpha_1\sigma_{\alpha_1}^2}{\sigma_{\alpha_1}^2\sigma_{\alpha_0}^2}\right] > 0.
\end{aligned}$$

Also, because $\alpha_1 \leq \alpha_0$ and $\sigma_{\alpha_1}^2 \leq \sigma_{\alpha_0}^2$, then $\alpha_1^2(\sigma_{\alpha_0}^2-\sigma_{\alpha_1}^2) \leq \alpha_0^2(\sigma_{\alpha_0}^2-\sigma_{\alpha_1}^2)$. This yields

$$\alpha_1^2\sigma_{\alpha_0}^2+\alpha_0^2\sigma_{\alpha_1}^2 \leq \alpha_0^2\sigma_{\alpha_0}^2+\alpha_1^2\sigma_{\alpha_1}^2$$

then

$$\exp\left[-\frac{\alpha_1^2\sigma_{\alpha_0}^2+\alpha_0^2\sigma_{\alpha_1}^2}{2\sigma_{\alpha_0}^2\sigma_{\alpha_1}^2}\right]\geq\exp\left[-\frac{\alpha_0^2\sigma_{\alpha_0}^2+\alpha_1^2\sigma_{\alpha_1}^2}{2\sigma_{\alpha_0}^2\sigma_{\alpha_0}^2}\right] > 0.$$

Therefore, $v(\alpha_1,\alpha_0) \geq 0$. Then, $h_2(r_i) \geq 0$. To this point, it has been shown rigorously that all the terms in (38) are positive or nonnegative under the assumptions made above. Thus, $h(r_i) \geq 0$.

More generally, when $\alpha_1^l = \alpha_0^l = 0$, $\alpha_1^u = \alpha_0^u = +\infty$, $h(r_i) = h_2(r_i) \geq 0$.

At this point, we have proven $\partial\lambda/\partial r_i \geq 0$. Next, consider $\partial z/\partial r_i$

$$\frac{\partial z}{\partial r_i}=e^{-\hat{\mu}_{\alpha_1}t_i}-e^{-\hat{\mu}_{\alpha_0}t_i} > 0. \qquad (42)$$

Given (28), $\partial\lambda/\partial r_I \geq 0$ and $\partial z/\partial r_i \geq 0$, then $d\lambda/dz$ is nonnegative as well. Therefore, this makes the relationship stated in the proposition proved, i.e., the LRT and the GLRT/matched filter are essentially the same for this problem and provide the same performance.

V. Requirements of Noise Variance for Different Detectors to Achieve a Particular Performance Level

It is of interest to determine what noise level is required to achieve a particular P_d and P_{fa} for different signal processing techniques using the appropriate sensor data. With this motivation, we investigated the required noise variance for each detector in order to achieve a pre-determined performance using multichannel EMI sensor.

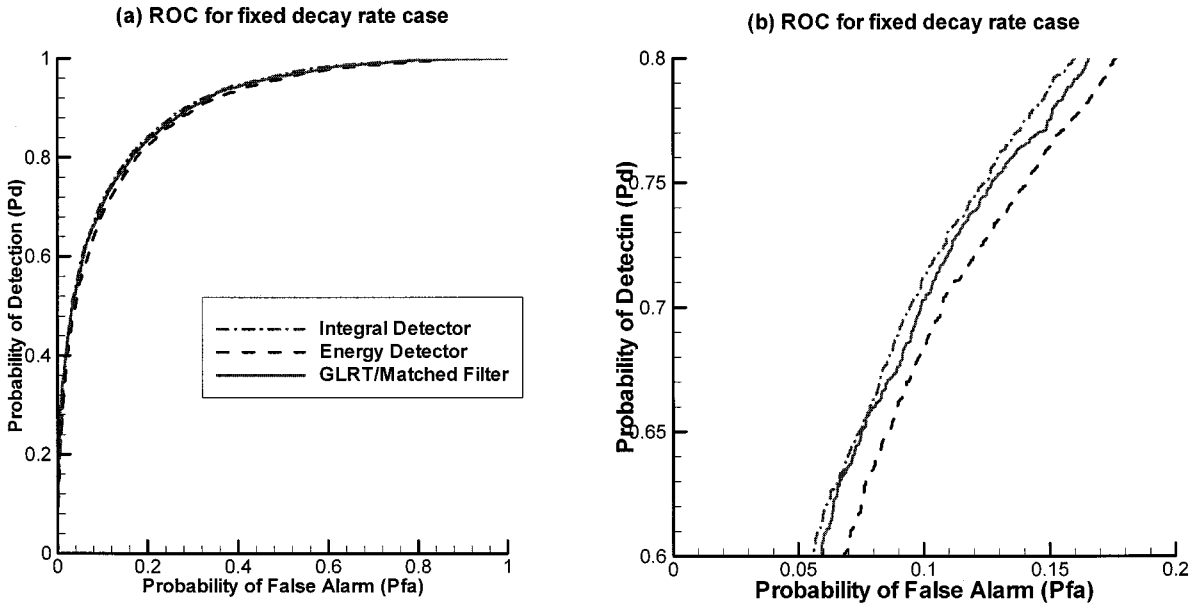


Fig. 8. ROC for various processors at the theoretically calculated noise variance required, achieving (0.1, 0.7) performance. (a) Entire scale. (b) Enlarged version of (a).

P_d and P_{fa} are defined as [13], [14]

$$P_d \equiv \int_{\gamma}^{\infty} f(x|H_1)\, dx$$
$$P_{fa} \equiv \int_{\gamma}^{\infty} f(x|H_0)\, dx. \tag{43}$$

If the outputs of the detectors follow a Gaussian distribution, it is possible to theoretically calculate the value of noise variance required for known parameters t, α_1, and α_0. For the matched filter in the known parameter case, x is Gaussian-distributed as derived above (Section III-A1). The distribution of x under H_1 is $x|H_1 \sim N(\mu_1, \sigma_1^2)$, where

$$\mu_1 = \sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i}) e^{-\alpha_1 t_i},$$
$$\sigma_1^2 = \sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i})^2 \sigma^2$$

and under H_0 $x|H_0 \sim N(\mu_0, \sigma_0^2)$, where

$$\mu_0 = \sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i}) e^{-\alpha_0 t_i})$$
$$\sigma_0^2 = \sum_i (e^{-\alpha_1 t_i} - e^{-\alpha_0 t_i})^2 \sigma^2$$

so [16]

$$P_d = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right] dx. \tag{44}$$

Upon letting $y = (x - \mu_1)/\sigma_1$

$$P_d = \int_{(\gamma \cdot \mu_1/\sigma_1)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(y^2/2)}\, dy = \text{cerf}\left(\frac{\gamma - \mu_1}{\sigma_1}\right) \tag{45}$$

where cerf$(\cdot)$ is the complementary error function defined by

$$\text{cerf}(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(x^2/2)}\, dx.$$

Similarly

$$P_{fa} = \text{cerf}\left(\frac{\gamma - \mu_0}{\sigma_0}\right). \tag{46}$$

The value of noise variance required to achieve a particular P_d and P_{fa} can be obtained by table look-up [16]. For the other detectors, (45) and (46) are still applicable. Corresponding to each detector, appropriate μ_1, σ_1^2, μ_0, and σ_0^2 values are substituted into (45) and (46).

For the random parameter case, no explicit expression for P_d and P_{fa} could be derived, thus no calculation of the theoretical noise variance was made. However, the variance can be estimated through simulation. Table I lists the noise variance required to achieve $P_{fa} = 0.1$ and $P_d = 0.7$ for both a fixed and a random parameter case.

The values in Table I were obtained by using signal $A_1 e^{-\alpha_1 \mathbf{t}}$ for H_1 and $A_0 e^{\alpha_0 \mathbf{t}}$ for H_0. Here, $\mathbf{t}$ is a 20×1 vector of [.3525, .4275, .525, .6475, .8025, 1.003, 1.258, 1.583, 1.998, 2.525, 3.198, 4.055, 5.148, 6.543, 8.323, 10.59, 13.49, 17.19, 21.90, 27.92] measured in ms and $A_1 = A_0 = 220$ (equivalent to using $A_1 = A_0 = 1$). In the fixed parameter case, $\alpha_1 = 2.1$, and $\alpha_0 = 2.5$. These values are chosen based on the DARPA experimental data [5]. In the random environment case, $\alpha_1 \sim \mathcal{N}(2.1, 0.2^2)$ and $\alpha_0 \sim \mathcal{N}(2.5, 0.2^2)$. These values and distributions were again chosen based on the statistics of real data. From Table I, it can been concluded that to achieve a particular level of performance for either the fixed or random parameter case, the matched filter can always withstand the highest noise level among the four algorithms investigated. Additionally, an energy detector can achieve the same performance as an integral detector under a higher noise level, or lower SNR. In addition, as we expected if the environment is known exactly, the algorithms can afford lower SNR. In practice, however, the environment is always uncertain.

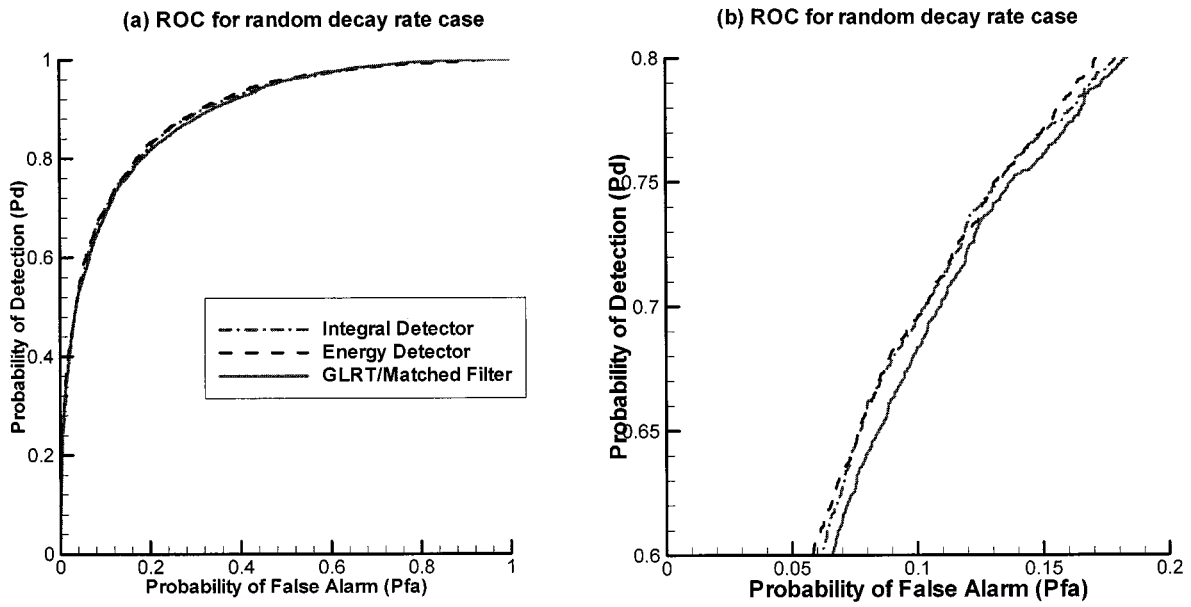


Fig. 9. ROC for various processors at the noise variance estimated from simulations to achieve (0.1, 0.7) performance. (a) Entire scale. (b) Enlarged version of (a).

VI. Simulations

Simulations for EMI sensor data using different detection strategies were implemented. The synthetic data was obtained by adding Gaussian white noise to an exponentially damped signal. In the parameter-known case, the decay rate of the target α_1 is 2.1, and that of clutter α_0 is 2.5. Those values were chosen by inspecting histograms of the decay rates from targets and clutter obtained from DARPA data, and they are the empirical MLE. For the random parameter case, the distributions of the decay rate and noise variance were also estimated from the DARPA experiment data.

The various signal processing approaches have been implemented on synthetic data. Figs. 3 and 4 show the ROC's for the integral, energy and matched filter detectors and the threshold detector of single time sample data (at the sample time that maximizes the performance) for both fixed parameter and random parameter cases, respectively. As expected, the performance of the LRT/GLRT is better than that of energy detector, which in turn is better than the integral detector, and the performance of threshold detector using single time sample data is the worst among all the detectors considered, since the remaining detectors use multiple time samples of EMI data as their inputs. Additionally, Fig. 4 shows the performance of the LRT when parameters are uncertain, in which Monte Carlo integration was used to compute the integrals in (15). $f(r|H_1) = \sum_{k=1}^{M} f(r|H_1, \alpha_{ik})/M$ was used to evaluate the integral in the numerator and denominator of (15), where i is 1 or 0, α_{ik} is chosen from Gaussian distributions with means and variances stated above. One thousand iterations were used when calculating the integral.

It is clear that for the random parameter case, the ROC for the LRT is equivalent to that of the matched filter (GLRT), which agrees the theoretical derivation presented in Section III. Furthermore, a simulation obtained when both the decay rate as well as the initial amplitude are random is shown in Fig. 5. Again, the LRT and the GLRT exhibit the same performance.

Next, the theoretical calculation of d'^2 and simulation results for the processors were compared. Table II lists the d'^2 values for each of the detection algorithms under both fixed and random parameter cases. Figs. 6 and 7 compare the ROC's generated for the Gaussian detection problem with the calculated d'^2 values and simulation results. ROC's for the theoretical values match the simulation results consistently in all cases. Note that when the variances of the decay rates of target and clutter are small, the uncertainty has almost no effects on the performance.

Again, Table I lists the theoretical noise variance required to achieve $P_{fa} = 0.1$ and $P_d = 0.7$ for the fixed parameter case and the estimated noise variance required to achieve the same performance for the random parameter case. These estimated values were determined by adjusting the noise variance during the simulations. Fig. 8 shows that the ROC curves for the various algorithms at the theoretical required noise variances for the fixed parameter case. It can be seen that the ROC curves approximately cross the point (0.1, 0.7), indicating the simulations verify the theoretical calculations. Fig. 9 shows similar ROC curves for random decay rates with the simulated noise.

VII. Discussion

We have presented results which indicate that a detection theory based approach (LR test) can be used to improve the detectability of land mines and small UXO objects using EMI sensors over standard thresholding techniques. In [3], the results were verified by implementing the processor on real mine and clutter field data, however, only a few targets were present in that case. In this paper, simulation results for each of the processors also verifies the improvement obtained using a LRT detector. Figs. 3 and 4 show that regardless of whether or

not the parameters are known constants or random variables, an LRT/GLRT/matched filter of multichannel time-domain EMI data always provides the best performance among the algorithms investigated here, and an energy detector performs better than an integral detector. The optimal detector on single time sample of time-domain EMI data is the threshold test, and its performance is worst among all the detectors considered in the paper. This indicates that developing a new sensor which provides more of the time-domain decay curve could prove advantageous for improved landmine detection.

To avoid computational expense, the GLRT is often used as a suboptimal detector. In some cases, the performance of the GLRT can be as good as that of the LRT. In [2], [3], we have shown that for a single-channel (integrated) time-domain EMI sensor, the LRT reduces to the GLRT. In this study, we have shown that for a multichannel time-domain EMI sensor, the GLRT is optimal under certain assumptions on the distributions of the decay rates.

Future work will involve incorporating more environmental parameters into the EMI model. To do so, a more complicated model is required, and a further performance improvement is expected.

ACKNOWLEDGMENT

The authors would like to thank Dr. L. Carin, Dr. R. Dugan, Dr. T. Altshuler, V. George, Dr. L. Nolte, and Dr. S. Tantum for helpful discussions regarding this research. They are particularly grateful to Dr. A. Barnes for his generous help on the mathematical legality.

REFERENCES

[1] A. M. Andrews, V. George, and T. W. Altshuler, "Quantifying performance of mine detectors with fewer than 10 000 targets," in *Proc. SPIE*, vol. 3079, Orlando, FL, Apr. 1997, pp. 273–281.

[2] P. Gao, "Improved approaches to land mine remediation using signal detection and estimation theory," M.S. thesis, Duke Univ., Durham, NC, Dec. 1997.

[3] L. M. Collins, P. Gao, and L. Carin, "An improved Bayesian decision theoretic approach for land mine detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 811–819, Mar. 1999.

[4] P. Gao and L. M. Collins, "Improved signal processing approaches for landmine detection," in *Proc. SPIE: Detection Technologies for Mines and Minelike Targets*, Orlando, FL, Apr. 1998.

[5] V. George, T. Altshuler, A. Andrew, J. Nicoll, E. Cespedes, D. Butler, T. Broach, and R. Mehta, Background Data Collection Plan, Defense Adv. Res. Projects Agency (DARPA), Def. Sci. Office, Arlington, VA, Dec. 1996.

[6] G. D. Sower and S. P. Cave, "Detection and identification of mines from natural magnetic and electromagnetic resonances," in *Proc. SPIE*, vol. 2496, Orlando, FL, Apr. 1995, pp. 1015–1024.

[7] N. Geng, P. Garber, L. Collins, L. Carin, D. Hansen, D. Keiswetter, and L. J. Won, "Wideband electromagnetic induction for metal-target identification: Theory, measurement and signal processing," Tech. Rep., Duke Univ., Durham, NC, 1998.

[8] C. E. Baum, "Low-frequency near-field magnetic scattering from highly, but not perfectly, conducting bodies,", Albuquerque, NM, Phillips Lab. Interaction Note 499, Nov. 1993.

[9] Y. Das, J. E. McFee, and R. H. Cherry, "Time-domain response of a sphere in the field of a coil: Theory and experiment," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-22, pp. 360–367, July 1984.

[10] Y. Das, J. E. McFee, J. Toews, and G. C. Short, "Analysis of an electromagnetic induction detector for real-time location of buried objects," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 278–288, May 1990.

[11] A. D. Whalen, *Detection of Signal in Noise*. New York: Academic, 1971, p. 352.

[12] N. Geng, C. E. Baum, and L. Carin, "On the low-frequency natural response of conducting and permeable targets," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 347–359, Jan. 1999.

[13] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.

[14] L. Selin, *Detection Theory*. Princeton, NJ: Princeton Univ. Press, 1965.

[15] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*. New York: Wiley, 1984.

[16] M. H. DeGroot, *Probability and Statistics*. Reading, MA: Addison-Wesley, 1989.

Ping Gao (S'97) receved the B.S. and M.E. degrees (honors) in electrical engineering from Beijing University, Beijing, China, in 1993 and 1996, respectively, and the M.S.E.E. from Duke University, Durham, NC, in 1997, where she is currently pursuing the Ph.D. degree in electrical and computer engineering.

Since 1997, she has been with Duke University as a Research Assistant. Her main research interests are in signal detection and estimation theory, statistical signal modeling and processing, and applied electromagnetics.

Leslie M. Collins (S'83–M'96) [redacted] She received the B.S.E.E. degree from the University of Kentucky, Lexington, in 1985, and the M.S.E.E. and Ph.D. degrees in electrical engineering from the University of Michigan, Ann Arbor, in 1986 and 1995, respectively.

She was a Senior Engineer with Westinghouse Research and Development Center, Pittsburgh, PA, from 1986 to 1990. Since 1995, she has been with the Electrical and Computer Engineering Department, Duke University, Durham, NC. Her research interests include incorporating physics-based models into statistical signal processing algorithms, and she is pursuing applications in subsurface sensing and applied electromagnetics, as well as enhancing speech understanding by hearing impaired individuals.

Dr. Collins is a member of the Tau Beta Pi, Eta Kappa Nu, and Sigma Xi societies.

On the Interaction of Elastic Waves with Buried Land Mines: an Investigation Using the Finite-Difference Time-Domain Method

A Thesis
Presented to
The Academic Faculty

by

Christoph T. Schröder

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy in Electrical and Computer Engineering

Georgia Institute of Technology
July 2001

On the Interaction of Elastic Waves with Buried Land Mines: an Investigation Using the Finite-Difference Time-Domain Method

Approved:

Waymond R. Scott, Jr., Chairman

Glenn S. Smith

Andrew F. Peterson

W. Marshall Leach

Peter H. Rogers

Date Approved _____

ACKNOWLEDGMENTS

This dissertation is the result of my studies at the Georgia Institute of Technology between the fall of 1996 and the summer of 2001. Many people have, in one way or the other, contributed to this work, and I am grateful to all of them.

I owe my deepest gratitude to Prof. Waymond R. Scott, Jr., who served as my thesis advisor and has guided my way at Georgia Tech from the very first moments. His expertise in both electromagnetics and elasticity, his curiosity and his interest have taught me, have kept my spirit up through all these years and have made my work truly worthwhile.

I wish to thank the members of my dissertation committee, Prof. Glenn S. Smith, Prof. Andrew F. Peterson, Prof. W. Marshall Leach, and Prof. Peter. H. Rogers, together with Prof. Scott, for their time, effort, and constructive criticism. I am especially grateful to Prof. Smith, who inspired me with his vast knowledge in electromagnetics and his understanding of wave propagation. He always had an open ear for discussions of any kind, making long hours at school enjoyable.

Dr. Gregg D. Larson and James S. Martin, together with Prof. Scott, have developed the elastic/electromagntic sensor system. They performed numerous measurements specifically for this work, results of which will be shown within this text. Their insight and knowledge of elasticity were extremely helpful.

Benny Venkatesan has built and administered the Beowulf computer cluster that was used for most of the computations in this work. The cluster proved to be invaluable providing the computational resources required for this work.

My parents, Detlef and Hannelore Schröder, and my brother, Achim Schröder,

have lovingly supported me throughout these years. Without their help and encouragement nothing of this would have been possible. I regret having been parted from them for so long, and look forward to being closer to them in the future.

I furthermore want to express my gratitude to the many friends I have found during my time in Atlanta. Padma Rao has been closest to me, and I am grateful for her friendship and her faith in me. Thorsten Hertel has known me the longest, and has come a long way with me since we started our studies in the city of Braunschweig in Germany in 1993. I appreciate his friendship and his kind helpfulness. I would also like to thank Peter Knobel, Stefan Galler and Harald Seckel for being good friends and the World Student Fund, skillfully and warmly steered by Carlton Parker, for providing the opportunity for so many international students to study in the United States.

This work has been supported in part under the OSD MURI program by the US Army Research Office under contract DAAH04-96-1-0448, by a grant from the US Office of Naval Research under contract N00014-99-1-0995, and by an equipment grant from the Intel Corporation.

Contents

ACKNOWLEDGMENTS ii

LIST OF TABLES ix

LIST OF FIGURES x

SUMMARY xv

1 Introduction 1

2 Background 5

3 Elastic Wave Propagation in Solids 10

3.1 Governing Equations 10

3.2 Elastic Waves in Isotropic Media 13

3.2.1 The Wave Equation in Isotropic Media 13

3.2.2 Propagating Waves 14

4 Numerical Model 18

4.1 The Finite-Difference Scheme 18

4.1.1 Problem Statement 18

4.1.2 Governing Equations 20

4.1.3 Discretization of the Governing Equations 23

4.2 Boundaries 30

4.2.1 Source 30
4.2.2 Internal Boundaries 30
4.2.3 Free-Surface Boundary 32
4.2.4 Perfectly-Matched Layer Absorbing Boundary 36
4.3 Injection of Plane Waves 39
4.3.1 Procedure 41
4.3.2 Incident Field 44
4.4 Parallelization 45

5 Stability of the FDTD Algorithm at a Material Interface 47
5.1 Stability Analysis: Theory 48
5.1.1 1-D Longitudinal Wave Incident onto a Material Interface 49
5.2 Stability Analysis: Numerical Results 61
5.2.1 1-D Case 61
5.2.2 2-D Case 67
5.3 Averaging 71
5.3.1 Material Density 74
5.3.2 Lamé's Constants 77
5.4 The Courant Condition 79
5.5 Concluding Remarks 80

6 Elastic Surface Waves 81
6.1 The Rayleigh Equation 82
6.2 The Roots of the Rayleigh Equation 86
6.2.1 Materials with a Poisson Ratio larger than 0.263 87
6.2.2 Materials with a Poisson Ratio smaller than 0.263 95
6.3 Waves due to a Line Source on the Surface 101
6.3.1 General Considerations 103
6.3.2 Steepest-Descent Approximation 106
6.3.3 Example 114

7 Propagation and Scattering of Elastic Waves in the Ground 120

7.1 Surface Waves 124

7.2 Interaction of Elastic Waves with Buried Land Mines 133

7.2.1 2-D Analysis 135

7.2.2 3-D Analysis 142

7.3 Mines in the Presence of Clutter 151

7.4 A Moving-Source/Moving-Receiver System 155

8 Modeling of Resonant Structures: the Resonance Behavior of a Buried Land Mine 162

8.1 Discretization of Elastic Objects 163

8.1.1 Long Thin Bar 163

8.1.2 Thin Circular Plate 167

8.1.3 Tuning Fork 170

8.2 Resonance Behavior of a Buried Land Mine 174

8.3 Excitation by a Pressure Wave 185

9 Experiments 191

9.1 Measuring the Water Content in the Ground 192

9.2 Measuring the Wave Speeds in the Ground 196

9.2.1 Experimental Set-Up 197

9.2.2 Measurement Procedure 199

9.2.3 Results 203

9.3 Measurements of the Mine-Wave Interaction 215

10 Summary and Conclusions 217

APPENDIX A 3-D FDTD Equations 221

APPENDIX B Total-Field/Scattered-Field Formulation: Correction Terms 226

APPENDIX C Software Documentation **229**

C.1 Input File 230

C.1.1 General Simulation Parameters 230

C.1.2 Sources 232

C.1.3 Objects 234

C.1.4 Background 236

C.1.5 Materials 237

C.1.6 Sample Input File 237

APPENDIX D Elastic Constants for Isotropic Solids **239**

BIBLIOGRAPHY **240**

VITA **247**

List of Tables

4.1 Positions of the field components of the (i, j, k)-th cell in the finite-difference grid. 27

6.1 Solutions to the Rayleigh equation for $\nu > 0.263$. 89
6.2 Solutions to the Rayleigh equation for $\nu = 0.4$: wave numbers. 90
6.3 Solutions to the Rayleigh equation for $\nu < 0.263$. 100
6.4 Solutions to the Rayleigh equation for $\nu = 0.2$: wave numbers. 101

7.1 Material properties. 121

8.1 Wave numbers and wave length of the first two harmonics of a bar with free ends according to the Bernoulli-Euler theory. 164
8.2 Resonant frequencies of a long thin bar. 165
8.3 Resonant frequencies of the first three modes of a thin circular disc that is free along its edge. 169
8.4 Resonant frequencies of the thin circular plate as computed with the numerical model. 169
8.5 Resonant frequencies of a tuning fork. 172
8.6 Material properties of the parts of the detailed mine model. 174

9.1 Measurement of the water content as a function of depth. 194
9.2 Measured arrival times and wave speeds: pressure wave front. 205

B.1 Field components and correction terms at Interface I and II. 227
B.2 Field components and correction terms at Interface III and IV. 228
B.3 Field components and correction terms at Interface V and VI 228

List of Figures

1.1 Schematic drawing of the elastic-electromagnetic sensor. 3

3.1 Infinitesimally small cube in a cartesian coordinate system with the stresses acting onto its surfaces. 12
3.2 Elastic waves due to a point source on a free surface. 17

4.1 Problem geometry; top: isometric view, bottom: cross sectional view. 19
4.2 Three-dimensional finite-difference model. 21
4.3 Three-dimensional finite-difference basis cell. 25
4.4 A portion of the three-dimensional finite-difference grid. 26
4.5 Schematic drawing of the finite-difference grid at the interface between two media. The normal particle velocity components are always located on the interface. 31
4.6 Finite difference grid and its field components at the free surface; cross section at $y = (j - 0.5)\Delta y$. An extra row at $k = -1$ must be inserted to satisfy the stress-free boundary condition at the free surface. . . . 34
4.7 Schematic drawing of a PML layer in the x-direction: loss profile according to Eq. (4.64). The crosses and circles indicate the loss values of the field components at the center and at the edges of a cell, respectively. 40
4.8 Division of the solution space into a *Total-Field Region* and a *Scattered-Field Region*. The scattering objects must be completely embedded in the total-field region. 41
4.9 Both the field components just outside and just inside the total-field box must be corrected with the known incident fields. 42
4.10 Parallelization; the solution space is divided into several cubes, which are assigned to different processor. 46

5.1 1-D finite-difference grid for a longitudinal wave normally incident onto a material boundary. 50
5.2 Stability bounds due to the presence of a material interface. 56
5.3 Stability bound for the actual matrix (A) and estimate from the eigenvalues of the modified matrix (B). 59
5.4 Normalized stability condition; 1-D case. The material properties are not averaged on the boundary. 63

5.5 Normalized stability condition; 1-D case. The material properties are averaged on the boundary. 64
5.6 Normalized stability condition; 1-D case; for four constant velocity ratios. 65
5.7 Boundary arrangements used for the 2-D analysis. (a) *Boundary A*, (b) *Boundary B*. 68
5.8 Normalized stability condition; 2-D case. The material parameters are not averaged on the boundary. The boundary is placed as in Fig. 5.7 (a). 70
5.9 Normalized stability condition; 2-D case; boundary placed according to Fig. 5.7 (b). 72
5.10 Comparison of 1-D and 2-D stability behavior; four velocity ratios. . 73
5.11 Boundary between two media in a 2-D finite-difference grid; the normal particle velocities and the shear stress are located on the interface. . . 75
5.12 Averaging of the material density ρ; the normal stress components are integrated along the contour. 76
5.13 Averaging of the shear stiffness μ; the tangential velocity components are integrated along the contour. 78

6.1 Schematical arrangement of the roots in the complex ξ-plane for $\nu > 0.263$. 88
6.2 Horizontal and vertical displacements according to the five solutions of the Rayleigh equation. 91
6.3 Allowed wave vector surfaces in the k-space for the pressure wave (α) and the shear wave (β); lossless case. 93
6.4 Allowed wave vector surfaces for the five solutions to the Rayleigh equation. 96
6.4 continued. 97
6.4 continued. 98
6.5 Schematical arrangement of the roots in the complex ξ-plane for $\nu < 0.263$. 99
6.6 Magnitude of the displacements according to the four solutions due to the two additional real roots for $\nu < 0.263$. 102
6.7 Line source on the surface. 103
6.8 Location of the poles and branch cuts in the complex ξ-plane for the line-source problem. 105
6.9 Location of the poles and branch cuts (a) in the complex w^S-plane and (b) in the complex w^P-plane. 108
6.10 Steepest descent paths for (a) the shear wave terms in the complex w^S-plane and (b) the pressure wave terms in the complex w^P-plane. . 111
6.11 Waves due to a point source on the surface. From top to bottom: shear wave, pressure wave, Rayleigh surface wave, leaky surface wave, lateral wave. 116
6.12 Finite-difference results; comparison to asymptotic solution. 117

6.13 Hodograms of the particle motion at the surface. Plots for the vertical displacement vs. the horizontal displacement for (a) the Rayleigh surface wave, (b) the leaky surface wave, and (c) the lateral wave. . . 119

7.1 Variation of the material properties with depth. 122
7.2 Waterfall graphs of the vertical particle displacement on the surface according to (a) experiment and (b) numerical simulation. The ground is assumed to be homogeneous in the numerical model. 125
7.3 Normal particle displacement on the surface (top) and on a cross section through the ground (bottom). The ground is assumed to be homogeneous. T_1 and T_2 correspond to the vertical lines in Fig. 7.2 (b). 127
7.4 Top: particle motion on the surface at a distance of 60 cm from the source as a function of time. Bottom: vertical vs. horizontal displacement for the leaky surface wave and the Rayleigh surface wave. Homogeneous case. 129
7.5 Waterfall graphs of the vertical particle displacement on the surface according to (a) experiment and (b) numerical simulation. The material properties in the numerical model are assumed to vary with depth. . 131
7.6 Normal particle displacement on the surface (top) and on a cross section through the ground (bottom). T_1 and T_2 correspond to the vertical lines in Fig. 7.5 (b). Depth-varying material properties are assumed. 132
7.7 Top: particle motion on the surface at a distance of 60 cm from the source as a function of time. Bottom: vertical vs. horizontal displacement for the leaky surface wave and the Rayleigh surface wave. Depth-varying material properties are assumed. 134
7.8 Schematic drawing of possible set-ups for the experimental system. . 135
7.9 Simple model for the TS-50 anti-personnel mine; cross-sectional drawing and photograph of a real TS-50 mine. 136
7.10 2-D model for an anti-personnel mine. 137
7.11 Interaction of elastic waves with a buried anti-personnel mine; 2-D results. 138
7.12 Model to determine the nature of the resonance: elongated mine. . . 139
7.13 Flexural waves propagating above the mine. 141
7.14 Wave paths through and around the mine. The incident Rayleigh wave (R) and leaky surface wave (LS) couple into a longitudinal wave (L) and a transverse flexural wave (TF) in the thin plate above the mine and into a wave passing through the plastic body of the mine (M). . . 141
7.15 Interaction of elastic waves with a buried anti-personnel mine; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instants in time. 143

7.16 (a) Experimental results for a buried TS-50 anti-personnel mine and (b) numerical results for a simple anti-personnel mine model; waterfall graphs of the vertical particle displacement on the surface. Land mine buried at 1 cm, 2 cm, 3 cm, and 6 cm (from left to right). 145
7.17 Interaction of elastic waves with a buried spherical object; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instants in time. 147
7.18 Numerical results for a buried sphere; waterfall graphs of the vertical particle displacement on the surface. Sphere buried at 1 cm, 2 cm, 3 cm, and 6 cm (from left to right). 148
7.19 Simple model for the VS 1.6 anti-tank mine; schematic picture of the model and photograph of a VS 1.6 mine. 148
7.20 Interaction of elastic waves with a buried anti-tank mine; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instants in time. 150
7.21 (a) Experimental results for a buried VS 1.6 anti-tank mine and (b) numerical results for a simple anti-tank mine model; waterfall graphs of the vertical particle displacement on the surface. Land mine buried at 2 cm, 4 cm, 6 cm, and 10 cm (from left to right). 152
7.22 Configuration of the mines and rocks in the ground; the rectangles indicate the rocks, the circles indicate the mines. The boxed numbers correspond to the burial depths. 153
7.23 Interaction of elastic waves with four anti-personnel mines, one anti-tank mine and four rocks buried in the ground; pseudo color plots of the normal particle displacement on the surface at four instants in time. 154
7.24 Image formed from the interaction of elastic waves with the buried mines using an imaging algorithm; (a) numerical data, (b) experimental data. The image describes a surface area of 120 cm by 80 cm. . . 156
7.25 Schematic drawing of the moving-source/moving-receiver system. . . 157
7.26 Moving-source/moving-receiver system; water fall graphs of the vertical particle displacement at the receiver location; for four distances between source and receiver: 0 cm, 10 cm, 20 cm, 30 cm (from left to right). (a) Scaled to the maximum value of each graph; (b) not scaled. The peaks in (b) are clipped for the 0 cm- and the 10 cm-case. 160
7.27 Signal paths for three positions of the source-receiver system. 161

8.1 Long thin bar. 164
8.2 (a) Resonant frequencies of the bar and (b) error relative to the theoretical prediction as a function of the number of nodes across the width of the bar. 166
8.3 Thin circular plate. 167
8.4 Modes of a circular plate with free edge. The bold lines indicate the nodal diameters and nodal circles. 168

8.5 Tuning fork. 170
8.6 Schematic drawing of the first four modes of a tuning fork. 171
8.7 The motion of a tuning fork. 173
8.8 Refined mine model that includes the structural details of an actual TS-50 mine such as the explosives, the case, the rubber pressure plate, and the air chambers. 175
8.9 A vertically polarized shear wave is incident on the buried mine. The elastic wave fields are computed only in the surrounding of the mine. 177
8.10 Resonant oscillations at the location of a buried land mine; mine buried 2 cm beneath the surface. The vertical particle displacement in the surrounding of the mine as computed with the detailed mine model is shown at four instants in time on the surface plane (top) and on a cross section through the ground (bottom). 178
8.11 Transfer function of a point on the surface above the mine; mine buried at 1 cm; $c_S = 40$ m/s. The two resonant peaks correspond to two distinct modes of the vertical surface motion. 180
8.12 Surface motion at a point above the mine; transfer functions. (a) Mine buried 2 cm beneath the surface, variable shear wave speed; (b) shear wave speed $c_S = 40$ m/s, variable depth. 182
8.13 Parametric graphs of the resonant frequency and the quality factor of the first resonant peak as a function of the shear wave speed in the ground with the burial depth as a parameter. 183
8.14 Experimental results for a mine buried at 1 cm; transfer function. . . 185
8.15 A pressure wave is vertically incident on the buried mine. 187
8.16 Transfer functions of the surface motion due to the excitation with shear wave from the side and a pressure wave from above. 188
8.17 Surface motion at a point above the mine; tranfer functions. Excitation with a pressure wave from above. (a) Mine buried 2 cm beneath the surface, variable shear wave speed; (b) shear wave speed $c_S = 40$ m/s, variable depth. 189
8.18 Parametric graphs of the resonant frequency and the quality factor of the first resonant mode as a function of the shear wave speed in the ground with the burial depth as a parameter. Pressure wave excitation. 190

9.1 Water content (mass percent) as a function of depth. 193
9.2 Material density as a function of depth. 195
9.3 Pressure wave speed and shear wave speed as a function of depth, if the stiffness in the ground is assumed to be constant and the density changes as in Fig. 9.2. 196
9.4 Set-up for accelerometer measurements. (a) Transducer and accelerometer array; (b) measurement set-up. 198

9.5 Signal processing: the accelerometer signal is Fourier transformed, divided by the Fourier transform of the drive signal (chirp), multiplied by the reference signal (differentiated Gaussian pulse), and then inverse-Fourier transformed. 200
9.6 Accelerometer and transducer positions in the laboratory tank. . . . 202
9.7 Horizontal and vertical particle acceleration as a function of time and depth; accelerometer measurements. The distance from transducer to accelerometers is 1.5 m. P: pressure wave, LS: leaky surface wave, S: shear wave, GS: guided shear wave, R: Rayleigh wave. 204
9.8 Pressure wave speed and shear wave speed as a function of depth. . . 207
9.9 Horizontal and vertical particle acceleration as a function of time and depth; numerical simulation. The distance from transducer to accelerometers is 1.5 m. P: pressure wave, LS: leaky surface wave, S: shear wave, GS: shear wave, R: Rayleigh wave. 208
9.10 Horizontal and vertical particle acceleration as a function of time and depth; for three transducer distances: 0.5 m, 1 m, 1.5 m; left: accelerometer measurements, right: numerical simulation. 209
9.11 Frequency-time decomposition. The transfer function is convolved with a Gaussian pulse shifted in frequency, and thus the signal excited in a limited frequency band is determined. 212
9.12 Frequency-time decomposition; particle motion 1.5 m from source, at a depth of 5 cm. Top: horizontal component; bottom: vertical component. Left: measurement (acceleration); right: numerical simulation (displacement). 213
9.13 Transfer functions for the horizontal and the vertical field component; distance to source is 1.5 m, at a depth of 5 cm. Left: measurement (acceleration), right: numerical simulation (displacement). 214
9.14 Schematic drawing of the radar measurement. 216

B.1 Numbering of the interfaces. 227

SUMMARY

A three-dimensional finite-difference time-domain model has been developed to simulate the propagation of elastic waves in an elastic half space. The model incorporates a free-surface boundary condition for the surface of the half-space and a perfectly-matched layer absorbing boundary condition. This thesis includes a detailed description of the numerical model, a theoretical study of the stability of the finite-difference algorithm at a material interface, an analytical derivation of the elastic surface waves that arise at the surface of a solid medium, an extensive analysis of the interaction of elastic waves with buried land mines, a description of the resonant behavior of various mechanical objects, and results of experimental measurements of the elastic wave motion in a sand tank. The numerical model is primarily used to investigate the elastic wave motion in the ground and to explore the interaction of elastic waves with buried land mines. Various aspects of the mine-wave interaction have been studied. Simple mine models are used to explore the interaction with buried anti-personnel and anti-tank mines on a large scale. A more detailed mine-model is utilized to analyze the resonant behavior of buried mines. When elastic waves interact with a buried land mine, resonant oscillations occur at the location of the buried mine. These resonant oscillations enhance the mine's signature and distinguish it from clutter objects such as rocks. The resonance has been found to be largely dependent on the soil properties in the vicinity of the mine and on the burial depth of the mine. Although the numerical model is primarily used for the the mine-detection problem, it is far more versatile and applicable to general problems in the field of elasticity. As an example for the model's versatility, the resonant motion of a tuning fork, as computed with

the numerical model, is illustrated in this text.

CHAPTER 1

Introduction

In today's technology, elastic waves in solid media are utilized for various applications in a wide range of different fields. Surface Acoustic Wave devices are found in many TV-sets and mobile telephones as frequency filters with excellent characteristics. Quartz resonators are used in watches and computers and wherever a fixed clock rate must be maintained. Ultrasound imaging is used for medical diagnosis, and the number of applications for acousto-optic devices is steadily growing. In this work, the elastic wave motion in solids is revisited, and a numerical model is presented that simulates the elastic wave propagation in a solid medium.

Recently, the use of elastic waves to detect buried land mines has been proposed [1] – [12]. With more than 100 Million land mines buried throughout the world causing an estimated 26,000 injuries and deaths each year, the public attention on the problem of buried land mines has increased dramatically in recent years, resulting in extensive research efforts to develop more effective detection systems. Acoustic (elastic) sensor systems might, under certain circumstances, bear significant advantages over systems that are in use today. Ground-penetrating radars (GPR), for example, utilize electromagnetic waves to locate buried mines and are, thus, dependent on the contrast in the dielectric properties between a mine and the surrounding soil. In other words, ground-penetrating radar systems often fail if the metal content of a buried mine is low, as it is the case, for example, for many plastic-cased land mines. With a GPR, plastic mines are almost impossible to detect in dry soils, if the soil

and the mine have similar dielectric properties. Elastic systems, on the other hand, exploit the differences in the elastic properties of a buried mine and its surrounding soil, which are, in general, substantial. Furthermore, elastic systems and ground-penetrating radar systems might complement one another in certain scenarios. For example, the electromagnetic loss in a wet soil is significant, degrading the functionality of a ground-penetrating radar. The elastic loss in a wet soil, however, in general is reduced, thus enhancing the response in an elastic sensor.

An elastic land mine detection system utilizes elastic waves that propagate in the soil. To investigate the elastic wave motion in solid media, a numerical model has been developed. The model is based on the equation of motion and the stress-strain relation, which, together with a constitutive relation, form a set of first-order partial differential equations that completely describes the elastic wave motion in a medium. Introducing finite differences, this set of equations is discretized and adapted to the finite-difference time-domain (FDTD) modeling scheme. The numerical model has been implemented in two and three dimensions. In the model, the ground is approximated as an infinite half space, the surface of the ground is modeled as a stress-free boundary, and an absorbing boundary condition is implemented to avoid artificial reflections at the outer faces of the numerical grid.

The model has been developed to supplement the development of an experimental sensor system that simultaneously uses both elastic and electromagnetic waves to detect buried land mines [12] – [15]. Fig. 1.1 illustrates the experimental set-up for the elastic-electromagnetic sensor. An electrodynamic transducer is placed on the surface of the ground and excites elastic waves in the ground. The elastic waves propagate mainly along the surface and interact with the buried land mine, causing both the mine and the ground to vibrate. Due to the presence of the mine, the ground will move differently around the mine than elsewhere. A radar, mounted above the surface of the ground, scans the surface, records the vibrations and, thus, detects the mine.

Although the numerical model is primarily applied to investigate the elastic

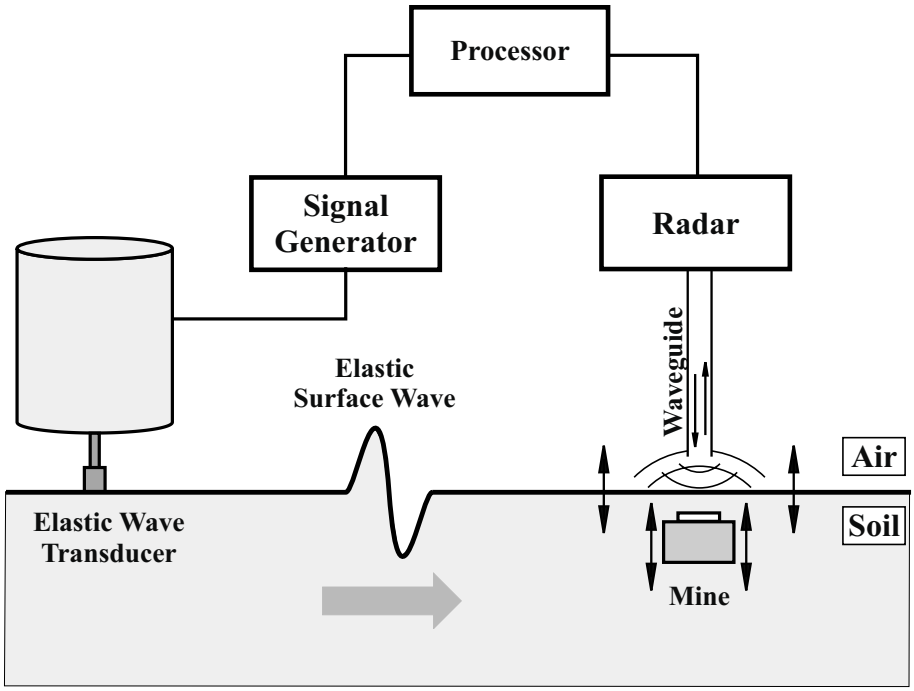


Figure 1.1: Schematic drawing of the elastic-electromagnetic sensor.

wave motion in the ground, it is applicable to a wide range of general elastic problems. With minor adjustments, the numerical model could be used, for example, to simulate ultrasonic imaging systems, to compute the wave motion in Acoustic-Wave devices, or to determine the acoustic wave fields of loud speaker systems. As an example for the applicability of the model, the resonant motion of a tuning fork, as computed with the numerical model, is presented in this text.

This dissertation gives a detailed description of elastic wave motion in solids and explains various aspects of the interaction of elastic waves with buried mines. The text is largely divided into two parts. Chapters 2 – 6 are theoretical in nature and describe the numerical model and its theoretical foundation. Chapters 7 – 9, on the other hand, describe the practical application of the model and explain the elastic wave propagation in the ground.

In Chapter 2, the background of numerical modeling using the finite-difference

time-domain method is described, and the history of the algorithm and the fundamental literature that lead to the numerical model as presented in this text are outlined. In Chapter 3, the fundamental equations governing the elastic wave motion in solids are explained. In Chapter 4, the numerical model and its implementation are described in detail. In Chapter 5, the stability of the finite-difference algorithm at a material interface is studied and results are presented that indicate that the finite-difference algorithm will only be stable if the material properties are averaged on the interface between two media. In Chapter 6, the theory of elastic surface waves is revisited. The waves that are excited at a surface are identified and their characteristics are explained. In Chapter 7, the interaction of elastic waves with buried land mines is described. Results are presented for various studies, considering buried anti-personnel mines as well as anti-tank mines. In Chapter 8, the resonant behavior of buried land mines is investigated. In Chapter 9, experimental procedures to measure the elastic wave fields in the ground are described. Results are presented, and the material properties in the ground are derived. Finally, in Chapter 10, the work is summarized and conclusions are drawn.

Four sections are appended to the text. In Appendix A, the complete set of the three-dimensional finite-difference equations is given, incorporating the split-field formulation that is necessary for the implementation of the absorbing boundary condition. In Appendix B, a listing of the correction terms required for the total-field/scattered-field formulation as described in Chapter 4 is given. In Appendix C, the computer program that has been developed to implement the numerical model is documented. And in Appendix D, the definitions of some fundamental elastic constants are listed.

CHAPTER 2

Background

The *finite-difference time-domain* (FDTD) method has long been applied to solve problems in the field of elasticity and electromagnetics. With advances in the computer technology, the FDTD method has gained importance especially since the end of the 1980's and the beginning of the 1990's. Numerous publications on the FDTD method and its application have been published throughout the years. A few of these are described and referenced in this chapter, briefly outlining the history of the method and the fundamental work that led up to the model as presented in this text.

Various algorithms are widely used for numerical modeling in both elasticity and electromagnetics. Some algorithms, such as the well-known and frequently employed *Method of Moments* (MoM), solve integral equations in the frequency-domain. Other algorithms, such as the finite-difference and the *finite-element* (FEM) method, operate on the partial-differential equations (PDE) and solve the equations either in the time-domain or in the frequency domain. Each of these algorithms has its own advantages and disadvantages, and some algorithms are more suitable for certain problems than others. The development of PDE algorithms was motivated by certain properties that these algorithms exhibit. PDE algorithms have been found to be robust and accurate. They yield either sparse-matrices (FEM) or no matrices at all (FDTD), reducing constraints on the size of the numerical models. These constraints limit the MoM algorithm, because the MoM, in general, requires operations on full matrices. Furthermore, PDE algorithms generally handle inhomgeneities without any significant

additional computational cost, whereas the inclusion of inhomogeneities in the MoM in general induces a drastic increase in computational cost.

For the particular modeling problem presented in this text, the FDTD method bears several advantages over other PDE algorithms, such as the finite-element method. The finite-difference formulation, on the one hand, is simple and straightforward. Boundary conditions can be implemented easily, and the numerical equations can be solved efficiently. The FDTD method can attack large problems involving large numbers of unknowns. Due to the structure of the finite-difference grid and the computation procedure being simple, the adaptation of a FDTD model to a parallel computer is straightforward [16, 17]. The parallelization is efficient, and an almost linear speed-up with the number of processors can be achieved. In the FEM, on the other hand, a matrix must be inverted. Current linear-algebra technology limits the size of the matrix that can be inverted and, thus, limits the number of unknowns that the FEM can handle. Furthermore, the parallelization of FEM models is not as straightforward as for the FDTD algorithm. Contrary to the first-order FDTD scheme, the FEM operates on the wave equation. A common problem of the FEM is that spurious, non-physical modes are predicted. These spurious modes must be excluded in a post-processing step, which can be a formidable task. Essentially no differences between the FDTD algorithm and the FEM arise when comparing their accuracy, as was pointed out by Marfurt [18]. Concluding, the FDTD method has been chosen for the numerical model that is described in this thesis. The deciding factor for this choice has been that a problem of the size considered here (involving up to 20,000,000 unknowns) either cannot be handled at all by the FEM or can be implemented only in a most tedious way, forfeiting the simplicity inherent to the FDTD algorithm.

The finite-difference time-domain algorithm has been used in the area of elasticity since the late 1960's, when Alterman and Karal introduced a second-order finite-difference formulation based on the elastic wave equation to compute the propagation of elastic waves in a layered medium in 1968 [19]. Following Alterman and Karal, the

second-order FDTD formulation has been advanced and applied to a wide variety of different problems in the field of elasticity [20] – [18]. In the late 1970's and the 1980's, various other elastic finite-difference schemes have been proposed. Emerman *et al.* presented an implicit finite-difference scheme based on the wave equation [24]. And Madariaga (1976) and Virieux (1984, 1986) introduced a first-order finite-difference formulation which used the two-dimensional first-order equations of elasticity [25] – [27]. Madariaga's and Virieux's first-order formulation was based on the early work of Yee, who had introduced a first-order scheme for electromagnetic finite-difference modeling in 1966 [28]. Yee was the first to introduce the characteristic staggered finite-difference grid, in which the field components, unlike in the second-order formulations, were not known at the same points in space and time. The first-order formulation by Madariaga and Virieux was second-order accurate in space and time. Based on Madariaga's and Virieux's first-order formulation, Levander developed a scheme that was fourth-order accurate in space by employing a four-point differencing approximation for the spatial derivatives in the governing equations, providing an improvement in accuracy on the expense of higher computational cost [29]. In today's finite-difference modeling, the first-order scheme by Madariaga and Virieux is the most common approach. The first-order formulation exhibits several advantages over the second-order formulation. Specifically, due to the staggering of the field components in the grid, the first-order scheme is more accurate, reducing grid dispersion effects. Furthermore, boundary conditions are easier implemented in the first-order formulation. In today's research and technology, the FDTD method is applied to a wide range of different problems in the field of elasticity. For example, the FDTD method is applied to model the elastic wave propagation in porous media [30, 31], to compute elastic waves in boreholes [32], or to simulate the ultrasonic pulse propagation in human tissue [33].

The numerical model described in this text computes the elastic wave fields in a linear, isotropic, lossless, heterogeneous half-space. The model simplifies the elastic behavior of the ground. In reality, the ground might, dependent on the condition and

composition of the ground and the amplitude of the excitation, act in a strongly nonlinear way, and loss will play a role in the wave propagation. Furthermore, the ground only approximately behaves as an elastic half-space. Sand, for example, is a porous medium, and the existence of pores, formed by the grains of the sand, might alter the wave motion significantly. Various schemes have been proposed to include these and other effects into a numerical finite-difference model. Krebes and Quiroga-Goode (1994) [34] and Robertson *et al.* (1994) [35] described the modeling of viscoelastic materials. The inclusion of loss effects in the FDTD formulation, however, involves the solution of a convolution integral, thus complicating the formulation. Faria and Stoffa (1994) [36] developed a finite-difference formulation for transversely isotropic media, which represent a certain class of anisotropic media. The finite-difference modeling of porous media has been treated by Dai *et al.* (1995) [30] and by Zeng and Liu (2001) [31]. Poroelastic modeling is based on the poroelastic equations developed by Biot (1956) [37]. In poroelastic materials, in addition to the shear and the pressure phase, a third phase – a slow pressure wave – might arise, which is due to the relative motion of the solid frame and the fluid filling the pores. It has been found that, for this work's purpose, the assumptions that have been made (i. e., linear, isotropic, lossless half-space) are reasonable. The simplicity in the formulation that is achieved far outweighs the gain in accuracy that would be obtained by incorporating some or all of the effects mentioned above.

In a numerical finite-difference model, boundary conditions must be satisfied explicitly. The model as presented here incorporates a free-surface boundary condition to approximate the surface of the ground and a *Perfectly-Matched Layer* absorbing boundary condition to reduce artificial reflections at the outer grid faces. The treatment of boundaries has been described extensively in the literature. The free-surface boundary condition is well-known and is commonly used to approximate the surface of the ground [19, 21, 22, 29, 38]. Absorbing boundary conditions have been a topic of vast research for a long time. In the 1980's, one commonly applied absorbing boundary condition in elastic finite-difference modeling was a boundary condition presented

by Clayton and Engquist (1977) for the second-order formulation [39]. The Clayton-Engquist absorbing boundary condition was based on a paraxial approximation of the wave equation, allowing to separate the wave fields into inward and outward traveling wave components. In this way, the outward traveling waves could be absorbed, and a boundary arose which appeared transparent to outward traveling waves. One drawback of the Clayton-Engquist condition was that, due to the validity of the paraxial approximations, it worked well only over a limited range of incident angles. Various authors subsequently proposed improvements to the Clayton-Engquist condition [40, 41] or presented modified or different schemes [42] – [48], all with somewhat similar limitations as the Clayton-Engquist condition. In 1994, Berenger presented an entirely new concept of absorbing boundary conditions. Berenger developed the so-called *Perfectly-Matched Layer* (PML) absorbing boundary condition for finite-difference modeling in electromagnetics [49]. This novel boundary condition used a non-physical splitting of the wave fields to introduce a lossy boundary layer which was perfectly matched to the solution space. The PML reduced the reflections at the outer grid faces by as much as 100 dB and its performance was roughly independent from the angle of incidence. Following Berenger, the PML formulation was improved by various authors [50] – [53] and adapted to elastic finite-difference modeling [54] – [56]. Due to its superior performance, the PML absorbing boundary is now the commonly used absorbing boundary condition in both elastic and electromagnetic finite-difference modeling. For the numerical model presented here, the PML formulation of Chew and Liu has been applied [55].

CHAPTER 3

Elastic Wave Propagation in Solids

3.1 Governing Equations

The equation of motion, better known as Newton's law, and the strain-displacement relation, combined by an elastic constitutive relation, form a fundamental set of equations which completely describes the elastic wave motion in a linear medium. This fundamental set of equations is comparable to Maxwell's equations in electromagnetics. The field quantities describing an elastic field are the vector of the particle displacement, $\mathbf{u}$, and the tensors of the mechanical stress, $\boldsymbol{\tau}$, and the mechanical strain, $\mathbf{S}$.[1] In electromagnetics, the elastic field quantities find their analogues in the vectors of the electric field $\mathbf{E}$, the electrical displacement $\mathbf{D}$, the magnetic field $\mathbf{B}$ and the magnetic excitation $\mathbf{H}$. The fundamental governing equations are well-known and shall be explained only briefly in this text. A detailed description of the theory of elasticity can be found, for example, in [58] and [57].

In a lossless medium, the equations that govern the wave motion are the *equation of motion*

$$\nabla \cdot \boldsymbol{\tau} = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} - \mathbf{F} \tag{3.1}$$

and the *strain-displacement relation*

$$\mathbf{S} = \frac{1}{2}(\nabla(\mathbf{u}^{\mathrm{T}}) + (\nabla(\mathbf{u}^{\mathrm{T}}))^{\mathrm{T}}), \tag{3.2}$$

[1] The terminology used is by no means standard. All symbols are chosen according to [57], except for the stress tensor, which is marked with $\boldsymbol{\tau}$ throughout this text, but is denoted with $\mathbf{T}$ in [57].

combined by the *elastic constitutive relation*

$$\boldsymbol{\tau} = \mathbf{c} \cdot \mathbf{S}, \tag{3.3}$$

also known as *Hooke's Law* [57]. Here, ρ is the density of the medium and is in general a function of position. The vector $\mathbf{F}$ describes the inner body forces. The tensor $\mathbf{c}$ is called the *stiffness matrix*. It is a $3 \times 3 \times 3 \times 3$ matrix and describes the medium and its characteristics. ∇ is the Nabla operator,

$$\nabla = \frac{\partial}{\partial x} \cdot \hat{\mathbf{x}} + \frac{\partial}{\partial y} \cdot \hat{\mathbf{y}} + \frac{\partial}{\partial z} \cdot \hat{\mathbf{z}}. \tag{3.4}$$

In cartesian coordinates, the displacement vector is defined as

$$\mathbf{u} = u_x \cdot \hat{\mathbf{x}} + u_y \cdot \hat{\mathbf{y}} + u_z \cdot \hat{\mathbf{z}}, \tag{3.5}$$

and the tensors are in matrix notation

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \tau_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \tau_{zz} \end{bmatrix}, \tag{3.6}$$

$$\mathbf{S} = \begin{bmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{bmatrix}. \tag{3.7}$$

In Eq. (3.2), ∇ is assumed to be a column vector and $\mathbf{u}^{\mathrm{T}}$ is a row vector. In that case, the vector product $\nabla(\mathbf{u}^{\mathrm{T}})$ represents a tensor:

$$\nabla(\mathbf{u}^{\mathrm{T}}) = \begin{bmatrix} \frac{\partial u_x}{\partial x} & \frac{\partial u_y}{\partial x} & \frac{\partial u_z}{\partial x} \\ \frac{\partial u_x}{\partial y} & \frac{\partial u_y}{\partial y} & \frac{\partial u_z}{\partial y} \\ \frac{\partial u_x}{\partial z} & \frac{\partial u_y}{\partial z} & \frac{\partial u_z}{\partial z} \end{bmatrix}. \tag{3.8}$$

The superscript T in Eq. (3.2) denotes the transpose.

The notation of the tensor elements is such that the first subscript of each element describes the plane in which the component is effective and the second subscript

marks its direction. Thus, τ_{xy} is the stress in the y-direction acting in the y-z-plane (described by its normal vector in the x-direction). Figure 3.1 shows an infinitesimally small cube in a cartesian coordinate system with the stresses that act on its surfaces. Within a medium, this cube must be in both rotational and lateral equilib-

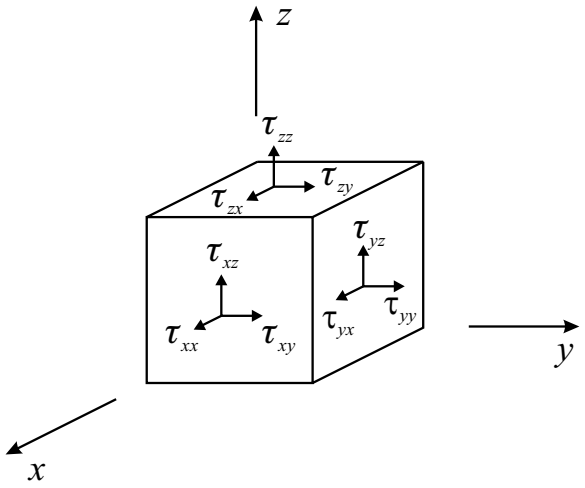


Figure 3.1: Infinitesimally small cube in a cartesian coordinate system with the stresses acting onto its surfaces.

rium. The lateral equilibrium implies that the stresses acting on the cube's faces are point-symmetric to the cube center. Hence, the stresses on the sides of the cube that are not visible are equivalent to the stresses on their opposing sides, but rotated by 180 degrees. From the rotational equilibrium condition it follows that the sum of the moments acting onto the cube must be zero. Then, the stress tensor $\boldsymbol{\tau}$ must be symmetric, i. e., $\tau_{yx} = \tau_{xy}$, $\tau_{zx} = \tau_{xz}$ etc. Consequently, there are only six independent stress components in the stress tensor. From Eq. (3.2) it can be seen, that the strain tensor $\mathbf{S}$ is always symmetric, because it is just the symmetric part of the matrix $\nabla(\mathbf{u}^{\mathrm{T}})$. Thus, the strain tensor is comprised of only six independent components. The (i, j)-th element of the strain tensor can also be written as

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{3.9}$$

3.2 Elastic Waves in Isotropic Media

3.2.1 The Wave Equation in Isotropic Media

In isotropic media, the entries of the stiffness matrix $\mathbf{c}$ can be completely described in terms of two independent constants. These constants are called *Lame's constants* and are denoted by λ and μ. Using Lame's constants, the constitutive relation for the six unknown stress components reduces to

$$\tau_{xx} = (\lambda + 2\mu)S_{xx} + \lambda S_{yy} + \lambda S_{zz} \tag{3.10}$$

$$\tau_{yy} = \lambda S_{xx} + (\lambda + 2\mu)S_{yy} + \lambda S_{zz} \tag{3.11}$$

$$\tau_{zz} = \lambda S_{xx} + \lambda S_{yy} + (\lambda + 2\mu)S_{zz} \tag{3.12}$$

$$\tau_{yz} = 2\mu S_{yz} \tag{3.13}$$

$$\tau_{xz} = 2\mu S_{xz} \tag{3.14}$$

$$\tau_{xy} = 2\mu S_{xy}. \tag{3.15}$$

Using equation (3.2), the strain can be expressed in terms of the displacement:

$$\tau_{xx} = (\lambda + 2\mu)\frac{\partial u_x}{\partial x} + \lambda\frac{\partial u_y}{\partial y} + \lambda\frac{\partial u_z}{\partial z} \tag{3.16}$$

$$\tau_{yy} = \lambda\frac{\partial u_x}{\partial x} + (\lambda + 2\mu)\frac{\partial u_y}{\partial y} + \lambda\frac{\partial u_z}{\partial z} \tag{3.17}$$

$$\tau_{zz} = \lambda\frac{\partial u_x}{\partial x} + \lambda\frac{\partial u_y}{\partial y} + (\lambda + 2\mu)\frac{\partial u_z}{\partial z} \tag{3.18}$$

$$\tau_{yz} = \mu(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}) \tag{3.19}$$

$$\tau_{xz} = \mu(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}) \tag{3.20}$$

$$\tau_{xy} = \mu(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}). \tag{3.21}$$

Finally, combining Eq. (3.1) with Eqs. (3.16)–(3.21) and assuming that no body forces are present ($\mathbf{F} = 0$), the elastic wave equation for the three displacement components in x, y and z is obtained:

$$\mu\nabla^2 u_x + (\lambda+\mu)(\frac{\partial^2 u_x}{\partial x^2} + \frac{\partial^2 u_y}{\partial x \partial y} + \frac{\partial^2 u_z}{\partial x \partial z}) = \rho\frac{\partial^2 u_x}{\partial t^2} \tag{3.22}$$

$$\mu\nabla^2 u_y + (\lambda+\mu)(\frac{\partial^2 u_y}{\partial y^2} + \frac{\partial^2 u_z}{\partial y \partial z} + \frac{\partial^2 u_x}{\partial y \partial x}) = \rho\frac{\partial^2 u_y}{\partial t^2} \tag{3.23}$$

$$\mu\nabla^2 u_z + (\lambda+\mu)(\frac{\partial^2 u_z}{\partial z^2} + \frac{\partial^2 u_x}{\partial z \partial x} + \frac{\partial^2 u_y}{\partial z \partial y}) = \rho\frac{\partial^2 u_z}{\partial t^2}, \tag{3.24}$$

where ∇^2 is the *Laplace operator*: $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$.

Regrouping Eqs. (3.22)–(3.24) and using vector notation, the vector wave equation is obtained:

$$(\lambda+2\mu)\nabla\cdot\nabla\cdot\mathbf{u} - \mu\nabla\times\nabla\times\mathbf{u} = \rho\frac{\partial^2\mathbf{u}}{\partial t^2}. \tag{3.25}$$

3.2.2 Propagating Waves

Generally, the elastic wave equation gives rise to two different kinds of waves propagating in an elastic solid: a longitudinal wave referred to as pressure wave and a transverse wave called shear wave. Both waves are propagating independently of each other with different phase velocities. The wave number for the pressure wave can be written as

$$k_P^2 = \frac{\omega^2\rho}{(\lambda+2\mu)} \tag{3.26}$$

with the corresponding phase velocity $c_P = \sqrt{\frac{\lambda+2\mu}{\rho}}$. The wave number for the shear wave, similarly, can be expressed as

$$k_S^2 = \frac{\omega^2\rho}{\mu} \tag{3.27}$$

with its phase velocity $c_S = \sqrt{\frac{\mu}{\rho}}$. Since λ and μ are both always positive, the phase velocity of the longitudinal wave is always larger than the phase velocity of the transverse wave.

When a boundary or discontinuity is introduced, the existence of both longitudinal and transverse waves in an elastic medium gives rise to additional waves. The most important of these is the *Rayleigh surface wave*[2] and appears at a free surface boundary.

At a free surface (i. e., at the interface between an elastic medium and vacuum), the normal stress components vanish due to the continuity of the normal stress at an interface. Both the longitudinal and the transverse wave traveling along a free surface cannot fulfill this boundary condition by themselves. However, a combination of the two can, giving rise to the Rayleigh surface wave.

Rayleigh surface waves have characteristic properties that distinguish them from longitudinal and transverse waves. Rayleigh surface waves travel along the surface and decay exponentially into the medium. The reach of the surface wave into the medium is dependent on frequency. The higher the frequency, the faster the surface wave decays into the medium. Due to their origin, surface waves are neither purely longitudinal nor purely transverse. The particles that are subjected to a surface wave instead undergo an elliptic motion, due to the superposition of longitudinal and transverse field components. A surface wave travels with a phase velocity slightly smaller than the phase velocity of the transverse wave and, thus, is slower than both the longitudinal and the transverse wave. The wave speed of the surface wave cannot be determined explicitly, but is defined by a transcendental equation. Since the surface wave is traveling along a plane (i. e., the surface of a medium), it has a somewhat two-dimensional character. While longitudinal and transverse waves due to a point source in a three-dimensional medium decay proportionally to the inverse of the distance from the source($\propto r^{-1}$), a surface wave decays with the inverse of

[2] After Lord Rayleigh, who discovered the existence of elastic surface waves in the late nineteenth century.

the square root of the distance ($\propto r^{-1/2}$). Thus, its energy is in some way confined to the surface. Moreover, if a two-dimensional case is considered (e. g., an infinite line source in a three dimensional medium), the surface wave has a one-dimensional character and does not decay at all with distance from the source. This property of confinement of energy to a surface is most successfully used in *Surface Acoustic Wave* (SAW) devices.

Figure 3.2 shows the elastic waves in the ground due to a point-like excitation on the ground's surface.[3] The particle displacement on a cross section through the ground is plotted in a pseudo color plot. A logarithmic scale is used, where dark red corresponds to the largest magnitude (0 dB) and blue to the smallest (-70 dB). The top edge of the plot coincides with the free surface. The excitation is a differentiated Gaussian pulse. The pressure wave and the shear wave are visible. Both describe spherical wave fronts. The Rayleigh surface wave propagates along the surface. At the surface, plane waves arise. One is a *lateral wave* and is due to the passage of the pressure wave along the surface , inducing a plane shear wave. The other one is a *leaky surface wave*. The origin of this wave will be discussed in detail in a later chapter.

[3]The wave fields are computed with the numerical model which is described in the subsequent chapter.

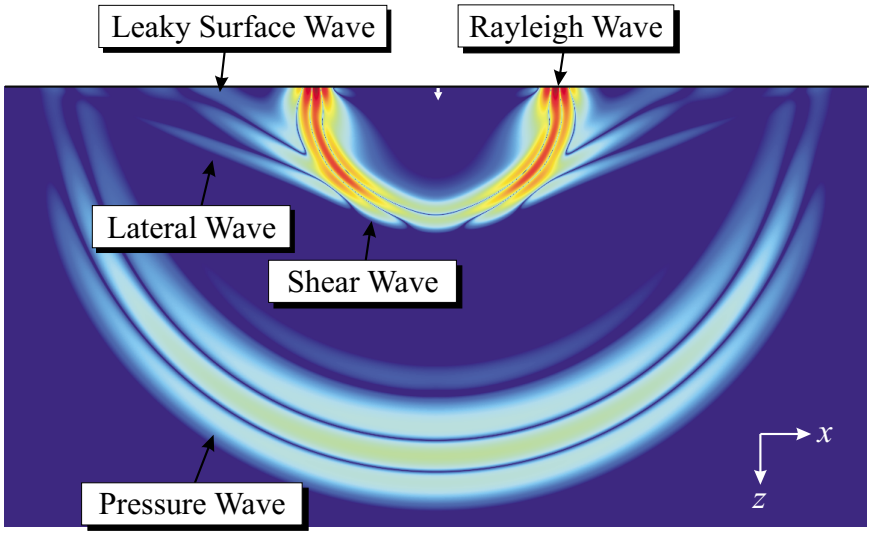


Figure 3.2: Elastic waves due to a point source on a free surface.

CHAPTER 4

Numerical Model

During the development of the elastic/electromagntic sensor, experiments have been performed with mines buried in a large sand tank [12]. In these experiments, elastic waves are excited by an electrodynamic transducer placed on the surface of the ground. The waves propagate in the ground and along its surface and interact with the buried land mines. To investigate the interactions of the elastic waves with the buried mines, a three-dimensional finite-difference model has been developed. In this chapter, the finite-difference model and its implementation are described.

4.1 The Finite-Difference Scheme

4.1.1 Problem Statement

Fig. 4.1 outlines the problem geometry. A transducer (labeled *Source*) is placed on the surface of the ground. The transducer launches elastic waves, which interact with an *Object* that is buried in the ground at some distance from the transducer. When adapting this problem to a numerical model, assumptions must be made to simplify the model and to make its implementation feasible. For the numerical model, the ground is assumed to be a semi-infinite half-space, bounded only by the surface at $z = 0$. Thus, the tank with its walls as used in the experiments is not modeled and its effects are neglected. The ground is approximated to be linear, isotropic, and lossless.

Within a numerical model, the elastic wave fields are computed at a number of

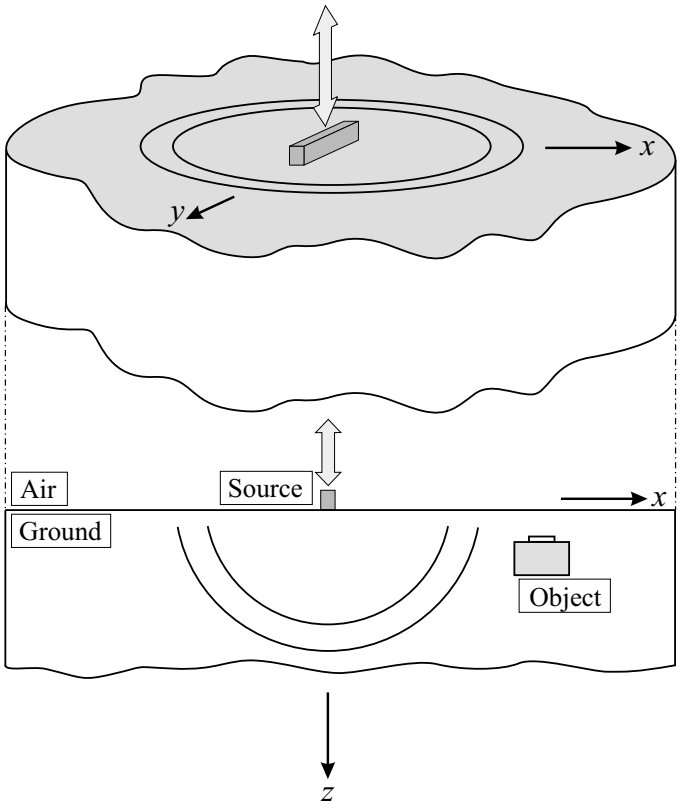


Figure 4.1: Problem geometry; top: isometric view, bottom: cross sectional view.

discrete points in space. For the finite-difference modeling scheme, a discrete grid of regular cubic shape is introduced. Within the numerical model, the infinite half-space must be truncated and appropriate boundary conditions must be implemented to avoid artificial reflections at the outer grid faces. Figure 4.2 shows a schematic drawing of the three-dimensional finite-difference model. The solution space is bounded on top by a free-surface boundary and on all other sides by a *Perfectly-Matched Layer* (PML) absorbing boundary. The source is located on the free surface.

4.1.2 Governing Equations

As described in Chapter 3, elastic wave propagation in a linear medium is governed by three fundamental partial differential equations: the equation of motion, the strain-displacement relation, and the elastic constitutive relation. By eliminating the strain, a system of partial differential equations is obtained that describes the wave fields in terms of solely the displacement and the stress:

$$\rho\frac{\partial^2 u_x}{\partial t^2} = \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} \tag{4.1}$$

$$\rho\frac{\partial^2 u_y}{\partial t^2} = \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} \tag{4.2}$$

$$\rho\frac{\partial^2 u_z}{\partial t^2} = \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} \tag{4.3}$$

$$\tau_{xx} = (\lambda + 2\mu)\frac{\partial u_x}{\partial x} + \lambda\frac{\partial u_y}{\partial y} + \lambda\frac{\partial u_z}{\partial z} \tag{4.4}$$

$$\tau_{yy} = \lambda\frac{\partial u_x}{\partial x} + (\lambda + 2\mu)\frac{\partial u_y}{\partial y} + \lambda\frac{\partial u_z}{\partial z} \tag{4.5}$$

$$\tau_{zz} = \lambda\frac{\partial u_x}{\partial x} + \lambda\frac{\partial u_y}{\partial y} + (\lambda + 2\mu)\frac{\partial u_z}{\partial z} \tag{4.6}$$

$$\tau_{yz} = \mu(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}) \tag{4.7}$$

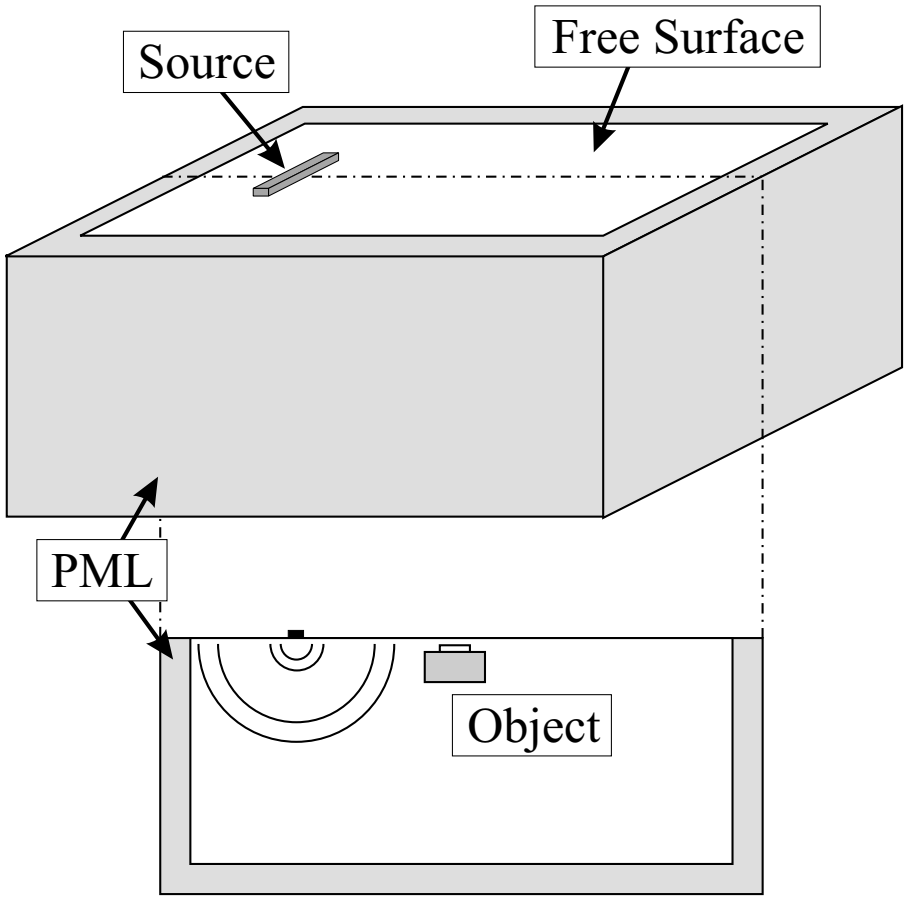


Figure 4.2: Three-dimensional finite-difference model.

$$\tau_{xz} = \mu(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}) \tag{4.8}$$

$$\tau_{xy} = \mu(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}). \tag{4.9}$$

Two approaches can be chosen to adapt the governing equations to the finite-difference scheme. First, the wave equations for the displacement field components, as described in Eqs. (3.22)–(3.24), can be discretized [19]. This results in only three equations with three unknowns. However, due to the second order character of the wave equations, the finite-difference equations are complicated. A different approach is to discretize Eqs. (4.1)–(4.9) [26, 27]. The resulting system of equations includes nine equations with nine unknowns. The equations contain only first-order derivatives, if the particle velocity is introduced for the particle displacement. These equations are simple and, hence, can be implemented in a straightforward way. With the first-order approach, the implementation of boundary conditions, for example, is considerably easier than with the second-order approach. Because of its simplicity, the first-order approach is now the most common approach for finite-difference modeling in elastodynamics.

To eliminate the second-order time-derivatives from Eqs. (4.1)–(4.9), the particle velocity is introduced for the displacement. The particle velocity is the time derivative of the displacement:

$$\mathbf{v} = \frac{\partial \mathbf{u}}{\partial t}. \tag{4.10}$$

Taking the time derivative of Eq. (4.4) – (4.9) and replacing $\frac{\partial u_x}{\partial t}$ with v_x etc., one obtains a system of equations with only derivatives of first order:

$$\rho\frac{\partial v_x}{\partial t} = \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z} \tag{4.11}$$

$$\rho\frac{\partial v_y}{\partial t} = \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z} \tag{4.12}$$

$$\rho\frac{\partial v_z}{\partial t} = \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{xz}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} \tag{4.13}$$

$$\frac{\partial \tau_{xx}}{\partial t} = (\lambda + 2\mu)\frac{\partial v_x}{\partial x} + \lambda\frac{\partial v_y}{\partial y} + \lambda\frac{\partial v_z}{\partial z} \tag{4.14}$$

$$\frac{\partial \tau_{yy}}{\partial t} = \lambda\frac{\partial v_x}{\partial x} + (\lambda + 2\mu)\frac{\partial v_y}{\partial y} + \lambda\frac{\partial v_z}{\partial z} \tag{4.15}$$

$$\frac{\partial \tau_{zz}}{\partial t} = \lambda\frac{\partial v_x}{\partial x} + \lambda\frac{\partial v_y}{\partial y} + (\lambda + 2\mu)\frac{\partial v_z}{\partial z} \tag{4.16}$$

$$\frac{\partial \tau_{yz}}{\partial t} = \mu(\frac{\partial v_y}{\partial z} + \frac{\partial v_z}{\partial y}) \tag{4.17}$$

$$\frac{\partial \tau_{xz}}{\partial t} = \mu(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x}) \tag{4.18}$$

$$\frac{\partial \tau_{xy}}{\partial t} = \mu(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}). \tag{4.19}$$

In these equations, the elastic wave fields are described in terms of nine unknowns: the three particle displacement components v_x, v_y, and v_z, and the six stress tensor components τ_{xx}, τ_{yy}, τ_{zz}, τ_{yz}, τ_{xz}, and τ_{xy}.

4.1.3 Discretization of the Governing Equations

In the finite-difference scheme, the partial differential equations are commonly discretized using centered finite differences. Introducing the finite difference Δx, the derivative of a function $f = f(x)$ with respect to x at the point x_0 can be approximated by

$$\left.\frac{\partial f}{\partial x}\right|_{x_0} = \frac{f(x_0 + \frac{\Delta x}{2}) - f(x_0 - \frac{\Delta x}{2})}{\Delta x}. \tag{4.20}$$

Similarly, the time derivative of a function $f = f(t)$ at t_0 can be expressed introducing the finite difference Δt:

$$\left.\frac{\partial f}{\partial t}\right|_{t_0} = \frac{f(t_0 + \frac{\Delta t}{2}) - f(t_0 - \frac{\Delta t}{2})}{\Delta t}. \tag{4.21}$$

When applying the finite-difference discretization to the governing partial-differential equations, a discrete grid of regular cubic shape arises. In this grid each

field component is surrounded by the field components it is dependent on [27]. One important characteristic of the finite-difference grid is that the displacement and stress components are not known at the same points in space and time. The resulting finite-difference expressions are centered around the field components to be determined. The structure of the finite-difference grid is very similar to the one used for electromagnetic finite-difference modeling, where the so-called *Yee-cell* is introduced [17].

The finite-difference grid can be viewed as being composed of basis cells. Each basis cell is characterized by its dimensions in x, y, and z, i. e., the finite differences Δx, Δy, and Δz, and the position of the cell center in the grid. By introducing the indices i, j, and k for the position of the cell in x, y, and z, respectively, the position of the center is labeled and in that way uniquely identified within the grid. Fig. 4.3 shows the (i, j, k)-th basis cell of the three-dimensional finite-difference grid with its elastic field components. The numerical (discrete) velocity components are indicated by V and the numerical stress tensor components by T. In Fig. 4.4, a larger portion of the finite-difference grid is depicted.

The position of the cell center in real space is given by $x = i \cdot \Delta x$, $y = j \cdot \Delta y$, and $z = k \cdot \Delta z$. However, no field component is located at the cell center. The positions of the field components of the (i, j, k)-th cell in real space are given in Table 4.1. To uniquely identify the field components in the finite-difference grid, only the field components on the corners of the shaded region in Fig. 4.3 are assigned to the (i, j, k)-th basis cell.

The discrete time is labeled with the index l. Due to the centered finite-difference formulation, half indices must be introduced for the velocity equations. Assuming the incremental time of the finite-difference algorithm to be Δt, the values for the stress tensor components are arbitrarily set to be known at the full time step, $t = l \cdot \Delta t$, and the particle velocity components are calculated at the half time step, $t = (l + 0.5) \cdot \Delta t$.

Introducing finite differences in space and time, the partial-differential equations are discretized and discrete difference equations approximating the differential

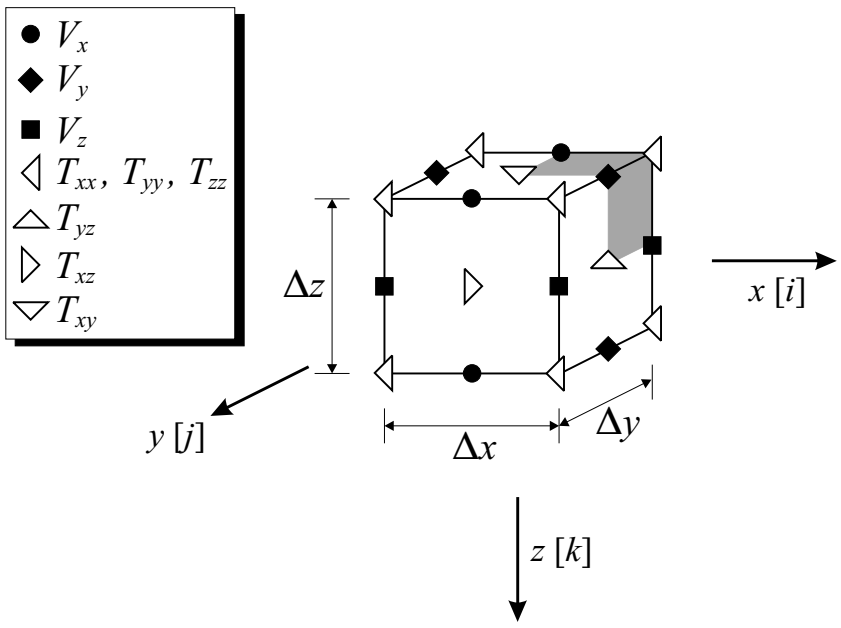


Figure 4.3: Three-dimensional finite-difference basis cell.

equations are obtained. For example, applying the discretization to Eq. (4.11) yields

$$\rho\frac{V_x^{l+0.5}|_{i,j,k} - V_x^{l-0.5}|_{i,j,k}}{\Delta t} = \frac{T_{xx}^{l}|_{i,j,k} - T_{xx}^{l}|_{i-1,j,k}}{\Delta x} +$$

$$+ \frac{T_{xy}^{l}|_{i,j,k} - T_{xx}^{l}|_{i,j-1,k}}{\Delta y} + \frac{T_{xz}^{l}|_{i,j,k} - T_{xz}^{l}|_{i,j,k-1}}{\Delta z}. \qquad (4.22)$$

Here, the capital letters mark the numerical values of the field components at their discrete locations in space and time. The notation is such that $V_x^{l+0.5}|_{i,j,k}$, for example, stands for the numerical value of the particle velocity v_x at the (i, j, k)-th cell at time $t = (l + 0.5)\Delta t$.

By rearranging the difference equation, an equation for V_x at the incremented time $t = (l+0.5)\Delta t$ can be obtained entirely in terms of field components at previous times. Thus, if the field values at and prior to $t = l\Delta t$ are known, V_x at the incremented time $t = (l+0.5)\Delta t$ can be computed. In the same way, each field component can be expressed in terms of the previous field values. The resulting equations are

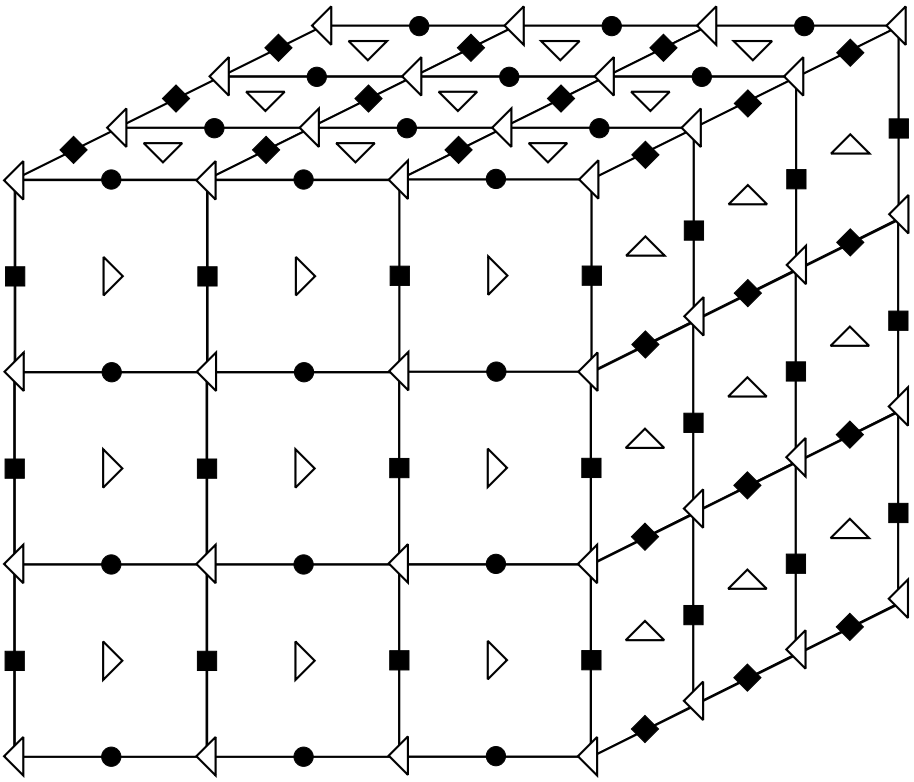


Figure 4.4: A portion of the three-dimensional finite-difference grid.

Table 4.1: Positions of the field components of the (i, j, k)-th cell in the finite-difference grid.

	x	y	z
V_x	$i\Delta x$	$(j-0.5)\Delta y$	$(k-0.5)\Delta z$
V_y	$(i+0.5)\Delta x$	$j\Delta y$	$(k-0.5)\Delta z$
V_z	$(i+0.5)\Delta x$	$(j-0.5)\Delta y$	$k\Delta z$
T_{xx}	$(i+0.5)\Delta x$	$(j-0.5)\Delta y$	$(k-0.5)\Delta z$
T_{yy}	$(i+0.5)\Delta x$	$(j-0.5)\Delta y$	$(k-0.5)\Delta z$
T_{zz}	$(i+0.5)\Delta x$	$(j-0.5)\Delta y$	$(k-0.5)\Delta z$
T_{yz}	$(i+0.5)\Delta x$	$j\Delta y$	$k\Delta z$
T_{xz}	$i\Delta x$	$(j-0.5)\Delta y$	$k\Delta z$
T_{xy}	$i\Delta x$	$j\Delta y$	$(k-0.5)\Delta z$

commonly called the finite-difference *update equations*. The complete set of update equations for the nine field components of the (i, j, k)-th grid cell is given by

$$V_x^{l+0.5}|_{i,j,k} = V_x^{l-0.5}|_{i,j,k} + \frac{\Delta t}{\rho} \cdot \left[\frac{T_{xx}^l|_{i,j,k} - T_{xx}^l|_{i-1,j,k}}{\Delta x} + \right. \tag{4.23}$$

$$\left. + \frac{T_{xy}^l|_{i,j,k} - T_{xy}^l|_{i,j-1,k}}{\Delta y} + \frac{T_{xz}^l|_{i,j,k} - T_{xz}^l|_{i,j,k-1}}{\Delta z} \right]$$

$$V_y^{l+0.5}|_{i,j,k} = V_y^{l-0.5}|_{i,j,k} + \frac{\Delta t}{\rho} \cdot \left[\frac{T_{xy}^l|_{i+1,j,k} - T_{xy}^l|_{i,j,k}}{\Delta x} + \right. \tag{4.24}$$

$$\left. + \frac{T_{yy}^l|_{i,j+1,k} - T_{yy}^l|_{i,j,k}}{\Delta y} + \frac{T_{yz}^l|_{i,j,k} - T_{yz}^l|_{i,j,k-1}}{\Delta z} \right]$$

$$V_z^{l+0.5}|_{i,j,k} = V_z^{l-0.5}|_{i,j,k} + \frac{\Delta t}{\rho} \cdot \left[\frac{T_{xz}^l|_{i+1,j,k} - T_{xz}^l|_{i,j,k}}{\Delta x} + \right. \tag{4.25}$$

$$+ \frac{T^l_{yz}|_{i,j,k} - T^l_{yz}|_{i,j-1,k}}{\Delta y} + \frac{T^l_{zz}|_{i,j,k+1} - T^l_{zz}|_{i,j,k}}{\Delta z}\Bigg]$$

$$T^{l+1}_{xx}|_{i,j,k} = T^l_{xx}|_{i,j,k} + \Delta t \cdot \Bigg[(\lambda + 2\mu)\frac{V^{l+0.5}_x|_{i+1,j,k} - V^{l+0.5}_x|_{i,j,k}}{\Delta x} + \tag{4.26}$$

$$+ \lambda\frac{V^{l+0.5}_y|_{i,j,k} - V^{l+0.5}_y|_{i,j-1,k}}{\Delta y} + \lambda\frac{V^{l+0.5}_z|_{i,j,k} - V^{l+0.5}_z|_{i,j,k-1}}{\Delta z}\Bigg]$$

$$T^{l+1}_{yy}|_{i,j,k} = T^l_{yy}|_{i,j,k} + \Delta t \cdot \Bigg[\lambda\frac{V^{l+0.5}_x|_{i+1,j,k} - V^{l+0.5}_x|_{i,j,k}}{\Delta x} + \tag{4.27}$$

$$+ (\lambda + 2\mu)\frac{V^{l+0.5}_y|_{i,j,k} - V^{l+0.5}_y|_{i,j-1,k}}{\Delta y} + \lambda\frac{V^{l+0.5}_z|_{i,j,k} - V^{l+0.5}_z|_{i,j,k-1}}{\Delta z}\Bigg]$$

$$T^{l+1}_{zz}|_{i,j,k} = T^l_{zz}|_{i,j,k} + \Delta t \cdot \Bigg[\lambda\frac{V^{l+0.5}_x|_{i+1,j,k} - V^{l+0.5}_x|_{i,j,k}}{\Delta x} + \tag{4.28}$$

$$+ \lambda\frac{V^{l+0.5}_y|_{i,j,k} - V^{l+0.5}_y|_{i,j-1,k}}{\Delta y} + (\lambda + 2\mu)\frac{V^{l+0.5}_z|_{i,j,k} - V^{l+0.5}_z|_{i,j,k-1}}{\Delta z}\Bigg]$$

$$T^{l+1}_{yz}|_{i,j,k} = T^l_{yz}|_{i,j,k} + \Delta t \cdot \mu\Bigg[\frac{V^{l+0.5}_z|_{i,j+1,k} - V^{l+0.5}_z|_{i,j,k}}{\Delta y} + \tag{4.29}$$

$$+ \frac{V^{l+0.5}_y|_{i,j,k+1} - V^{l+0.5}_y|_{i,j,k}}{\Delta z}\Bigg]$$

$$T^{l+1}_{xz}|_{i,j,k} = T^l_{xz}|_{i,j,k} + \Delta t \cdot \mu\Bigg[\frac{V^{l+0.5}_z|_{i,j,k} - V^{l+0.5}_z|_{i-1,j,k}}{\Delta x} + \tag{4.30}$$

$$+ \frac{V^{l+0.5}_x|_{i,j,k+1} - V^{l+0.5}_x|_{i,j,k}}{\Delta z}\Bigg]$$

$$T^{l+1}_{xy}|_{i,j,k} = T^l_{xy}|_{i,j,k} + \Delta t \cdot \mu\Bigg[\frac{V^{l+0.5}_y|_{i,j,k} - V^{l+0.5}_y|_{i-1,j,k}}{\Delta x} + \tag{4.31}$$

$$+ \frac{V_x^{l+0.5}|_{i,j+1,k} - V_x^{l+0.5}|_{i,j,k}}{\Delta y}\Bigg].$$

Note that the velocity components are computed at the half time steps, whereas the stress components are determined at the full time steps.

The finite-difference scheme, then, works as follows. Initially, i. e., at $l = 0$, the values of all field components are set to zero. At one point or in an entire region (i. e., in the *source region*) an elastic field is excited. The velocity components are updated throughout the grid for $l = 0.5$, and using the velocities at $l = 0.5$ the stress components are calculated for $l = 1$. Then the velocities are computed for $l = 1.5$ from the stresses at $l = 1$ and so on. Following this procedure, sequentially the velocity and stress fields can be calculated up to any desired time. Due to the gradual progressing in time by "hopping" from the stresses to the velocities, the algorithm is commonly referred to as *leapfrog* scheme [17].

General constraints on the size of the space increments Δx, Δy, and Δz and the time increment Δt arise. To compute the elastic wave fields with a reasonable accuracy, the space step must not exceed one tenth of the minimum wave length excited within the model [17]:

$$\Delta x, \Delta y, \Delta z < \lambda_{\min}/10. \tag{4.32}$$

By obeying this condition, artificial numerical dispersion effects are minimized. The time increment is linked to the space increment by the general condition for stability, the *Courant condition*:

$$c_{\max}\Delta t\sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}} \leq 1, \tag{4.33}$$

where $c_{\max}$ is the maximum wave speed occurring in the numerical model. If $\Delta x = \Delta y = \Delta z$, this reduces to

$$c_{\max}\frac{\Delta t}{\Delta x} \leq \frac{1}{\sqrt{3}}. \tag{4.34}$$

In plain words, the Courant condition states that the physical wave speed of any wave excited in the model must not exceed the speed information can travel with in the numerical grid.

4.2 Boundaries

When implementing the finite-difference scheme, boundary conditions have to be treated in a special manner. Three different kinds of boundaries arise: points or regions where the fields are excited (i. e., the *source*), internal boundaries (i. e., boundaries within the medium caused by a change in material properties), and external boundaries (i. e., the outer grid faces).

4.2.1 Source

Elastic wave fields can be excited at any point in space. For most of the results that are presented in this thesis, a source is used that emulates the transducer used in the experiments. For this, the vertical particle velocity component, v_z, is excited on the surface throughout a region of approximately the same shape and size as the transducer foot. A differentiated Gaussian pulse is used as excitation, which closely resembles the transducer foot motion. For some results, instead of modeling the transducer, a plane wave is injected. This requires a special formulation, which is described in Sec. 4.3.

4.2.2 Internal Boundaries

Internal boundaries that occur at the interface between two different media, are usually satisfied implicitly within the finite-difference formulation. Figure 4.5 shows a cross section of the finite-difference grid at an interface between two media. The interface between two media is represented naturally by a change in the material properties. To minimize the number of field components that are located on the interface, the boundary is placed such that it always passes through the cell center. In that case, the normal particle velocity components will always be placed on the interface.

It has been found that at a material interface the Courant condition might not

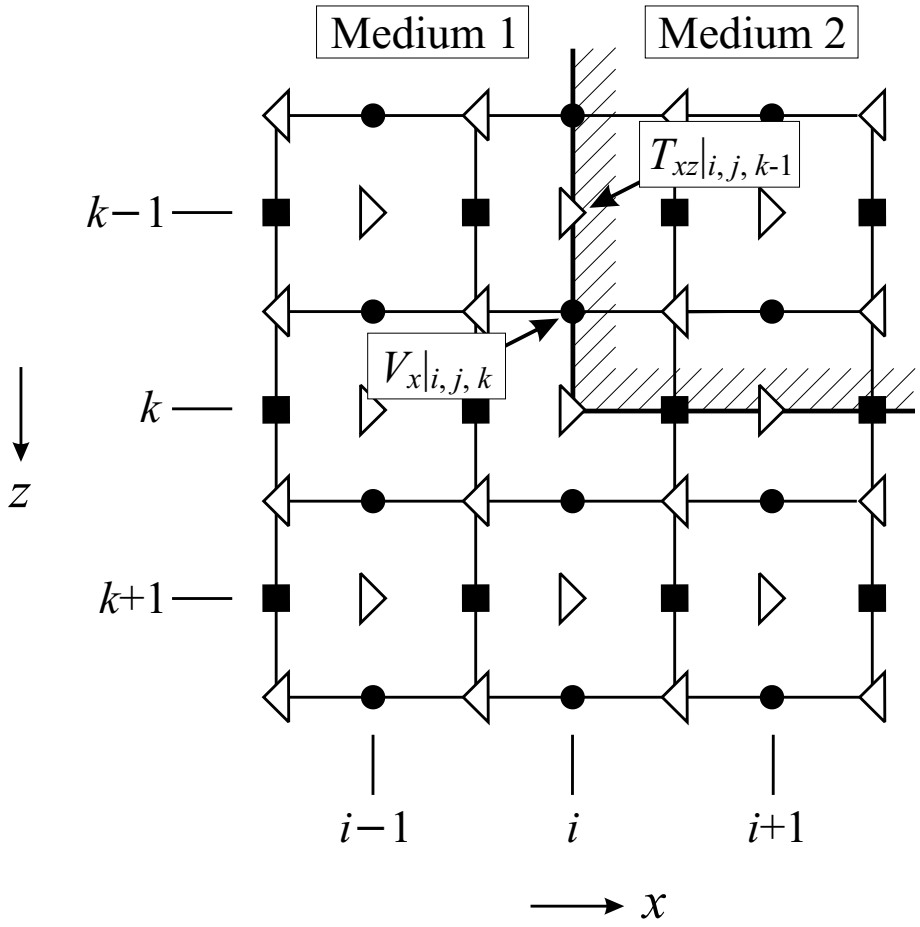


Figure 4.5: Schematic drawing of the finite-difference grid at the interface between two media. The normal particle velocity components are always located on the interface.

be a sufficient condition for stability of the finite-difference algorithm. To ensure that the Courant condition is a sufficient stability condition, the material properties must be averaged for all field components located on the boundary. A discussion of the stability of the finite-difference algorithm at a material interface and a derivation of the averaging procedure is given in Chapter 5.

At an interface between two media, say Medium 1 and Medium 2, the averaged

material density and the averaged shear stiffness are

$$\rho_{\text{avg}} = \frac{\rho_{\text{Medium 1}} + \rho_{\text{Medium 2}}}{2} \tag{4.35}$$

$$\mu_{\text{avg}} = \frac{2}{\frac{1}{\mu_{\text{Medium 1}}} + \frac{1}{\mu_{\text{Medium 2}}}}. \tag{4.36}$$

The longitudinal stiffness does not have to be averaged, because, due to the placement of the interface, the longitudinal stresses will never lie on the boundary. Note that the inverse of the shear stiffness is averaged.

In the discrete finite-difference grid, the material density and the shear stiffness are known at the locations of the particle velocity components and the shear stresses, respectively. For example, in the cross section in Fig. 4.5, $V_x|_{i,j,k}$ lies on the interface, and at its location the material density is averaged:

$$\rho_{\text{avg}}|_{i,j,k} = \left(\rho|_{i-1,j,k} + \rho|_{i,j,k}\right)/2. \tag{4.37}$$

At the location of the shear stress $T_{xz}|_{i,j,k-1}$, the averaged shear stiffness is

$$\mu_{\text{avg}}|_{i,j,k-1} = 2\left(\frac{1}{\mu|_{i,j,k-1}} + \frac{1}{\mu|_{i-1,j,k-1}}\right)^{-1}. \tag{4.38}$$

For the shear stress on the corner of the interface (which corresponds to an edge if the grid is expanded to its three dimensions), the shear stiffness must be averaged between all four adjacent cells:

$$\mu_{\text{avg}}|_{i,j,k} = 4\left(\frac{1}{\mu|_{i,j,k}} + \frac{1}{\mu|_{i-1,j,k}} + \frac{1}{\mu|_{i,j,k+1}} + \frac{1}{\mu|_{i-1,j,k+1}}\right)^{-1}. \tag{4.39}$$

The averaging is performed in an analogous manner for all field components on the interface.

4.2.3 Free-Surface Boundary

Six external boundaries arise at the six outer grid faces of the numerical model. Contrary to internal boundaries, external boundaries have to be treated explicitly to satisfy the required boundary conditions.

The numerical model is bounded at its upper face by a free surface (see Fig. 4.2). A free surface is the interface between a medium and vacuum. In vacuum, all elastic fields vanish. Consequently, due to the continuity of the normal stress, all normal stress components on the free surface must be zero:

$$\tau_{xz}(x, y, z = 0) = 0, \tag{4.40}$$

$$\tau_{yz}(x, y, z = 0) = 0, \tag{4.41}$$

$$\tau_{zz}(x, y, z = 0) = 0. \tag{4.42}$$

To enforce these conditions numerically, the field components on the free surface have to be determined separately.

Fig. 4.6 shows a cross section of the grid at the free surface. The normal stress T_{zz} is located on the surface, whereas T_{xz} and T_{yz} are not. All components within the medium are computed using the conventional finite-difference update equations. The components on the boundary are then computed by applying the boundary conditions according to Eqs. (4.40)–(4.42).

The stress-free boundary condition has to be satisfied during the computation of both the stress components at $t = l\Delta t$ and the velocity components at $t = (l+0.5)\Delta t$. The numerical field components of the $(i, j, 0)$-th cell located on the free surface are the velocities $V_x|_{i,j,0}$ and $V_y|_{i,j,0}$, the longitudinal stresses $T_{xx}|_{i,j,0}$,$T_{yy}|_{i,j,0}$ and $T_{zz}|_{i,j,0}$, and the shear stress $T_{xy}|_{i,j,0}$. At $t = l\Delta t$, the boundary conditions Eqs. (4.40)–(4.42) must be satisfied for the stress components. Eq. (4.42) is fulfilled by setting T_{zz} on the surface to zero:

$$T^l_{zz}|_{i,j,0} = 0. \tag{4.43}$$

Because T_{xz} and T_{yz} do not lie on the boundary, the boundary conditions Eqs. (4.40) and (4.41) cannot be applied directly. To make the shear stresses vanish on the surface, an extra row of basis cells is introduced one step beyond the free surface. With this extra row, T_{xz} and T_{yz} are averaged on the boundary, and the average is

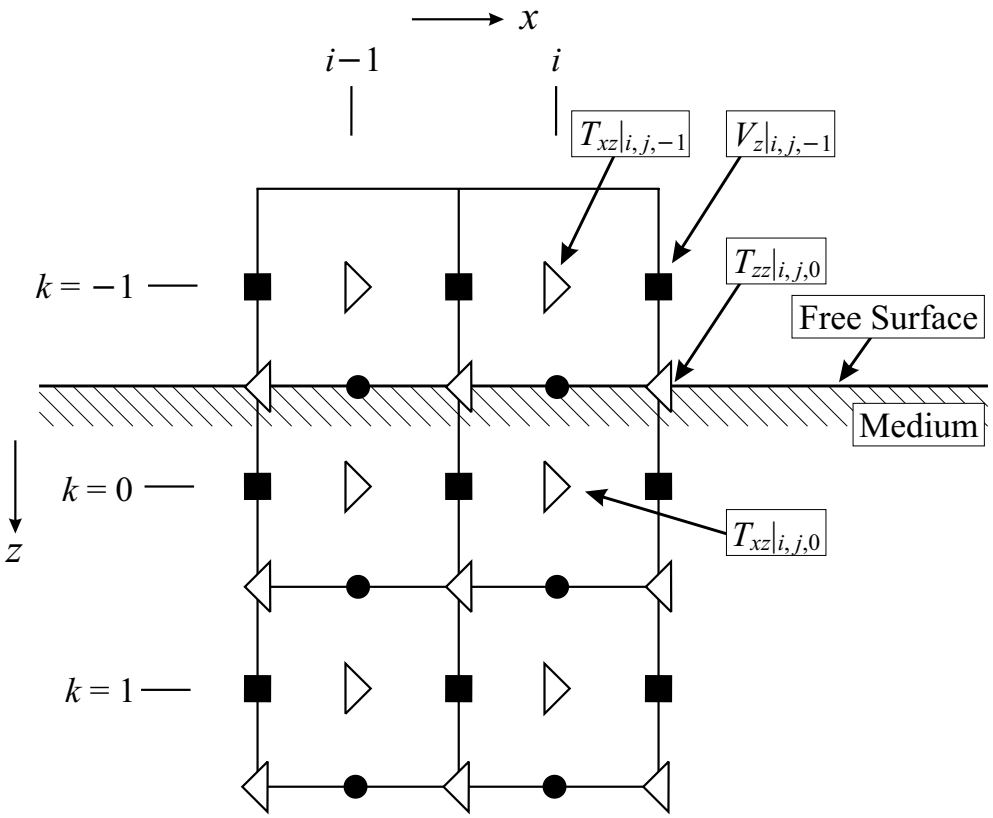


Figure 4.6: Finite difference grid and its field components at the free surface; cross section at $y = (j - 0.5)\Delta y$. An extra row at $k = -1$ must be inserted to satisfy the stress-free boundary condition at the free surface.

set to zero:

$$\left(T^l_{xz}|_{i,j,-1} + T^l_{xz}|_{i,j,0}\right)/2 \;=\; 0 \tag{4.44}$$

$$\left(T^l_{yz}|_{i,j,-1} + T^l_{yz}|_{i,j,0}\right)/2 \;=\; 0. \tag{4.45}$$

Because $T^l_{xz}|_{i,j,0}$ and $T^l_{yz}|_{i,j,0}$ lie within the medium, they are determined together with the field components in the medium using the normal update equations. Thus, Eqs. (4.44) and (4.45) can be used to determine the shear stresses on the extra row

beyond the boundary:

$$T^l_{xz}|_{i,j,-1} = -T^l_{xz}|_{i,j,0} \tag{4.46}$$

$$T^l_{yz}|_{i,j,-1} = -T^l_{yz}|_{i,j,0}. \tag{4.47}$$

Finally, knowing the velocity components on the free surface and on the extra row from the previous time step, $T^l_{xx}|_{i,j,0}$, $T^l_{yy}|_{i,j,0}$, and $T^l_{xy}|_{i,j,0}$ on the boundary are calculated using the normal update equations.

At $t = (l + 0.5)\Delta t$, the velocity components are determined. Using the stress components on the surface and on the inserted row, the velocity components, including the ones on the surface, can be updated using the normal update equations. The only velocity component that cannot be obtained in this manner is $V_z^{l+0.5}|_{i,j,-1}$ on the inserted row. Instead, $V_z^{l+0.5}|_{i,j,-1}$ is obtained by setting $T_{zz}^{l+1}|_{i,j,0}$ to zero within Eq. (4.28) according to (4.42). One obtains

$$\begin{aligned} 0 &= (\lambda + 2\mu)\frac{\Delta t}{\Delta z}(V_z^{l+0.5}|_{i,j,0} - V_z^{l+0.5}|_{i,j,-1}) \\ &\quad + \lambda\frac{\Delta t}{\Delta x}(V_x^{l+0.5}|_{i+1,j,0} - V_x^{l+0.5}|_{i,j,0}) \qquad (4.48) \\ &\quad + \lambda\frac{\Delta t}{\Delta y}(V_y^{l+0.5}|_{i,j,0} - V_y^{l+0.5}|_{i,j-1,0}). \qquad (4.49) \end{aligned}$$

By rearranging, V_z on the extra row is determined:

$$\begin{aligned} V_z^{l+0.5}|_{i,j,-1} &= V_z^{l+0.5}|_{i,j,0} + \frac{\lambda}{\lambda + 2\mu} \cdot \\ &\quad \cdot \left(\frac{\Delta z}{\Delta x}(V_x^{l+0.5}|_{i+1,j,0} - V_x^{l+0.5}|_{i,j,0}) + \right. \\ &\quad \left. + \frac{\Delta z}{\Delta y}(V_y^{l+0.5}|_{i,j,0} - V_y^{l+0.5}|_{i,j-1,0}) \right). \qquad (4.50) \end{aligned}$$

$V_z|_{i,j,-1}$ is necessary to calculate the stress components on the surface at the next time

step. Following this scheme, the stress-free boundary condition is satisfied across the entire surface, and all field components on the surface are determined.

4.2.4 Perfectly-Matched Layer Absorbing Boundary

To model the infinite extent of the half space, the waves that reach the outer grid faces must be absorbed. The boundary condition that does this most accurately is the *Perfectly Matched Layer* boundary condition. The particular formulation that is used here has been developed by Chew and Liu [55].

The Perfectly Matched Layer (PML) boundary condition was first introduced by Berenger [49] for the finite-difference modeling of electromagnetic wave propagation, then simplified for example by Sacks [50] and Gedney [51] and adapted to the elastic case by Chew and Liu [55]. A PML is a boundary layer consisting of a lossy material, in which waves are propagating with exactly the same phase velocity and the exact same wave impedance as waves in the solution space. Arbitrarily polarized waves incident on the PML are perfectly transmitted and then attenuated within the PML. In order to match the PML to the medium of the solution space, the field quantities must be split, leading to a set of non-physical equations.

Chew and Liu interpret the PML as coordinate stretching in the frequency domain [55]. This can be understood as follows: given a plane wave propagating in a lossless medium and described by the phasor $e^{jkx'}$, the stretching of coordinates represents a change of variables: $x' \longrightarrow (1+j\alpha)x$. The resulting plane wave is described by $e^{-\alpha x}e^{jkx}$ and, thus, is attenuated in the x-direction due to the complex coordinate stretching.

Introducing the nabla operator for stretched coordinates,

$$\nabla_e = \begin{bmatrix} \frac{1}{e_x}\frac{\partial}{\partial x} \\ \frac{1}{e_y}\frac{\partial}{\partial y} \\ \frac{1}{e_z}\frac{\partial}{\partial z} \end{bmatrix}, \tag{4.51}$$

where $e_x = 1+j\alpha_x$, $e_y = 1+j\alpha_y$ and $e_z = 1+j\alpha_z$ represent the stretching factors, it can

be shown that a perfectly matched layer exists simultaneously for both pressure (P) and vertical shear (SV) waves [55]. Considering an interface between Medium 1 and Medium 2 at $z = 0$ normal to the z-direction, e_z in (4.51) can be chosen differently in Medium 1 and Medium 2 while still maintaining a perfectly matched interface between the two media . For example, if Medium 1 is lossless in the z-direction ($e_z = 1$) and Medium 2 is lossy ($e_z = 1 + j\alpha_z$), then Medium 1 will still be perfectly matched to Medium 2, and a wave incident from Medium 1 onto Medium 2 will be transmitted without any reflection. A theoretical derivation of the PML for elastodynamics is given in [55].

The implementation of the PML for the discrete grid shall be explained here by deriving the formulation for the velocity component v_x from Eq. (4.11). Transforming Eq. (4.11) into the frequency domain and introducing the coordinate-stretching variables

$$e_x = 1 + j\frac{\Omega_x}{\omega} \tag{4.52}$$

$$e_y = 1 + j\frac{\Omega_y}{\omega} \tag{4.53}$$

$$e_z = 1 + j\frac{\Omega_z}{\omega}, \tag{4.54}$$

one obtains

$$-j\omega\rho\overline{v}_x = \frac{1}{1 + j\frac{\Omega_x}{\omega}}\frac{\partial\overline{\tau}_{xx}}{\partial x} + \frac{1}{1 + j\frac{\Omega_y}{\omega}}\frac{\partial\overline{\tau}_{xy}}{\partial y} + \frac{1}{1 + j\frac{\Omega_z}{\omega}}\frac{\partial\overline{\tau}_{xz}}{\partial z}, \tag{4.55}$$

where the bar indicates quantities in the frequency domain. In order to dampen the fields independently in the x-, y- and z-direction, the field components must be split. For each derivative in space, a new field component is introduced. Thus, three split field components arise for Eq. (4.55), one for each term on the right hand side of the equation:

$$-j\omega\rho\overline{v}_{x,x} = \frac{1}{1 + j\frac{\Omega_x}{\omega}}\frac{\partial\overline{\tau}_{xx}}{\partial x}, \tag{4.56}$$

$$-j\omega\rho\overline{v}_{x,y} = \frac{1}{1+j\frac{\Omega_y}{\omega}}\frac{\partial\overline{\tau}_{xy}}{\partial y}, \tag{4.57}$$

$$-j\omega\rho\overline{v}_{x,z} = \frac{1}{1+j\frac{\Omega_z}{\omega}}\frac{\partial\overline{\tau}_{xz}}{\partial z}, \qquad \text{and} \tag{4.58}$$

$$\overline{v}_x = \overline{v}_{x,x} + \overline{v}_{x,y} + \overline{v}_{x,z}. \tag{4.59}$$

Rearranging Eqs. (4.56)–(4.58) and transforming them back into the time domain, expressions for the split fields in the time domain are obtained:

$$(\rho\frac{\partial}{\partial t} + \rho\Omega_x)v_{x,x} = \frac{\partial\tau_{xx}}{\partial x}, \tag{4.60}$$

$$(\rho\frac{\partial}{\partial t} + \rho\Omega_y)v_{x,y} = \frac{\partial\tau_{xy}}{\partial y}, \tag{4.61}$$

$$(\rho\frac{\partial}{\partial t} + \rho\Omega_z)v_{x,z} = \frac{\partial\tau_{xz}}{\partial z} \tag{4.62}$$

$$v_x = v_{x,x} + v_{x,y} + v_{x,z} \tag{4.63}$$

The above set of equations can be discretized by introducing finite differences. Note that, when Eqs. (4.60)–(4.62) are discretized, formally the velocities must be known at the previous full and half time step. This is due to the fact that the velocities on the left hand side appear once in their time-differentiated form and once in a non-differentiated form. For example, in order to calculate $V_{x,x}^{l+0.5}$, both $V_{x,x}^{l}$ and $V_{x,x}^{l-0.5}$ must be known due to the fact that $\frac{\partial v_{x,x}}{\partial t}$ and $v_{x,x}$ appear in the governing equation (see Eq. (4.60)). However, $V_{x,x}^{l}$ can be approximated by $V_{x,x}^{l} = (V_{x,x}^{l+0.5} + V_{x,x}^{l-0.5})/2$. In that way, $V_{x,x}^{l}$ is eliminated and only $V_{x,x}^{l+0.5}$ and $V_{x,x}^{l-0.5}$ remain.

Altogether, 21 split field components arise from the original set of governing equations. A set of 21 equations is derived, which, together with the nine combining equations, form a complete set for the 21 split field components and the nine original field components. The finite-difference update equations for the split field formulation can be found in Appendix A.

Even though a perfectly matched layer exists in continuous space, it does not exist in discrete space and reflections will occur at the interface between the solution space and the PML. In order to keep the reflections low, a tapered loss profile is used within the PML, causing the loss to increase gradually. Typically, a second-, third- or fourth-order profile for the loss within the PML is chosen. In this way, reflections at the interface can be reduced by as much as 80 dB. For this work, the loss profile within the PML is chosen to be

$$\Omega(i) = \Omega_{\max} \cdot \left(\frac{n}{N_{\mathrm{PML}}}\right)^m, \tag{4.64}$$

where $m = 2.1$ and $\Omega_{\max} = \frac{0.1 N_{\mathrm{PML}}}{\Delta t}$; N_{PML} is the thickness of the PML in basis cells, and n indicates the position within the PML. A similar loss profile has been found to yield best performance in electromagnetic finite-difference modeling. Figure 4.7 shows a schematic drawing of a PML adjacent to the solution space. The crosses indicate the loss values for the field components at the center of a cell, whereas the circles mark the loss values of the field components on the cell edges.

4.3 Injection of Plane Waves

To excite plane waves in the numerical model, the *total-field/scattered-field* formulation is used. In [17], a detailed description of the total-field/scattered-field formulation for electromagnetic finite-difference modeling can be found. In this section, the adaption of the procedure to the elastodynamic case is explained.

The total-field/scattered-field formulation makes use of the linearity of the governing elastic equations by decomposing the total field into its incident and its scattered field component:

$$\mathbf{v}_{\mathrm{tot}} = \mathbf{v}_{\mathrm{inc}} + \mathbf{v}_{\mathrm{scat}}\,. \tag{4.65}$$

For the total-field/scattered-field formulation, the solution space is divided into two regions. Within a box, the total field, consisting of the incident and the scattered field, is computed, whereas outside of the box only the scattered field is determined.

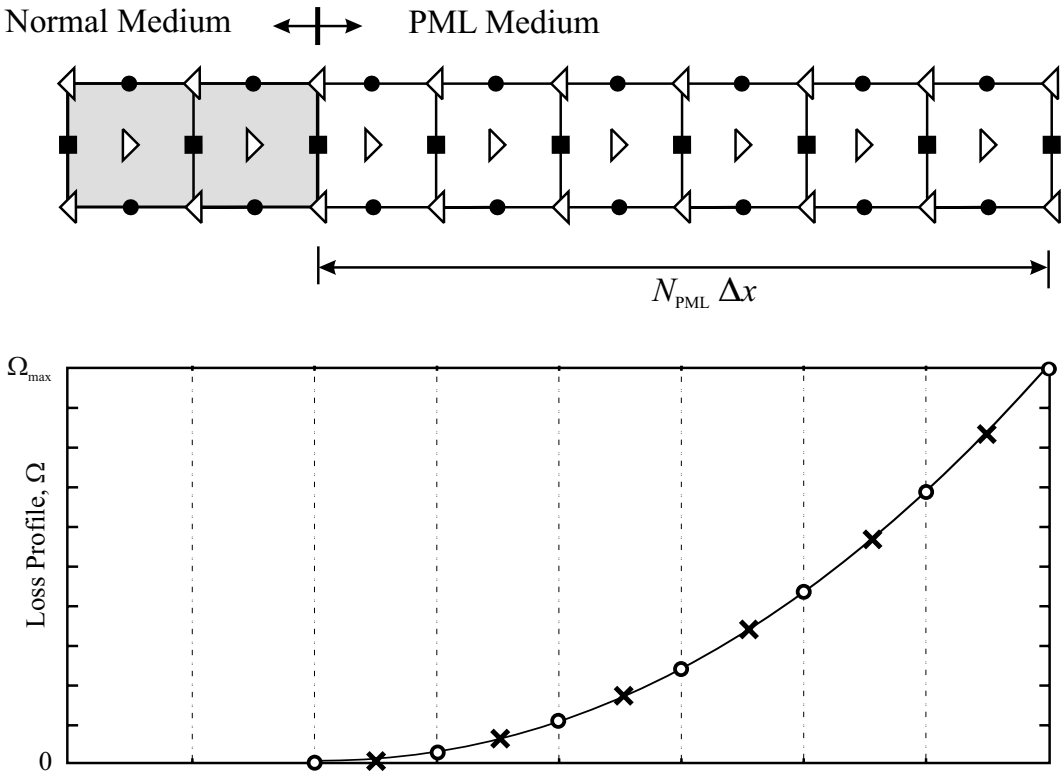


Figure 4.7: Schematic drawing of a PML layer in the x-direction: loss profile according to Eq. (4.64). The crosses and circles indicate the loss values of the field components at the center and at the edges of a cell, respectively.

Fig. 4.8 illustrates schematically the division of the solution space into the two regions, the inner *Total-Field Region* and the outer *Scattered-Field Region*. The structure that scatters the incident waves must necessarily be completely embedded in the total-field region. Within the total-field region, the finite-difference algorithm operates on the total-field components. In the outer region, no incident wave is present and the algorithm operates only on the scattered field. The scattered-field region is terminated by the boundary conditions at the outer grid faces, i. e., either by the free-surface boundary condition or the PML absorbing boundary condition.

An artificial, non-physical interface will arise between the total-field region and the scattered-field region. To provide the transition from the total-field region to the

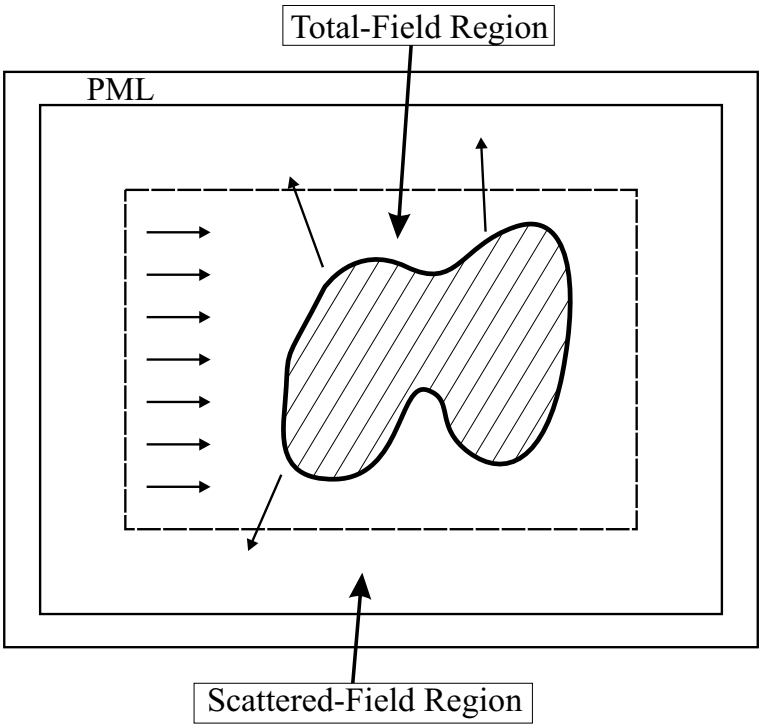


Figure 4.8: Division of the solution space into a *Total-Field Region* and a *Scattered-Field Region*. The scattering objects must be completely embedded in the total-field region.

scattered-field region, all field components adjacent to the interface must be adjusted. This is possible, if the incident field is known at all points in space and for all times. In that case, the scattered field can be computed from the total field for the components on the interface and vice versa.

4.3.1 Procedure

In each the total-field region and the scattered-field region, the conventional finite-difference update equations are applied to compute the field values. At the interface between the two regions, the field components are adjusted. In general, a correction term must be added to all field components that are dependent on field components lying *across* the boundary.

Fig. 4.9 shows the field components at the interface on a cross section representing the x-z-plane and cutting through the in-plane velocity components V_x and

and V_z. The field components at the interface are indicated in black. The interface

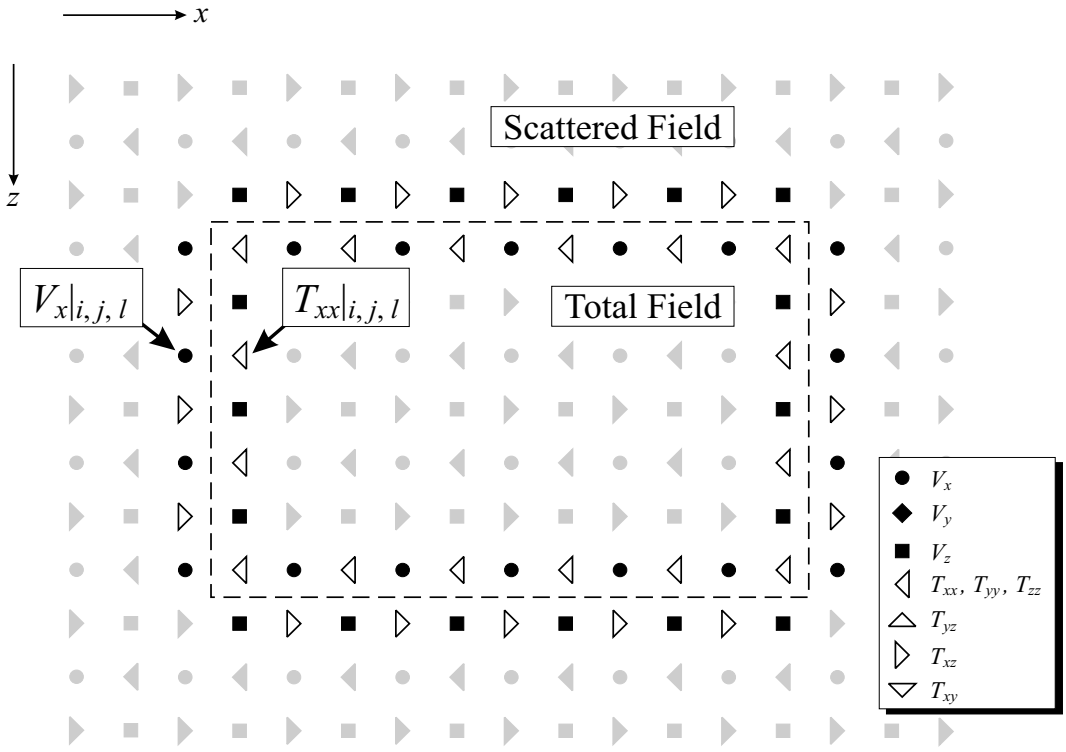


Figure 4.9: Both the field components just outside and just inside the total-field box must be corrected with the known incident fields.

is placed such that the longitudinal stresses will always lie on the edges and corners of the bounding box for the inner total-field region.

The adjustment procedure is illustrated for $V_x|_{i,j,k}$ and $T_{xx}^{l+1}|_{i,j,k}$ at the interface to the left. $V_x|_{i,j,k}$ lies in the scattered-field region just *outside* of the total-field region. To update the velocity components throughout the scattered-field region, the normal finite-difference equations are applied. The update equation for $V_x|_{i,j,k}$ at the interface then becomes:

$$V_{x,\mathrm{scat}}^{l+0.5}|_{i,j,k} = V_{x,\mathrm{scat}}^{l-0.5}|_{i,j,k} + \frac{\Delta t}{\rho} \cdot \left[\frac{T_{xx,\mathrm{scat}}^{l}|_{i,j,k} - T_{xx,\mathrm{scat}}^{l}|_{i-1,j,k}}{\Delta x} + \right. \tag{4.66}$$

$$\left. + \frac{T_{xy,\mathrm{scat}}^{l}|_{i,j,k} - T_{xy,\mathrm{scat}}^{l}|_{i,j-1,k}}{\Delta y} + \frac{T_{xz,\mathrm{scat}}^{l}|_{i,j,k} - T_{xz,\mathrm{scat}}^{l}|_{i,j,k-1}}{\Delta z} \right].$$

In this equation, $T_{xx,\mathrm{scat}}^{l}|_{i,j,k}$ is unknown, because T_{xx} at the (i,j,k)-th cell lies in the

total field region. However, $T^l_{xx,\mathrm{scat}}|_{i,j,k}$ can be computed from the total-field value, if the incident field value is known:

$$T^l_{xx,\mathrm{scat}}|_{i,j,k} = T^l_{xx,\mathrm{tot}}|_{i,j,k} - T^l_{xx,\mathrm{inc}}|_{i,j,k}. \tag{4.67}$$

Inserting this into Eq. (4.66), one obtains an update equation for $V_x|_{i,j,k}$ at the interface in which all field components are known:

$$\begin{aligned} V^{l+0.5}_{x,\mathrm{scat}}|_{i,j,k} &= V^{l-0.5}_{x,\mathrm{scat}}|_{i,j,k} + \frac{\Delta t}{\rho} \cdot \left[\frac{T^l_{xx,\mathrm{tot}}|_{i,j,k} - T^l_{xx,\mathrm{scat}}|_{i-1,j,k}}{\Delta x} + \right. \\ &\left. + \frac{T^l_{xy,\mathrm{scat}}|_{i,j,k} - T^l_{xy,\mathrm{scat}}|_{i,j-1,k}}{\Delta y} + \frac{T^l_{xz,\mathrm{scat}}|_{i,j,k} - T^l_{xz,\mathrm{scat}}|_{i,j,k-1}}{\Delta z} \right] - \\ &- \frac{\Delta t}{\rho \Delta x} \cdot T^l_{xx,\mathrm{inc}}|_{i,j,k}. \end{aligned} \tag{4.68}$$

The field components at the interface just *inside* the total-field region are treated in a similar way. By applying the conventional update equation, $T^{l+1}_{xx}|_{i,j,k}$ at the interface is obtained:

$$\begin{aligned} T^{l+1}_{xx,\mathrm{tot}}|_{i,j,k} &= T^l_{xx,\mathrm{tot}}|_{i,j,k} + \Delta t \cdot \left[(\lambda + 2\mu) \frac{V^{l+0.5}_{x,\mathrm{tot}}|_{i+1,j,k} - V^{l+0.5}_{x,\mathrm{tot}}|_{i,j,k}}{\Delta x} + \right. \\ &\left. + \lambda \frac{V^{l+0.5}_{y,\mathrm{tot}}|_{i,j,k} - V^{l+0.5}_{y,\mathrm{tot}}|_{i,j-1,k}}{\Delta y} + \lambda \frac{V^{l+0.5}_{z,\mathrm{tot}}|_{i,j,k} - V^{l+0.5}_{z,\mathrm{tot}}|_{i,j,k-1}}{\Delta z} \right]. \end{aligned} \tag{4.69}$$

In this equation, $V^{l+0.5}_{x,\mathrm{tot}}|_{i,j,k}$ is unknown, because it is located in the scattered-field region. Using the incident field, an update equation with only known field values is derived:

$$\begin{aligned} T^{l+1}_{xx,\mathrm{tot}}|_{i,j,k} &= T^l_{xx,\mathrm{tot}}|_{i,j,k} + \Delta t \cdot \left[(\lambda + 2\mu) \frac{V^{l+0.5}_{x,\mathrm{tot}}|_{i+1,j,k} - V^{l+0.5}_{x,\mathrm{scat}}|_{i,j,k}}{\Delta x} + \right. \\ &\left. + \lambda \frac{V^{l+0.5}_{y,\mathrm{tot}}|_{i,j,k} - V^{l+0.5}_{y,\mathrm{tot}}|_{i,j-1,k}}{\Delta y} + \lambda \frac{V^{l+0.5}_{z,\mathrm{tot}}|_{i,j,k} - V^{l+0.5}_{z,\mathrm{tot}}|_{i,j,k-1}}{\Delta z} \right] - \\ &- \frac{\Delta t(\lambda + 2\mu)}{\Delta x} \cdot V^{l+0.5}_{x,\mathrm{inc}}|_{i,j,k}. \end{aligned} \tag{4.70}$$

Usually, to simplify the implementation, the field components at the interface are first computed using the conventional update equations and are then adjusted after each time step. In three dimensions, six interfaces between the inner total-field region and the outer scattered-field region arise. By applying the normal update equations to the terms at the interface, equations containing both scattered- and total-field components arise. These are inconsistent and must be corrected by adding a term containing the incident-field component (see Eqs. (4.68) and (4.70)). A complete listing of the field components at each interface and their corresponding correction terms can be found in Appendix B.

4.3.2 Incident Field

To adjust the field components on the interface between the total-field and the scattered-field region, the incident field must be known at all points in space at all times. Various schemes are possible to predict the incident field. To describe an incident plane wave, a one-dimensional finite-difference formulation can be employed. With the 1-D finite-difference algorithm, the plane wave is computed separately using the same space and time increment as in the 3-D grid. By mapping the 1-D field onto the 3-D grid, a *look-up table* is created that contains the incident field values for all points in space and time. Using this look-up table, the field components at the interfaces between the total- and scattered-field region are adjusted. A plane pressure wave as well as a plane shear wave can be, in this way, injected into the grid.

Employing the 1-D finite-difference algorithm to compute the incident field has the further advantage that the incident field as a function of time and space is described as it appears in the three-dimensional grid. By using a numerical algorithm rather than an analytical description, the numerical wave speed in the grid, which will slightly differ from its analytical value, is emulated. Thus, a more accurate adjustment of the field components on the interface can be achieved.

4.4 Parallelization

The three-dimensional finite-difference model has been implemented using `FORTRAN 90`. The computational and memory requirements of the numerical model are immense, and the computations usually cannot be performed on a conventional personal computer. Therefore, the model is implemented in a fully parallel fashion using the *Message-Passing Interface* (`MPI`), and computers with a parallel architecture that contain multiple processors are used.

In the model, the elastic wave fields are computed on a three-dimensional discrete grid. When implemented on a parallel computer, this grid is divided into several sub-grids, and each sub-grid is assigned to one processor of the parallel machine [17]. Fig. 4.10 visualizes the distribution of the processors across the computational grid. Here, the grid is divided into eight sub-grids, and each sub-grid is assigned to one processor. The processors then compute the wave fields on their sub-grids independently from each other and share the data on the common interfaces after each time step. The procedure is algorithmically simple, but its implementation is rather tedious. To improve the computational efficiency of the parallel program, some important guide lines should be followed. Most importantly, the amount of data that is passed between the processors after each time step must be minimized. For this, the grid is divided into sub-grids whose dimensions in all directions are approximately equal. In that way, the surface area of the sub-grids, i. e., the face area that each sub-grid shares with its neighbors and thus the amount of shared data, is minimal.

The computations for this work have been performed on a Cray T3E parallel supercomputer at the *ERDC Massively Shared Resource Center* of the DoD, located in Vicksburg, Mississippi, and on a Beowulf computer cluster [60]. The Beowulf cluster is located at the Georgia Institute of Technology and contains 50 Pentium III processors and 6 GBytes of memory.

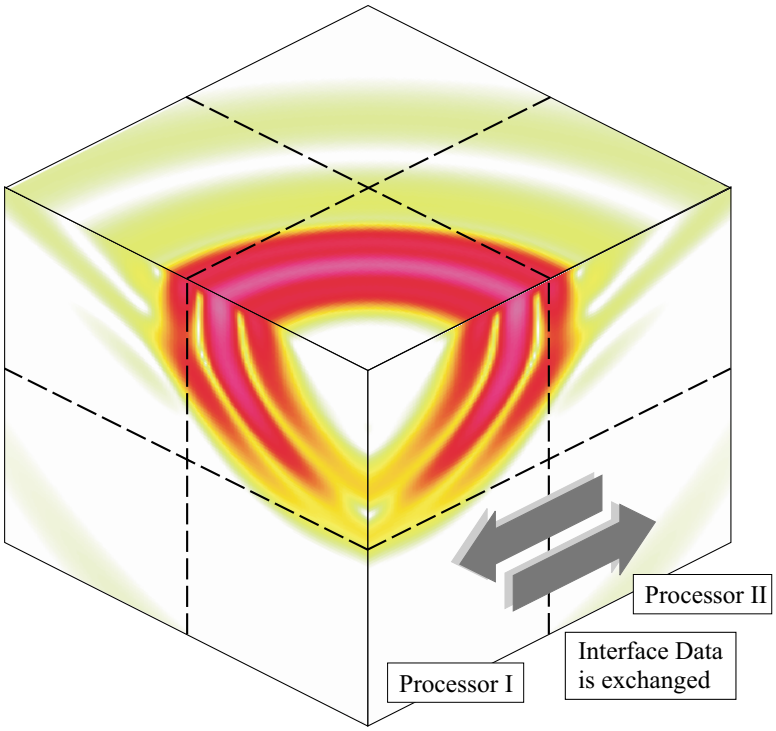


Figure 4.10: Parallelization; the solution space is divided into several cubes, which are assigned to different processor.

CHAPTER 5

Stability of the FDTD Algorithm at a Material Interface

As described in Chapter 4, a finite-difference model has been developed in two and three dimensions that simulates the elastic wave propagation in the ground. During the development of the finite-difference model, instabilities in the algorithm occurred.

In the finite-difference algorithm, internal boundaries, i. e., boundaries between different materials, are satisfied implicitly. The finite-difference algorithm is usually stable, if the Courant condition is satisfied. However, if media with greatly different material properties are considered, it turns out that the Courant condition is not necessarily a sufficient condition for stability, but a more restrictive condition arises. One way to avoid the more restrictive stability condition is to average the material parameters at the interface between two different media.

In this chapter, the stability behavior of the finite-difference scheme at a material interface is analyzed. First, a longitudinal wave normally incident onto a material interface is considered. The von-Neumann method and the matrix method are employed, and theoretical bounds for stability are established. It is shown that, if the material parameters are averaged on the boundary, the Courant condition poses a sufficient condition for stability, and no further restrictions on the stability condition due to the presence of the material interface arise. Second, a numerical study of stability in 1-D and 2-D is performed, and the theoretical bounds are validated.

Most importantly, it is shown that the FDTD algorithm in fact is stable whenever the material properties are averaged on the boundary. Finally, the averaging procedure is described.

The mathematical theorems and procedures presented in this chapter are well-known. The stability analysis closely follows Richtmyer [61], Smith [62], and Mitchell and Griffiths [63], who also provide a more thorough description of the mathematical theory. A similar analysis of stability has been performed, for example, by Ilan and Loewenthal, who investigated the stability of the second-order finite-difference formulation, based on the elastic wave equation, at a free-surface boundary [22].

5.1 Stability Analysis: Theory

The fundamental condition for stability of the finite-difference algorithm, i. e., the condition that relates the size of the time increment to the spacing of the discrete nodes in the FDTD grid and that must be fulfilled for the finite-difference algorithm to be stable, is the *Courant condition* [64, 17]. The Courant condition states that the physical wave speed of any wave excited in the model must not exceed the velocity that information can travel with in the discrete grid. Mathematically, for a space step of size Δl, the Courant condition in an n-dimensional grid is

$$\frac{\Delta t}{\Delta l} c_{max} \leq \frac{1}{\sqrt{n}}. \tag{5.1}$$

Here, Δt denotes the time step, and c_{max} is the maximum wave speed occurring in the numerical model. Note that the Courant condition becomes more restrictive when the number of dimensions is increased.

In this section, it is shown that the Courant condition by itself is not always a sufficient condition for stability, if a material interface is present. A theoretical bound in terms of the material properties is derived beyond which the material parameters must be averaged for the finite-difference scheme to be stable. The bound is derived using the *von-Neumann* method and the *Matrix* method for a 1-D longitudinal wave

incident onto a material interface.

5.1.1 1-D Longitudinal Wave Incident onto a Material Interface

The stability behavior of a 1-D finite-difference grid for a longitudinal wave normally incident onto a material interface is to be investigated. The propagation direction coincides with the x-direction, and the only non-zero velocity component is v_x. Figure 5.1 shows a portion of the discrete first-order finite-difference grid. The material interface is located between T_{xx} at $i-0.5$ and V_x at i. The first-order finite-difference system of equations can be combined to obtain a second-order formulation in terms of solely the particle velocity. By combining the discrete first-order equations rather than discretizing the second-order wave equation, the stability behavior of the first-order finite-difference formulation is preserved. In the second-order formulation, the stress T_{xx} is eliminated, but the Lamé constants, which are associated with the stress components, are still located half a step in between the velocity components. The second-order finite-difference update equation for the longitudinal particle velocity V_x becomes

$$\begin{aligned} V_x^{k+1}|_n &= V_x^k|_n \left[2 - \frac{\Delta t^2}{\Delta x^2} \frac{\lambda|_{n-0.5} + 2\mu|_{n-0.5} + \lambda|_{n+0.5} + 2\mu|_{n+0.5}}{\rho|_n}\right] \\ &\quad + V_x^k|_{n+1} \left[\frac{\Delta t^2}{\Delta x^2} \frac{\lambda|_{n+0.5} + 2\mu|_{n+0.5}}{\rho|_n}\right] \\ &\quad + V_x^k|_{n-1} \left[\frac{\Delta t^2}{\Delta x^2} \frac{\lambda|_{n-0.5} + 2\mu|_{n-0.5}}{\rho|_n}\right] \\ &\quad - V_x^{k-1}|_n, \end{aligned} \tag{5.2}$$

where $n = 1 \ldots N$ labels the nodes in the 1-D grid and k represents the discrete time step. Again, note that the Lamé constants are placed at the half steps in between the velocity nodes.

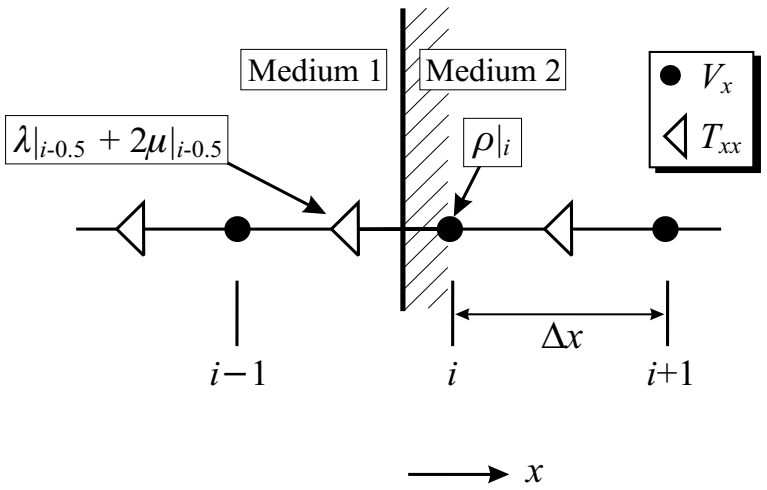


Figure 5.1: 1-D finite-difference grid for a longitudinal wave normally incident onto a material boundary.

The von-Neumann Method

The second-order finite-difference equation for V_x at the i-th node, i. e., the node closest to the interface, can be written as

$$\frac{V_x^{k+1}|_i - 2V_x^k|_i + V_x^{k-1}|_i}{\Delta t^2} = \tag{5.3}$$

$$\frac{1}{\rho_2}\frac{(\lambda_2 + 2\mu_2)V_x^k|_{i+1} - [(\lambda_1 + 2\mu_1) + (\lambda_2 + 2\mu_2)]V_x^k|_i + (\lambda_1 + 2\mu_1)V_x^k|_{i-1}}{\Delta x^2}.$$

Applying the *von-Neumann* (or *Fourier-Series*) method [63, 62], a solution of the form

$$V_x^k|_i = q^k e^{j\beta i} \tag{5.4}$$

is inserted into Eq. (5.46). The solution can be thought of as the Fourier series representation of an arbitrary error function which must not be amplified as it propagates through the grid. For the finite-difference algorithm to be stable all solutions for arbitrary β must be bounded at all times and, thus, their magnitude must be smaller than or equal to 1, yielding

$$|q^k| \leq |q| \leq 1. \tag{5.5}$$

Introducing the notation

$$L = \frac{\lambda_2 + 2\mu_2}{\rho_2}\frac{\Delta t^2}{\Delta x^2}$$

$$R = \frac{\lambda_1 + 2\mu_1}{\rho_2}\frac{\Delta t^2}{\Delta x^2}, \tag{5.6}$$

and inserting the von-Neumann solutions, Eq. (5.3) becomes

$$q^{k+1}e^{j\beta i} - 2q^k e^{j\beta i} + q^{k-1}e^{j\beta i} = Lq^k e^{j\beta(i+1)} - (L+R)q^k e^{j\beta i} + Rq^k e^{j\beta(i-1)}. \tag{5.7}$$

Dividing by $q^{k-1}e^{j\beta i}$, a quadratic equation for q is obtained:

$$q^2 - 2q + 1 = Lqe^{j\beta} - (L+R)q + Rqe^{-j\beta}. \tag{5.8}$$

In the worst case, $e^{j\beta}$ and $e^{-j\beta}$ are equal to -1 and the quadratic equation reads

$$q^2 - 2q(1 - (L+R)) + 1 = 0. \tag{5.9}$$

The roots of this equation are

$$q = 1 - (L+R) \pm \sqrt{L+R}\sqrt{L+R-2}. \tag{5.10}$$

To satisfy $|q| \leq 1$, it can be shown that $(L+R) \leq 2$ and, thus,

$$\frac{\Delta t^2}{\Delta x^2}\frac{(\lambda_1 + 2\mu_1) + (\lambda_2 + 2\mu_2)}{\rho_2} = \frac{\Delta t^2}{\Delta x^2}\left(\frac{\lambda_1 + 2\mu_1}{\rho_2} + c_2^2\right) \leq 2. \tag{5.11}$$

This is a sufficient condition for the local stability of the finite-difference algorithm at the node closest to the interface.

The analysis outlined above does not take the general medium into account, but derives a stability criterion for the boundary node viewed by itself. The von-Neumann method checks for *local* stability in the finite-difference grid. It is not yet clear how good an estimate this analysis provides for the stability of the entire system. However, in the next section it will be shown that the condition derived here is in fact a *bound* for stability and, furthermore, in a certain range represents a fairly good estimate for the stability of the entire system.

The Matrix Method

For a finite-sized problem, the finite-difference scheme can be formulated as a matrix equation [62]. Considering a one-dimensional grid with nodes $n = 1 \dots N$, and assuming $V_x|_0 = V_x|_{N+1} = 0$ on the outer grid edges at all times, the finite-difference scheme can be written as

$$\mathbf{V_x}^{k+1} = \mathbf{A} \cdot \mathbf{V_x}^{k} - \mathbf{I} \cdot \mathbf{V_x}^{k-1}, \tag{5.12}$$

where

$$\mathbf{V_x}^{k} = \begin{bmatrix} V_x^k|_1 \\ \vdots \\ V_x^k|_N \end{bmatrix}$$

is a vector containing the velocity values at the N nodes at time k, and $\mathbf{I}$ is the $N \times N$ identity matrix. $\mathbf{A}$ is a tri-diagonal matrix:

$$\mathbf{A} = \begin{bmatrix} \ddots & \ddots & & & & & \\ \ddots & 2-2\frac{\Delta t^2}{\Delta x^2}c_1^2 & \frac{\Delta t^2}{\Delta x^2}c_1^2 & & & 0 & \\ & \frac{\Delta t^2}{\Delta x^2}c_1^2 & 2-2\frac{\Delta t^2}{\Delta x^2}c_1^2 & \frac{\Delta t^2}{\Delta x^2}c_1^2 & & & \\ & & \frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1+2\mu_1}{\rho_2} & 2-\frac{\Delta t^2}{\Delta x^2}\left(c_2^2+\frac{\lambda_1+2\mu_1}{\rho_2}\right) & \frac{\Delta t^2}{\Delta x^2}c_2^2 & & \\ & & & \frac{\Delta t^2}{\Delta x^2}c_2^2 & 2-2\frac{\Delta t^2}{\Delta x^2}c_2^2 & \frac{\Delta t^2}{\Delta x^2}c_2^2 & \\ & 0 & & & \frac{\Delta t^2}{\Delta x^2}c_2^2 & 2-2\frac{\Delta t^2}{\Delta x^2}c_2^2 & \ddots \\ & & & & & \ddots & \ddots \end{bmatrix}, \tag{5.13}$$

where

$$c_1 = \sqrt{\frac{\lambda_1 + 2\mu_1}{\rho_1}}, \quad c_2 = \sqrt{\frac{\lambda_2 + 2\mu_2}{\rho_2}} \tag{5.14}$$

describe the longitudinal wave speeds in Medium 1 and Medium 2, respectively. In this matrix, rows $n = 1 \dots i-1$ describe Medium 1, the i-th row is the material

interface, and the rows $n = i+1 \ldots N$ represent Medium 2. Eq. (5.12) can be written in a more compact form:

$$\begin{bmatrix} \mathbf{V_x}^{k+1} \\ \mathbf{V_x}^{k} \end{bmatrix} = \mathbf{B} \cdot \begin{bmatrix} \mathbf{V_x}^{k} \\ \mathbf{V_x}^{k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{V_x}^{k} \\ \mathbf{V_x}^{k-1} \end{bmatrix}, \tag{5.15}$$

where each submatrix of $\mathbf{B}$ is $N \times N$ and, consequently, $\mathbf{B}$ is $2N \times 2N$.

For the finite-difference algorithm to be stable, the magnitude of all eigenvalues of $\mathbf{B}$ must be smaller than 1 [63] –[61]. However, the eigenvalues cannot be easily determined in closed form and, thus, they are estimated and bounds on the eigenvalues are established.

The analysis of the eigenvalues is made considerably easier considering the structure of the matrix $\mathbf{B}$. $\mathbf{B}$ consists of four blocks of size $N \times N$: $\mathbf{A}$, $\mathbf{I}$, $-\mathbf{I}$ and $\mathbf{0}$. In general, if a matrix can be divided into $M \times M$ square sub-blocks of equal size $N \times N$, and the sub-blocks have a common set of N linearly independent eigenvectors, then the eigenvalues of the entire matrix are given by the eigenvalues of the matrices

$$\begin{bmatrix} \lambda_{11}^{p} & \cdots & \lambda_{1M}^{p} \\ \vdots & & \vdots \\ \\ \lambda_{M1}^{p} & \cdots & \lambda_{MM}^{p} \end{bmatrix}, \qquad p = 1 \ldots N \tag{5.16}$$

where p labels the p-th eigenvector to each sub-block [62].

Since any vector is an eigenvector of $\mathbf{I}$ and $\mathbf{0}$, the four sub-blocks of $\mathbf{B}$ indeed have a common set of eigenvectors: all eigenvectors of $\mathbf{A}$ are also eigenvectors to the identity matrix as well as the zero matrix. The identity matrix has the N-fold eigenvalue one and the zero-matrix has the N-fold eigenvalue zero. Thus, letting λ_{A} indicate any eigenvalue of $\mathbf{A}$, the eigenvalues of $\mathbf{B}$ are given by the eigenvalues of the matrix

$$\begin{bmatrix} \lambda_{\mathrm{A}} & -1 \\ 1 & 0 \end{bmatrix}, \tag{5.17}$$

and are determined from the quadratic equation

$$1 - \lambda(\lambda_{\mathrm{A}} - \lambda) = 0. \tag{5.18}$$

The eigenvalues are then

$$\lambda = \frac{\lambda_A}{2} \pm \sqrt{\frac{\lambda_A^2}{4} - 1}. \tag{5.19}$$

Note here that all eigenvalues of $\mathbf{A}$ are real.[1]

The magnitude of all eigenvalues of $\mathbf{B}$ must be smaller than or equal to 1: $|\lambda| \leq 1$. For this to be true, it can be shown from Eq. (5.19) that the magnitude of all eigenvalues of $\mathbf{A}$ must be smaller than or equal to 2:

$$|\lambda_A| \leq 2. \tag{5.20}$$

It is then sufficient to find or estimate the eigenvalues of $\mathbf{A}$ and conclude the stability condition from these.

The eigenvalues of a matrix can most easily be approximated using *Gershgorin's Circle Theorem* [65]. Gershgorin's Circle theorem states that the eigenvalues of a matrix lie within circles in the complex plane whose centers are the elements of the matrix's main diagonal and whose radii are equal to the sum of the magnitude of the off-diagonal row elements:

$$|\lambda_n - a_{nn}| \leq \sum_{\substack{m \\ n \neq m}} |a_{nm}|. \tag{5.21}$$

According to Gershgorin's Circle Theorem, the eigenvalues of $\mathbf{A}$ lie within three circles:

$$\left|\lambda_A - \left(2 - 2\frac{\Delta t^2}{\Delta x^2} c_1^2\right)\right| \leq 2\frac{\Delta t^2}{\Delta x^2} c_1^2; \tag{5.22}$$

$$\left|\lambda_A - \left(2 - \frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1 + 2\mu_1}{\rho_2}\right)\right)\right| \leq \frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1 + 2\mu_1}{\rho_2}\right); \tag{5.23}$$

$$\left|\lambda_A - \left(2 - 2\frac{\Delta t^2}{\Delta x^2} c_2^2\right)\right| \leq 2\frac{\Delta t^2}{\Delta x^2} c_2^2. \tag{5.24}$$

Eq. (5.22) and Eq. (5.24) yield the Courant condition for Medium 1 and 2, respectively. Because all eigenvalues of $\mathbf{A}$ are real, λ_A according to Eq. (5.22) will lie in the

[1] A tridiagonal matrix with either all its off-diagonal elements positive or all its off-diagonal elements negative is diagonizable and has only real eigenvalues [62].

range

$$2 - 4\frac{\Delta t^2}{\Delta x^2}c_1^2 \leq \lambda_{\mathrm{A}} \leq 2. \tag{5.25}$$

Because $|\lambda_{\mathrm{A}}| \leq 2$ is to be satisfied, it follows that

$$\frac{\Delta t^2}{\Delta x^2}c_1^2 \leq 1, \tag{5.26}$$

which is the Courant condition for Medium 1. Similarly, from Eq. (5.24), the Courant condition for Medium 2 is obtained as

$$\frac{\Delta t^2}{\Delta x^2}c_2^2 \leq 1. \tag{5.27}$$

Eq. (5.23) gives a bound for the stability condition at the node between Medium 1 and Medium 2. For the boundary node, the eigenvalues lie in the range

$$2 - 2\frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1 + 2\mu_1}{\rho_2}\right) \leq \lambda_{\mathrm{A}} \leq 2. \tag{5.28}$$

This yields the stability criterion for the boundary node:

$$\frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1 + 2\mu_1}{\rho_2}\right) \leq 2, \tag{5.29}$$

which is the same result as obtained earlier with the von-Neumann method.

The most restrictive of the three conditions (Eq. (5.26), (5.27) and (5.29)) poses the stability condition for the entire system. Note that the overall stability condition is a *sufficient* condition, not a necessary one and, consequently, represents a bound for stability.

Eq. (5.29) can be written as

$$\frac{\Delta t^2}{\Delta x^2}c_2^2\left(1 + \frac{\lambda_1 + 2\mu_1}{\lambda_2 + 2\mu_2}\right) = \frac{\Delta t^2}{\Delta x^2}c_1^2\frac{\rho_1}{\rho_2}\left(1 + \frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1}\right) \leq 2. \tag{5.30}$$

For $c_1 > c_2$, the stability condition at the boundary will be most restrictive, if

$$\frac{\rho_1}{\rho_2}\left(1 + \frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1}\right) > 2. \tag{5.31}$$

On the other hand, for $c_2 > c_1$, the boundary stability criterion is most restrictive, if

$$\left(1 + \frac{\lambda_1 + 2\mu_1}{\lambda_2 + 2\mu_2}\right) > 2. \tag{5.32}$$

Expressed differently, the material interface does not impose an additional constraint on the stability of the finite-difference scheme, if

$$\sqrt{\frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1}} \geq 1 \qquad \text{or} \tag{5.33}$$

$$\sqrt{\frac{\lambda_2 + 2\mu_2}{\lambda_1 + 2\mu_1}} \leq \sqrt{2\frac{\rho_2}{\rho_1} - 1}. \tag{5.34}$$

This stability bound is plotted in Fig. 5.2 as a function of $\sqrt{(\lambda_2 + 2\mu_2)/(\lambda_1 + 2\mu_1)}$ and $\sqrt{\rho_2/\rho_1}$. The radial lines in Fig. 5.2 correspond to lines of constant velocity ratios. In the outer region, indicated by *Courant Region* in Fig. 5.2, the Courant condition is a sufficient condition for the stability of the finite-difference algorithm. In the region called *Boundary Region*, the stability condition at the boundary is more restrictive than the Courant condition and poses the decisive condition for stability of the finite-difference algorithm.

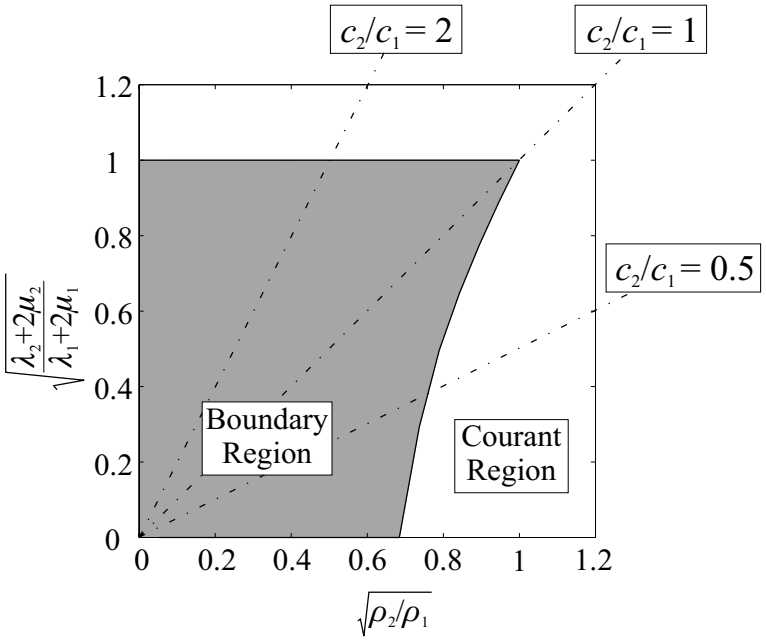


Figure 5.2: Stability bounds due to the presence of a material interface.

Note that, if the boundary in Fig. 5.1 is placed in between the velocity at node i and the stress at node $i + 0.5$, the bounds on stability will be exactly the same as

depicted in Fig. 5.2, but the ratios on the axes are inverted: $\sqrt{\rho_1/\rho_2}$ on the horizontal axis and $\sqrt{(\lambda_1 + 2\mu_1)/(\lambda_2 + 2\mu_2)}$ on the vertical axis. Thus, the stability condition at the boundary node now becomes the restrictive condition for stability of the finite-difference algorithm, if the density and the stiffness in Medium 1 are *smaller* than in Medium 2.

The analysis of a 1-D shear wave incident onto a material interface is very similar and the results are analogous to the results obtained for the pressure wave. In fact, by substituting μ_2/μ_1 for $(\lambda_2 + 2\mu_2)/(\lambda_1 + 2\mu_1)$ in Eq. (5.33) and Eq. (5.34), the bounds for the shear wave case are obtained.

The above description is not very intuitive but rather describes the derivation of a mathematical bound on the stability of the finite-difference algorithm at a boundary. To obtain a more intuitive picture of the actual reason why instabilities occur at the boundary, one can look at the boundary row of **A** in Eq. (5.1.1). It is seen that in the boundary row a mixed term appears which is composed of the stiffness in Medium 1 and the density in Medium 2: $(\lambda_1 + 2\mu_1)/\rho_2$. This mixed term has the units of a squared wave speed. Thus, by comparing the value of the mixed term to the wave speeds in the general media, a reason for the instabilities can be suggested. If the longitudinal stiffness in Medium 1 is large or the density in Medium 2 is small (which, for example, is true at the interface between a solid and air), the mixed term might be quite large and, thus, the square root of its value might exceed the wave speeds in both Medium 1 and Medium 2. In that case, the stability condition at the boundary can become more restrictive than the Courant condition in the general medium, and neither the Courant condition for Medium 1 nor the one for Medium 2 is a sufficient condition for stability.

Estimated Stability Condition

An estimate for the stability condition is achieved by modifying the matrix which describes the finite-difference algorithm. If a matrix is found for which the eigenvalues

can be calculated explicitly and which closely approximates the actual matrix, an estimate for the stability condition of the finite-difference algorithm is obtained.

In Eq. (5.1.1), the upper $i-1$ rows describe Medium 1, whereas the lower $N-i$ rows describe Medium 2. The intermediate row represents the node at the interface between the two media. If the entire tridiagonal matrix $\mathbf{A}$ is filled with the elements of the intermediate row, the new matrix will describe a non-physical problem. However, it approximates the original matrix, and its eigenvalues can be determined explicitly.

The matrix $\mathbf{A}$ is redefined as

$$\mathbf{A}' = \begin{bmatrix} \ddots & \ddots & & & \\ \ddots & 2-\frac{\Delta t^2}{\Delta x^2}\left(c_2^2+\frac{\lambda_1+2\mu_1}{\rho_2}\right) & \frac{\Delta t^2}{\Delta x^2}c_2^2 & & \\ & \frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1+2\mu_1}{\rho_2} & 2-\frac{\Delta t^2}{\Delta x^2}\left(c_2^2+\frac{\lambda_1+2\mu_1}{\rho_2}\right) & \frac{\Delta t^2}{\Delta x^2}c_2^2 & \\ & & \frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1+2\mu_1}{\rho_2} & 2-\frac{\Delta t^2}{\Delta x^2}\left(c_2^2+\frac{\lambda_1+2\mu_1}{\rho_2}\right) & \ddots \\ & & & \ddots & \ddots \end{bmatrix}. \tag{5.35}$$

Note that $\mathbf{A}'$ is, like $\mathbf{A}$, of size $N \times N$. The other submatrices of $\mathbf{B}$ remain the same and, as before, the magnitude of the eigenvalues of $\mathbf{A}'$ must be smaller than or equal to 2. According to [62], the eigenvalues of $\mathbf{A}'$ are defined as

$$\lambda_{\mathrm{A}'} = 2 - \frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1+2\mu_1}{\rho_2}\right) + 2\frac{\Delta t^2}{\Delta x^2}c_2\sqrt{\frac{\lambda_1+2\mu_1}{\rho_2}}\cos\left(\frac{p\pi}{N+1}\right), \tag{5.36}$$

where $p = 1 \ldots N$ indicates the p-th eigenvalue. Since the cosine can assume values in between -1 and $+1$, the eigenvalues lie in the range

$$\begin{aligned} 2 - \frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1+2\mu_1}{\rho_2} + 2c_2\sqrt{\frac{\lambda_1+2\mu_1}{\rho_2}}\right) &\leq \lambda_{\mathrm{A}'} \\ &\leq 2 - \frac{\Delta t^2}{\Delta x^2}\left(c_2^2 + \frac{\lambda_1+2\mu_1}{\rho_2} - 2c_2\sqrt{\frac{\lambda_1+2\mu_1}{\rho_2}}\right). \end{aligned} \tag{5.37}$$

This can also be written as

$$2 - \frac{\Delta t^2}{\Delta x^2}\left(c_2 + \sqrt{\frac{\lambda_1+2\mu_1}{\rho_2}}\right)^2 \leq \lambda_{\mathrm{A}'} \leq 2 - \frac{\Delta t^2}{\Delta x^2}\left(c_2 - \sqrt{\frac{\lambda_1+2\mu_1}{\rho_2}}\right)^2. \tag{5.38}$$

The magnitude of $\lambda_{A'}$ must be smaller or equal to 2. The right hand side of Eq. (5.38) is obviously always smaller than 2. For the left hand-side to be larger than -2,

$$\frac{\Delta t^2}{\Delta x^2}\left(c_2+\sqrt{\frac{\lambda_1+2\mu_1}{\rho_2}}\right)^2 \leq 4, \tag{5.39}$$

which is the condition for stability for the modified matrix. Note that Eq. (5.39) represents an *estimate*, not a bound, for the stability of the actual matrix. Figure 5.3 shows that Eq. (5.39) yields a less restrictive stability condition (dark gray area) than the derived bound for the actual matrix (light gray area).

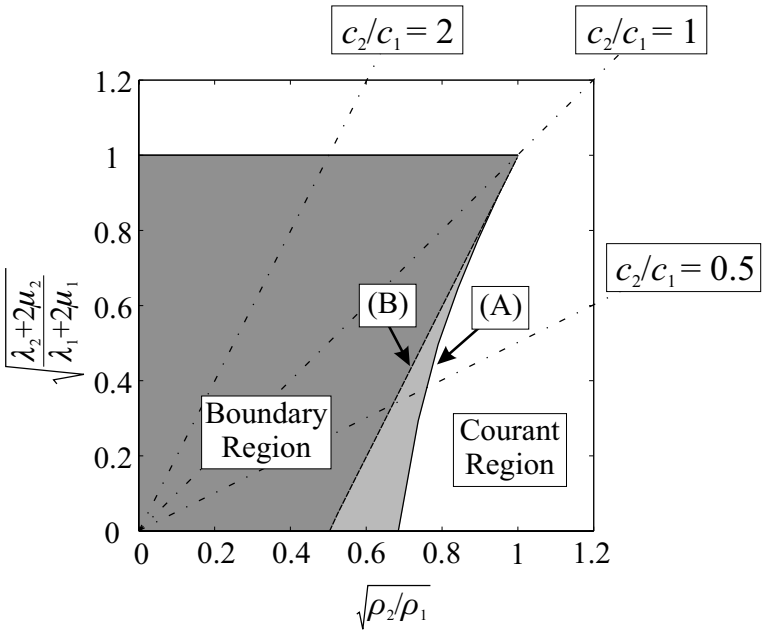


Figure 5.3: Stability bound for the actual matrix (A) and estimate from the eigenvalues of the modified matrix (B).

Averaging the Material Parameters on the Interface

To ensure that the Courant condition is a sufficient condition for stability, the material parameters are averaged for the field components on the boundary. The averaging procedure is described in Sec. 5.3. For the 1-D case discussed above, the velocity component V_x is placed on the boundary (see Fig. 5.1) and the material density is averaged according to Eq. (5.51). In the matrix equation (Eq. (5.1.1)), only the

boundary row of **A** will change, and **A** for the averaged case is obtained by replacing ρ_2 by the averaged density, $(\rho_1 + \rho_2)/2$, throughout the boundary row.

By applying Gershgorin's Circle Theorem to the interface row of the averaged matrix, the eigenvalue associated with this row is determined to lie in the range

$$2 - 2\frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1 + 2\mu_1 + \lambda_2 + 2\mu_2}{(\rho_1 + \rho_2)/2} \leq \lambda_{\mathrm{A}_{avg}} \leq 2. \tag{5.40}$$

As shown previously, the magnitude of the eigenvalues of **A** must be smaller than or equal to 2, and the stability condition is obtained as

$$\frac{\Delta t^2}{\Delta x^2}\frac{\lambda_1 + 2\mu_1 + \lambda_2 + 2\mu_2}{\rho_1 + \rho_2} \leq 1. \tag{5.41}$$

This can be rewritten by introducing the wave speeds:

$$\frac{\Delta t^2}{\Delta x^2}\frac{c_1^2\rho_1 + c_2^2\rho_2}{\rho_1 + \rho_2} = \frac{\Delta t^2}{\Delta x^2}c_1^2\frac{\rho_1 + \rho_2(c_2/c_1)^2}{\rho_1 + \rho_2} = \frac{\Delta t^2}{\Delta x^2}c_2^2\frac{\rho_1(c_1/c_2)^2 + \rho_2}{\rho_1 + \rho_2} \leq 1. \tag{5.42}$$

When having a closer look at this equation, it becomes clear that the stability condition for the averaged boundary is always less restrictive than the Courant condition in the general medium, and the finite-difference scheme will be stable as long as the Courant condition is satisfied. If $c_1 > c_2$, the stability condition for the entire system is

$$\frac{\Delta t^2}{\Delta x^2}c_1^2 \leq 1, \tag{5.43}$$

which is the Courant condition for Medium 1. From the second term in Eq. (5.42), it can be seen that the stability condition for the averaged boundary layer is always less restrictive than the Courant condition for Medium 1, because c_2 is smaller than c_1 and, consequently, the numerator of the second term will be smaller than the denominator. Thus, the boundary does not pose any further constraints on the stability of the system. Similarly, if $c_2 > c_1$, the stability condition for the system is

$$\frac{\Delta t^2}{\Delta x^2}c_2^2 \leq 1, \tag{5.44}$$

which is the Courant condition for Medium 2. From the third term in Eq. (5.42), it is clear that again the stability criterion for the averaged boundary layer is always satisfied, if the Courant condition for Medium 2 is fulfilled.

5.2 Stability Analysis: Numerical Results

A numerical study has been performed to investigate the stability behavior of the finite-difference scheme. Results are presented for the 1-D and 2-D finite-difference algorithm. For the 1-D case, a grid with 10 nodes is considered, whereas for the 2-D case a grid with 6×7 nodes for both V_x and V_z is chosen. The grid dimensions are far smaller than for any practical application, but a small grid will display a stability behavior very similar to a large grid, while the computation time for the stability analysis is greatly reduced. For the 1-D case, the matrix $\mathbf{B}$ describing the finite-difference algorithm (see Eq. (5.15)) has dimensions 20×20. The finite-difference scheme for the 2-D case can also be written as a matrix equation, similar to Eq. (5.15). For a 2-D grid with 6×7 nodes for each velocity component, the matrix has dimensions 84×84 ($2 \cdot 7 \cdot 6 \times 2 \cdot 7 \cdot 6$). The matrix for the 2-D case has a more complex structure, because it relates $\mathbf{V_x}$ and $\mathbf{V_z}$ at time k to $\mathbf{V_x}$ and $\mathbf{V_z}$ at time $k+1$.

For the stability analysis, the matrix $\mathbf{A}$ is generated and its eigenvalues are determined. A root finder seeks the value of the time step, Δt, for which the magnitude of the largest eigenvalue just exceeds 2 and, thus, determines the maximum size of the time step that must not be exceeded to ensure that the finite-difference algorithm is stable. The time step is determined as a function of the material density and Lamé's constants, and the results for the averaged and non-averaged case are compared.

5.2.1 1-D Case

The stability behavior of the 1-D finite-difference grid for a longitudinal wave (see Fig. 5.1) is analyzed. In Fig. 5.4, the required time step, Δt, as determined from the eigenvalues is plotted as a function of the square root of the ratios of material density and the longitudinal stiffness in the two media. The results are normalized to the value of the time step calculated from the Courant condition in the general medium (Eq. (5.1)). Due to the normalization, a value of 1 corresponds to the value predicted by the Courant condition. Values smaller than 1 represent a condition for stability

that is more restrictive than the Courant condition.

When the material parameters are not averaged on the material interface, the interface imposes a constraint on the stability condition. If the density ratio $\sqrt{\rho_2/\rho_1}$ and the stiffness ratio $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)}$ fall below some critical value (≈ 1), the time step must be chosen smaller than dictated by the Courant condition in order to achieve a stable behavior. When $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)}$ approach zero, the necessary stability condition also approaches zero, and, thus, the required time step becomes infinitely small. For $\sqrt{\rho_2/\rho_1} > 1$ or $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)} > 1$, the stability condition at the boundary is less restrictive than the Courant condition.

Fig. 5.5 shows the stability condition that arises when the material parameters are averaged on the interface. For the averaged case, the normalized required time step always has a value larger than 1, and, consequently, the boundary does not pose an additional stability constraint. Thus, the Courant condition is a sufficient condition for stability independent of the ratios for the material density and Lamé's constants.

Note that, if the material interface in Fig. 5.1 was shifted by half a step, the ratios on the axes of Fig. 5.4 would be inverted, and the stability condition at the boundary node would become more restrictive than the Courant condition for $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)}$ *larger* than 1. Regardless, when the material properties are averaged on the boundary, the Courant condition is a sufficient condition for stability.

To analyze the stability behavior in a more quantitative manner, the stability condition is plotted along the lines of constant velocity ratios (the radial lines in the density-/stiffness-plane in Fig. 5.4). In Fig. 5.6, the stability condition is shown for four ratios of pressure wave speeds in the two media. The plots include the normalized time step for the non-averaged case, for the averaged case, the lower bound for the stability condition at the boundary according to Eq. (5.29), an upper bound which has been found empirically to be the lower bound multiplied by $\sqrt{2}$ and the estimate according to Eq. (5.39).

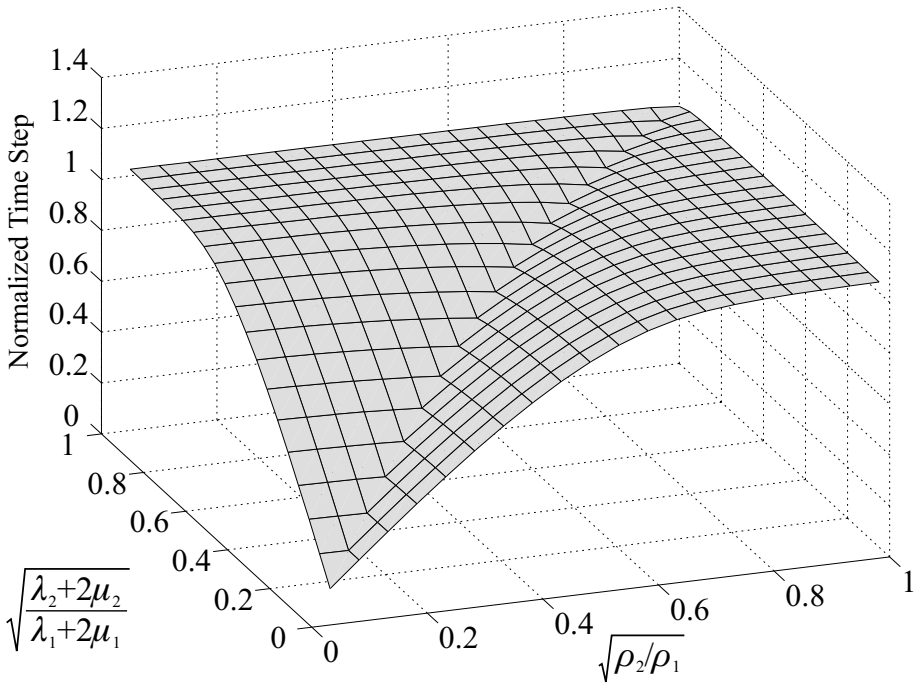


Figure 5.4: Normalized stability condition; 1-D case. The material properties are not averaged on the boundary.

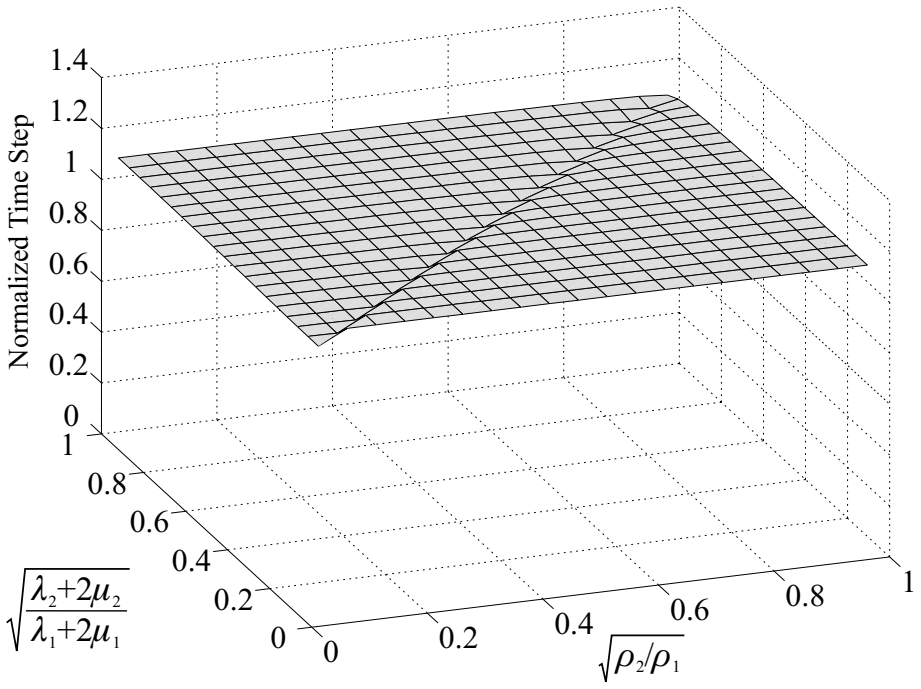


Figure 5.5: Normalized stability condition; 1-D case. The material properties are averaged on the boundary.

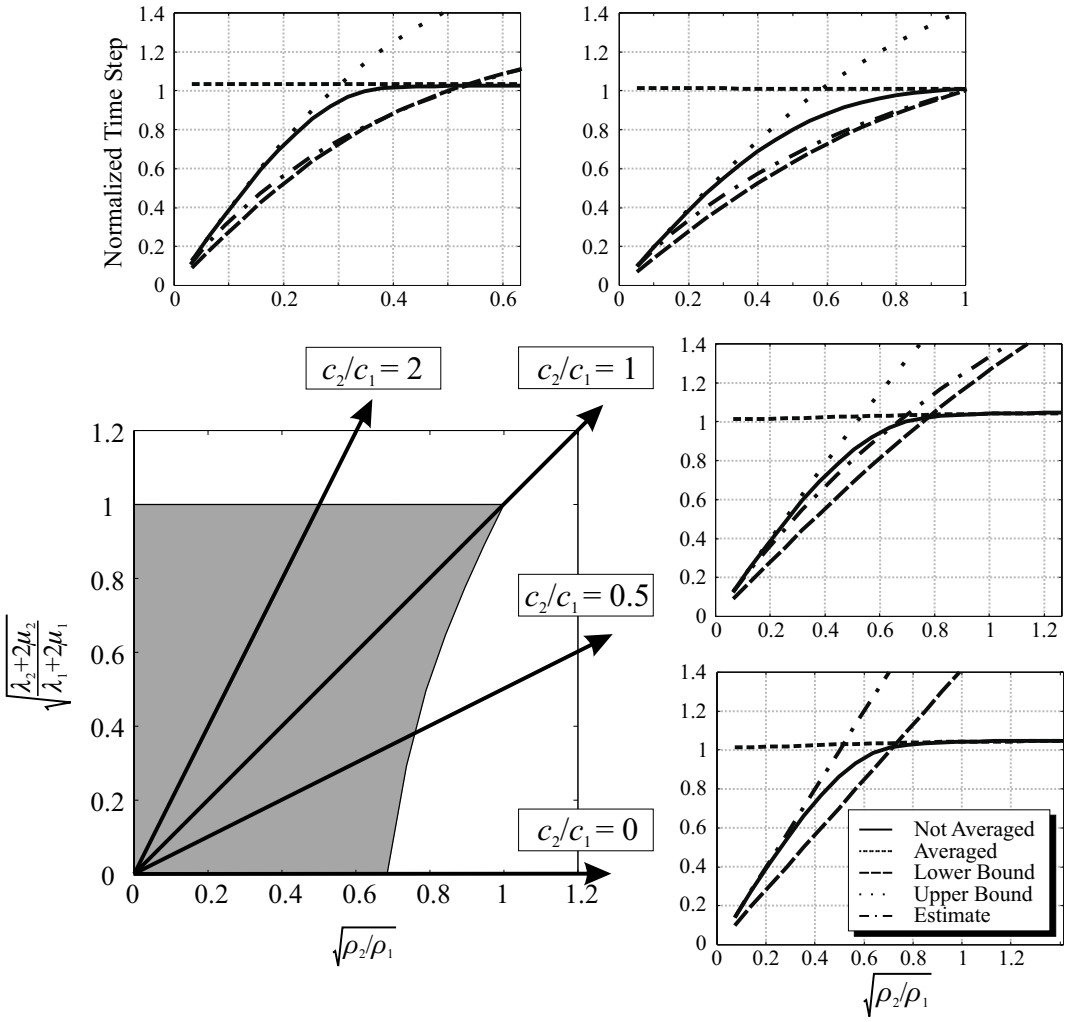


Figure 5.6: Normalized stability condition; 1-D case; for four constant velocity ratios.

The stability condition for the non-averaged matrix exhibits two regions. In the first region, the Courant condition for the general medium is the decisive stability condition (see Fig. 5.2, *Courant Region*). In the second region, the stability condition for the boundary will be the restrictive one, and, consequently, the finite-difference scheme would be unstable, if the time step satisfied only the Courant condition (Fig. 5.2, *Boundary Region*). Note that the bounds and the estimate in Fig. 5.6 apply only to the second region.

Fig. 5.6 shows that the actual stability condition that is determined numerically always lies in between the lower and the upper bound, as does the estimated stability condition. For $c_2/c_1 = 0$, the estimate coincides with the upper bound. As c_2/c_1 increases, the estimate approaches the lower bound.

From Fig. 5.6, the critical value of the material density ratio ρ_2/ρ_1, below which the Courant condition does not represent a sufficient stability condition, can be determined as a function of the wave speed ratio. For $c_2/c_1 = 0$, the critical value of the density ratio is approximately $K_{\text{crit}} \approx 0.7$. This value is predicted also by the lower bound. For $\sqrt{\rho_2/\rho_1} < 0.7$, the time step must be chosen smaller than required by the Courant condition, or the material properties must be averaged on the boundary. Note that the case $c_2/c_1 = 0$ describes a non-physical situation. For $c_2/c_1 = 0$, the pressure wave speed is zero in the second medium. Only in vacuum the pressure wave speed becomes zero, and thus $c_2/c_1 = 0$ only for $\rho_2/\rho_1 = 0$. However, the case of $c_2/c_1 = 0$ represents the limiting case for two materials with greatly different pressure wave speeds and is therefore included here. For $c_2/c_1 = 0.5$, the critical value is shifted slightly up, and for $c_2/c_1 = 1$, $K_{\text{crit}} \approx 1$. Thus, for $c_2/c_1 = 1$, the material properties have to be averaged whenever the material density in the second medium is smaller than the density in the first medium ($\rho_2 < \rho_1$). Overall, for $c_2/c_1 < 1$, the critical value is predicted accurately by the theoretically determined lower bound. For $c_2/c_1 = 2$, i. e., the velocity in the second medium is larger than in the first medium, the critical value predicted by the lower bound ($K_{\text{crit}} = 0.5$) exceeds the actual value ($K_{\text{crit}} \approx 0.35$). However, the actual stability condition still lies in between lower and

upper bound.

Note that the Courant condition can be slightly exceeded in the first region (*Courant Region*), indicated by a value for the normalized time step larger than 1. This is due to the limited size of the problem dimensions used for the stability analysis. An explanation for this can be found in Sec. 5.4.

5.2.2 2-D Case

A similar analysis has been performed for the 2-D finite-difference algorithm. Two different placements of the material interface are considered. First, the boundary is placed such that a plane pressure wave normally incident onto the boundary sees the same node arrangement as the 1-D pressure wave in the 1-D grid according to Fig. 5.1. This arrangement will be referred to as *Boundary A* and is shown schematically in Fig. 5.7 (a). Second, the boundary will be shifted by half a space step in both the x- and the z-direction . This case will be referred to as *Boundary B*, Fig. 5.7 (b).

For the 1-D analysis, the shear wave case and the pressure wave case were treated separately. Thus, by varying the density ratio and the ratio of the longitudinal stiffness, all possible cases for a longitudinal wave in 1-D normally incident onto a material interface have been considered. However, for the 2-D analysis, the shear wave case and the pressure wave case cannot be treated separately, because only one grid exists that supports both shear and pressure waves. Thus, not only the variation of the longitudinal stiffness $\lambda + 2\mu$ is of interest, but also the size of λ relative to μ, i. e., the speed of the pressure wave relative to the shear wave. Hence, more degrees of freedom arise, and the results cannot be summarized easily. The 2-D analysis is performed for different absolute shear wave speed values in the two media, and the stability condition is, as before, determined as a function of the square root of the ratios of density and longitudinal stiffness. The results presented by no means cover all possible scenarios. Rather than striving for completeness, it will be shown that the Courant condition is a sufficient condition for stability whenever the material

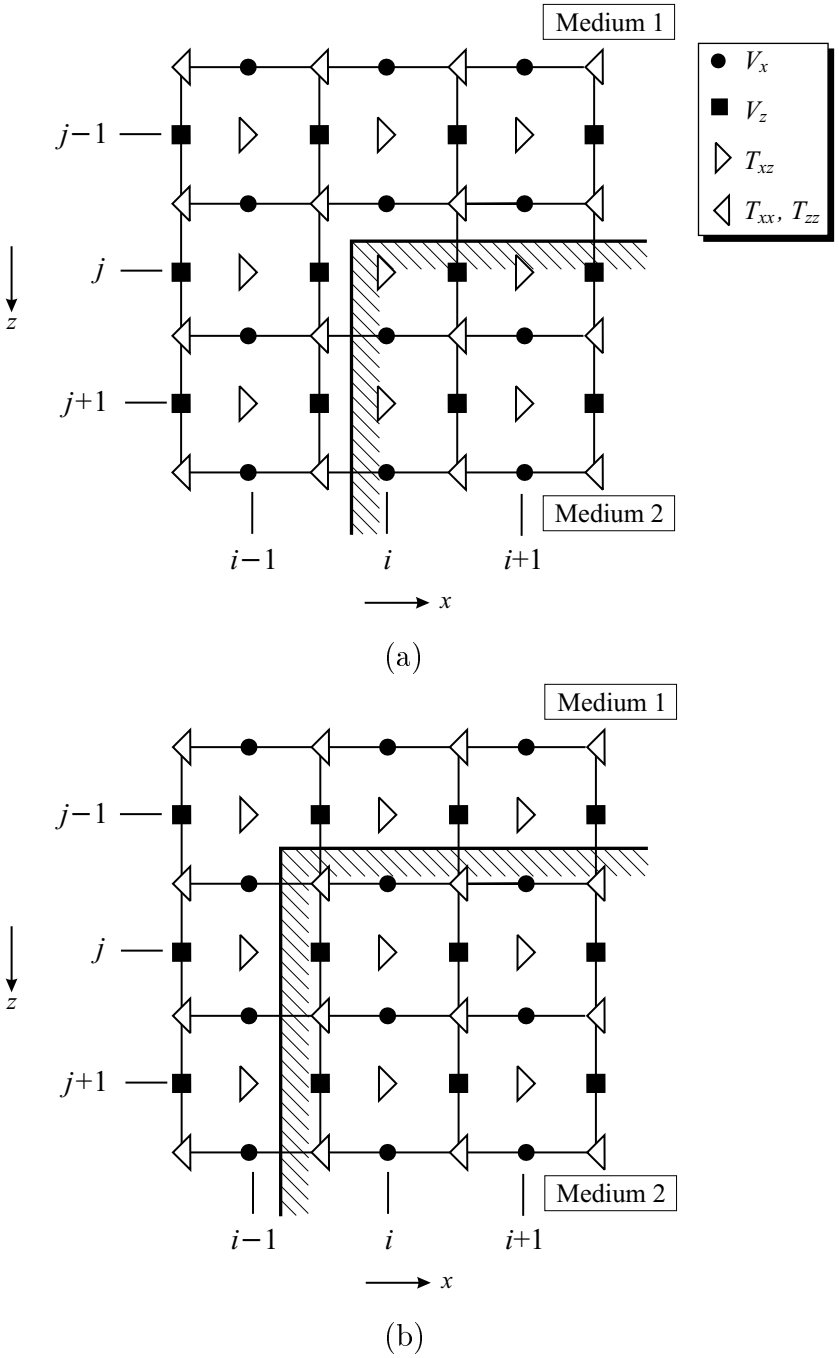


Figure 5.7: Boundary arrangements used for the 2-D analysis. (a) *Boundary A*, (b) *Boundary B*.

properties are averaged on the boundary.

In Fig. 5.8, the required time step for *Boundary A* is shown as a function of the square root of the ratios of density and longitudinal stiffness in the two media. The pressure wave speed in the first medium is $c_{P,1} = 250$ m/s, and the shear wave speed is $c_{S,1} = 87$ m/s. The material density is $\rho_1 = 1400$ kg/m^3. The shear wave speed in the second medium is $c_{S,2} = 0$ m/s (i. e., no shear waves can propagate, as for example in air), and by changing the density ratio and the longitudinal stiffness the pressure wave speed in the second medium is varied. The time step is normalized to the time step for the 2-D Courant condition for the general medium. As observed in the 1-D analysis, if the material properties are not averaged, the stability condition for the boundary becomes more restrictive than the Courant condition in the general medium. If the material properties are averaged on the boundary, the Courant condition is a sufficient condition for stability, and no additional stability constraints arise due to the presence of the boundary. The normalized stability condition then is flat and always larger than one.

The same behavior is obtained for other values of the shear wave speeds in the two media. This indicates that for this placement of the boundary and for $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2 + 2\mu_2)/(\lambda_1 + 2\mu_1)}$ smaller than one, the decisive quantity for stability is the pressure modulus, whereas the shear modulus does not have an impact. Unlike the 1-D case, the 2-D finite-difference scheme also exhibits instabilities, when $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2 + 2\mu_2)/(\lambda_1 + 2\mu_1)}$ are *larger* than one. The shape of the resulting stability curves now depends on the shear modulus. The nature of these instabilities can be explained by shifting the position of the boundary.

The boundary is shifted by half a space step in both the x- and z-direction (Fig. 5.7 (b), *Boundary B*). The material properties are assumed to be the same as earlier. Results for the non-averaged case are shown in Fig. 5.9. It is seen that the stability condition drops below 1 for small values of $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2 + 2\mu_2)/(\lambda_1 + 2\mu_1)}$. This can be explained by considering the stability behavior of waves normally incident onto the boundary and, thus, by drawing conclusions from the 1-D results

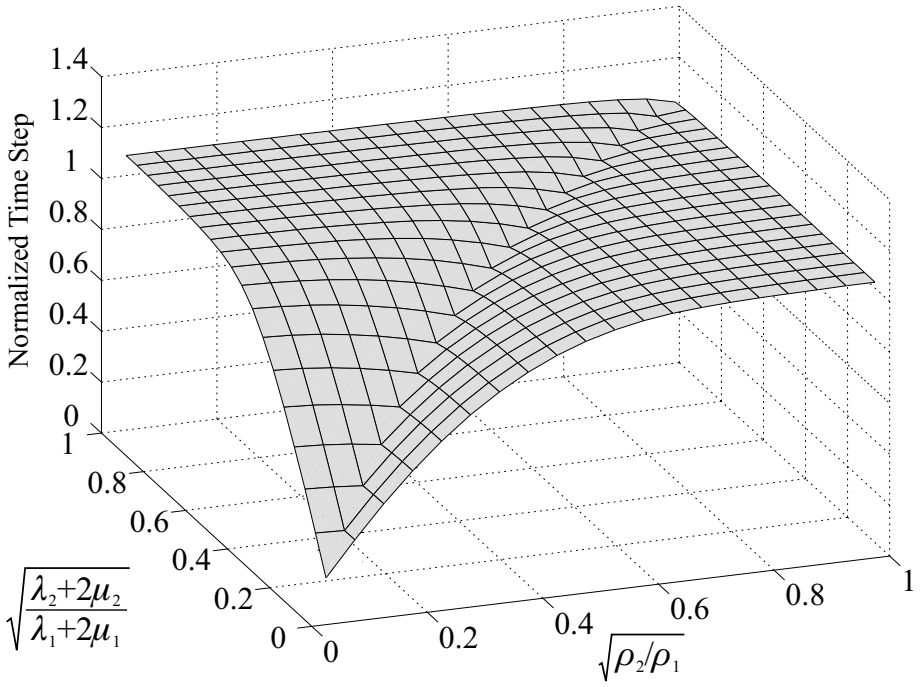


Figure 5.8: Normalized stability condition; 2-D case. The material parameters are not averaged on the boundary. The boundary is placed as in Fig. 5.7 (a).

we obtained earlier. One can argue that in the 2-D grid as in *Boundary A* a normally incident longitudinal wave will experience the same instabilities as described for the 1-D case, and the boundary will be the restrictive condition for stability, if $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)}$ are *smaller* than 1. However, for an incident shear wave that propagates in the same grid, instabilities will occur for $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)}$ *larger* than 1. For *Boundary B*, the situation is just reversed: now, for an incident shear wave the stability condition at the boundary will become more restrictive for $\sqrt{\rho_2/\rho_1}$ and $\sqrt{(\lambda_2+2\mu_2)/(\lambda_1+2\mu_1)}$ *smaller* than 1. Due to the shear modulus being smaller than the pressure modulus, the stability constraints for the boundary due to the shear modulus will occur for smaller density and stiffness ratios. This effect is visible in Fig. 5.9.

The 2-D results are summarized in Fig. 5.10. Here, the stability condition is plotted along the lines of three constant velocity ratios. The time step necessary for stability is graphed for the 1-D lower bound, the 1-D non-averaged case, the 2-D non-averaged case for *Boundary A*, the 2-D non-averaged case for *Boundary B*, and for the 2-D averaged case. It is evident, that in all cases the Courant condition is *not* sufficient for stability, unless the material properties are averaged on the boundary. However, if the material properties are averaged, the Courant condition poses a sufficient condition for the stability of the finite-difference algorithm, independent of the material properties. Note that the 1-D results are normalized to the 1-D Courant condition, whereas the 2-D results are normalized to the 2-D Courant condition.

5.3 Averaging

To ensure the stability of the finite-difference scheme at a material interface, the material parameters are averaged for the field components on the interface. In this section, the averaging procedure is derived for a two-dimensional grid. The procedure can be easily extended to three dimensions.

Figure 5.11 shows a portion of a 2-D finite-difference grid at the interface be-

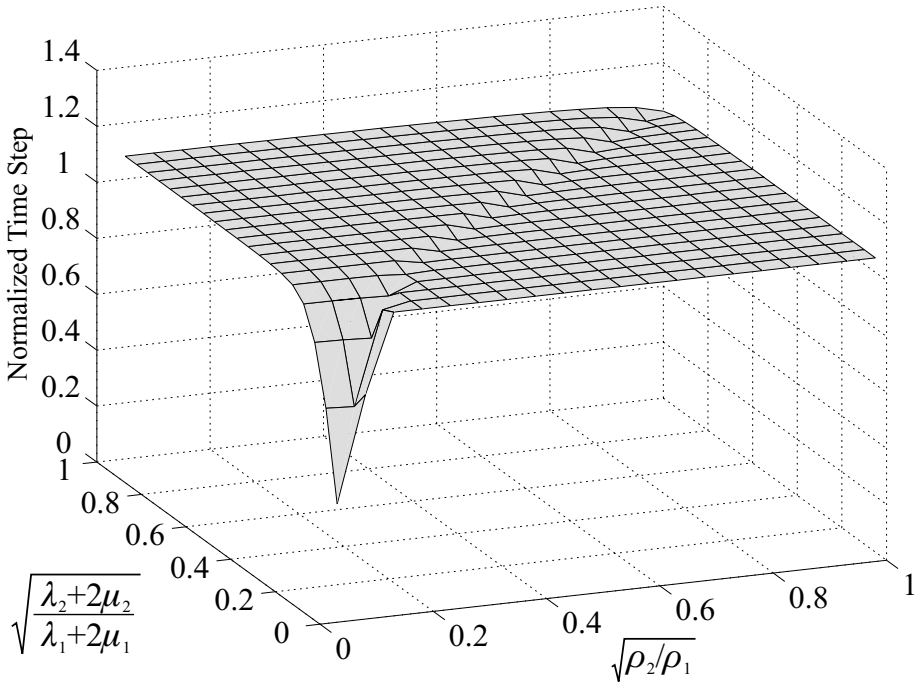


Figure 5.9: Normalized stability condition; 2-D case; boundary placed according to Fig. 5.7 (b).

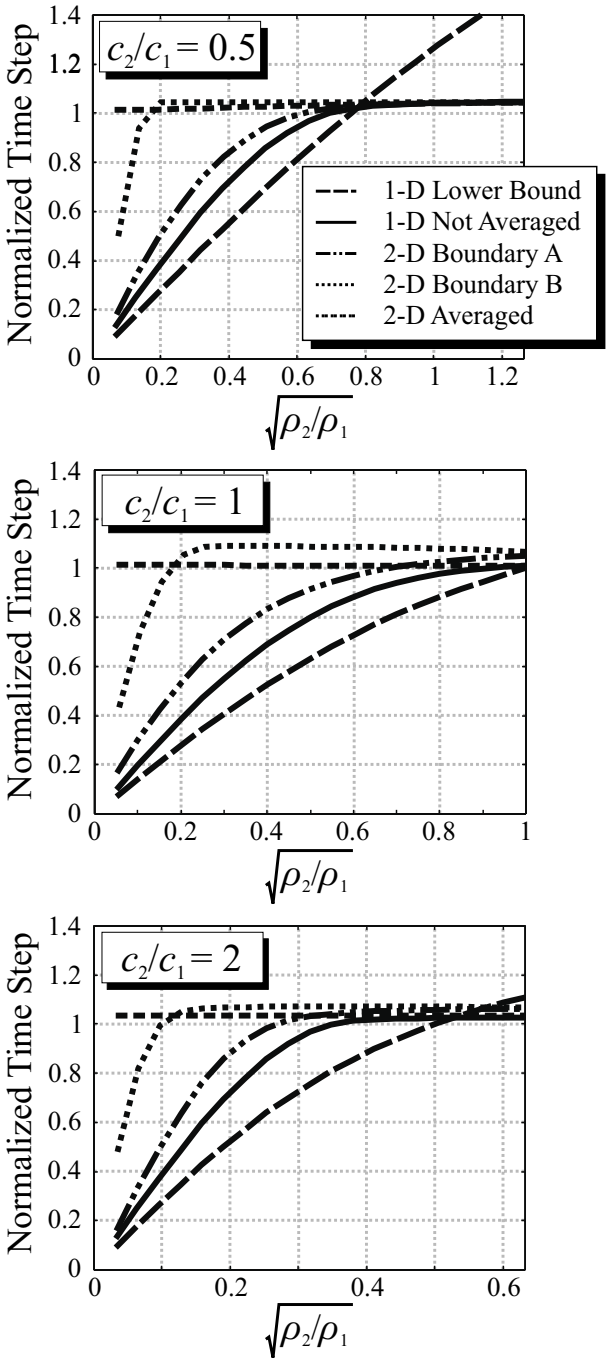


Figure 5.10: Comparison of 1-D and 2-D stability behavior; four velocity ratios.

tween two different media. The field components in x and z are labeled by i and j, respectively. The boundary between the two media is placed such that the normal particle velocity components are located on the interface. Thus, V_x is placed on a vertical boundary, whereas V_z is placed on a horizontal boundary. The only stress component, in that case, that lies on the boundary is the shear stress T_{xz}. First the procedure for averaging the material density in the equation of motion is derived, then the averaging of the shear stiffness μ is explained.

5.3.1 Material Density

The procedure for averaging the material density on the boundary is derived from the equation of motion. The derivation of the averaging is analogous for both the horizontal and vertical boundary and, thus, the averaging procedure is presented only for the horizontal particle velocity on a vertical interface. The equation of motion,

$$\nabla \cdot \boldsymbol{\tau} = \rho \frac{\partial \mathbf{v}}{\partial t}, \tag{5.45}$$

is expanded for the 2-D Shear-Vertical (S-V) case in the x-z plane:

$$\frac{\partial}{\partial x}\tau_{xx} + \frac{\partial}{\partial z}\tau_{xz} = \rho \frac{\partial v_x}{\partial t} \tag{5.46}$$

$$\frac{\partial}{\partial x}\tau_{xz} + \frac{\partial}{\partial z}\tau_{zz} = \rho \frac{\partial v_z}{\partial t}. \tag{5.47}$$

The only non-zero velocity components are v_x and v_z. Considering the horizontal particle velocity v_x, the procedure for averaging the material density on the vertical boundary is derived by integrating Eq. (5.46) over a box enclosing a portion of the boundary as depicted in Fig. 5.12:

$$\iint_{Box} \left[\frac{\partial}{\partial x}\tau_{xx} + \frac{\partial}{\partial z}\tau_{xz} \right] dx\, dz = \oint \left(\tau_{xx}\hat{\mathbf{x}} + \tau_{xz}\hat{\mathbf{z}} \right) \hat{\mathbf{n}}\, ds = \iint_{Box} \rho \frac{\partial v_x}{\partial t}\, dx\, dz, \tag{5.48}$$

where the density ρ is a function of position, and Gauss' integral theorem has been applied to convert the surface integral on the left hand side into a contour integral;

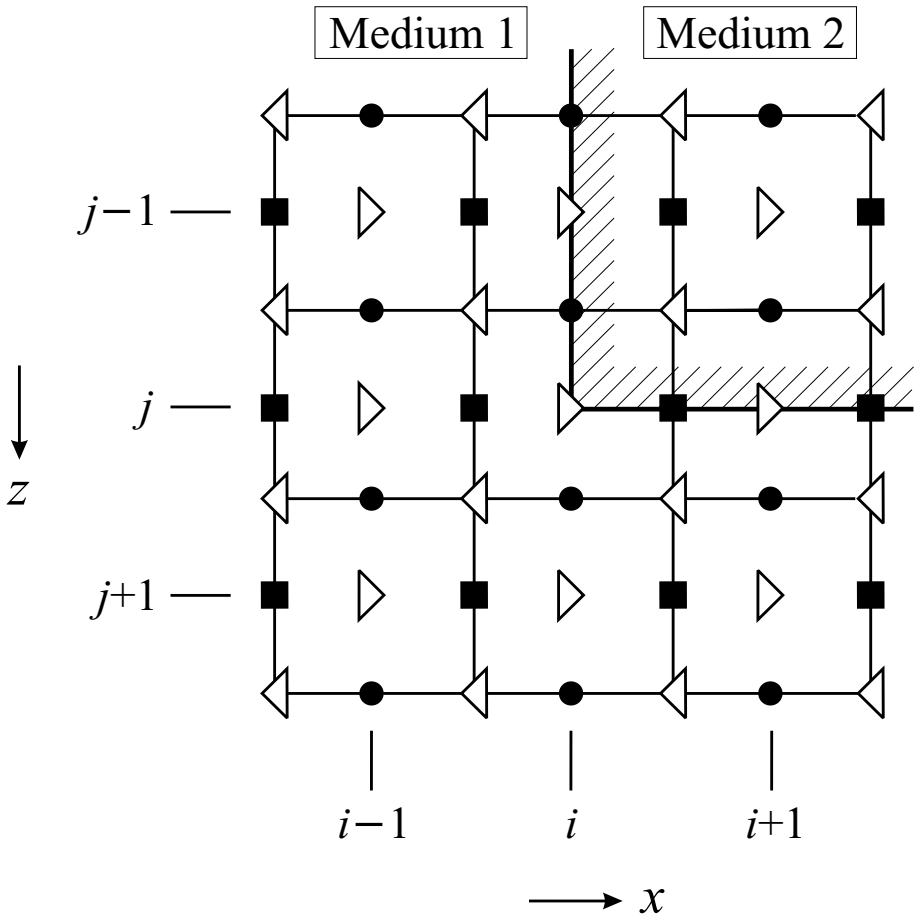


Figure 5.11: Boundary between two media in a 2-D finite-difference grid; the normal particle velocities and the shear stress are located on the interface.

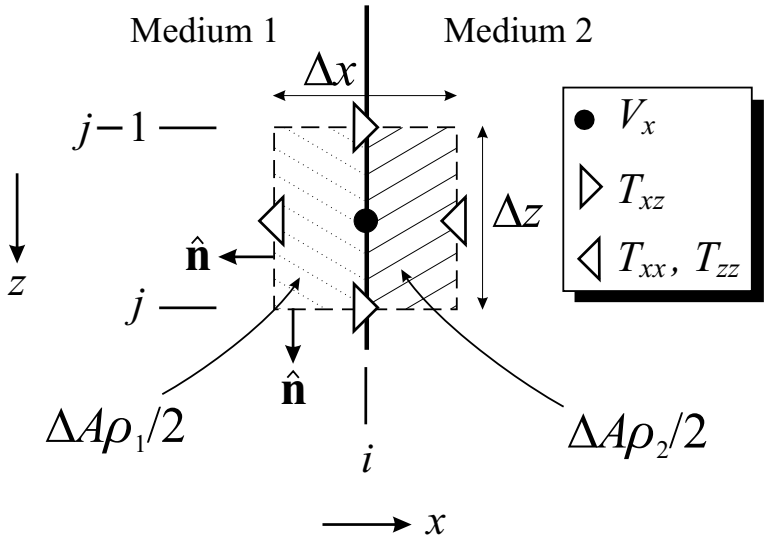


Figure 5.12: Averaging of the material density ρ; the normal stress components are integrated along the contour.

$\hat{\mathbf{n}}$ is the normal vector to the contour, and $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ are normal vectors in the x- and z-direction, respectively.

Using the discrete finite-difference formulation, the integration is carried out over a box of area $\Delta A = \Delta x \cdot \Delta z$, equivalent to the size of one cell of the finite-difference grid (see Fig. 5.12). Assuming the boundary between Medium 1 and Medium 2 divides the area of integration in half, the integral comes out to be

$$\frac{\rho_1 + \rho_2}{2} \Delta A \cdot \frac{V_x^{k+0.5}|_{i,j-0.5} - V_x^{k-0.5}|_{i,j-0.5}}{\Delta t} =$$

$$\left(T_{xx}^k|_{i+0.5,j-0.5} - T_{xx}^k|_{i-0.5,j-0.5} \right) \cdot \Delta z +$$

$$+ \left(T_{xz}^k|_{i,j} - T_{xz}^k|_{i,j-1} \right) \cdot \Delta x. \tag{5.49}$$

After dividing by ΔA, the finite-difference approximation to Eq. (5.46) is obtained:

$$\frac{\rho_1 + \rho_2}{2} \cdot \frac{V_x^{k+0.5}|_{i,j-0.5} - V_x^{k-0.5}|_{i,j-0.5}}{\Delta t} =$$

$$\frac{T_{xx}^k|_{i+0.5,j-0.5} - T_{xx}^k|_{i-0.5,j-0.5}}{\Delta x} +$$

$$+ \frac{T_{xz}^k|_{i,j} - T_{xz}^k|_{i,j-1}}{\Delta z}. \tag{5.50}$$

Thus, for the velocity components on the boundary, the material densities of the two media are averaged arithmetically:

$$\rho_{avg} = (\rho_1 + \rho_2)/2. \tag{5.51}$$

The same result is obtained for the vertical particle velocity on a horizontal interface.

5.3.2 Lamé's Constants

The procedure for averaging Lamé's constants for the stress components on the boundary is derived by using the elastic constitutive relation together with the strain-velocity relation. The elastic constitutive relation combines the mechanic strain and the stress and is usually expressed as

$$\boldsymbol{\tau} = \mathbf{c} \cdot \mathbf{S}, \tag{5.52}$$

where the stress $\boldsymbol{\tau}$ and the strain $\mathbf{S}$ are 3×3 tensors and the stiffness matrix $\mathbf{c}$ is a four-dimensional tensor of size $3 \times 3 \times 3 \times 3$. For an isotropic and lossless medium, the entries of the stiffness matrix are completely described by two independent constants, Lamé's constants λ and μ.

In the 2-D finite-difference gird (Fig. 5.11), the shear stress T_{xz} is placed on the boundary and, consequently, Lamé's constants are averaged for only T_{xz}. By combining the strain-velocity relation and the constitutive relation, an equation relating v_x and v_z to τ_{xz} arises, from which, in the finite-difference formulation, the shear stress is determined:

$$\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} = \frac{1}{\mu} \frac{\partial \tau_{xz}}{\partial t}. \tag{5.53}$$

The averaging of μ is achieved by integrating over a box enclosing parts of the boundary, and Eq. (5.53) becomes

$$\iint_{Box} \left[\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right] dx\, dz = \iint_{Box} \frac{1}{\mu} \frac{\partial \tau_{xz}}{\partial t}\, dx\, dz. \tag{5.54}$$

Fig. 5.13 shows the integration contour around τ_{xz}. Applying Stoke's integral theorem, the left hand side of Eq. (5.54) can be written as

$$\iint_{Box} \left[\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right] dx\, dz = \iint_{Box} \nabla \times (v_x \cdot \hat{\mathbf{x}} - v_z \cdot \hat{\mathbf{z}})\, \vec{dA} = \oint (v_x \cdot \hat{\mathbf{x}} - v_z \cdot \hat{\mathbf{z}})\, \vec{ds}. \tag{5.55}$$

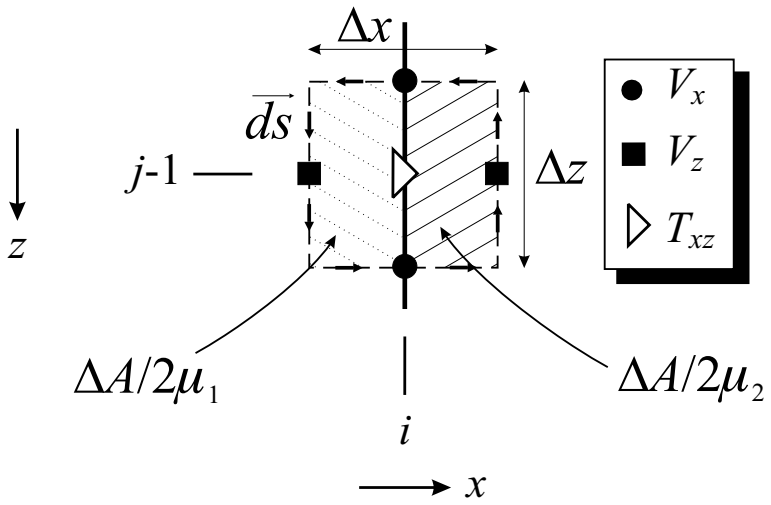


Figure 5.13: Averaging of the shear stiffness μ; the tangential velocity components are integrated along the contour.

Using the discrete finite-difference formulation and Eq. (5.55), Eq. (5.54) becomes

$$\left(V_x^{k+0.5}|_{i,j-0.5} - V_x^{k+0.5}|_{i,j-1.5}\right) \cdot \Delta x + \left(V_z^{k+0.5}|_{i+0.5,j-1} - V_z^{k+0.5}|_{i-0.5,j-1}\right) \cdot \Delta z$$

$$= \frac{\Delta A}{2} \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} \right) \frac{T_{xz}^{k+1}|_{i,j-1} - T_{xz}^{k}|_{i,j-1}}{\Delta t}. \tag{5.56}$$

Thus, the inverse of the shear stiffness μ is averaged on the boundary between two media as

$$\frac{1}{\mu_{avg}} = \frac{1}{2}\left(\frac{1}{\mu_1} + \frac{1}{\mu_2}\right). \tag{5.57}$$

The same is obtained for the field components on a horizontal boundary. For the shear stress on a corner of the boundary (i. e., $T_{xz}(i,j)$ in Fig. 5.11), the integration yields

$$\frac{1}{\mu_{avg}} = \frac{1}{4}\left(\frac{3}{\mu_1} + \frac{1}{\mu_2}\right). \tag{5.58}$$

5.4 The Courant Condition

By analyzing the eigenvalues of $\mathbf{A}$ (see Eq. (5.15)) for a homogeneous material, it can be shown that the Courant condition itself is only a bound on the stability and can be exceeded. For a homogeneous medium with wave speed c and density ρ, $\mathbf{A}$ can be written as

$$\mathbf{A} = \begin{bmatrix} \ddots & \ddots & & & \\ \ddots & 2-2\frac{\Delta t^2}{\Delta x^2}c^2 & \frac{\Delta t^2}{\Delta x^2}c^2 & 0 & \\ & \frac{\Delta t^2}{\Delta x^2}c^2 & 2-2\frac{\Delta t^2}{\Delta x^2}c^2 & \frac{\Delta t^2}{\Delta x^2}c^2 & \\ & 0 & \frac{\Delta t^2}{\Delta x^2}c^2 & 2-2\frac{\Delta t^2}{\Delta x^2}c^2 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}, \tag{5.59}$$

where $\mathbf{A}$ has size $N \times N$. The eigenvalues of this matrix can be determined explicitly:

$$\lambda_{\mathrm{A}} = 2 - 2\frac{\Delta t^2}{\Delta x^2}c^2 + 2\frac{\Delta t^2}{\Delta x^2}c^2 \cos\left(\frac{p\pi}{N+1}\right), \tag{5.60}$$

where $p = 1 \ldots N$ and $|\lambda_{\mathrm{A}}| \leq 2$ for the system to be stable. The Courant condition is obtained by assuming the smallest possible value for the cosine, i. e., -1. However, from Eq. (5.60) it becomes clear that the cosine will never become -1, because the

argument of the cosine will never reach π. Only in the limit as N goes to infinity the argument of the cosine will approach π for $k = N$, and the cosine will become -1. However, if for example $N = 5$, in the worst case $\cos(\frac{5\pi}{6}) = -0.8660$, and the stability condition becomes

$$2 - 2\frac{\Delta t^2}{\Delta x^2}c^2(1 + 0.8660) \geq -2,$$

and from this

$$\frac{\Delta t^2}{\Delta x^2}c^2 \leq \frac{2}{1.8660} = 1.0718. \tag{5.61}$$

Thus, the time step can be chosen larger than dictated by the Courant condition.

Obviously, this observation is of only theoretical value, because actual problem dimensions will never be small enough for this observation to be relevant. For example, if $N = 50$, the factor by which the Courant condition can be exceeded reduces to 1.0004.

5.5 Concluding Remarks

The type of instabilities described in this section are not limited to only the FDTD algorithm for elastodynamics. Due to the similar structure of the FDTD algorithm, the same effects will also be observed in electromagnetics at a material interface with greatly different properties, but can be avoided if the material parameters are averaged on the boundary.

CHAPTER 6

Elastic Surface Waves

Elastic surface waves have been the object of extensive research, since Lord Rayleigh first discovered the existence of elastic waves confined to the superficial region of an infinite homogeneous isotropic solid more than a century ago [66]. Elastic surface waves play an important role in various fields. In seismology, surface waves have been found to carry the bulk of the energy among the waves excited by an earthquake. In electroacoustics, surface waves are utilized to make filters and resonators. In geoscience, the propagation characteristics of surface waves are used to obtain information about the physical properties of the ground. In this chapter, the theory of elastic surface waves is revisited.

A point source that is placed on the surface of a homogeneous isotropic medium excites five different kinds of waves: a pressure wave and a shear wave propagating in the medium, a Rayleigh surface wave that is confined to the surface of the medium, a lateral wave that is induced by the pressure wave at the surface, and a leaky surface wave that travels along the surface with a wave speed smaller than the pressure wave but larger than the shear wave. The first four of these wave types are well-known and have been treated extensively in the literature. However, the existence and theoretical foundation of the leaky surface wave has not been discussed much.

The leaky surface wave arises from the complex conjugate roots of the Rayleigh equation. The leaky surface wave is an inhomogeneous wave that propagates along the surface with a phase velocity larger than the shear wave but smaller than the

pressure wave. It couples into a plane shear wave that propagates in the medium. Due to the coupling, the surface wave loses energy and, thus, decays in its propagation direction.

The leaky surface wave has been observed by various authors in experimental and numerical studies. For example, Roth *et al.* noticed a rapidly decaying seismic surface wave in an environment with a very high Poisson ratio that had a phase velocity larger than the Rayleigh wave, but smaller than the pressure wave [67]. Smith *et al.* identified prograde and retrograde surface wave modes in a geologic study conducted on the shore of the gulf of Mexico [68]. And Glass and Maradudin found a leaky surface wave to exist in the flat-surface limit of a corrugated crystal surface [69].

In this chapter, the theoretical derivation of the leaky surface wave is described. In Sec. 6.1, the governing equations are briefly outlined, leading to the Rayleigh equation. The derivation of the Rayleigh equation is adopted from Graff [58] and is described for the sake of completeness. In Sec. 6.2, the various roots of the Rayleigh equation are discussed. It will be shown that for materials with a high Poisson ratio a leaky surface wave exists, due to the complex conjugate roots of the Rayleigh equation. In Sec. 6.3, the waves excited by a line source on the surface are derived analytically. The method of steepest descent is applied to obtain closed-form expressions for the various waves in the far field. The line-source case has been treated by many authors [70] – [72]. The procedure that is described here closely follows Felsen and Marcuvitz [73] and is mainly aimed at showing that a leaky surface wave is excited by a source on the surface. The asymptotic expressions for all other wave types in the far field are well-known and are given only for completeness reasons.

6.1 The Rayleigh Equation

The elastic wave fields at the surface of a semi-infinite, isotropic, lossless, homogeneous half space are to be determined. The half-space is bounded at $z = 0$ by a free-surface

boundary. The fields are assumed to be invariant in the y-direction and non-zero only in the x-z plane (*plane-strain case*, $u_y = \partial/\partial y = 0$). Thus, the originally three-dimensional problem reduces to a two-dimensional one.

The elastic wave fields in a medium may be expressed in terms of their potential functions [58]:

$$\mathbf{u} = \nabla \cdot \Phi + \nabla \times \mathbf{H}, \tag{6.1}$$

where $\mathbf{u}$ is the displacement vector, Φ is a scalar potential describing the longitudinal pressure wave, and $\mathbf{H}$ is a vector potential describing the transverse shear wave. The potentials satisfy the wave equations

$$\nabla^2 \Phi = \frac{1}{c_P^2} \frac{\partial^2 \Phi}{\partial t^2} \tag{6.2}$$

$$\nabla^2 \mathbf{H} = \frac{1}{c_S^2} \frac{\partial^2 \mathbf{H}}{\partial t^2}, \tag{6.3}$$

for the pressure and shear wave, respectively, with their corresponding wave speeds, c_P and c_S. In the plane-strain case, the only non-zero vector potential component is H_y, and the only non-zero displacement components are u_x and u_z:

$$u_x = \frac{\partial \Phi}{\partial x} - \frac{\partial H_y}{\partial z} \tag{6.4}$$

$$u_z = \frac{\partial \Phi}{\partial z} + \frac{\partial H_y}{\partial x}. \tag{6.5}$$

The only independent stress components are τ_{xx}, τ_{zz}, and τ_{xz}. The stresses are expressed in terms of the displacements as

$$\tau_{xx} = (\lambda + 2\mu) \frac{\partial u_x}{\partial x} + \lambda \frac{\partial u_z}{\partial z} \tag{6.6}$$

$$\tau_{yy} = \lambda \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_z}{\partial z} \right) \tag{6.7}$$

$$\tau_{zz} = \lambda \frac{\partial u_x}{\partial x} + (\lambda + 2\mu) \frac{\partial u_z}{\partial z} \tag{6.8}$$

$$\tau_{xz} = \mu \left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x} \right). \tag{6.9}$$

Note that τ_{yy} is non-zero, but is dependent on τ_{xx} and τ_{zz} and can be written as:

$$\tau_{yy} = \frac{\lambda}{2\lambda + 2\mu} \left(\tau_{xx} + \tau_{zz} \right). \tag{6.10}$$

Introducing the potentials, the stresses become

$$\tau_{xx} = (\lambda + 2\mu)\frac{\partial^2 \Phi}{\partial x^2} + \lambda \frac{\partial^2 \Phi}{\partial z^2} - 2\mu \frac{\partial^2 H_y}{\partial x \partial z} \tag{6.11}$$

$$\tau_{zz} = \lambda \frac{\partial^2 \Phi}{\partial x^2} + (\lambda + 2\mu)\frac{\partial^2 \Phi}{\partial z^2} + 2\mu \frac{\partial^2 H_y}{\partial x \partial z} \tag{6.12}$$

$$\tau_{xz} = \mu \left(2\frac{\partial^2 \Phi}{\partial x \partial z} - \frac{\partial^2 H_y}{\partial z^2} + \frac{\partial^2 H_y}{\partial x^2} \right). \tag{6.13}$$

Assuming harmonic time-dependence, the plane wave solutions for Φ and H_y satisfying Eqs. (6.2) and (6.3) are given by

$$\Phi = A e^{j\xi x + j\alpha z - j\omega t} \tag{6.14}$$

$$H_y = B e^{j\xi x + j\beta z - j\omega t}, \tag{6.15}$$

and the wave numbers are defined by

$$\alpha^2 = \omega^2 / c_P^2 - \xi^2 \tag{6.16}$$

$$\beta^2 = \omega^2 / c_S^2 - \xi^2 \tag{6.17}$$

$$\xi^2 = \omega^2 / c^2. \tag{6.18}$$

Inserting Eqs. (6.14) and (6.15) into Eqs. (6.4)–(6.5) and (6.11) – (6.13), the displacement and stress components are obtained:

$$u_x = \left(j\xi A e^{j\alpha z} - j\beta B e^{j\beta z} \right) e^{j\xi x} \tag{6.19}$$

$$u_z = \left(j\alpha A e^{j\alpha z} + j\xi B e^{j\beta z} \right) e^{j\xi x} \tag{6.20}$$

$$\tau_{xx} = \mu \left((2\alpha^2 - \beta^2 - \xi^2) A e^{j\alpha z} + 2\beta\xi B e^{j\beta z} \right) e^{j\xi x} \tag{6.21}$$

$$\tau_{zz} = \mu\left((\xi^2-\beta^2)Ae^{j\alpha z} - 2\beta\xi Be^{j\beta z}\right)e^{j\xi x} \tag{6.22}$$

$$\tau_{xz} = \mu\left(-2\alpha\xi Ae^{j\alpha z} + (\beta^2-\xi^2)Be^{j\beta z}\right)e^{j\xi x}. \tag{6.23}$$

Eqs. (6.19)–(6.23) describe the wave fields that are comprised of two and only two wave components: one pressure wave component, defined by its potential function Φ (according to Eq. (6.14)) with the wave number α and the amplitude A, and one shear wave component, characterized by its potential function H_y (Eq. (6.15)) with the wave number β and the amplitude B. By applying the appropriate boundary conditions to Eqs. (6.19)–(6.23), the waves that arise at the free surface and that can be described by the two components are obtained. As a consequence, multiple reflected waves that arise from an incident wave due to mode conversion at the surface will not be part of the solution, because the multiple reflected waves are defined by more than two components. For example, in general a plane pressure wave that is incident on the surface at an angle θ will give rise to both a reflected pressure wave and a reflected shear wave. However, three wave components with three distinct wave numbers and amplitudes are necessary to describe this scenario, and therefore this case will not be a solution to Eqs. (6.19)–(6.23).

At the surface, the normal stress vanishes, and thus $\tau_{zz}|_{z=0} = \tau_{xz}|_{z=0} = 0$. Using this condition, the ratio of the coefficients is determined from Eqs. (6.22) and (6.23) to be

$$\frac{A}{B} = \frac{2\xi\beta}{\xi^2-\beta^2} = -\frac{\xi^2-\beta^2}{2\xi\alpha}. \tag{6.24}$$

Re-inserting the amplitude ratio into Eqs. (6.22) and (6.23), the frequency equation is obtained:

$$(\xi^2-\beta^2)^2 + 4\xi^2\alpha\beta = 0. \tag{6.25}$$

Eq. (6.25) is commonly called the *Rayleigh Equation*, because it gives rise to the well-known Rayleigh surface wave.

Using Eqs. (6.16)–(6.18), the frequency equation can be rewritten in terms of

the wave speeds:

$$\left(2-\frac{c^2}{c_S^2}\right)^2+4\sqrt{\frac{c^2}{c_P^2}-1}\sqrt{\frac{c^2}{c_S^2}-1}=0. \tag{6.26}$$

By rationalizing, this equation may be expressed as

$$\left(\frac{c}{c_S}\right)^2\left[\left(\frac{c}{c_S}\right)^6-8\left(\frac{c}{c_S}\right)^4+\left(24-16(c_S/c_P)^2\right)\left(\frac{c}{c_S}\right)^2-\right.$$

$$\left.-16\left(1-(c_S/c_P)^2\right)\right]=0. \tag{6.27}$$

Eq. (6.27) always has three solutions for c^2 (when neglecting the trivial solution). Dependent on the Poisson Ratio ν of a material, different kinds of roots arise. For $\nu > 0.263$, Eq. (6.27) has one real root and two complex conjugate roots. For $\nu < 0.263$, Eq. (6.27) has three real roots. In each case, the real root that is smallest in magnitude gives rise to the Rayleigh surface wave, which propagates along the surface and decays into the medium. The other roots have often been classified as erroneous or non-physical roots of the Rayleigh equation. However, it will be shown here that the complex conjugate roots of the Rayleigh equation for $\nu > 0.263$ in fact give rise to a *leaky surface wave* and the two additional real roots for $\nu < 0.263$ describe the angles of incidence at which *complete mode conversion* occurs at the surface.

6.2 The Roots of the Rayleigh Equation

Let the roots of Eq. (6.27) be denoted by c. In general, c will be complex:

$$c = c_r + jc_i, \tag{6.28}$$

where c_r is the real part and c_i is the imaginary part of c. If c is purely real ($c = c_r$), ξ will be real (see Eq. (6.18)). If additionally c is smaller than both the pressure wave speed and the shear wave speed, i. e., $c = c_r < c_S < c_P$, α and β will be purely imaginary. For the solution to be physical, $\mathcal{I}m\{\alpha\} = \alpha_i > 0$ and $\mathcal{I}m\{\beta\} = \beta_i > 0$, thus the waves described by these wave numbers *propagate* in the x-direction and *decay* in the z-direction. This solution represents the well-known Rayleigh surface

wave, first explored by Lord Rayleigh more than a century ago [66]. A solution of this form always exists, independent of the Poisson ratio of a material.

6.2.1 Materials with a Poisson Ratio larger than 0.263

If $\nu > 0.263$, a real root and two complex conjugate roots of the Rayleigh equation arise. The real root gives rise to the Rayleigh surface wave. For the complex conjugate roots, the wave speed is complex, $c = c_r + jc_i$, and consequently also the wave numbers are complex: $\xi = \xi_r + j\xi_i$, $\alpha = \alpha_r + j\alpha_i$ and $\beta = \beta_r + j\beta_i$. It can be shown that for the complex conjugate roots the real part of the wave speed is always smaller than the pressure wave speed, but larger than the shear wave speed, $c_S < \mathcal{R}e\{c\} < c_P$.

Although c may be a solution to Eq. (6.27), it does not necessarily follow that also Eq. (6.25) is fulfilled. This is due to the manipulation of Eq. (6.25) to arrive at Eq. (6.27). In fact, the complex conjugate roots of Eq. (6.27) do not represent solutions to Eq. (6.26). It can be shown, however, that they do represent solutions to Eq. (6.25), if the signs of the wave numbers α and β are picked correctly. It may be recalled that according to Eqs. (6.16) and (6.17) the wave numbers α and β are functions of the square root of c^2,

$$\alpha = \pm\frac{\omega}{c}\sqrt{\frac{c^2}{c_P^2} - 1} = \pm(\alpha_r + j\alpha_i) \tag{6.29}$$

$$\beta = \pm\frac{\omega}{c}\sqrt{\frac{c^2}{c_S^2} - 1} = \pm(\beta_r + j\beta_i). \tag{6.30}$$

The sign in front of the square roots must be chosen according to physical and causal constraints of the underlying problem. To obtain Eq. (6.26), the positive sign has been assumed for both α and β. However, it turns out that Eq. (6.25) is only satisfied if, for the complex conjugate roots, both the real part and the imaginary part of α and β have opposite signs, i. e., $\text{sign}(\alpha_r) \neq \text{sign}(\beta_r)$ and $\text{sign}(\alpha_i) \neq \text{sign}(\beta_i)$.

Fig. 6.1 schematically shows the arrangement of the roots in the complex ξ-plane. The possible solutions of the Rayleigh equation are summarized in Table 6.1, giving all possible combinations of ξ, α and β. Only waves propagating in the positive

x-direction are considered. Five possible solutions arise. The first solution describes

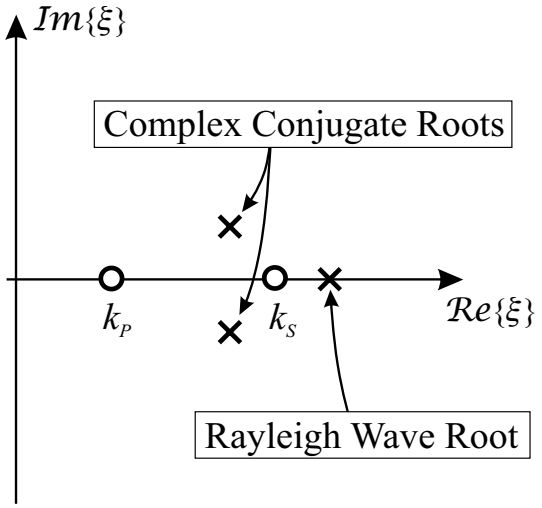


Figure 6.1: Schematical arrangement of the roots in the complex ξ-plane for $\nu > 0.263$.

the Rayleigh surface wave. For the second solution, both the pressure wave potential and the shear wave potential, Φ and H_y, propagate and increase in the positive x-direction ($\xi_r > 0$, $\xi_i < 0$). However, Φ propagates and decays in the positive z-direction ($\alpha_r > 0$, $\alpha_i > 0$), whereas H_y propagates and decays in the negative z-direction ($\beta_r < 0$, $\beta_i < 0$). For the third solution, the potentials decay in the x-direction ($\xi_i > 0$). Φ now propagates and increases in the negative z-direction ($\alpha_r < 0$, $\alpha_i > 0$), whereas H_y propagates and increases in the positive z-direction ($\beta_r > 0$, $\beta_i < 0$). For the fourth and fifth solution, the signs of α and β are reversed.

The behavior of the five possible solutions is best demonstrated by calculating the wave fields for a medium with a specific value of Poisson's ratio. Assuming Poisson's ratio to be $\nu = 0.4$, the elastic wave fields are computed using Eqs. (6.19)–(6.23). The amplitude of the shear wave potential is chosen to be unity, and the amplitude of the pressure wave potential is computed with Eq. (6.24). For $\nu = 0.4$,

Table 6.1: Solutions to the Rayleigh equation for $\nu > 0.263$.

(1)	$\xi_r > 0,\ \xi_i = 0$	$\alpha_r = 0,\ \alpha_i > 0$	$\beta_r = 0,\ \beta_i > 0$
		$\Phi \sim e^{-\lvert\alpha_i\rvert z} e^{j\lvert\xi_r\rvert x}$	$H_z \sim e^{-\lvert\beta_i\rvert z} e^{j\lvert\xi_r\rvert x}$
(2)	$\xi_r > 0,\ \xi_i < 0$	$\alpha_r > 0,\ \alpha_i > 0$	$\beta_r < 0,\ \beta_i < 0$
		$\Phi \sim e^{(j\lvert\alpha_r\rvert - \lvert\alpha_i\rvert)z} e^{(j\lvert\xi_r\rvert + \lvert\xi_i\rvert)x}$	$H_z \sim e^{(-j\lvert\beta_r\rvert + \lvert\beta_i\rvert)z} e^{(j\lvert\xi_r\rvert + \lvert\xi_i\rvert)x}$
(3)	$\xi_r > 0,\ \xi_i > 0$	$\alpha_r < 0,\ \alpha_i > 0$	$\beta_r > 0,\ \beta_i < 0$
		$\Phi \sim e^{(-j\lvert\alpha_r\rvert - \lvert\alpha_i\rvert)z} e^{(j\lvert\xi_r\rvert - \lvert\xi_i\rvert)x}$	$H_z \sim e^{(j\lvert\beta_r\rvert + \lvert\beta_i\rvert)z} e^{(j\lvert\xi_r\rvert - \lvert\xi_i\rvert)x}$
(4)	$\xi_r > 0,\ \xi_i < 0$	$\alpha_r < 0,\ \alpha_i < 0$	$\beta_r > 0,\ \beta_i > 0$
		$\Phi \sim e^{(-j\lvert\alpha_r\rvert + \lvert\alpha_i\rvert)z} e^{(j\lvert\xi_r\rvert + \lvert\xi_i\rvert)x}$	$H_z \sim e^{(j\lvert\beta_r\rvert - \lvert\beta_i\rvert)z} e^{(j\lvert\xi_r\rvert + \lvert\xi_i\rvert)x}$
(5)	$\xi_r > 0,\ \xi_i > 0$	$\alpha_r > 0,\ \alpha_i < 0$	$\beta_r < 0,\ \beta_i > 0$
		$\Phi \sim e^{(j\lvert\alpha_r\rvert + \lvert\alpha_i\rvert)z} e^{(j\lvert\xi_r\rvert - \lvert\xi_i\rvert)x}$	$H_z \sim e^{(-j\lvert\beta_r\rvert - \lvert\beta_i\rvert)z} e^{(j\lvert\xi_r\rvert - \lvert\xi_i\rvert)x}$

the pressure wave speed exceeds the shear wave speed by a factor of about 2.45: $c_P = 2.4495 c_S$. The roots of Eq. (6.27) in terms of the shear wave speed c_S are

$$c = \begin{cases} 0.9422 c_S \\ (1.9276 + j0.3996) c_S \\ (1.9276 - j0.3996) c_S \end{cases} . \tag{6.31}$$

The resulting wave numbers are shown in Table 6.2. All wave numbers are expressed in terms of the longitudinal wave number, $k_P = \omega / c_P$.

In Fig. 6.2, the displacements according to the five solutions of the Rayleigh equation are plotted versus x and z. The distance on the axes is normalized to the wave length of the Rayleigh surface wave, λ_R. The upper edge of each plot corresponds to the surface. The two columns show the horizontal and vertical displacements, u_x and u_z. Pseudo-color plots are used to display the wave fields, employing a logarithmic scale with a dynamic range of 50 dB. Superimposed with the horizontal displacement component is the real part of the complex Poynting vector, thus indicating the direction of the energy flow. The complex Poynting vector is defined as

Table 6.2: Solutions to the Rayleigh equation for $\nu = 0.4$: wave numbers.

	$\xi =$	$\alpha =$	$\beta =$
(1)	$2.5998k_P$	$j2.3997k_P$	$j0.8711k_P$
(2)	$(1.2184 - j0.2526)k_P$	$(0.4030 + j0.7636)k_P$	$(-2.1448 - j0.1435)k_P$
(3)	$(1.2184 + j0.2526)k_P$	$(-0.4030 + j0.7636)k_P$	$(2.1448 - j0.1435)k_P$
(4)	$(1.2184 - j0.2526)k_P$	$(-0.4030 - j0.7636)k_P$	$(2.1448 + j0.1435)k_P$
(5)	$(1.2184 + j0.2526)k_P$	$(0.4030 - j0.7636)k_P$	$(-2.1448 + j0.1435)k_P$

[57]

$$\mathbf{P} = \frac{j\omega \mathbf{u} \cdot \boldsymbol{\tau}^*}{2}, \tag{6.32}$$

where the star denotes the complex conjugate. The real part of the complex Poynting vector describes the average power per unit area.

First, the Rayleigh surface wave is shown. For the Rayleigh surface wave, the energy flow is seen to be parallel to the surface. Both the horizontal and the vertical displacements decay away from the surface. The second solution describes a shear wave in which energy is carried towards the surface. Close to the surface, the energy flows parallel to the surface. The waves grow exponentially in the x- and z-direction. The third solution is the converse of the second one. This time, the energy flows from the surface into the medium, and the waves decay in the x-direction. Again, close to the surface the energy flow is parallel to the surface. The fourth and fifth solution are similar to the second and third. However, now a pressure wave propagates in the medium, and the energy flows at a different angle with respect to the surface. Also, the exponential growth is enhanced.

All of the five possible solutions described above can, in certain cases, represent physical solutions. For example, if a field distribution is created on the surface that matches the field distribution of Solution (5) on the surface, waves similar to the ones described by Solution (5) will be induced in the medium. If a field distribution is

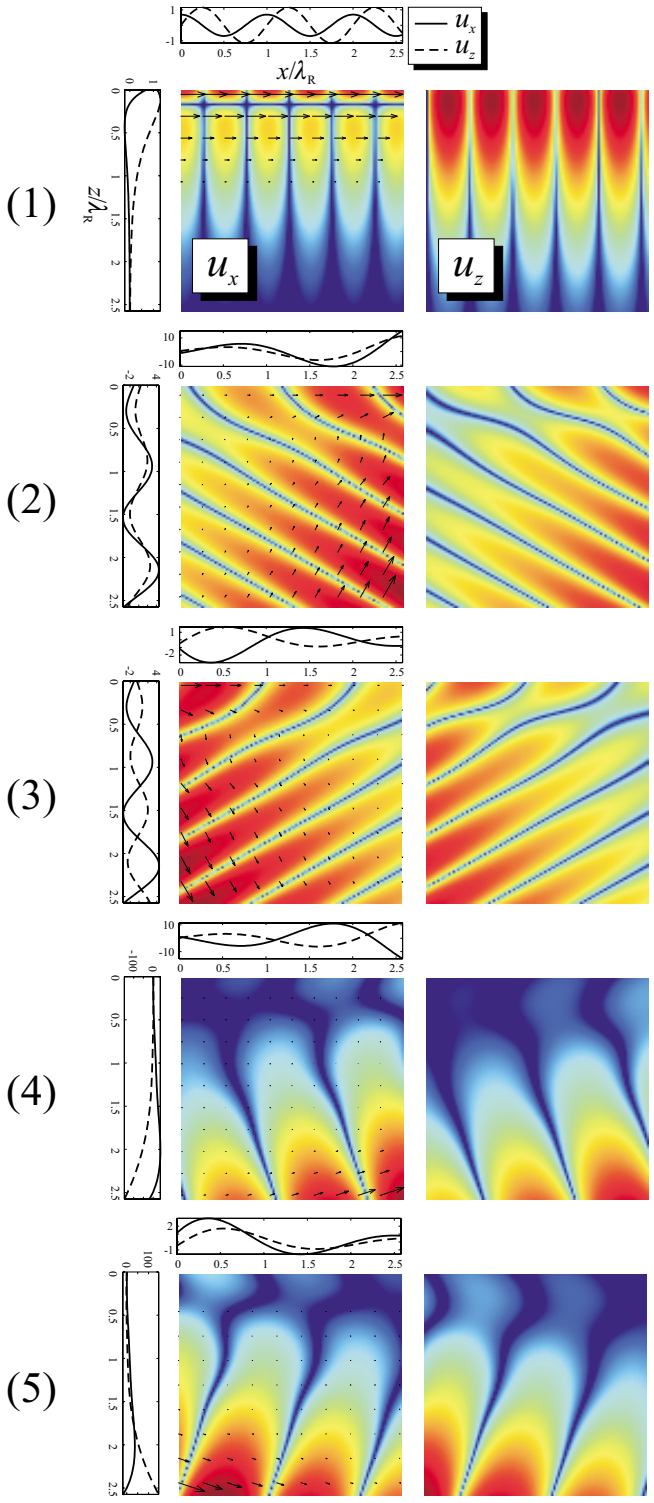


Figure 6.2: Horizontal and vertical displacements according to the five solutions of the Rayleigh equation.

generated within the medium that is equal to the field distribution of Solution (2), waves propagating towards the surface will be excited that perfectly couple into a surface wave. Of course, the solutions as described here would require an infinite medium and wave fields of infinite extent that are non-zero at infinity, which violates physical as well as causal constraints. However, over a finite range all of these solutions can be excited with the appropriate field distributions on the boundaries.

Interpretation in the k-Space

Further physical insight can be gained by looking at the wave vectors for the solutions in the k-space. The wave vectors in the k-space will lie on well-defined surfaces, the *allowed wave-vector surfaces*. By looking at the allowed wave-vector surfaces, the coupling between the surface waves and the waves in the medium can be explained.

Fig. 6.3 shows the allowed wave vector surfaces for the pressure wave and the shear wave in the k_x-k_z-plane. For the wave vector surfaces, the wave number of the pressure wave and the shear wave in the z-direction, α and β according to Eqs. (6.16) and (6.17), are plotted as a function of the wave number in the x-direction, ξ. The allowed wave vector surfaces for the pressure and the shear wave are circles with radii k_P and k_S, respectively. The allowed wave vector surfaces define the wave vectors that describe physical waves. The wave vectors indicate the propagation direction of the waves and are inversely proportional to their wave speed. In Fig. 6.3, the wave vectors of the surface waves that arise from the Rayleigh equation are indicated by the bold solid arrows along the k_x-axis. The gray arrow corresponds to the real part of the wave vector of the Rayleigh surface wave according to Solution (1), whereas the black arrow describes the real part of the wave vector in the x-direction of Solutions (2) – (5).

From the wave vectors in the k-space, conclusions about the coupling between the different waves can be drawn. At the surface, two waves will be coupled, if their wave vector components tangential to the surface match both in the real part and the imaginary part. The roots of the Rayleigh equation give rise to two different kinds

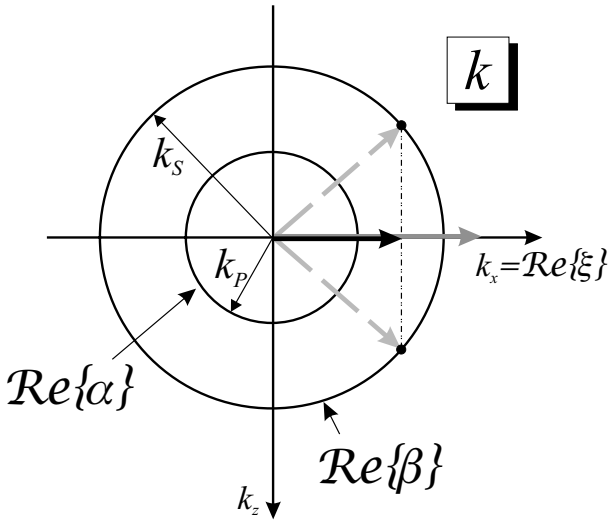


Figure 6.3: Allowed wave vector surfaces in the k-space for the pressure wave (α) and the shear wave (β); lossless case.

of surface waves: the Rayleigh surface wave, described by Solution (1), and leaky surface waves, described by Solutions (2) – (5). The surface waves are defined by their wave vectors, which will be parallel to the k_x-axis and, thus, tangential to the surface. If the wave vector of a surface wave matches the projection onto the surface of the wave vector describing a wave propagating in the medium, the two waves will be coupled. The Rayleigh surface wave cannot be coupled to any wave propagating in the medium. The wave number of the Rayleigh surface wave is larger than both the wave number of the pressure wave, k_P, and the wave number of the shear wave, k_S, and therefore the tangential wave vectors of the Rayleigh wave and a shear wave or a pressure wave in the medium can never match. This is evident in Fig. 6.3: the wave vector of the Rayleigh wave, indicated by the gray arrow, exceeds both k_P and k_S. The wave number for the leaky surface wave, on the other hand, is larger than k_P, but smaller than k_S. Thus, there exists a shear wave that propagates in the medium

at such an angle that its tangential component matches the wave vector of the leaky surface wave. This is shown in Fig. 6.3: the tangential wave vector components of the shear waves described by the dashed arrows match the wave vector of the leaky surface wave (solid arrow), and consequently the waves are coupled.

Fig. 6.3 describes the wave vector surfaces that arise when ξ is purely real. If ξ is complex, the wave vector surfaces for the pressure and the shear wave will not be circles, but they will be deformed. If ξ is real, the pressure wave number in the z-direction, α, is purely real for $|k_x| < k_P$ and purely imaginary for $|k_x| > k_P$. Similarly, the shear wave number in the z-direction, β, is real for $|k_x| < k_S$ and imaginary for $|k_x| > k_S$. If ξ is complex, however, α and β will also both be complex. Specifically, the real part of α and β will be non-zero for $|k_x| > k_P$ and $|k_x| > k_S$, respectively. In this case, a tangential wave vector match is possible between the leaky surface wave and both a pressure wave and a shear wave.

Fig. 6.4 shows the wave vector surfaces that arise for Solutions (1) – (5). Except for Solution (1), which describes the Rayleigh wave, the wave vector surfaces are deformed due to ξ being complex. For the Rayleigh wave (Solution (1)), ξ is purely real and exceeds k_P and k_S, whereas α and β are both purely imaginary. The imaginary parts of α and β both must be positive, because, due to physical constraints, the Rayleigh wave must decay into the medium. Also in Fig. 6.4, the energy flow of the Rayleigh wave in the x-z-plane is indicated schematically. For the Rayleigh wave, the energy flows parallel to the surface, and the Rayleigh wave does not decay in its propagation direction. Solutions (2) – (5) describe the leaky surface waves. For Solution (2), a shear wave propagates towards the surface and couples into a surface wave. Due to the coupling, both the surface wave and the shear wave in the medium are inhomogeneous. The wave vectors of both the surface wave and the shear wave in the medium are complex, and the allowed wave vector surfaces are deformed. The surface wave has to components: a pressure wave component, which propagates and decays into the medium, and a shear wave component, which propagates and decays towards the surface. This is indicated by the wave vectors for the pressure and the

shear wave component (Solution (2), dashed arrows): the real and imaginary part of α are positive, whereas the real and imaginary part of β are negative. In the medium, the energy flows towards the surface. Close to the surface, the energy flow is parallel to the surface. The surface wave grows exponentially in its propagation direction, because it is constantly fed by the shear wave in the medium. For Solution (3), the surface wave feeds a shear wave that propagates away from the surface. The surface wave "leaks" energy into the medium. Consequently, the wave vector for the shear wave component points into the medium. As for Solution (2), the shear wave increases exponentially into the medium, indicated by the imaginary part of β being negative. For Solutions (4) and (5), a pressure wave propagates in the medium. For Solution(4), the pressure wave propagates and decays towards the surface, indicated by the real and imaginary part of α being negative. For Solution (5), the pressure wave is fed from the surface wave and propagates into the medium. In this case, the real part of α is positive, and the imaginary part is negative.

6.2.2 Materials with a Poisson Ratio smaller than 0.263

For $\nu < 0.263$, three purely real roots of the Rayleigh equation exist. The root smallest in magnitude gives rise to the Rayleigh surface wave. For the additional two roots, the phase speeds along the surface are larger than the pressure wave speed, c_P. Thus, their wave numbers both in the x- and the z-direction are real and smaller in magnitude than the wave number of the pressure wave speed, k_P. The arrangement of the roots in the complex ξ-plane is shown schematically in Fig. 6.5. The possible solutions of the Rayleigh equation are summarized in Table 6.3, giving all possible combinations of ξ, α and β. Except for the Rayleigh wave (Solution (1)), all wave numbers are purely real, and therefore their imaginary part is zero. Only waves propagating in the positive x-direction are considered. Five possible solutions arise. The first solution again describes the Rayleigh surface wave. The other four solutions depict cases of *complete mode conversion* for a plane wave incident on the surface.

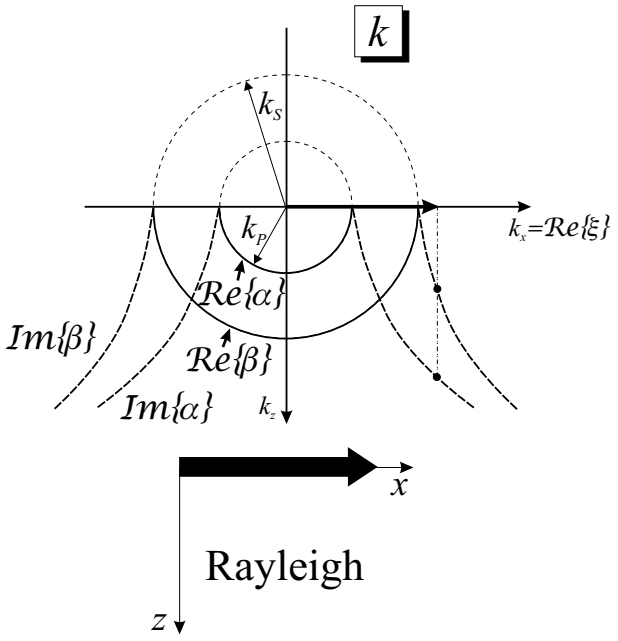
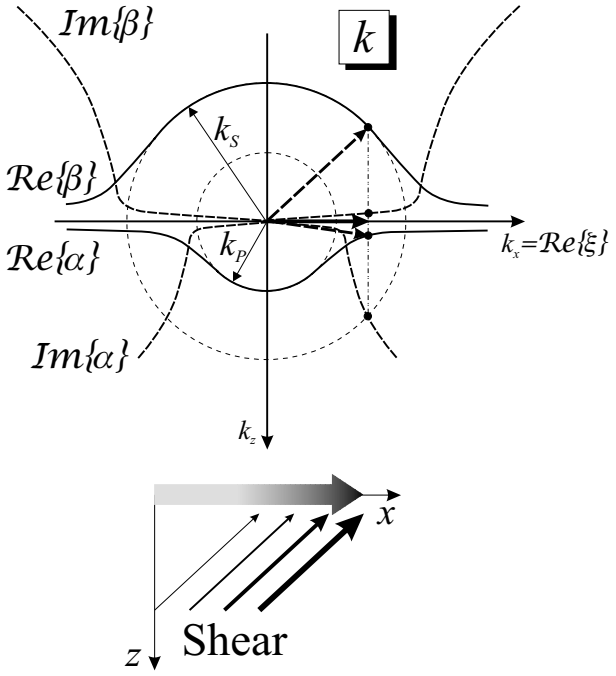
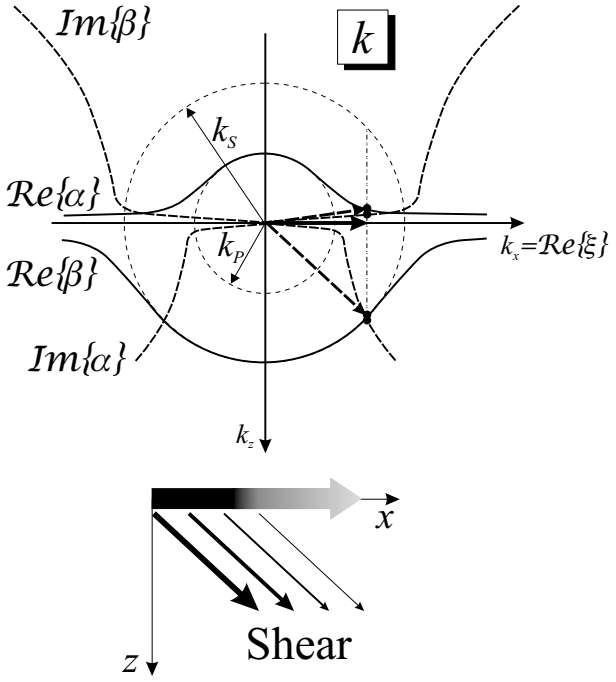


Figure 6.4: Allowed wave vector surfaces for the five solutions to the Rayleigh equation.

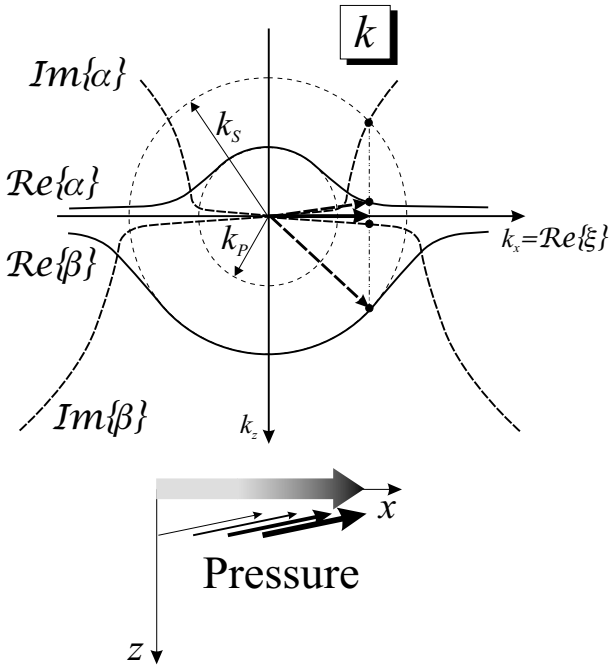


Solution (2)

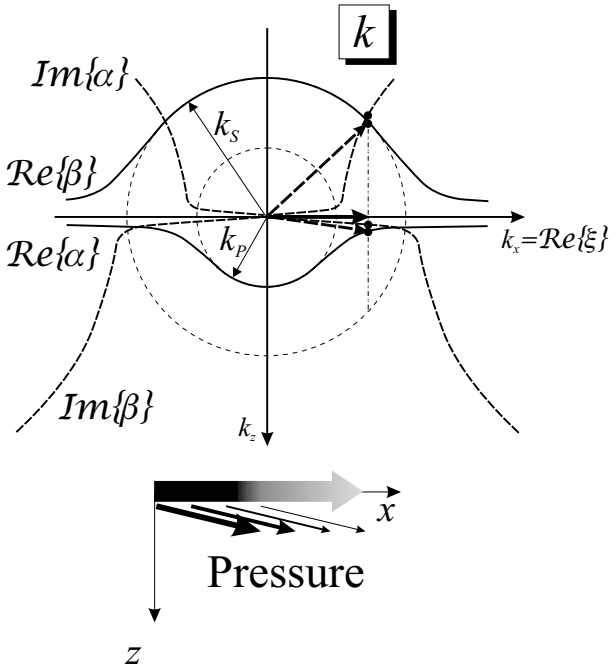


Solution (3)

Figure 6.4: continued.



Solution (4)



Solution (5)

Figure 6.4: continued.

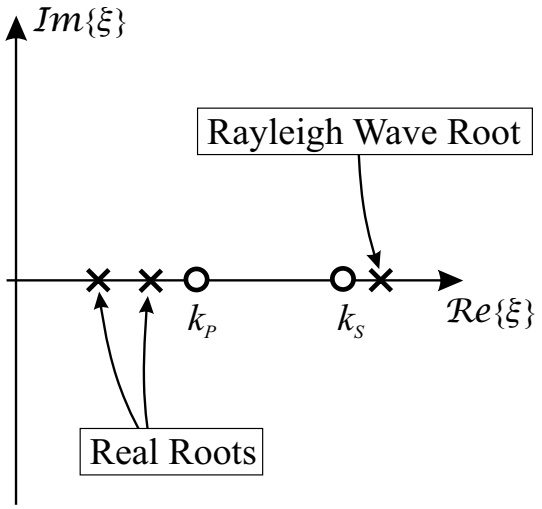


Figure 6.5: Schematical arrangement of the roots in the complex ξ-plane for $\nu < 0.263$.

Complete mode conversion occurs for both pressure waves and shear waves at certain angles of incidence. At these angles of incidence, a plane shear wave, for example, will be reflected solely as a plane pressure wave, and, thus, the shear wave mode is completely converted into a pressure wave mode. The angles of incidence are dependent on Poisson's ratio. Complete mode conversion is physically possible only for materials with a Poisson ratio smaller than 0.263, i. e., if the Rayleigh equation has three real roots. The second solution in Table 6.3 describes a shear wave incident on the surface that completely couples into a reflected pressure wave. The third solution describes the converse case: an incident pressure wave is completely converted into a reflected shear wave. The fourth and fifth solution have the same shape as the second and third solution, but are for different angles of incidence.

The three real roots for a material with a Poisson ratio of $\nu = 0.2$ in terms of

Table 6.3: Solutions to the Rayleigh equation for $\nu < 0.263$.

(1)	$\xi_r > 0$	$\alpha_i > 0$	$\beta_i > 0$
		$\Phi \sim e^{-\lvert\alpha_i\rvert z} e^{j\lvert\xi_r\rvert x}$	$H_z \sim e^{-\lvert\beta_i\rvert z} e^{j\lvert\xi_r\rvert x}$
(2)	$\xi_r > 0$	$\alpha_r > 0$	$\beta_r < 0$
		$\Phi \sim e^{j\lvert\alpha_r\rvert z} e^{j\lvert\xi_r\rvert x}$	$H_z \sim e^{-j\lvert\beta_r\rvert z} e^{j\lvert\xi_r\rvert x}$
(3)	$\xi_r > 0$	$\alpha_r < 0$	$\beta_r > 0$
		$\Phi \sim e^{-j\lvert\alpha_r\rvert z} e^{j\lvert\xi_r\rvert x}$	$H_z \sim e^{j\lvert\beta_r\rvert z} e^{j\lvert\xi_r\rvert x}$
(4)	$\xi_r > 0$	$\alpha_r > 0$	$\beta_r < 0$
		$\Phi \sim e^{j\lvert\alpha_r\rvert z} e^{j\lvert\xi_r\rvert x}$	$H_z \sim e^{-j\lvert\beta_r\rvert z} e^{j\lvert\xi_r\rvert x}$
(5)	$\xi_r > 0$	$\alpha_r < 0$	$\beta_r > 0$
		$\Phi \sim e^{-j\lvert\alpha_r\rvert z} e^{j\lvert\xi_r\rvert x}$	$H_z \sim e^{j\lvert\beta_r\rvert z} e^{j\lvert\xi_r\rvert x}$

the shear wave speed are

$$c = \begin{cases} 0.9110 c_S \\ 1.6398 c_S \\ 2.1169 c_S \end{cases} . \tag{6.33}$$

For $\nu = 0.2$, the pressure wave speed exceeds the shear wave speed by a factor of 1.63: $c_P = 1.6330 c_S$. Table 6.4 shows the wave numbers arising from the three real roots.

The solutions due to the real roots are best illustrated by plotting the resulting wave fields due to an excitation in the time domain rather than for harmonic time dependence. For this, the wave fields are computed using Eqs. (6.19)–(6.23), are Fourier transformed into the time-domain and convolved with a specific time-domain excitation. A differentiated Gaussian pulse with a center frequency of 400 Hz is used. The results are shown in Fig. 6.6. The magnitude of the displacement fields according to the four solutions arising from the two additional real roots are plotted. A logarithmic color scale with a dynamic range of 40 dB is used. Shear waves are labeled by S and pressure waves by P. The bold gray arrows indicate the propagation

Table 6.4: Solutions to the Rayleigh equation for $\nu = 0.2$: wave numbers.

	$\xi =$	$\alpha =$	$\beta =$
(1)	$1.7925k_P$	$j2.3997k_P$	$j0.8711k_P$
(2)	$0.9959k_P$	$0.0909k_P$	$-1.2942k_P$
(3)	$0.9959k_P$	$-0.0909k_P$	$1.2942k_P$
(4)	$0.7714k_P$	$0.6363k_P$	$-1.4393k_P$
(5)	$0.7714k_P$	$-0.6363k_P$	$1.4393k_P$

direction of the plane waves. As indicated in Table 6.3, Solutions (2) and (4) describe cases of complete mode conversion for an incident shear wave, whereas Solutions (3) and (5) are for an incident pressure wave. The angle of incidence at which complete mode conversion occurs can be determined easily from the wave numbers:

$$\theta = \tan^{-1}(\xi_r/|\alpha_r|) \tag{6.34}$$

for an incident pressure wave and

$$\theta = \tan^{-1}(\xi_r/|\beta_r|) \tag{6.35}$$

for an incident shear wave.

6.3 Waves due to a Line Source on the Surface

In the previous section, the solutions to the wave equation at a free-surface boundary have been described in a general form. In this section, the wave fields due to a specific excitation, a line source on the surface, are determined.

Fig. 6.7 shows the underlying geometry. A line source is placed on the surface at $x = 0$ and extends into the y-direction. The line source excites the normal stress component τ_{zz}. If harmonic time dependence is assumed, the displacement fields due

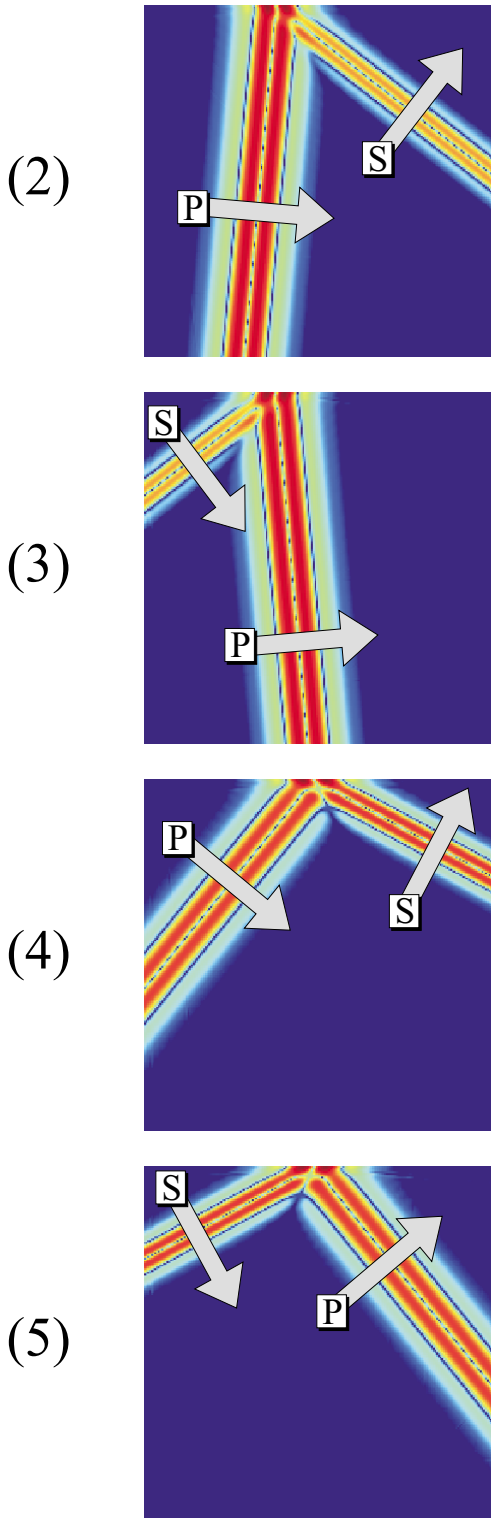


Figure 6.6: Magnitude of the displacements according to the four solutions due to the two additional real roots for $\nu < 0.263$.

to the line source can be written in form of an integral equation [58]:

$$u_x(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\xi}{F_0(\xi)}\left[-2\alpha\beta e^{j\beta z}+(\beta^2-\xi^2)e^{j\alpha z}\right]e^{j\xi x}d\xi \tag{6.36}$$

$$u_z(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\alpha}{F_0(\xi)}\left[2\xi^2 e^{j\beta z}+(\beta^2-\xi^2)e^{j\alpha z}\right]e^{j\xi x}d\xi, \tag{6.37}$$

where

$$F_0(\xi) = (\xi^2-\beta^2)^2+4\xi^2\alpha\beta \tag{6.38}$$

is the Rayleigh equation. The amplitude of the excitation is assumed to be unity. For the analysis in the following, a medium with a Poisson ratio larger than 0.263 is assumed, and, thus, F_0 has one real root and two complex conjugate roots.

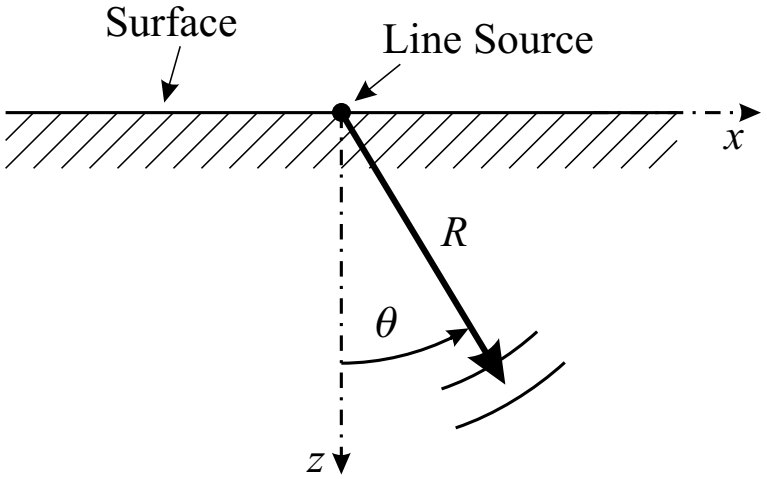


Figure 6.7: Line source on the surface.

6.3.1 General Considerations

The integrals in Eqs. (6.36) and (6.37) each represent an inverse Fourier transform from the wave-number domain into the spatial domain. The integrands contain poles and branch points. The poles are due to the roots of the Rayleigh equation in the denominator. The branch points arise because of the square-root dependence of α

and β on ξ. They are located at the roots of α and β, at $k_P = \pm(\omega/c_P)^2$ and $k_S = \pm(\omega/c_S)^2$ (see Eqs. (6.16) and (6.17)).

To compute the integrals, contour integration in the complex ξ-plane must be applied. The integration must be performed along the real ξ-axis. Fig. 6.8 shows the location of the poles and branch points in the complex ξ-plane. To determine the waves propagating in the positive x-direction, Eqs. (6.36) and (6.37) must be integrated along the path P'. The contour is closed at infinity. Only the poles and branch cuts for $\mathcal{R}e\{\xi\} > 0$ are included in the integration contour (indicated by the indentations of P'), whereas the poles for $\mathcal{R}e\{\xi\} < 0$ are excluded and, therefore, do not contribute to the integral.

Due to the branch points, the integrands are not single-valued. To make the integrands unique, a multi-sheeted Riemann surface for the ξ-plane is necessary, with branch cuts providing the transition from one Riemann sheet to the other [73]. The location of the branch cuts in general is arbitrary, but defines the disposition of those regions in the complex ξ-plane in which, for example, $\mathcal{R}e\{\alpha\} > 0$ or $\mathcal{R}e\{\alpha\} < 0$. Fig. 6.8 shows the top Riemann sheet for Eqs. (6.36) and (6.37). The signs of the wave numbers on the top Riemann sheet must be chosen for physical and causal reasons. The integration along the real axis determines the shear and pressure waves excited by the line source. For the shear and the pressure waves to be causal, they must propagate away from the source and vanish at infinity. For this to be true, the wave numbers along the real axis must be chosen such that $\mathcal{R}e\{\alpha\} > 0$, $\mathcal{I}m\{\alpha\} > 0$, $\mathcal{R}e\{\beta\} > 0$ and $\mathcal{I}m\{\beta\} > 0$. It can be easily shown that in that case the wave numbers in the entire second and fourth quadrant must behave in the same way. In the first and third quadrant, the branch cuts must then be chosen such that $\mathcal{R}e\{\alpha\} < 0$, $\mathcal{I}m\{\alpha\} > 0$, $\mathcal{R}e\{\beta\} > 0$ and $\mathcal{I}m\{\beta\} < 0$. This is true because α and β must be continuous across the real axis. Thus, in the first and third quadrant, the pressure wave potential propagates and increases in the negative z-direction, whereas the shear wave potential propagates and increases in the positive z-direction. It is evident that in the first quadrant of the top Riemann sheet the wave numbers behave as described

for Solution (3) of the Rayleigh equation as indicated in Table 6.1. The poles on the top Riemann sheet correspond to physically existing waves and, therefore, the pole associated with Solution (3) of the Rayleigh equation represents a physical solution to the line-source problem. The pole in the third quadrant is the equivalent to the pole in the first quadrant, but describes a wave traveling in the negative x-direction. The two poles on the real ξ-axis are present on all sheets and, consequently, also represent physical waves. All other poles of the Rayleigh equation lie on different sheets and, thus, are non-physical for the line-source case.

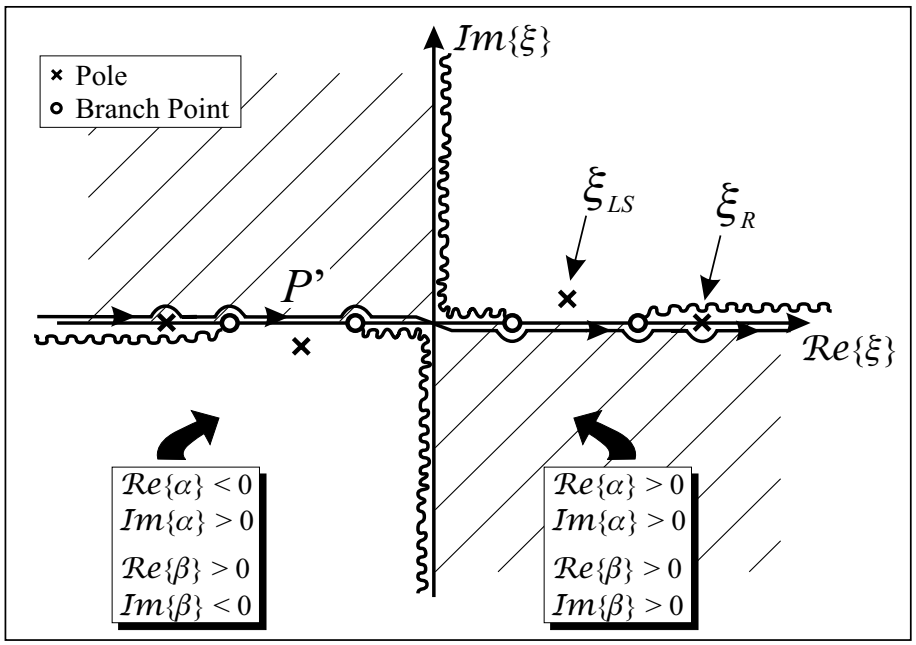


Figure 6.8: Location of the poles and branch cuts in the complex ξ-plane for the line-source problem.

It can be seen in Fig. 6.8 that four poles and four branch cuts exist on the top Riemann sheet. The poles at $\pm\xi_R$ on the real ξ-axis give rise to the well-known Rayleigh surface wave. The complex poles at $\pm\xi_{LS}$ in the first and third quadrant describe leaky surface waves propagating to the right and left, respectively. As de-

scribed earlier, the leaky surface wave couples into a plane shear wave. Both the leaky surface wave and the shear wave that is fed from the surface wave are inhomogeneous, which is indicated by the pole being complex.

6.3.2 Steepest-Descent Approximation

To evaluate the integrals asymptotically in the far field, the method of steepest descent shall be applied. To simplify the procedure, the two terms of the integral are treated separately. Dividing the integrals each into a pressure wave term and a shear wave term, Eqs. (6.36) and (6.37) are rewritten as

$$u_x^S(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\xi}{F_0(\xi)}(-2\alpha\beta)e^{j\beta z}e^{j\xi x}d\xi \tag{6.39}$$

$$u_x^P(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\xi}{F_0(\xi)}(\beta^2-\xi^2)e^{j\alpha z}e^{j\xi x}d\xi \tag{6.40}$$

$$u_z^S(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\alpha}{F_0(\xi)}2\xi^2 e^{j\beta z}e^{j\xi x}d\xi \tag{6.41}$$

$$u_z^P(x,z) = -\frac{j}{\mu\pi}\int_{P'}\frac{\alpha}{F_0(\xi)}(\beta^2-\xi^2)e^{j\alpha z}e^{j\xi x}d\xi. \tag{6.42}$$

The total displacements equal the superposition of the pressure wave component and the shear wave component:

$$u_x(x,z) = u_x^S(x,z)+u_x^P(x,z) \tag{6.43}$$

$$u_z(x,z) = u_z^S(x,z)+u_z^P(x,z). \tag{6.44}$$

To facilitate the evaluation of the integrals, the complex ξ-plane is transformed into the complex w^S-plane for the two shear wave terms, and into the w^P-plane for the pressure wave terms. The transformations are defined by

$$\xi = k_S \sin w^S \tag{6.45}$$

$$\xi = k_P \sin w^P. \tag{6.46}$$

These transformations are single-valued [73]. From the periodicity of $\sin w^S$ and $\sin w^P$ it is evident that multiple values for w^S and w^P correspond to a single value of ξ. Thus, the transformations can be used to map the entire ξ-plane with its multiple Riemann sheets into adjacent strips of width 2π in the w^S- or w^P-plane. The arrangement of the poles and branch cuts of the top Riemann sheet in the complex w^S-plane and w^P-plane are shown in Fig. 6.9. Here, the top Riemann sheet is mapped into a strip reaching from $-\pi$ to π in the complex w^S-plane for the shear wave terms, and similarly in the complex w^P-plane for the pressure wave terms. The positions of the transformed Rayleigh wave pole and the leaky surface wave pole in the complex w^S- and w^P-plane are indicated by w_R^S, w_{LS}^S, and w_R^P, w_{LS}^P, respectively. The transformed integration paths are denoted by P^S and P^P.

The separate transformations for the shear wave terms and the pressure wave terms become necessary, because, when the method of steepest descent is applied, the different terms will give rise to different steepest-descent paths. By applying the different transformations, the steepest-descent paths will have a rather simple shape for both the shear wave terms and the pressure wave terms, thus, making the steepest-descent approximation considerably easier.

Applying the transformations, two of the branch cuts are eliminated in each of the integrals in Eqs. (6.39)–(6.42). For the shear wave terms the branch cuts at $\pm k_S$ vanish, whereas for the pressure wave terms the branch cuts at $\pm k_P$ are removed. For the shear wave terms, β then reduces to

$$\beta = k_S \cos w^S, \tag{6.47}$$

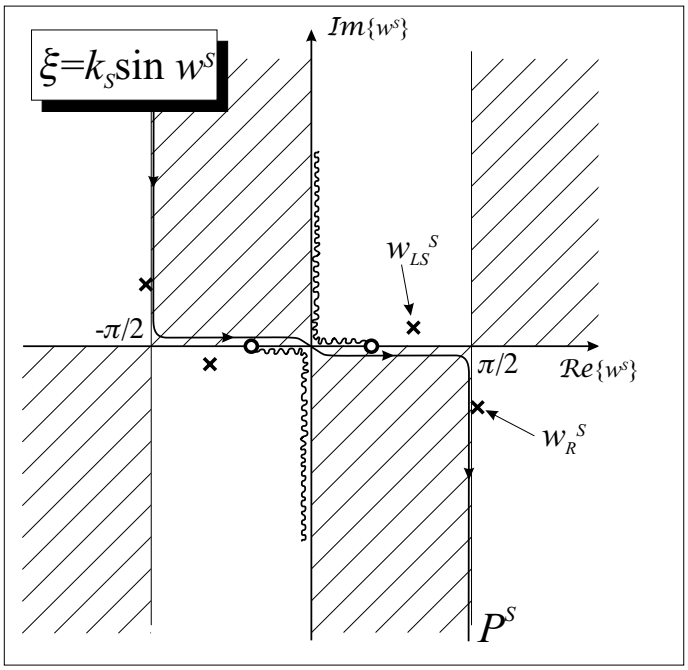
and for the pressure wave terms α becomes

$$\alpha = k_P \cos w^P. \tag{6.48}$$

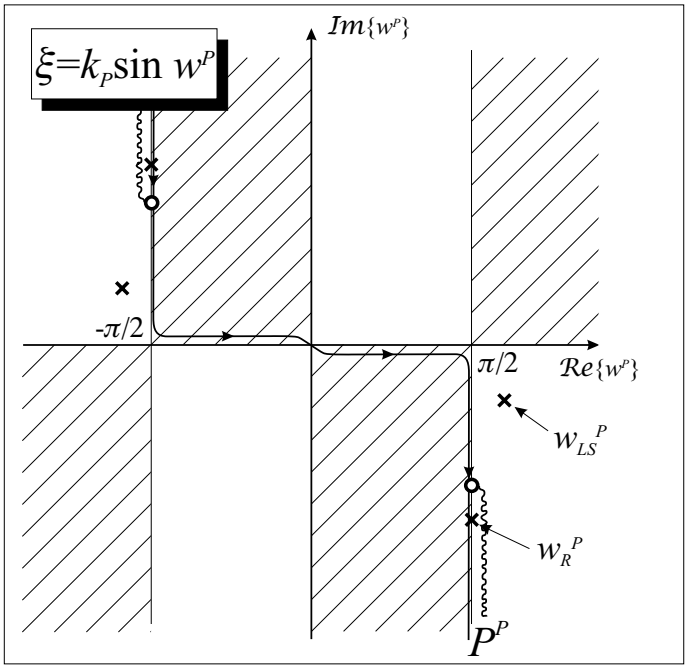
Introducing polar coordinates,

$$x = R \sin\theta \tag{6.49}$$

$$z = R \cos\theta, \tag{6.50}$$



(a)



(b)

Figure 6.9: Location of the poles and branch cuts (a) in the complex w^S-plane and (b) in the complex w^P-plane.

Eqs. (6.39)–(6.42) are rewritten as integrals in the w^S- and w^P-plane:

$$u_x^S(x,z) = -\frac{j}{\mu\pi}\int_{P^S}\frac{\xi}{F_0(\xi)}(-2\alpha\beta)e^{jk_S R\cos(w^S-\theta)}\beta dw^S \tag{6.51}$$

$$u_x^P(x,z) = -\frac{j}{\mu\pi}\int_{P^P}\frac{\xi}{F_0(\xi)}(\beta^2-\xi^2)e^{jk_P R\cos(w^P-\theta)}\alpha dw^P \tag{6.52}$$

$$u_z^S(x,z) = -\frac{j}{\mu\pi}\int_{P^S}\frac{\alpha}{F_0(\xi)}2\xi^2 e^{jk_S R\cos(w^S-\theta)}\beta dw^S \tag{6.53}$$

$$u_z^P(x,z) = -\frac{j}{\mu\pi}\int_{P^P}\frac{\alpha}{F_0(\xi)}(\beta^2-\xi^2)e^{jk_P R\cos(w^P-\theta)}\alpha dw^P, \tag{6.54}$$

where θ describes the polar angle measured from the surface normal towards the propagation direction (see Fig. 6.7). The wave numbers in terms of w^S are

$$\xi(w^S) = k_S\sin w^S \tag{6.55}$$

$$\beta(w^S) = k_S\cos w^S \tag{6.56}$$

$$\alpha(w^S) = \pm\sqrt{k_P^2-\xi(w^S)^2} \tag{6.57}$$

and in terms of w^P

$$\xi(w^P) = k_P\sin w^P \tag{6.58}$$

$$\beta(w^P) = \pm\sqrt{k_S^2-\xi(w^P)^2} \tag{6.59}$$

$$\alpha(w^P) = k_P\cos w^P. \tag{6.60}$$

The signs of $\alpha(w^S)$ and $\beta(w^P)$ must be chosen as described earlier for the complex ξ-plane. Thus, in the shaded and non-shaded regions of Fig. 6.9, α behaves just as in the shaded and non-shaded regions of Fig. 6.8.

With the integrals transformed as described above, it is relatively straightforward to apply the method of steepest descent. For the method of steepest descent, the integration paths P^S and P^P are deformed into new paths, the *steepest-descent*

paths $\bar{P}^S$ and $\bar{P}^P$, respectively. The new path is chosen such that the dominant contribution to the integral arises from only a small section of the path. To achieve this, the path is deformed such that it passes through the saddle point of the integrand. Away from the saddle point it follows the direction in which the integrand decays most rapidly. Along this path, the integrand will then be negligible everywhere but around the saddle point, and the integral can be approximated by the contribution from the integrand in the vicinity of the saddle point.

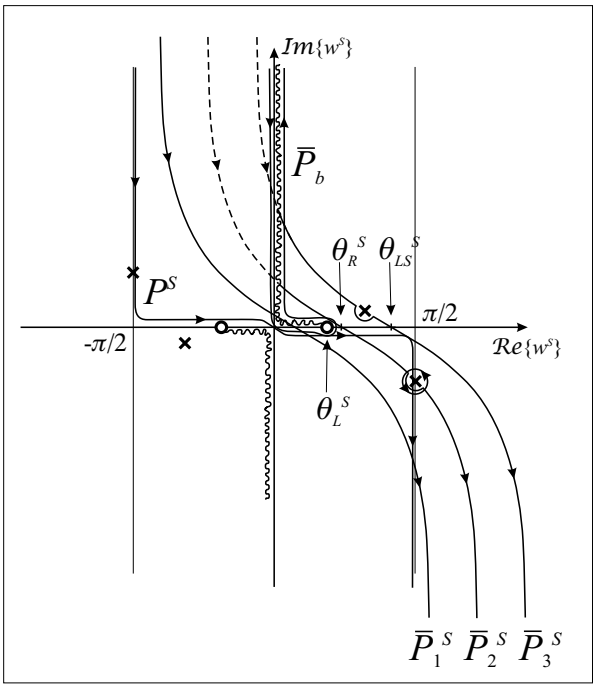
The path of steepest descent is a path of constant phase [73]. For integrals in the form of the ones in Eqs. (6.51)–(6.54), the steepest-descent path is given by

$$\mathcal{R}e\{w^{S,P}\} - \theta = \cos^{-1}(\mathrm{sech}(\mathcal{I}m\{w^{S,P}\})). \tag{6.61}$$

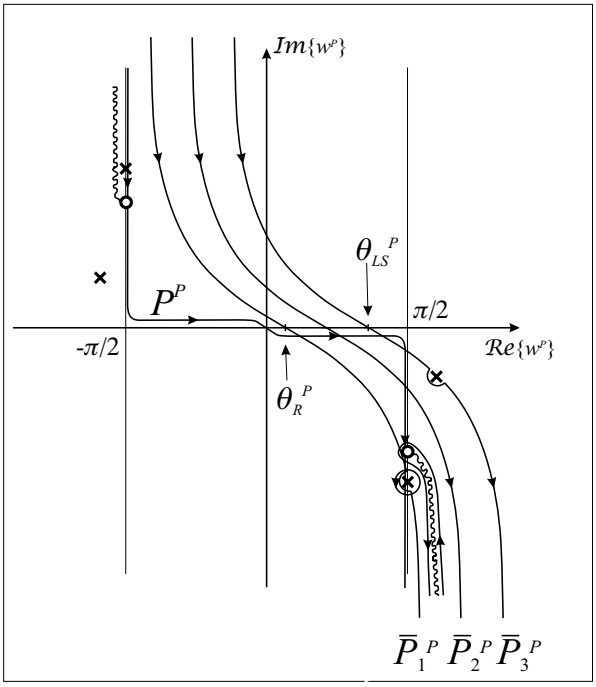
The procedure is the same for the shear wave terms and the pressure wave terms. In Fig. 6.10, three steepest-descent paths are shown each for the shear wave terms in the w^S-plane, $\bar{P}_1^S$, $\bar{P}_2^S$, $\bar{P}_3^S$, and for the pressure wave terms in the w^P-plane, $\bar{P}_1^P$, $\bar{P}_2^P$, $\bar{P}_3^P$. Each steepest descent path corresponds to a different propagation (polar) angle. The saddle point in each case is located at the intersection of the steepest-descent path with the real w^S- or w^P-axis, respectively. Physically, the contributions from the saddle points describe the pressure and the shear waves in the far field .

When the original integration path is deformed into the steepest-descent path, care has to be taken whether poles or branch cuts are crossed during the deformation. According to Cauchy's theorem, if a singularity is crossed during the deformation from one integration path into another, the contribution from the contour integral around the singularity must be included into the total integral. For example, when P^S in Fig. 6.10 is deformed into the steepest-descent path $\bar{P}_1^S$, no singularities are crossed during the deformation. However, for $\bar{P}_2^S$, the integrals around the branch cut, $\bar{P}_b$, and around the pole at w_R^S must be included. For $\bar{P}_3^S$, the branch cut integral as well as the integrals around the poles at w_R^S and w_{LS}^S contribute to the total integral.

The contour integrals around the singularities in both the w^S- and the w^P-plane give rise to different types of waves. The integral around the branch cut in the



(a)



(b)

Figure 6.10: Steepest descent paths for (a) the shear wave terms in the complex w^S-plane and (b) the pressure wave terms in the complex w^P-plane.

w^S-plane describes a lateral wave. The lateral wave is a plane shear wave induced by the pressure wave propagating along the surface. It appears only if the polar angle exceeds $\theta_L^S = \sin^{-1}(k_P/k_S)$, because, mathematically, the branch cut integral contributes to the total integral only for $\theta > \theta_L^S$. It can be shown that the integral around the branch cut in the w^P-plane is approximately zero and, thus, it does not contribute to the total integral. The integrals around the poles give rise to the Rayleigh surface wave and the leaky surface wave. They exist only for $\theta > \theta_R^S$ and $\theta > \theta_{LS}^S$ in the w^S-plane, and for $\theta > \theta_R^P$ and $\theta > \theta_{LS}^P$ in the w^P-plane. The total Rayleigh surface wave and the total leaky surface wave are comprised of the superposition of the contributions from the integrals around the singularities both in the w^S-plane and w^P-plane. The angles $\theta_R^{S,P}$ and $\theta_{LS}^{S,P}$ are easily obtained by inserting $w_R^{S,P} = \sin^{-1}(\xi_R/k_{S,P})$ and $w_{LS}^{S,P} = \sin^{-1}(\xi_{LS}/k_{S,P})$ into Eq. (6.61).

The integrals are now approximately determined using the method of steepest-descent. Five separate wave types arise: the bulk shear wave, the bulk pressure wave, the Rayleigh surface wave, the leaky surface wave and the lateral wave. A detailed description of the steepest-descent method is given, for example, by Felsen and Marcuvitz [73]. Using the method of steepest-descent, the shear wave in the far field comes out to be

$$\begin{aligned} u_x(R,\theta)|_{\text{Shear}} &= \frac{j}{\mu\pi}\sqrt{\frac{2\pi}{k_S R}}\,\frac{e^{j(k_S R-\pi/4)}}{F_0(k_S\sin\theta)}\cdot \\ &\quad \cdot 2\sqrt{k_P^2 - k_S^2\sin^2\theta}\;k_S^3\sin\theta\cos^2\theta \end{aligned} \tag{6.62}$$

$$\begin{aligned} u_z(R,\theta)|_{\text{Shear}} &= \frac{-j}{\mu\pi}\sqrt{\frac{2\pi}{k_S R}}\,\frac{e^{j(k_S R-\pi/4)}}{F_0(k_S\sin\theta)}\cdot \\ &\quad \cdot 2\sqrt{k_P^2 - k_S^2\sin^2\theta}\;k_S^3\sin^2\theta\cos\theta. \end{aligned} \tag{6.63}$$

For the pressure wave,

$$u_x(R,\theta)|_{\text{Pressure}} = \frac{-j}{\mu\pi}\sqrt{\frac{2\pi}{k_P R}}\,\frac{e^{j(k_P R-\pi/4)}}{F_0(k_P\sin\theta)}\cdot$$

$$\cdot \left[k_S^2 - 2k_P^2 \sin^2\theta\right]\ k_P^2 \sin\theta\cos\theta \tag{6.64}$$

$$u_z(R,\theta)|_{\text{Pressure}} = \frac{-j}{\mu\pi}\sqrt{\frac{2\pi}{k_P R}}\,\frac{e^{j(k_P R-\pi/4)}}{F_0(k_P\sin\theta)}\cdot$$

$$\cdot \left[k_S^2 - 2k_P^2 \sin^2\theta\right]\ k_P^2 \cos^2\theta. \tag{6.65}$$

The Rayleigh surface wave and the leaky surface wave are derived from the contour integral around the respective poles of the Rayleigh equation. Using contour integration, the Rayleigh wave is described by

$$u_x(R,\theta)|_{\text{Rayleigh}} = \frac{2\xi_R}{\mu F_0'|_{\xi_R}}\Big[U(\theta-\theta_R^S)\cdot(-2\alpha_R\beta_R)e^{j\beta_R R\cos\theta}$$

$$+U(\theta-\theta_R^P)\cdot(\beta_R^2-\xi_R^2)e^{j\alpha_R R\cos\theta}\Big]\cdot e^{j\xi_R R\sin\theta} \tag{6.66}$$

$$u_z(R,\theta)|_{\text{Rayleigh}} = \frac{2\alpha_R}{\mu F_0'|_{\xi_R}}\Big[U(\theta-\theta_R^S)\cdot 2\xi_R^2 e^{j\beta_R R\cos\theta}$$

$$+U(\theta-\theta_R^P)\cdot(\beta_R^2-\xi_R^2)e^{j\alpha_R R\cos\theta}\Big]\cdot e^{j\xi_R R\sin\theta}, \tag{6.67}$$

where

$$F_0'|_{\xi_R} = 4\xi_R(\xi_R^2-\beta_R^2) + 8\xi_R\beta_R\alpha_R - 4\xi_R^3\left(\frac{\beta_R}{\alpha_R}+\frac{\alpha_R}{\beta_R}\right). \tag{6.68}$$

$F_0'|_{\xi_R}$ is the derivative of the Rayleigh equation with respect to ξ at $\xi=\xi_R$, and $U(\theta-\theta_R^{S,P})$ is the Heaviside unit step function; α_R, β_R, and ξ_R describe the wave numbers of the Rayleigh wave (see Table 6.2). The result for the leaky surface wave is determined in exactly the same way and is obtained by simply replacing $\theta_R^{S,P}$, ξ_R, α_R and β_R by $\theta_{LS}^{S,P}$, ξ_{LS}, α_{LS} and β_{LS}, i. e., inserting the wave numbers for the leaky-wave pole instead of the Rayleigh wave pole.

The lateral wave is defined by the integral around the branch cut, $\bar{P}_b$ in Fig. 6.10. Following Felsen and Marcuvitz [73], the integral is asymptotically approximated to

become

$$u_x(R,\theta)|_{\text{Lateral}} = U(\theta-\theta_L^S)\cdot\frac{-2^{3/2}/(\mu\sqrt{\pi})}{[k_S R|\sin(\theta-\theta_L^S)|]^{3/2}}\cdot\frac{(\sin\theta_L^S)^{3/2}(\cos\theta_L^S)^{5/2}}{(2\sin^2\theta_L^S-1)^2}$$
$$\cdot\, e^{jk_S R\cos(\theta-\theta_L^S)+j3/4\pi} \tag{6.69}$$

$$u_z(R,\theta)|_{\text{Lateral}} = U(\theta-\theta_L^S)\cdot\frac{2^{3/2}/(\mu\sqrt{\pi})}{[k_S R|\sin(\theta-\theta_L^S)|]^{3/2}}\cdot\frac{(\sin\theta_L^S)^{5/2}(\cos\theta_L^S)^{3/2}}{(2\sin^2\theta_L^S-1)^2}$$
$$\cdot\, e^{jk_S R\cos(\theta-\theta_L^S)+j3/4\pi}. \tag{6.70}$$

6.3.3 Example

Eqs. (6.62)–(6.70) give the asymptotic far-field approximations for the wave fields excited by a harmonic line source on the surface. To determine the wave fields for a specific excitation in the time domain, the results must be transformed from the frequency domain into the time domain and convolved with the excitation function. The inverse-Fourier transform is given by

$$\hat{\mathbf{u}}(R,\theta,t) = \int_{\omega=-\infty}^{+\infty} G(\omega)\cdot\mathbf{u}(R,\theta,\omega)\cdot e^{-j\omega t}d\omega. \tag{6.71}$$

Here, $G(\omega)$ represents the Fourier transform of the excitation function. To obtain the particle velocity rather than the displacement, the displacement is differentiated with respect to time:

$$\hat{\mathbf{v}}(R,\theta,t) = \int_{\omega=-\infty}^{+\infty} G(\omega)\cdot\mathbf{u}(R,\theta,\omega)\cdot(-j\omega)e^{-j\omega t}d\omega. \tag{6.72}$$

In the following, a differentiated Gaussian pulse is used as excitation, with its Fourier transform

$$G(\omega) = -j\sqrt{2\pi}t_0^2\omega\cdot e^{0.5-0.5(\omega t_0)^2}, \tag{6.73}$$

where t_0 describes the width of the pulse.

The wave fields excited by a differentiated Gaussian pulse are computed for a material with a Poisson ratio of $\nu = 0.4$. The wave fields are calculated according to Eqs. (6.62)–(6.70) and then transformed into the time domain using Eq. (6.72). A differentiated Gaussian pulse with a center frequency of 400 Hz is used as excitation. In Fig. 6.11, the separate wave fields in the medium are plotted for one instant in time, ten milliseconds after their excitation. The plots show, from top to bottom, the shear wave, the pressure wave, the Rayleigh surface wave, the leaky surface wave, and the lateral wave. A logarithmic color scale is used, ranging from dark red (0 dB) over yellow and green to blue (-40 dB). The top of each plot coincides with the surface of the medium. The source is located on the surface, at the center of each plot.

The shear wave and the pressure wave exhibit cylindrical wave fronts. They both vanish at the surface. The Rayleigh surface wave is confined to the surface and decays into the z-direction. The leaky surface wave propagates along the surface with a speed greater than the one of the shear wave, but smaller than the speed of the pressure wave. It feeds an inhomogeneous plane shear wave. The angle that the shear wave makes with the surface is approximately defined by

$$\gamma_{LS} = \sin^{-1}(\mathcal{R}e\{\xi_{LS}\}/k_S). \tag{6.74}$$

Due to the coupling into the shear wave, the leaky wave leaks energy into the medium and decays in its propagation direction. The lateral wave propagates at an angle of approximately

$$\gamma_L = \theta_L^S = \sin^{-1}(k_P/k_S). \tag{6.75}$$

The artifacts that are especially visible for the lateral wave are due to the Fourier transform algorithm that is being used.

Fig. 6.12 shows the wave fields due to a line source on the surface as determined numerically with the finite-difference time-domain (FDTD) algorithm. Again, a material with a Poisson ratio of 0.4 is assumed, and the fields are plotted 10 milliseconds after their excitation. Clearly, the five different wave types are distinguishable. The differences between the FDTD result and the asymptotic approximation are mainly

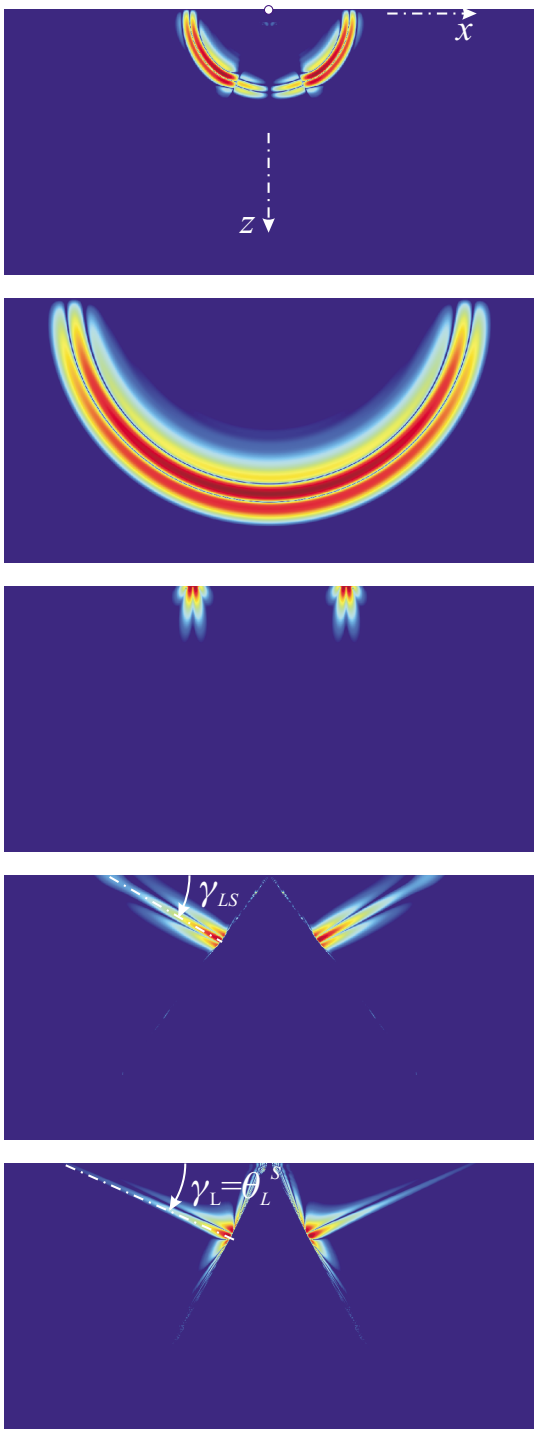


Figure 6.11: Waves due to a point source on the surface. From top to bottom: shear wave, pressure wave, Rayleigh surface wave, leaky surface wave, lateral wave.

due to the fact that the asymptotic approximation describes the waves in the far field, whereas the FDTD computations show the waves in the near field. Results essentially identical to the FDTD results have been obtained when integrating Eqs. (6.36) and (6.37) numerically rather than approximating the integrals asymptotically.

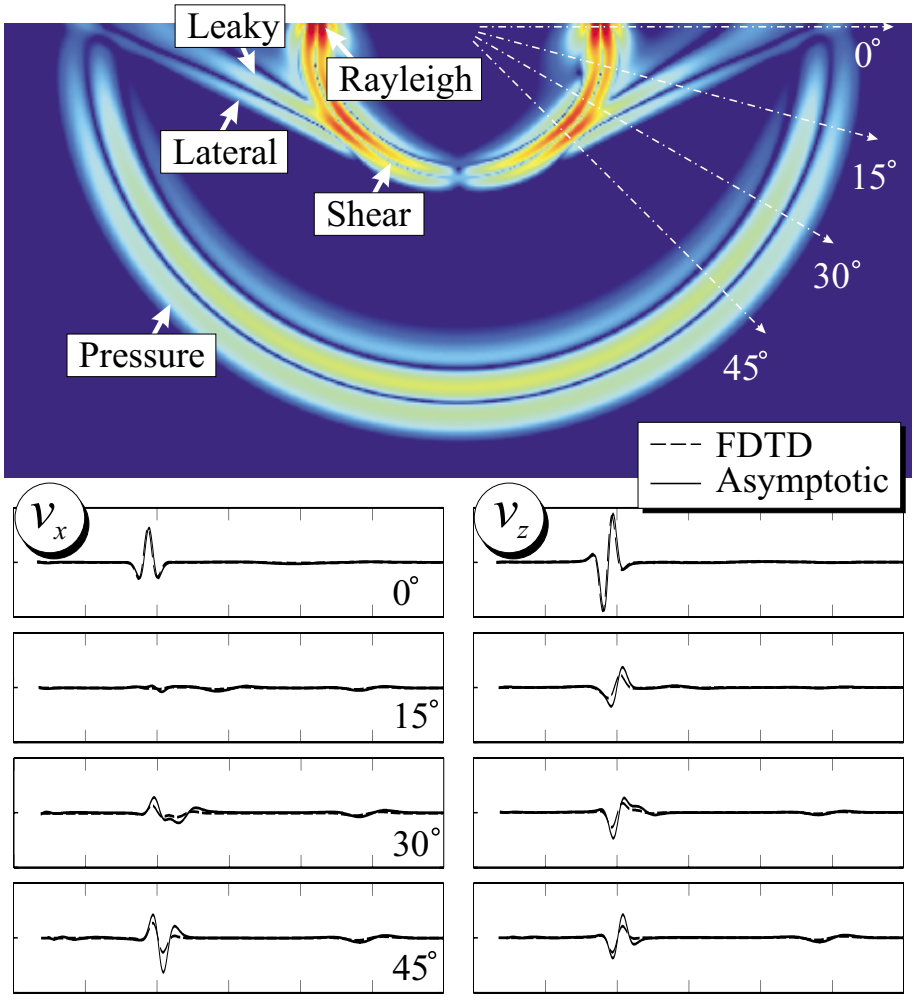


Figure 6.12: Finite-difference results; comparison to asymptotic solution.

To obtain a better picture of the behavior of the various waves at the surface, the particle motion due to the different surface waves is analyzed. For this, the wave

fields of the Rayleigh surface wave, the leaky surface wave and the lateral wave are computed using Eqs. (6.66)–(6.70) for harmonic time-dependence, and hodograms of the particle motion at the surface are generated. In these hodograms, the vertical displacement along the surface is plotted versus the horizontal displacement. The hodograms are shown in Fig. 6.13. As it is well-known, the particle motion due to a Rayleigh surface wave is retrograde (counterclockwise) in nature (Fig. 6.13 (a)). This is caused by a phase shift between the horizontal and the vertical displacement component: the horizontal displacement is leading the vertical component by 90 degrees in phase. The hodogram indicates that the Rayleigh wave does not decay as it propagates along the surface. The particle motion due to the leaky surface wave is prograde (clockwise), caused by the horizontal displacement lagging the vertical displacement in phase (Fig. 6.13 (b)). Clearly, the leaky surface wave decays as it travels along the surface. For the lateral wave, the displacement components are in phase, and the hodogram shows a diagonal line (Fig. 6.13 (c)). The lateral wave also decays as it propagates along the surface. The prograde and retrograde particle motions of surface waves have also been observed experimentally by, for example, Smith *et al.* [68].

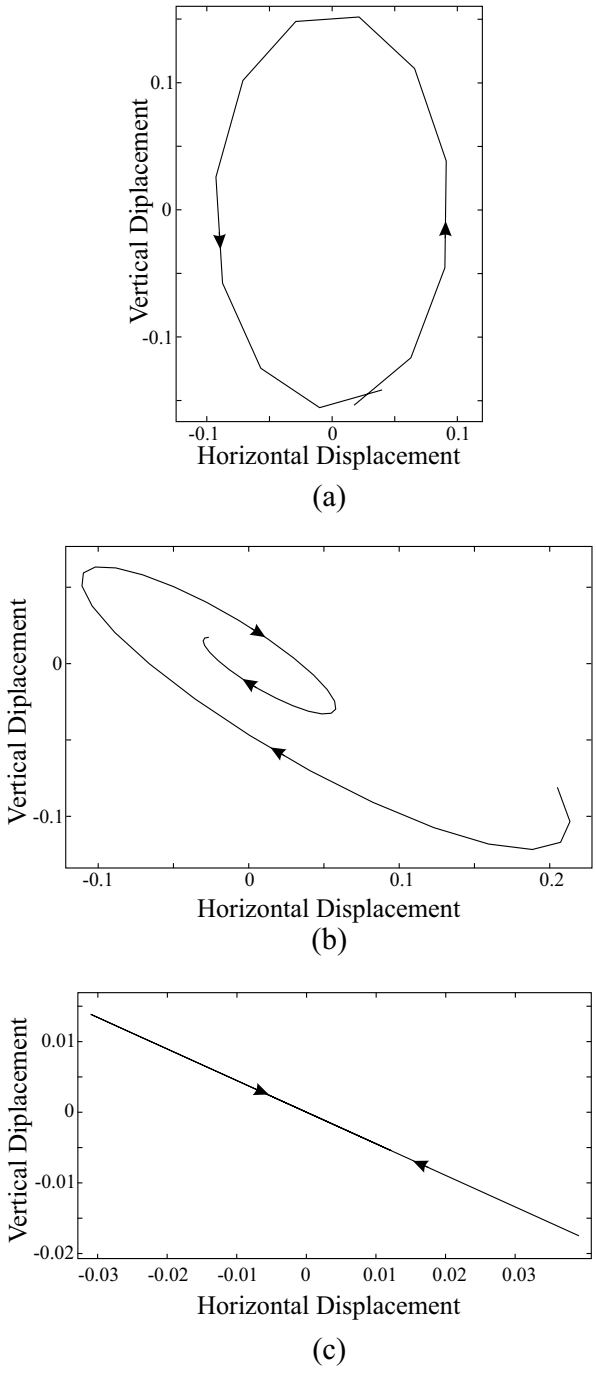


Figure 6.13: Hodograms of the particle motion at the surface. Plots for the vertical displacement vs. the horizontal displacement for (a) the Rayleigh surface wave, (b) the leaky surface wave, and (c) the lateral wave.

CHAPTER 7

Propagation and Scattering of Elastic Waves in the Ground

In laboratory experiments, inert land mines are buried in a large sand-filled tank. Elastic waves are excited, and the displacement of the surface above the mine is measured using a radar system. To analyze the experimental results, the experiments are duplicated with the numerical model. In this chapter, results obtained with the numerical model are presented. The chapter is divided into two main sections. In the first section, the numerical model is used to explore the propagation of elastic waves along the surface of the ground. In the second section, the interaction of elastic waves with buried anti-personnel and anti-tank mines is investigated.

For the numerical model, the soil is assumed to be linear, isotropic and lossless. In that case, the elastic properties of the soil can be completely described by three independent parameters: the material density, the shear wave speed, and the pressure wave speed. In order to model the soil and its behavior accurately within the numerical model, these parameters have been measured as a function of depth in the sand used in the experiments.

The measurements of the material properties and their results are described in detail in Chapter 9 and shall be summarized only briefly at this point. The wave speeds in the ground have been estimated by burying a vertical array of accelerometers. By using the accelerometers and displacing the transducer, the arrival times of

the wave fronts as a function of depth and distance are determined. From the arrival times, the wave speeds of the various wave types in the ground can be estimated.

The material density, the pressure wave speed and the shear wave speed as a function of depth are shown in Fig. 7.1. All three quantities have been found to be depth dependent. The material density has a value of roughly 1430 kg/m^3 at the surface and increases to about 1670 kg/m^3 at a depth of half a meter. The variation in the material density is due to changes in the water content within the soil. At the surface, the sand is fairly dry. Within the ground, the water content increases steadily, until the water table is reached at a depth of about half a meter. The pressure wave speed and the shear wave speed both also vary with depth. At the surface, a layer with fairly slow wave speeds exists. Beyond this layer, the wave speeds increase rapidly. At a depth of half a meter, the pressure wave speed and the shear wave speed have values of 210 m/s and 130 m/s, respectively.[1] The material properties of all other materials used in the following are summarized in Table 7.1. Note that the shear wave speed in air is zero, and, consequently, shear waves cannot propagate in air.

Table 7.1: Material properties.

	Shear Wave Speed c_S [m/s]	Pressure Wave Speed c_P [m/s]	Material Density ρ [kg/m^3]
Plastic	1100	2700	1200
Air	0	330	1.3
Rubber	500	800	1000

[1] The material properties are measured only to a depth of half a meter and are assumed to be constant deeper in the ground. This is a vague assumption and probably not true within the water table, but is justifiable, because, within the frequency range of interest, the variation of the material properties deeper in the ground will not have a significant impact onto the wave propagation at and close to the surface.

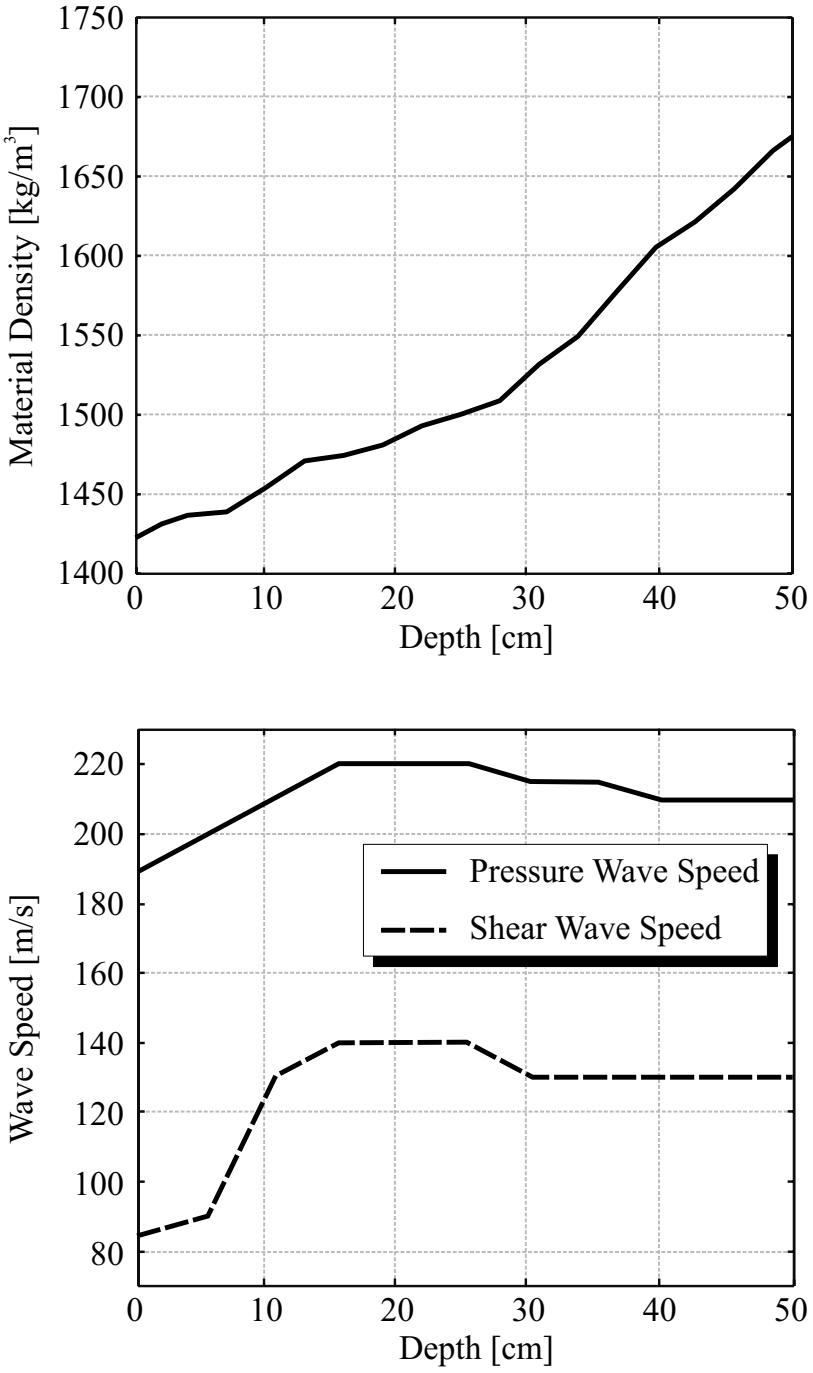


Figure 7.1: Variation of the material properties with depth.

In the experiments, the elastic waves are excited by a transducer placed on the surface of the ground. The transducer foot, which is in contact with the soil, has the shape of a bar. In the numerical model, the transducer is emulated by exciting the particle displacement component normal to the surface throughout an area that approximately equals the shaker foot in size and shape. The motion of the shaker foot has been measured using accelerometers and has been found to closely resemble a differentiated Gaussian pulse. For all results in this chapter, a differentiated Gaussian pulse with a center frequency of about 450 Hz is used as excitation.

Formally, the particle velocity is excited in the numerical model, and the particle velocity is computed as a function of time. Due to the linearity of the numerical model, the velocity fields due to the excitation of the particle velocity will be equal to the displacement fields, if the displacement rather than the velocity is excited. The model can be viewed as a linear system: the input is the particle velocity excitation, and the output are the computed particle velocity fields. If both input and output are integrated with respect to time, i. e., if transforming both the input and the output to the particle displacement, the results will not change. In other words, the particle displacement fields due to the excitation of the particle displacement with a differentiated Gaussian pulse are equivalent to the particle velocity fields due to the excitation of the particle velocity with a differentiated Gaussian pulse, because the displacement is just the time integral of the velocity. In the experiments, the particle displacement is measured and compared to the particle displacement fields as determined with the numerical model.

In the following, results are presented that are obtained with the numerical model in two and three dimensions. For both, a spacing of 0.5 cm between the nodes of the finite-difference grid is chosen, which satisfies the 10-nodes-per-wavelength requirement as stated in Eq. (4.32) for the frequency range of interest. The time increment is determined accordingly to satisfy the Courant condition. Because the maximum time increment according to the Courant condition depends on the maximum wave speed occurring in the model, different time steps are used for the different studies.

When modeling a mine, the maximum time increment is $\Delta t = 1.309 \cdot 10^{-6}$ s for the 2-D model and $\Delta t = 1.069 \cdot 10^{-6}$ s for the 3-D model. To compute the wave fields up to 30 ms, roughly 23000 time steps are necessary for the 2-D model and about 28000 for the 3-D model. The computations took about 45 minutes for the 2-D model (using a grid with 280×170 cells) on a 400-MHz single-processor PC, and about 92 minutes for the 3-D model (with $294 \times 50 \times 70$ cells) on a Beowulf computer cluster containing 28 500-MHz processors connected in parallel.

In this chapter, results obtained with the numerical model will be compared to experimental measurements. All measurements that are described here have been performed in an indoor laboratory at the Georgia Institute of Technology by W.R. Scott, G.D. Larson, and J.S. Martin. The measurement procedure is described in more detail in Chapter 9.

7.1 Surface Waves

Using the numerical model, the propagation of elastic waves in the ground and along the surface is studied. The simulation space spans a surface area of 1.67 m by 0.8 m and has a depth of 0.6 m. Two cases are considered: a homogeneous ground and an inhomogeneous ground. For the homogeneous case, the material properties are assumed to be constant throughout the ground. For the inhomogeneous case, the material properties are assumed to exhibit the depth dependence that has been measured experimentally (see Fig. 7.1).

First, the elastic wave propagation is determined assuming the ground to be homogeneous. The material density in the ground is set to 1400 kg/m^3, and the pressure wave speed and the shear wave speed are 250 m/s and 87 m/s, respectively. Fig. 7.2 shows waterfall graphs of the vertical particle displacement at the surface as measured experimentally and as determined numerically. In these graphs, the vertical particle displacement at a number of points along a line on the surface is plotted as a function of time. The time traces are offset along the vertical axis by the distance

from the source. The source is located at $x = 0$ m. Due to the fact that distance is plotted vs. time, the slopes of the progressing waves in the graph indicate the wave speeds. Thus, by determining the slope, the different wave types can be distinguished. A large slope corresponds to a fast wave, whereas a small slope indicates a slow wave.

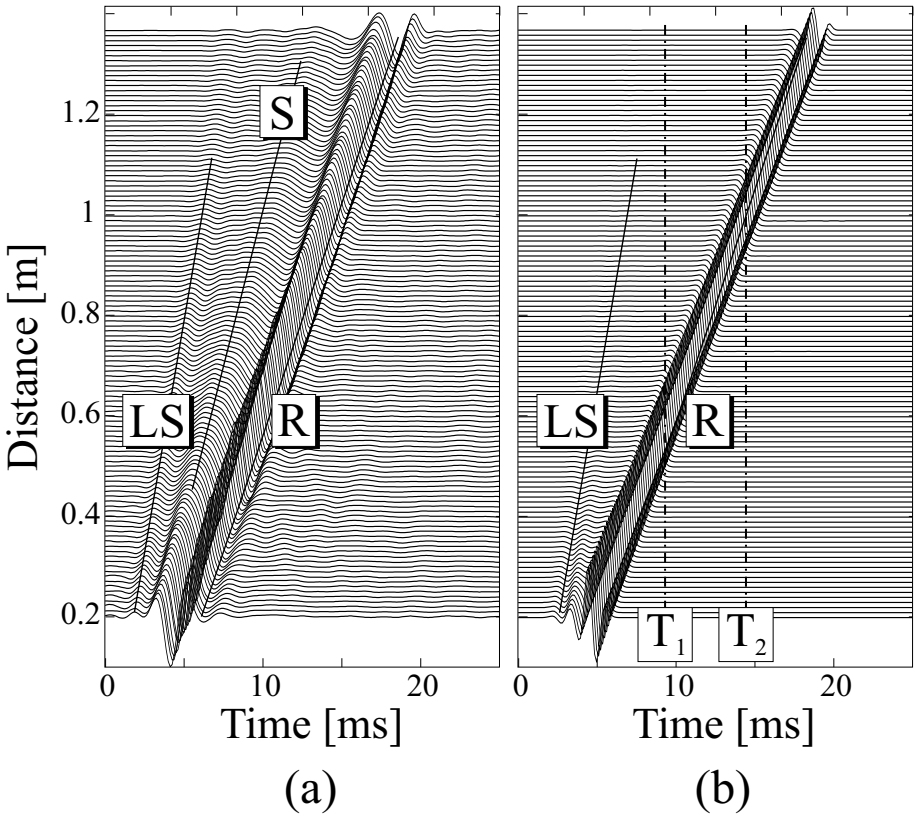


Figure 7.2: Waterfall graphs of the vertical particle displacement on the surface according to (a) experiment and (b) numerical simulation. The ground is assumed to be homogeneous in the numerical model.

The numerical model predicts two types of waves at the surface, indicated by LS and R (Fig. 7.2 (b)). The faster wave (LS) decays quickly. Its speed is determined from the graph to be roughly 180 m/s. The slower wave (R), on the other hand, barely decays, indicating that its energy is confined to the surface. This wave corresponds to the Rayleigh surface wave. From the graph, the average speed of the Rayleigh

wave comes out to be approximately 80 m/s. The experimental result (Fig. 7.2 (a)) looks quite different. A third type of wave is observed, having a wave speed along the surface larger than the Rayleigh wave, but slower than the fast wave. Also, the Rayleigh wave in the experiment is seen to be dispersed as it propagates along the surface, an effect that is not predicted by the numerical model if the ground is assumed to be homogeneous.

To distinguish and explain the waves that are observed on the surface of the ground, the particle displacement beneath the surface is studied. Fig. 7.3 shows the vertical particle displacement on the surface and on a cross section through the ground as computed with the numerical model. Again, the ground is assumed to be homogeneous. Pseudo-color plots are used to represent the wave fields, with a logarithmic scale ranging from red (0 dB) to blue (−50 dB). The wave fields are shown at two instants in time, T_1 and T_2, corresponding to the vertical lines in Fig. 7.2 (b). The source is located on the surface, to the left of the plot. At T_1, a Rayleigh surface wave (R), a shear wave (S) and a pressure wave (P) are seen to propagate. On the cross section an additional surface wave (LS) is observed. This surface wave corresponds to the leaky surface wave that has been discussed in length in Chapter 6. It travels along the surface with a wave speed faster than the shear wave, but slower than the pressure wave. Because it is faster than the shear wave, the leaky surface wave couples into a plane shear wave. Due to these coupling effects, the leaky surface wave decays rapidly in its propagation direction. On the surface plane, the leaky surface wave is clearly visible, whereas the pressure wave is not recognizable on the 50-dB scale. This suggests that the fast wave that has been observed in Fig. 7.2 in both experiment and numerical simulation corresponds to the leaky surface wave. At a later time, at T_2, the waves have propagated farther. The leaky surface wave has mainly decayed, and the pressure wave has left the range of the plot. On the surface only the Rayleigh wave is visible, whereas on the cross section both the Rayleigh wave and the shear wave are evident.

In Chapter 6, the elastic waves at and close to the surface have been determined

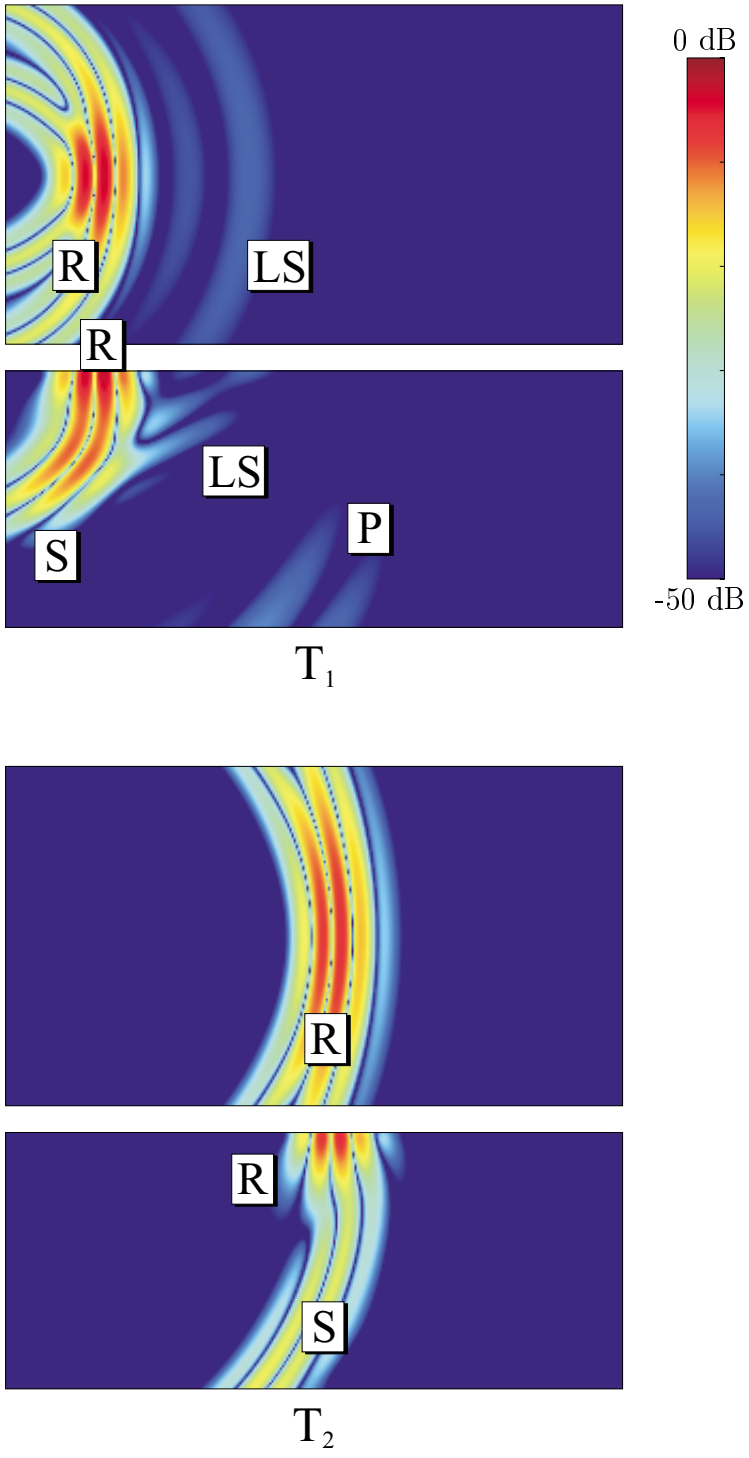


Figure 7.3: Normal particle displacement on the surface (top) and on a cross section through the ground (bottom). The ground is assumed to be homogeneous. T_1 and T_2 correspond to the vertical lines in Fig. 7.2 (b).

analytically. It has been found that a leaky surface wave, similar to the one observed in Fig. 7.3, will be excited by a line source on the surface of the ground. The leaky surface wave will cause a prograde particle motion on the surface, distinguishing it from the Rayleigh surface wave, which causes a retrograde motion. To verify this for the results shown in Fig. 7.3, the particle motion due to the waves on the surface is analyzed. In Fig. 7.4, the particle displacement is plotted at one point on the surface (60 cm from the source) as a function of time. Two hodograms are shown beneath the graph. In these hodograms, the vertical displacement is plotted versus the horizontal displacement and, thus, the actual particle motion is traced. The hodogram on the left describes the leaky surface wave, the hodogram on the right corresponds to the Rayleigh surface wave. As predicted, the leaky surface wave describes a prograde (clockwise) motion, whereas the Rayleigh surface wave causes a retrograde (counterclockwise) motion. The prograde particle motion is caused by the vertical displacement leading the horizontal displacement in phase. The retrograde motion, on the other hand, is due to the vertical displacement lagging behind in phase. Note that the pressure wave arrives prior to the leaky wave (see the magnified portion in Fig. 7.4). Both the pressure wave and the leaky wave are much weaker on the surface than the Rayleigh wave. The particle motion due to the pressure wave is linear, i. e., the horizontal and vertical component are in phase.

Experimental observations indicate that the Rayleigh wave is dispersed as it propagates along the surface. This effect is believed to be mainly due to the material properties in the ground varying with depth. The Rayleigh wave decays exponentially into the ground. The rate of decay depends on frequency: high-frequency components will decay fairly quickly, whereas low-frequency components reach deep into the ground. If the material properties vary with depth, the different frequency components will experience different material properties and, thus, different wave speeds. Consequently, the Rayleigh wave is dispersed.

Fig. 7.5 shows results that are obtained when implementing the depth variation that has been measured for the soil used in the experiments. Both in the experiment

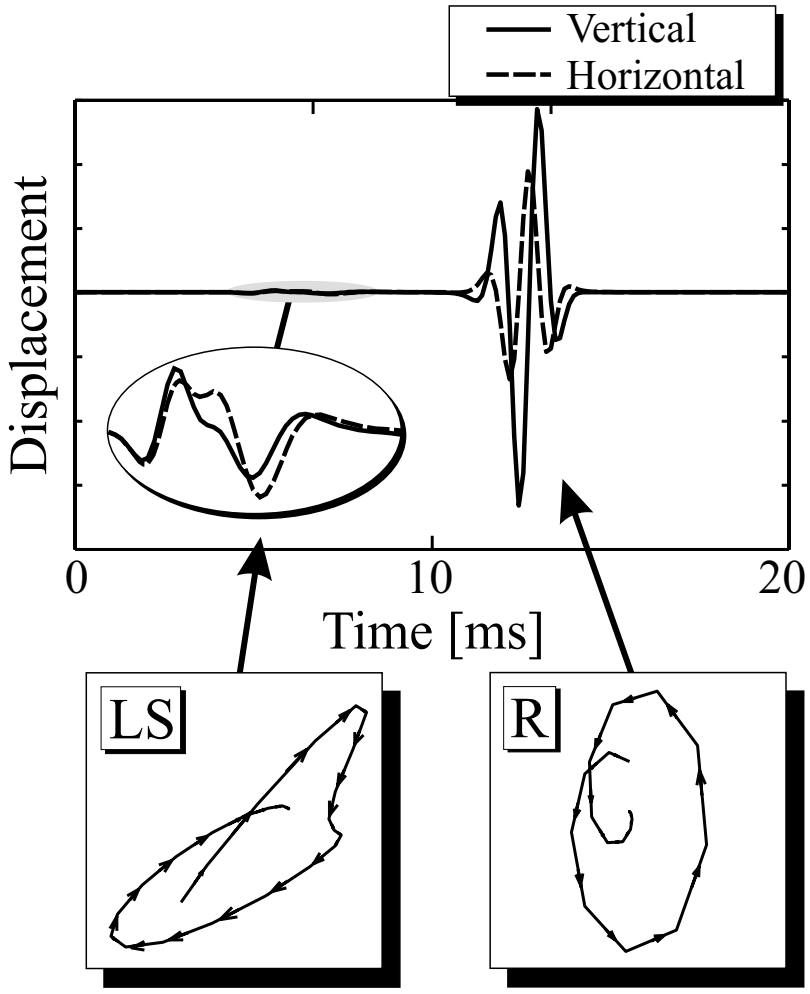


Figure 7.4: Top: particle motion on the surface at a distance of 60 cm from the source as a function of time. Bottom: vertical vs. horizontal displacement for the leaky surface wave and the Rayleigh surface wave. Homogeneous case.

and in the numerical model, three types of waves are observed, indicated by LS, S and R. The fastest wave of the three, the leaky surface wave (LS), decays quickly. Its speed is determined to be approximately 180 m/s. The second-fastest wave (S) has a speed of roughly 120 m/s and is fairly weak at the surface. Its origin cannot be clearly identified from this graph. The slowest wave is the Rayleigh surface wave (R) with a wave speed of about 80 m/s. Clearly, the Rayleigh wave is dispersed as it propagates. The agreement of experiment and numerical simulation is fairly good. The differences in Fig. 7.5 are believed to be mostly due to non-linear effects occurring in the sand. The non-linearities dampen the high-frequency components and, in this way, cause the pulse to broaden and to change in shape. By decreasing the amplitude of the excitation in the experiment, the agreement of experiment and numerical simulation can be improved. However, the signal-to-noise ratio of the radar output in the experiment will also decrease and, thus, the noise content of the data will increase.

Fig. 7.6 shows the vertical particle displacement on the surface and on a cross section through the ground as computed with the numerical model for the inhomogeneous ground. The wave fields are shown at two instants in time, T_1 and T_2. It is evident that, by assuming the material properties in the ground to be inhomogeneous, the wave fields within the ground are significantly altered. At T_1, the Rayleigh surface wave (R), the shear wave (S), the pressure wave (P) and the leaky surface wave (LS) are seen. As suggested by the depth profile, the shear wave is faster within the ground than for the homogeneous case. At a later time, at T_2, the leaky surface wave has decayed. Due to the rapid increase of the shear wave speed beyond the surface layer, a portion of the shear wave is refracted back towards the surface, and a guided shear wave arises. It is this guided shear wave that is the second wave evident at the surface in Fig. 7.5. The guided shear wave is strongly visible at the surface. It is faster than the Rayleigh surface wave, but slower than the shear wave within the ground.

In Fig. 7.7, the particle motion due to the waves on the surface is analyzed for

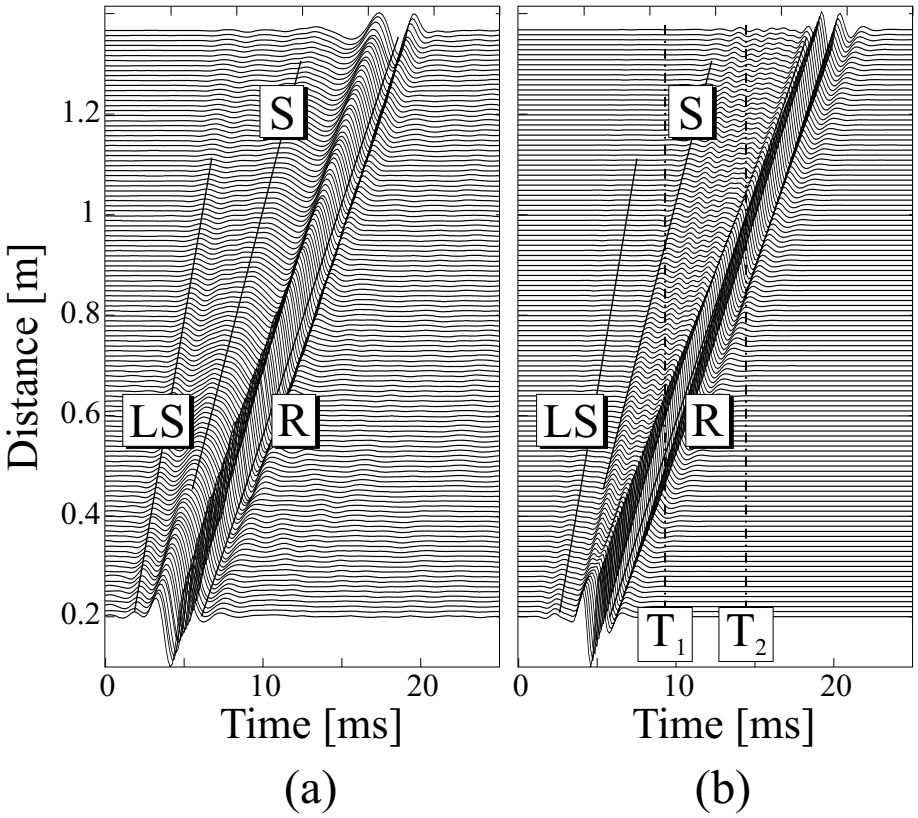


Figure 7.5: Waterfall graphs of the vertical particle displacement on the surface according to (a) experiment and (b) numerical simulation. The material properties in the numerical model are assumed to vary with depth.

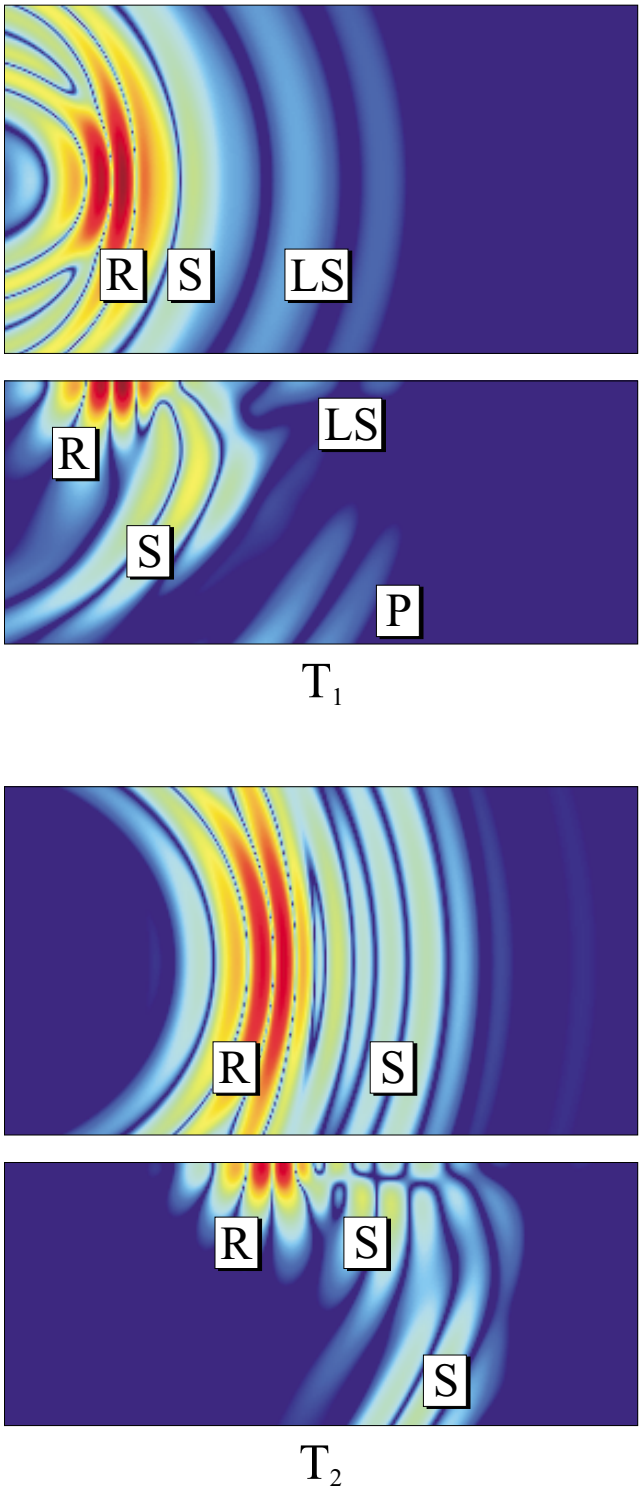


Figure 7.6: Normal particle displacement on the surface (top) and on a cross section through the ground (bottom). T_1 and T_2 correspond to the vertical lines in Fig. 7.5 (b). Depth-varying material properties are assumed.

the inhomogeneous case. The particle displacement is again plotted at one point on the surface (60 cm from the source) as a function of time. As for the homogeneous case, the leaky surface wave causes a prograde particle motion, whereas the Rayleigh surface wave induces a retrograde motion. The guided shear wave is clearly visible. It arrives ahead of the Rayleigh wave and has a predominantly vertical particle motion.

7.2 Interaction of Elastic Waves with Buried Land Mines

The interaction of elastic waves with buried land mines is to be investigated. For this, the mine-wave interaction is computed with the numerical model both in 2-D and in 3-D and compared to experimental results.

In the experimental system, the waves are excited from a transducer on the surface and detected by a radar system mounted above the surface. Fig. 7.8 shows a schematic drawing of two possible set-ups for such a system. In the first set-up, the waves are excited by a transducer (the *Source*) and recorded by an array of receivers placed at some distance from the transducer (top of Fig. 7.8). The array consists of a number of receivers, which simultaneously record their signal, yielding a fast scan rate and a small detection time. In another set-up, the array is realized as a synthetic receiver array. Hence, only one receiver is moved across the surface and records the signal that is excited repeatedly from the transducer. This second set-up (bottom of Fig. 7.8) yields a much more cost-effective system, but prolongs the scan time enormously. Currently, the second set-up is implemented in the experimental sensor system [14, 12]. A radar is used as receiver. The radar scans across the surface and measures the displacement of the surface as a function of time.

To investigate the interaction of elastic waves with a buried anti-personnel mine, a very simple model is used to approximate the mine's structure within the numerical

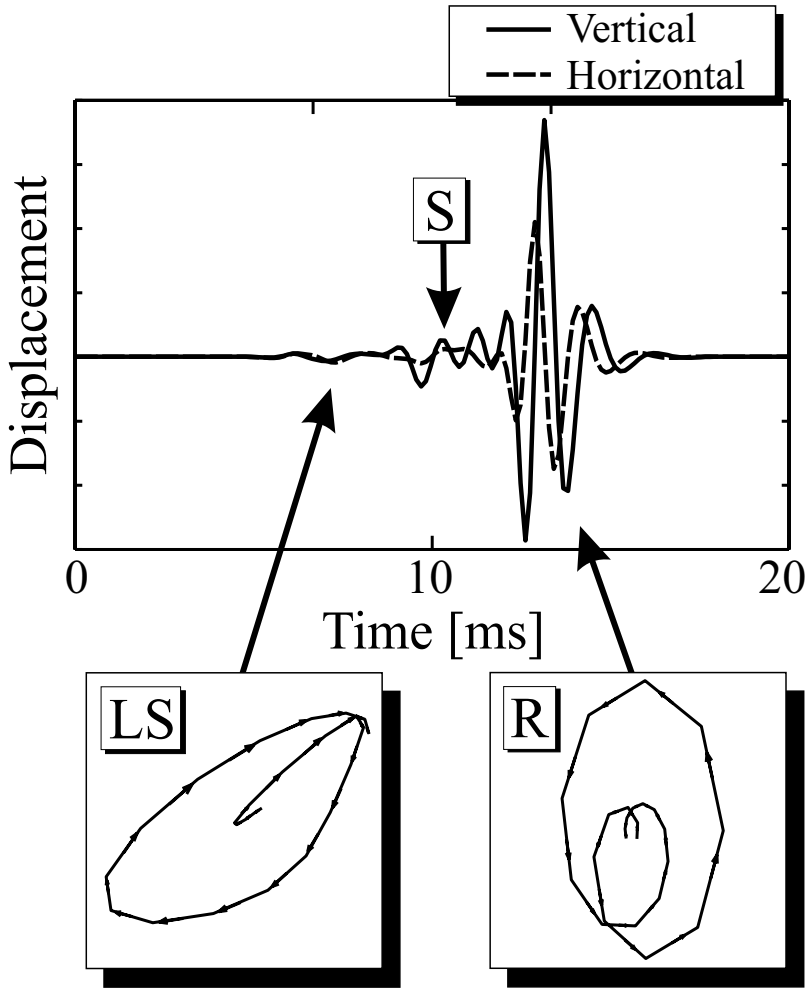


Figure 7.7: Top: particle motion on the surface at a distance of 60 cm from the source as a function of time. Bottom: vertical vs. horizontal displacement for the leaky surface wave and the Rayleigh surface wave. Depth-varying material properties are assumed.

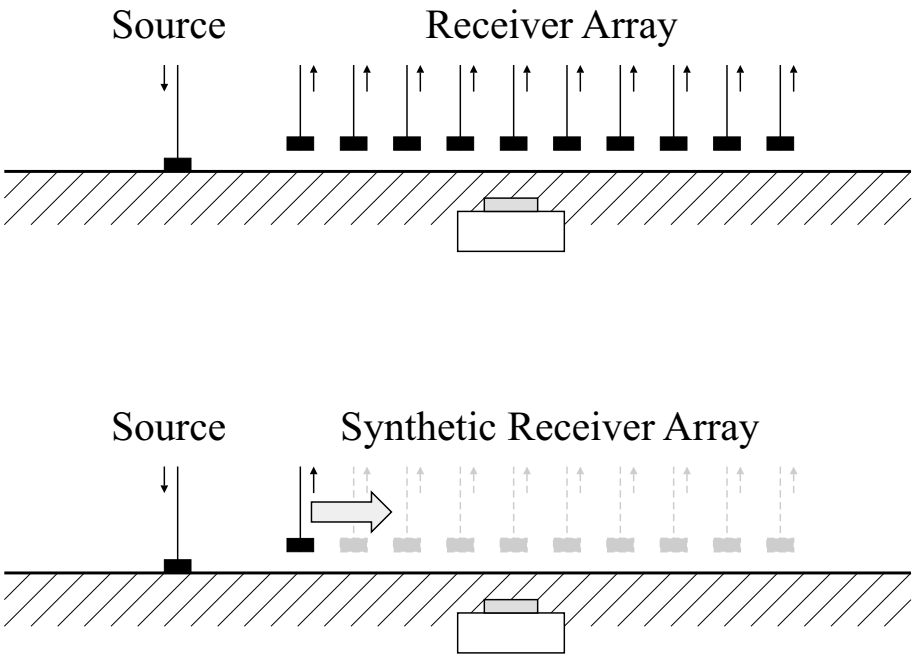


Figure 7.8: Schematic drawing of possible set-ups for the experimental system.

model. Fig. 7.9 shows the simple mine model, together with a simplified cross sectional drawing and a photograph of a TS-50 anti-personnel mine. The simple mine model consists of a large plastic chamber, and a small air-filled chamber. In the cubic finite-difference grid, the cylindrical shape of the mine is approximated by cubes, leading to the stair-case form evident in Fig. 7.9. The air-filled chamber is inserted into the model to approximate the structure of the TS-50 mine. A real land mine has a fairly complex structure, containing explosives, a trigger mechanism and several chambers.

7.2.1 2-D Analysis

As a first approach, the 2-D model is used to study the interaction of elastic waves with a buried anti-personnel mine. The model for the mine used in the 2-D model, as

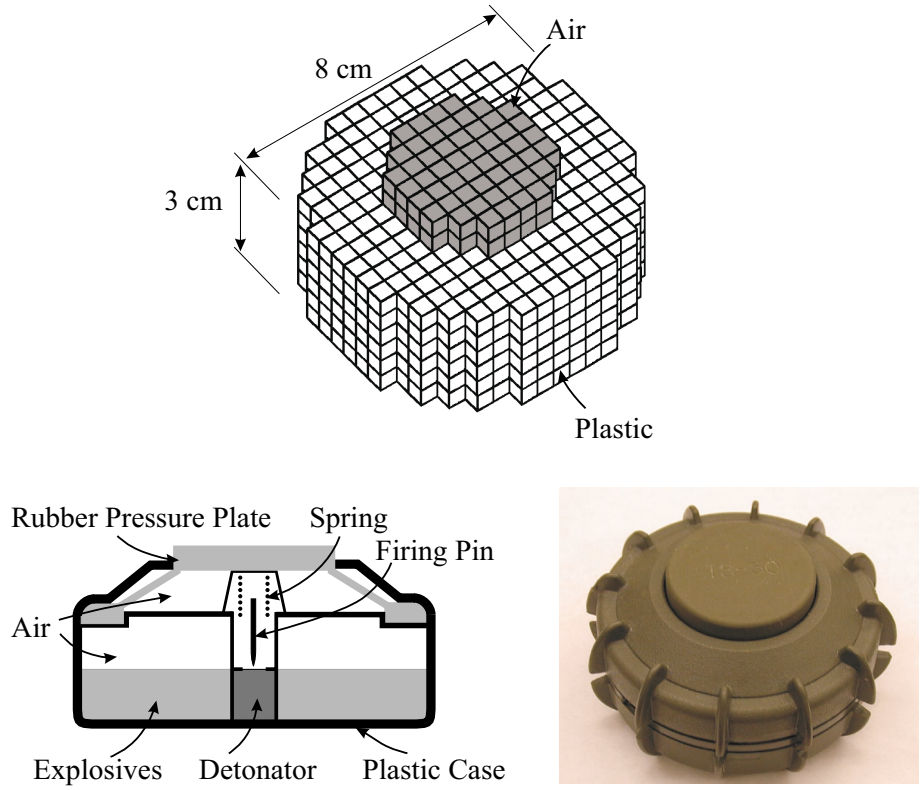


Figure 7.9: Simple model for the TS-50 anti-personnel mine; cross-sectional drawing and photograph of a real TS-50 mine.

shown in Fig. 7.10, is the 2-D projection of the 3-D mine model. The mine is buried 2 cm beneath the surface of the ground, at a distance of 60 cm from the source. A point source placed on the surface excites the vertical particle displacement using a differentiated Gaussian pulse with a center frequency of 450 Hz. The elastic wave fields are computed on a grid containing 280 by 170 cells, including a PML with a thickness of 10 cells and, consequently, yielding an effective solution space of 130 cm by 80 cm. For the 2-D model, the material properties in the ground are assumed to be homogeneous: $\rho = 1400$ kg/m^3, $c_P = 250$ m/s, and $c_S = 87$ m/s.

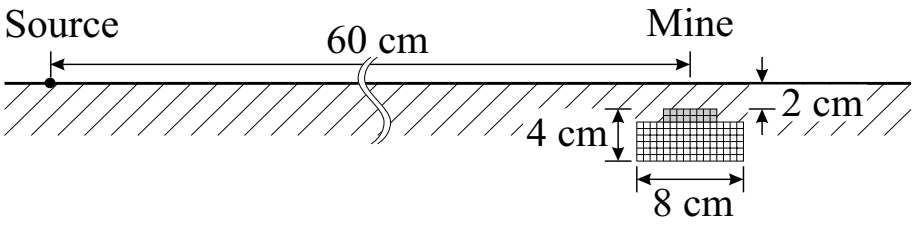


Figure 7.10: 2-D model for an anti-personnel mine.

Figure 7.11 shows the elastic wave fields on a cross section through the ground as computed with the 2-D model. As earlier, the vertical particle displacement is plotted using a logarithmic color scale with a dynamic range of 50 dB. The wave fields are shown at four different times. At T_1, a pressure wave (P), a shear wave (S), a Rayleigh surface wave (R) and a leaky surface wave (LS) are observed. The pressure wave, the fastest of the waves, just hits the mine. At T_2, the Rayleigh wave (R) has reached the mine. At T_3, the incident waves are seen to be scattered. Cylindrical wave fronts arise. Scattered Rayleigh waves as well as shear waves are induced. At T_4, the incident waves have passed the mine. However, some energy appears to be trapped above the mine, and continuous resonant oscillations occur. Due to these oscillations, waves are radiated from the mine location.

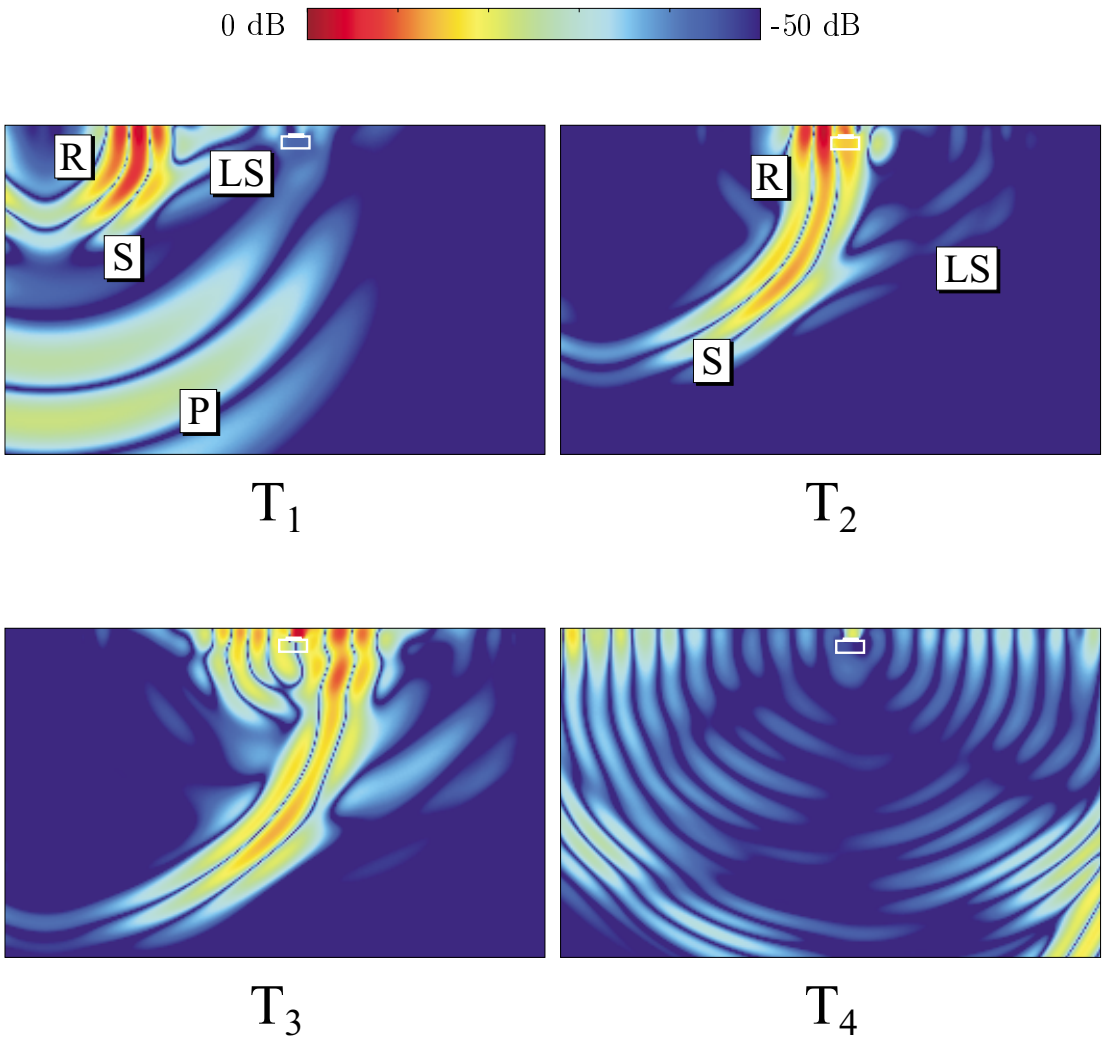


Figure 7.11: Interaction of elastic waves with a buried anti-personnel mine; 2-D results.

To explore the origin of the resonance and its nature, the mine is elongated within the numerical model, as shown in Fig. 7.12. The modified mine model has a length of 54 cm and an air chamber that is 50 cm long. Fig. 7.13 shows the

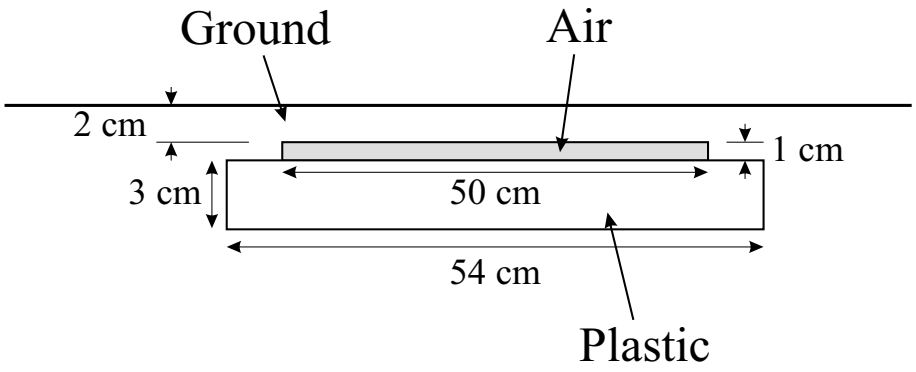


Figure 7.12: Model to determine the nature of the resonance: elongated mine.

vertical particle displacement on the surface in a waterfall graph as obtained with the modified model. A leaky surface wave (LS) and a Rayleigh wave (R) are seen to be incident on the mine. Both waves hit the mine and are partially reflected and partially transmitted. Above the mine, a complex pattern of propagating waves arises. This can be explained as follows. When the incident waves interact with the mine, they couple primarily into three different wave portions: into reflected waves (rR), into waves which are transmitted through the plastic body underneath the air chamber (M), and into waves which propagate within the thin soil layer between the surface and the air chamber (L and TF). Because the plastic body lies underneath the air chamber, the waves transmitted through the plastic are not visible on the surface. This is evident in Fig. 7.13: a wave arises beyond the mine, apparently without being caused by any wave visible on the surface above the mine. The path the wave takes through the plastic is indicated by M. The Rayleigh wave couples a significant amount of energy into the thin layer above the air chamber. Two different kinds of waves

arise within the thin layer: a longitudinal wave (L) and a transverse flexural wave (TF). The longitudinal wave is the faster one of the two. It propagates within the thin layer between surface and air chamber (L), is reflected at the edge of the air chamber, travels back within the layer and is reflected again. The transverse flexural wave also travels within the layer between surface and air chamber (TF), is reflected, travels back and is again reflected. Note that the flexural waves are almost completely reflected at the edges of the air chamber, and the energy remains within the thin layer above the mine. Figure 7.14 explains schematically the paths the various waves follow through and around the mine.

Using the modified model, the resonant behavior of the mine in the numerical model can be explained. When the waves interact with the mine, they are partially reflected and partially transmitted. However, a large portion of the energy couples into longitudinal and transverse-flexural waves in the thin soil layer above the mine. Due to the small size of the mine, these waves form standing waves, giving rise to the resonant oscillations observed in Fig. 7.11.

From the slopes of the traveling waves in Fig. 7.13, the wave speeds of the longitudinal wave (L) and the transverse flexural wave (TF) are approximately determined to be $c_L \approx 160$ m/s and $c_{TF} \approx 50$ m/s, respectively. A simple theory for waves propagating in a thin plate with clamped edges predicts the wave speeds of the longitudinal wave to be [58]

$$c_L = \left(\frac{E}{\rho(1-\nu^2)} \right)^{1/2}, \tag{7.1}$$

and the wave speed of the transverse flexural wave to be

$$c_{TF} = \sqrt{h\omega \left(\frac{E}{12\rho(1-\nu^2)} \right)^{1/2}}, \tag{7.2}$$

where E and ν are Young's modulus and Poisson's ratio (see Appendix D); ρ is the material density, and h is the thickness of the thin plate. For the longitudinal wave, this yields a wave velocity of $c_L = 163$ m/s. For the transverse flexural wave with a center frequency $f_c = 450$ Hz and a plate thickness of $h = 2$ cm, a wave speed of

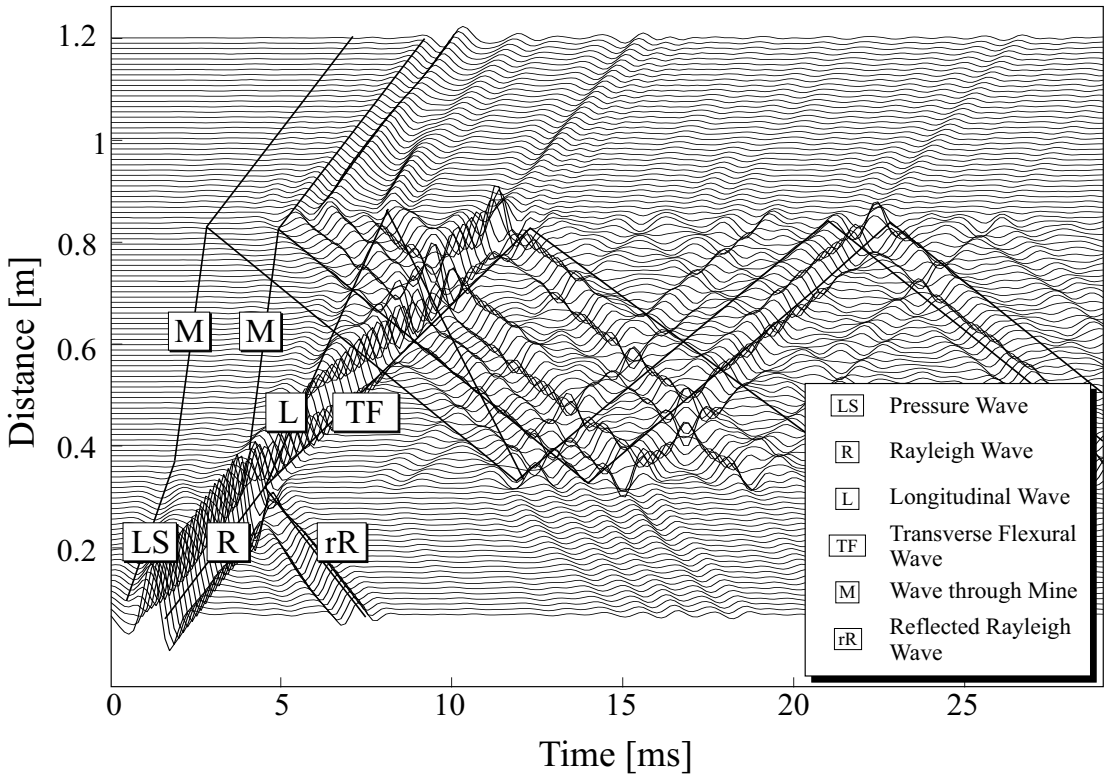


Figure 7.13: Flexural waves propagating above the mine.

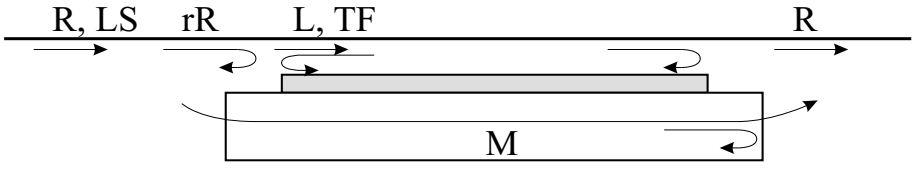


Figure 7.14: Wave paths through and around the mine. The incident Rayleigh wave (R) and leaky surface wave (LS) couple into a longitudinal wave (L) and a transverse flexural wave (TF) in the thin plate above the mine and into a wave passing through the plastic body of the mine (M).

$c_{TF} = 52$ m/s is predicted. This is in good agreement with the wave speeds obtained from the numerical simulation. Note that the wave speed of the transverse flexural wave is dependent on frequency, while the speed of the longitudinal wave is not. Thus, the transverse wave is strongly dispersive. The dispersion of the transverse flexural wave can be clearly seen in Fig. 7.13.

7.2.2 3-D Analysis

The experimental model is inherently three-dimensional, and, thus, a 3-D model is needed to adequately simulate the experiments. Using the 3-D model, the interaction of elastic waves with buried mines is investigated as a function of burial depth, and the results obtained with the numerical model are compared to experimental results.

Anti-Personnel Mine

With the 3-D model, the wave fields are computed throughout a solution space with a surface area of 1.67 m by 0.8 m and a depth of 0.4 m. The mine model as shown in Fig. 7.9 is used, and the numerical results are compared to experimental results. Both in experiments and numerical simulations, the mine is placed at a distance of 97 cm from the source. The material properties are assumed to vary with depth according to Fig. 7.1.

Fig. 7.15 shows pseudo-color plots of the vertical particle displacement on the surface plane (top) and on a cross section through the ground (bottom). The color scale ranges from red (0 dB) to blue (-50 dB). The wave fields are shown at four instants in time. The mine is placed 2 cm beneath the surface. At T_1, the guided shear wave (S) just hits the mine. On the cross section, the shear wave (S) and the Rayleigh wave (R) are seen. The leaky surface wave is faintly visible on the surface plane (LS). The pressure wave has already left the range of the plot. At T_2, the Rayleigh wave has reached the mine. Scattered cylindrical wave fronts appear at the surface, which are reflections from the incident guided shear wave. At T_3, scattered

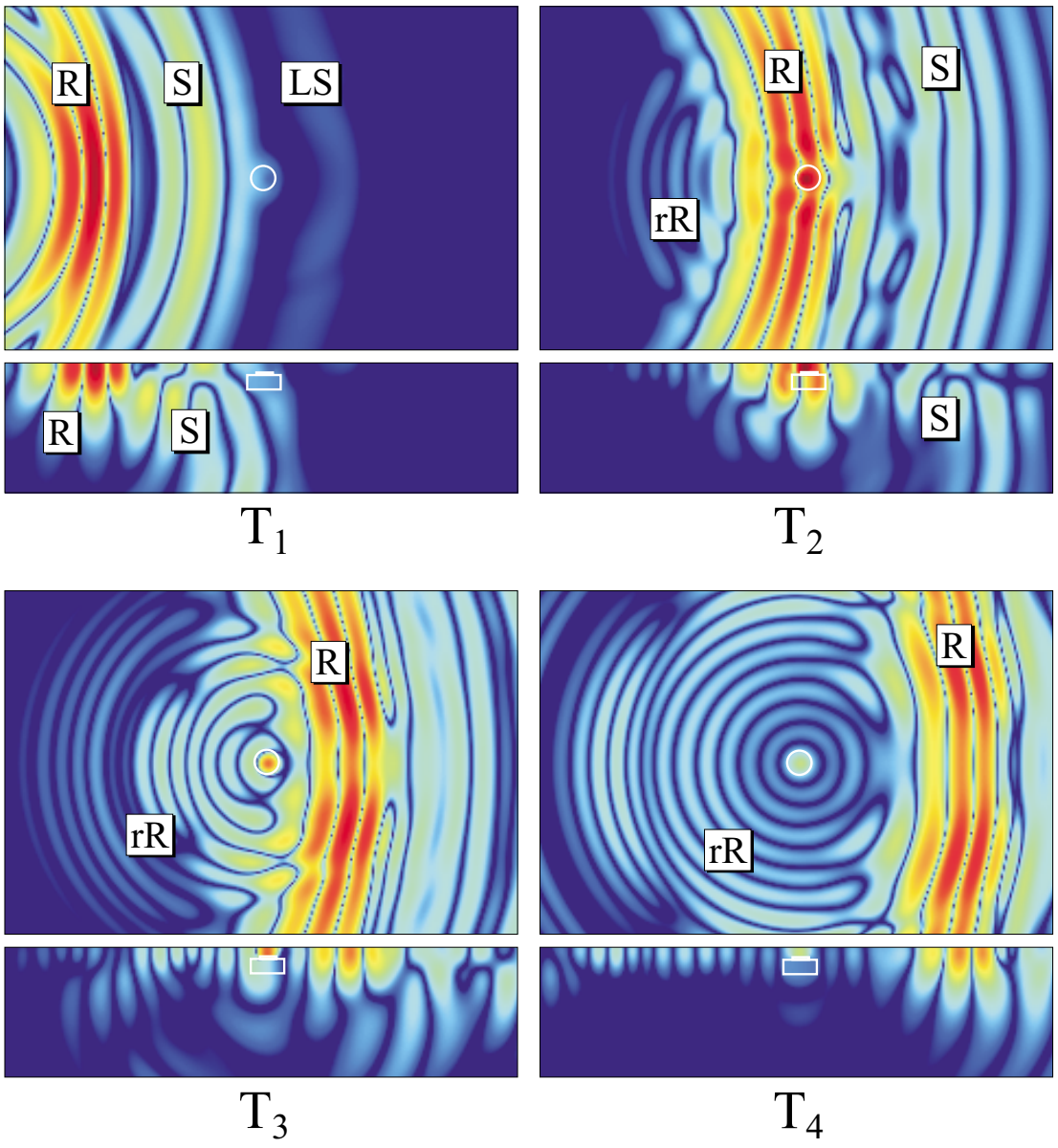
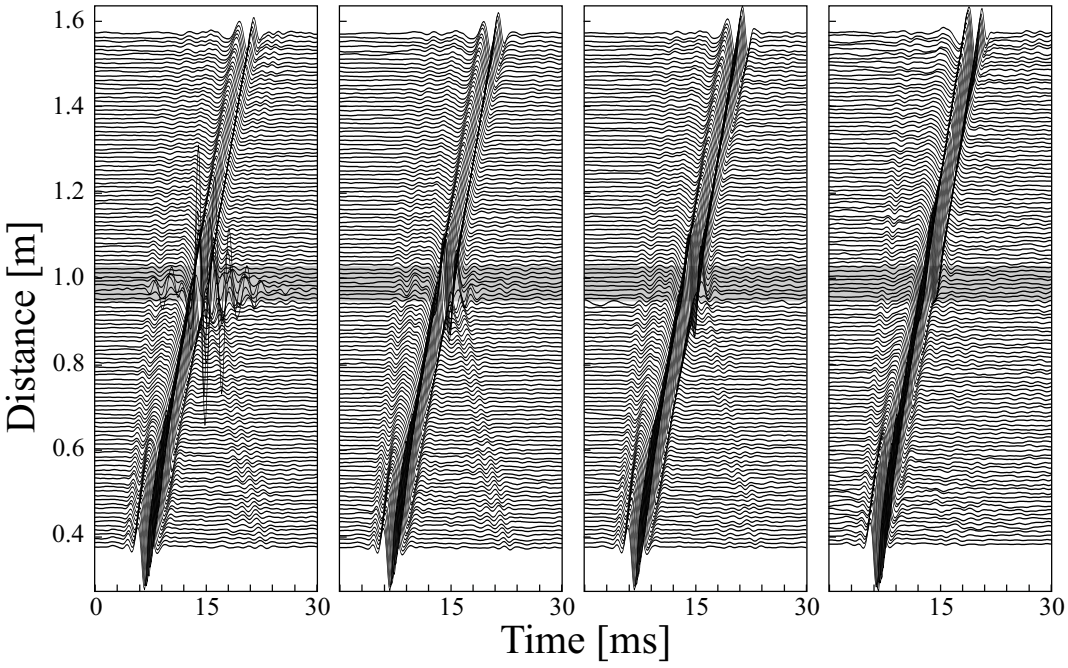


Figure 7.15: Interaction of elastic waves with a buried anti-personnel mine; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instants in time.

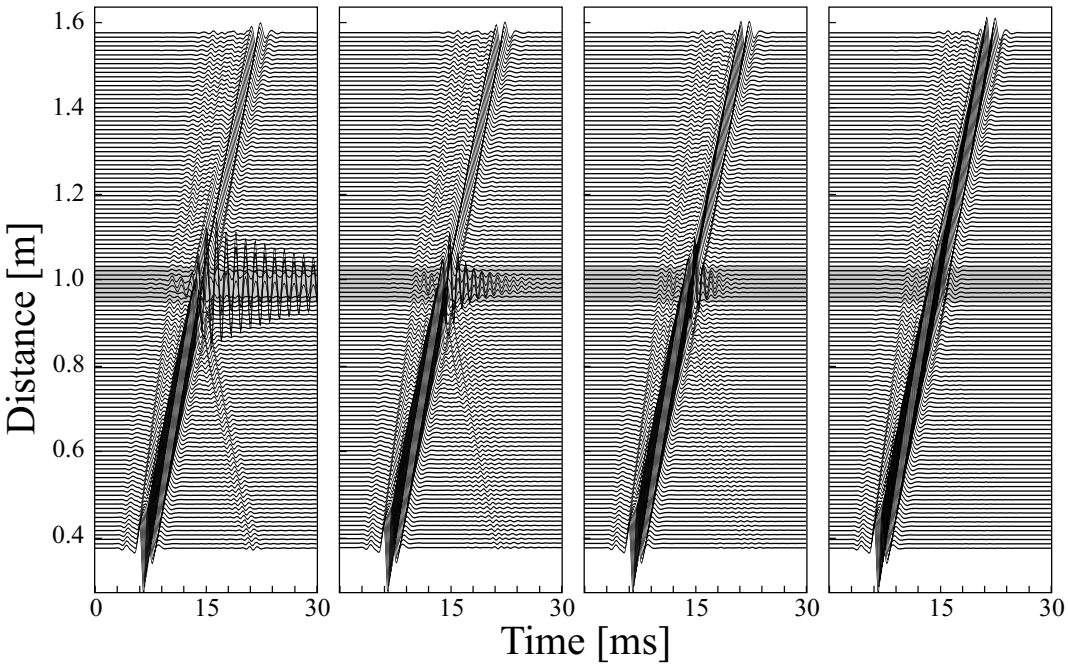
waves from both the guided shear wave and the Rayleigh wave are visible. The leading wave fronts are reflections from the incident guided shear wave. The strong waves that follow are due to the Rayleigh wave. At T_4, resonant oscillations are visible at the location of the buried mine. In cylindrical wave fronts, energy is radiated out. As observed with the 2-D model, the resonant oscillations are confined to the thin soil layer above the mine. Due to these oscillations, the mine location can be clearly identified on the surface.

To compare numerical to experimental results, the vertical surface displacement is plotted in a set of waterfall graphs. Fig. 7.16 shows these waterfall graphs for (a) experimental and (b) numerical results. Mines at four different depths are considered, at 1 cm, 2 cm, 3 cm, and 6 cm (from left to right). The mine, indicated by the gray shading, is located at a distance of 97 cm from the source. For the mine at 1 cm, strong oscillations are visible at the mine location. In the experiment, the amplitude of the oscillations appears larger, but the oscillations decay faster than in the numerical simulation. For the mines at 2 and 3 cm, the oscillations are weaker, but still clearly visible. For the mine at 6 cm, the oscillations are very weak and cannot be seen in this graph. However, by applying signal processing techniques the effects of the mine can be enhanced, making it easier to locate the buried mine [12].

The agreement between the numerical and experimental results is surprisingly good considering the uncertainties of the material parameters for the ground in the experimental model and the simplifications made in the numerical model. For the numerical model, it has been shown with the 2-D model that the oscillations at the mine location are caused by the incident waves coupling into flexural waves in the thin soil layer above the mine. These flexural waves are confined to the thin layer and form a standing wave pattern, giving rise to the resonant oscillations. As the burial depth of the mine is increased, the soil layer becomes thicker, causing the flexural waves to be dampened, because more energy is radiated out. The resonance in the experimental results has a more complex shape. In the experimental model, the resonance is strongly influenced by the resonant behavior of the actual buried



(a)



(b)

Figure 7.16: (a) Experimental results for a buried TS-50 anti-personnel mine and (b) numerical results for a simple anti-personnel mine model; waterfall graphs of the vertical particle displacement on the surface. Land mine buried at 1 cm, 2 cm, 3 cm, and 6 cm (from left to right).

mine. A TS-50 mine has several chambers, it has a flexible case that can support both flexural and longitudinal waves, and it contains springs that can also give rise to resonances. Furthermore, the soil above the mine might act in a strongly non-linear way, thus altering the shape of the resonance.

To compare the interaction of the elastic waves with a buried mine to, for example, the interaction with a buried rock, a solid body is inserted into the numerical model. The body has a spherical shape and a diameter of 8 cm. Its properties are assumed to be the same as the ones of plastic (see Table 7.1). The results are shown in Fig. 7.17 in some pseudo-color plots for the sphere buried 2 cm beneath the surface and in Fig. 7.18 in some waterfall graphs for the sphere at different depths. When the incident waves interact with the solid sphere, reflections occur which have about the same strength and shape as the reflections that have been observed for the buried anti-personnel mine. However, no continuous oscillations are visible at the location of the sphere, thus clearly distinguishing the solid body from the buried mine. Very similar results have also been observed experimentally.

Anti-Tank Mine

Next, the interaction of elastic waves with an anti-tank mine is to be analyzed. Fig. 7.19 shows a simple model for an anti-tank mine and a photograph of a VS 1.6 mine. The mine has a diameter of about 23 cm and, thus, is substantially larger than the anti-personnel mine. The simple mine model contains three parts: a large plastic body, an air-filled chamber, and a rubber plate simulating the pressure plate of the mine.

In Fig. 7.20, the interaction of elastic waves with an anti-tank mine is shown in some pseudo-color plots. The mine is buried 4 cm beneath the surface, and its center is placed at a distance of 77 cm from the source. As before, a logarithmic color scale with a dynamic range of 50 dB is used. As for the anti-personnel mine, the incident waves are seen to be scattered by the anti-tank mine. Upon incidence onto the mine, the waves are partially transmitted and partially reflected. The incident guided shear

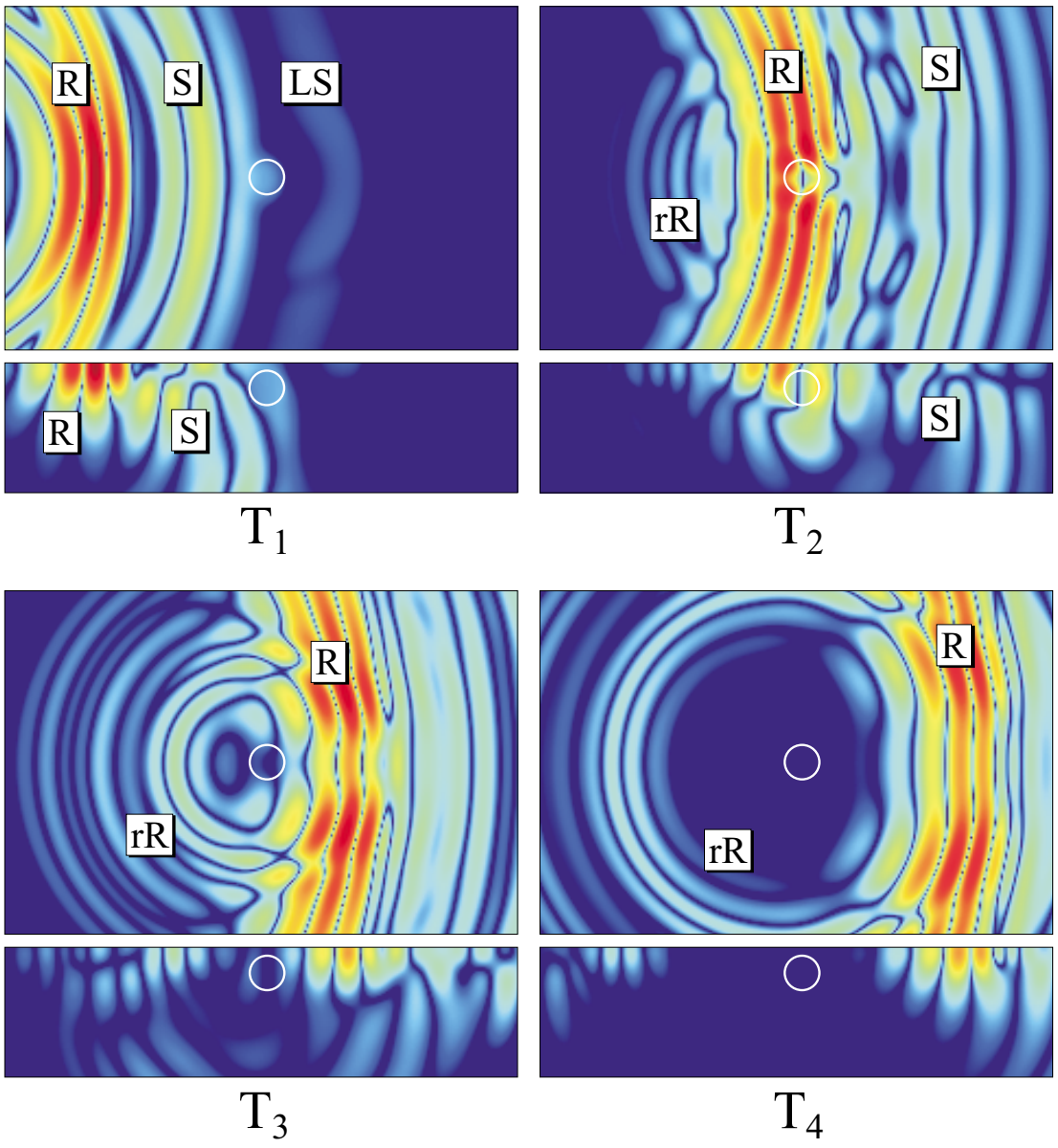


Figure 7.17: Interaction of elastic waves with a buried spherical object; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instants in time.

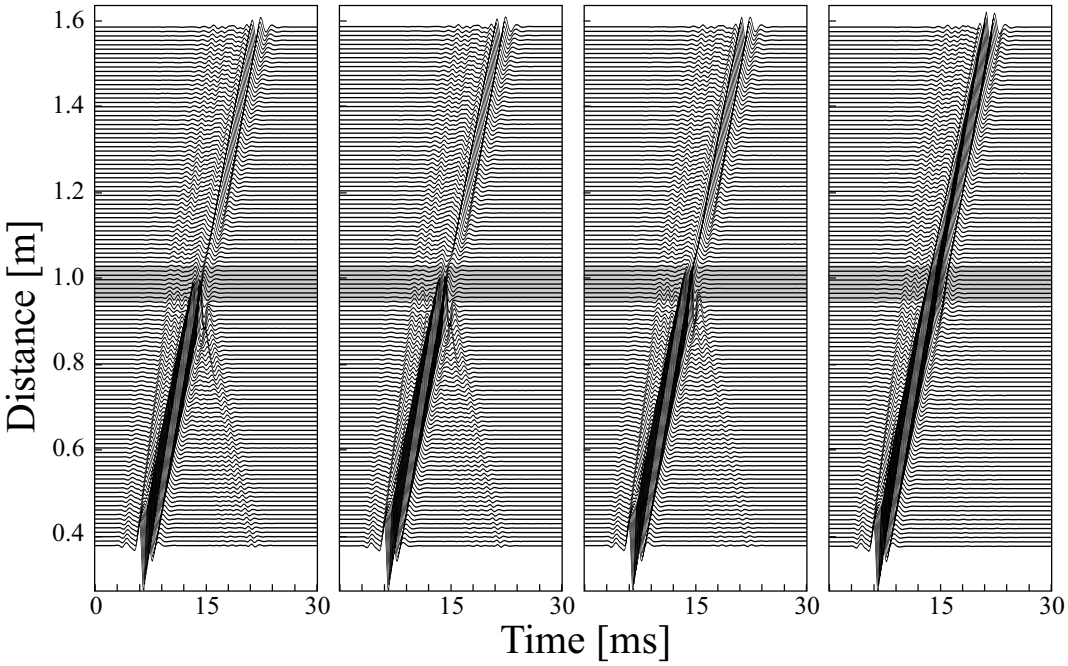


Figure 7.18: Numerical results for a buried sphere; waterfall graphs of the vertical particle displacement on the surface. Sphere buried at 1 cm, 2 cm, 3 cm, and 6 cm (from left to right).

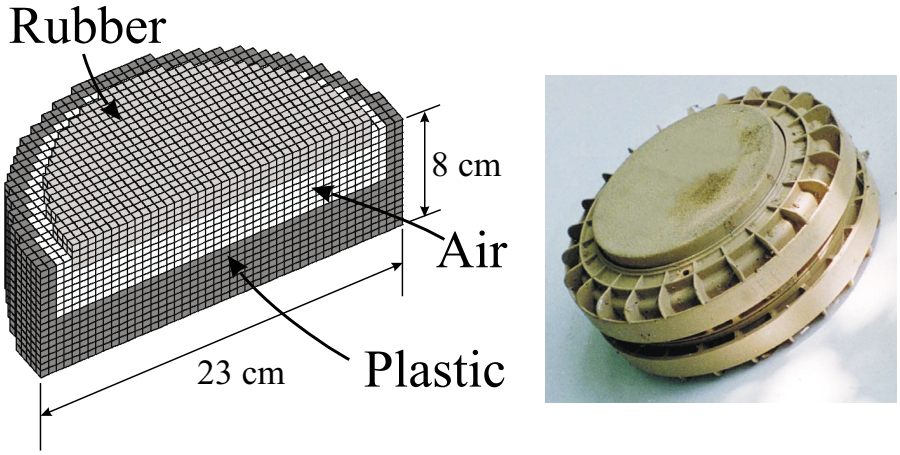


Figure 7.19: Simple model for the VS 1.6 anti-tank mine; schematic picture of the model and photograph of a VS 1.6 mine.

wave (at T_1) and the Rayleigh surface wave (at T_2) couple into the thin soil layer above the mine. The pattern of waves that arises above the mine and the shape of the radiated wave fronts are fairly complex. The waves above the mine, rather than forming standing waves, propagate in the thin layer and radiate energy as they travel back and forth across the layer.

Fig. 7.21 shows waterfall graphs according to experimental measurements and numerical simulations. The location of the mine is indicated by the gray shading. The mines are buried at four depths: at 2 cm, 4 cm, 6 cm and 10 cm (from left to right). For the mine at 2 cm and 4 cm, the interaction with the elastic waves is strong. At the mine location, a complex pattern of waves traveling back and forth above the mine is visible. This can be explained by using the argument given before to interpret the oscillations observed for an anti-personnel mine. The incident wave partially couples into flexural waves in the thin soil layer above the mine. For the anti-personnel mine, these flexural waves formed a standing wave pattern. For the larger anti-tank mine, however, the flexural waves propagate in the thin layer, are reflected at the edge of the mine, propagate back, are again reflected at the other edge of the mine, and so forth. Due to the large size of the anti-tank mine, the waves do not form standing waves, but are propagating back and forth. For the anti-tank mine at 2 cm, the flexural waves are almost completely reflected at the edges of the mine and are confined to the soil layer. For the mine at 4 cm, more energy leaks out of the thin layer and is radiated away. And for the mine at 6 cm, the waves are reflected only very weakly at the mine's edges. For the mine at 10 cm, barely any interaction of the incident waves with the buried mine is visible. The results of experimental measurements and numerical simulations are in fairly good agreement. The waves above the mine in the numerical results seem to travel faster than in the experimental results. This may be due to the depth profile of the shear wave speed being especially inaccurate close to the surface, and to the sand behaving in a strongly non-linear way at the surface. However, in spite of all uncertainties in the characterization of the soil and the simplicity of the mine model, the same qualitative behavior for the

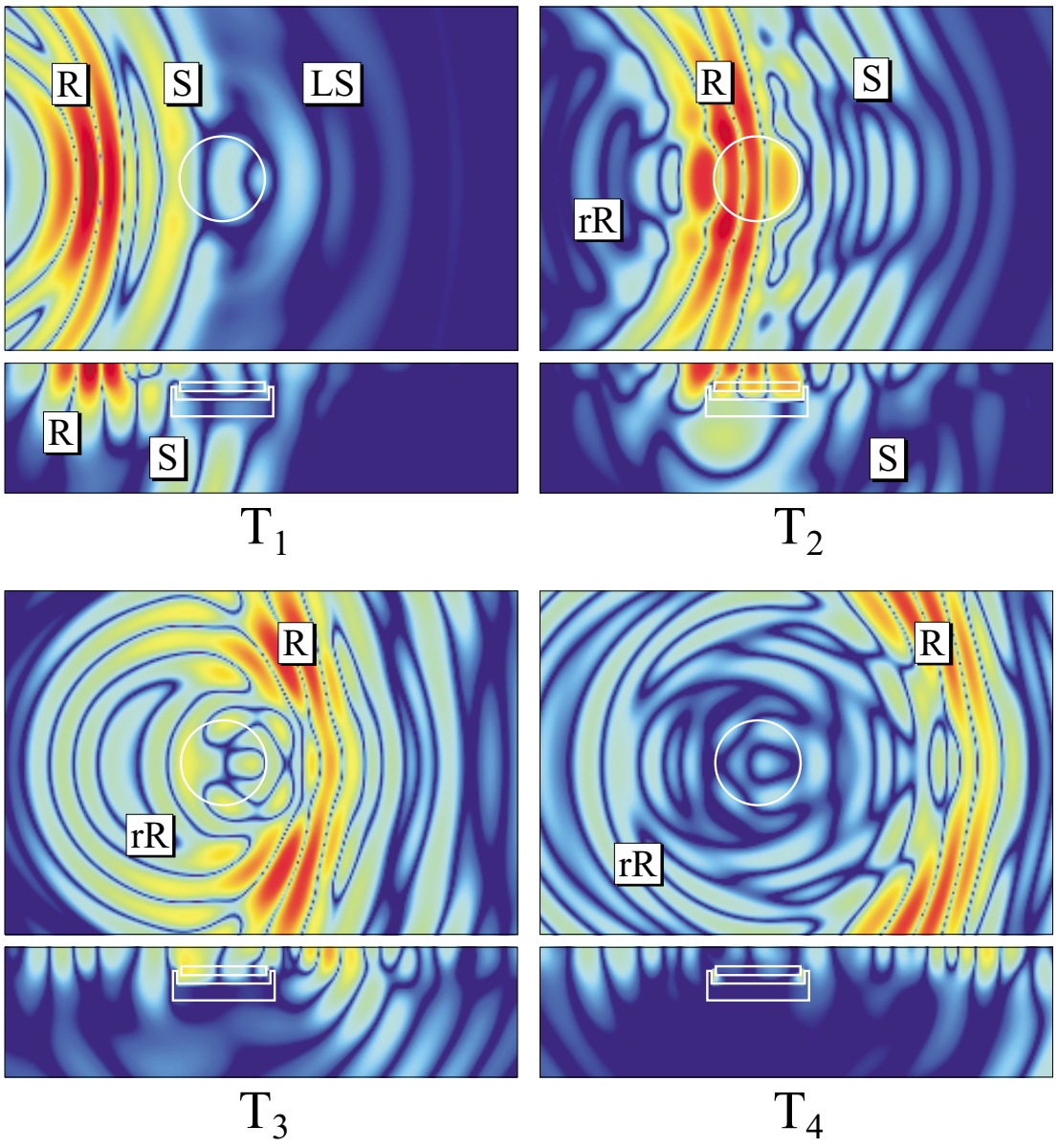


Figure 7.20: Interaction of elastic waves with a buried anti-tank mine; pseudo color plots of the normal particle displacement on the surface (top) and on a cross section through the ground (bottom) at four instants in time.

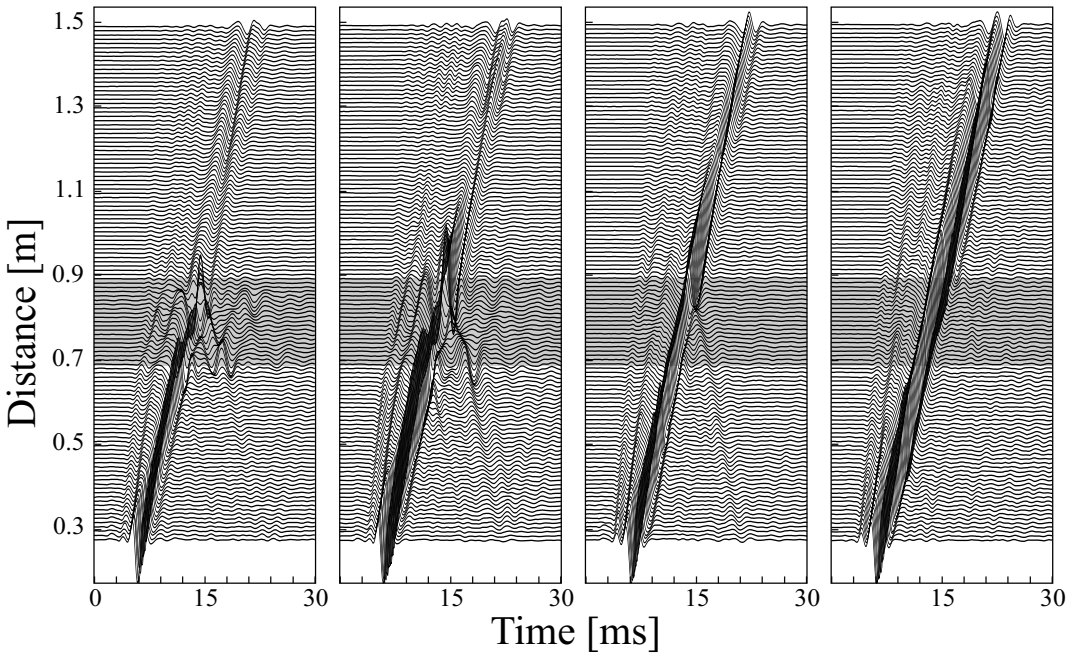
interaction of the elastic waves with the buried mine is observed in experiment and numerical model.

7.3 Mines in the Presence of Clutter

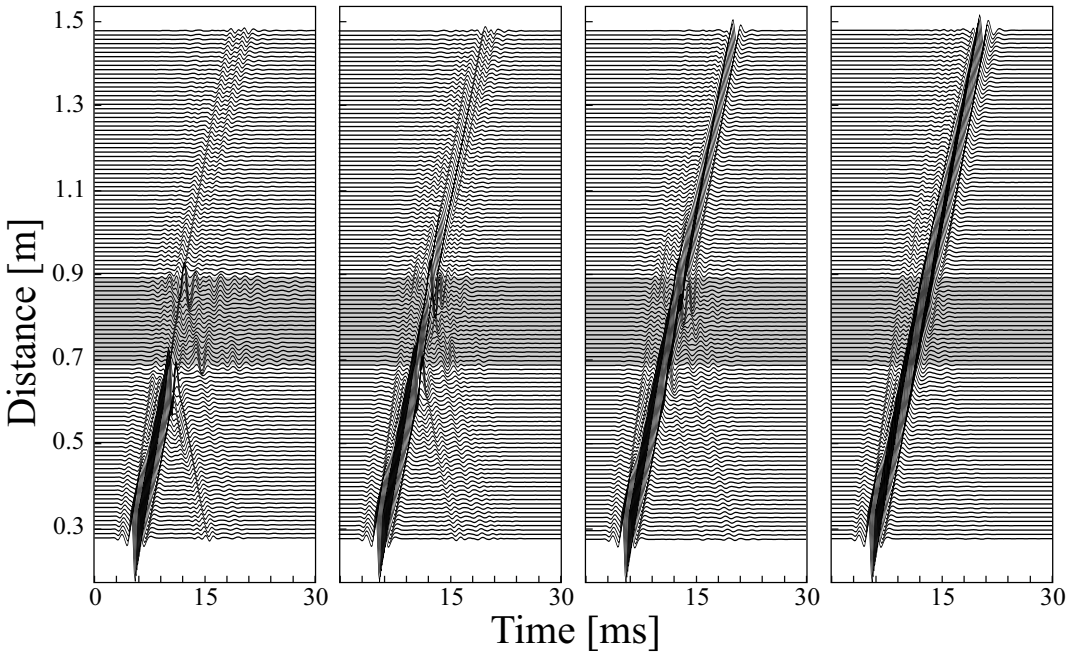
One of the most crucial requirements for a land mine detection system is the capability to accurately differentiate buried mines from surrounding clutter objects. In the last section it has been shown that resonant oscillations occur at the location of a buried mine, when the mine interacts with elastic waves. These resonant effects can be used to enhance the mine's signature, because the resonance appears to be inherent to minelike objects. It has also been shown that buried rocks will not exhibit such resonant oscillations. However, the scattering effects from buried rocks are comparable to those from buried mines, and the presence of rocks might significantly alter the response from the buried mines.

To explore the behavior of mines in the presence of clutter, the interaction of elastic waves with various mines being surrounded by rocks is computed. Fig. 7.22 shows schematically the configuration of mines and rocks in the ground and indicates their sizes and burial depths. The circles indicate the mines, the rectangles mark the rocks. Four anti-personnel mines are modeled surrounding an anti-tank mine. The mine closest to the source has a diameter of 6 cm and represents a small anti-personnel mine. Its distance to the source is 72 cm. The other three anti-personnel mine have a diameter of 8 cm and lie at distances of 92 cm and 112 cm from the source. The anti-tank mine, at the center of the scenario, has a diameter of 23 cm. The rocks, surrounding the mines, are cuboids with dimensions of 6 cm by 8 cm by 4 cm and lie at different tilt angles with respect to the source. Note that the rock in the lower left corner is tilted by 30° with respect to the surface plane.

Fig. 7.23 illustrates the interaction of the elastic waves with the buried mines and the rocks. In pseudo-color plots, the vertical particle displacement on the surface plane is plotted at four times. The source lies to the left of the plots. At T_1, the



(a)



(b)

Figure 7.21: (a) Experimental results for a buried VS 1.6 anti-tank mine and (b) numerical results for a simple anti-tank mine model; waterfall graphs of the vertical particle displacement on the surface. Land mine buried at 2 cm, 4 cm, 6 cm, and 10 cm (from left to right).

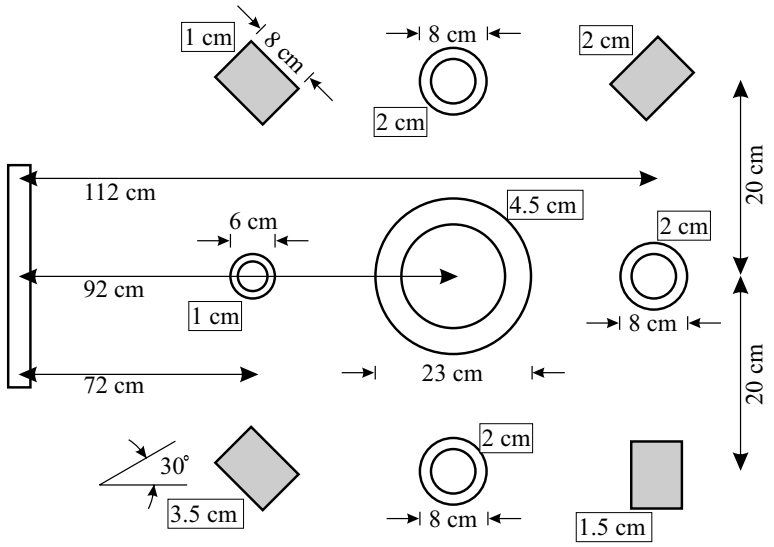


Figure 7.22: Configuration of the mines and rocks in the ground; the rectangles indicate the rocks, the circles indicate the mines. The boxed numbers correspond to the burial depths.

guided shear wave interacts with the mine and the two rocks closest to the source and is scattered. The leaky surface wave, although weak, has already reached the mine at the far right. At T_2, the Rayleigh surface wave has arrived at the mine closest to the source. The guided shear wave has passed the anti-tank mine and is seen to strongly interact with all mines. At T_3, the Rayleigh wave has reached the mine farthest from the source. Resonant oscillations are excited at all mines. At T_4, the incident waves have passed the mine formation. Due to the resonant oscillations, the anti-personnel mines are clearly discernible on the surface. The interaction of the waves with the buried mines strongly dominates over the interaction with the clutter. The locations of the mines can be easily identified, whereas the clutter objects are not apparent on the surface.

The anti-personnel mines can easily be distinguished from their surroundings, because resonant oscillations occur at the locations of the buried mines, visible as bright spots on the surface. From Fig. 7.23, however, the location of the anti-tank mine cannot be easily identified. Obviously, the interaction of the incident waves with

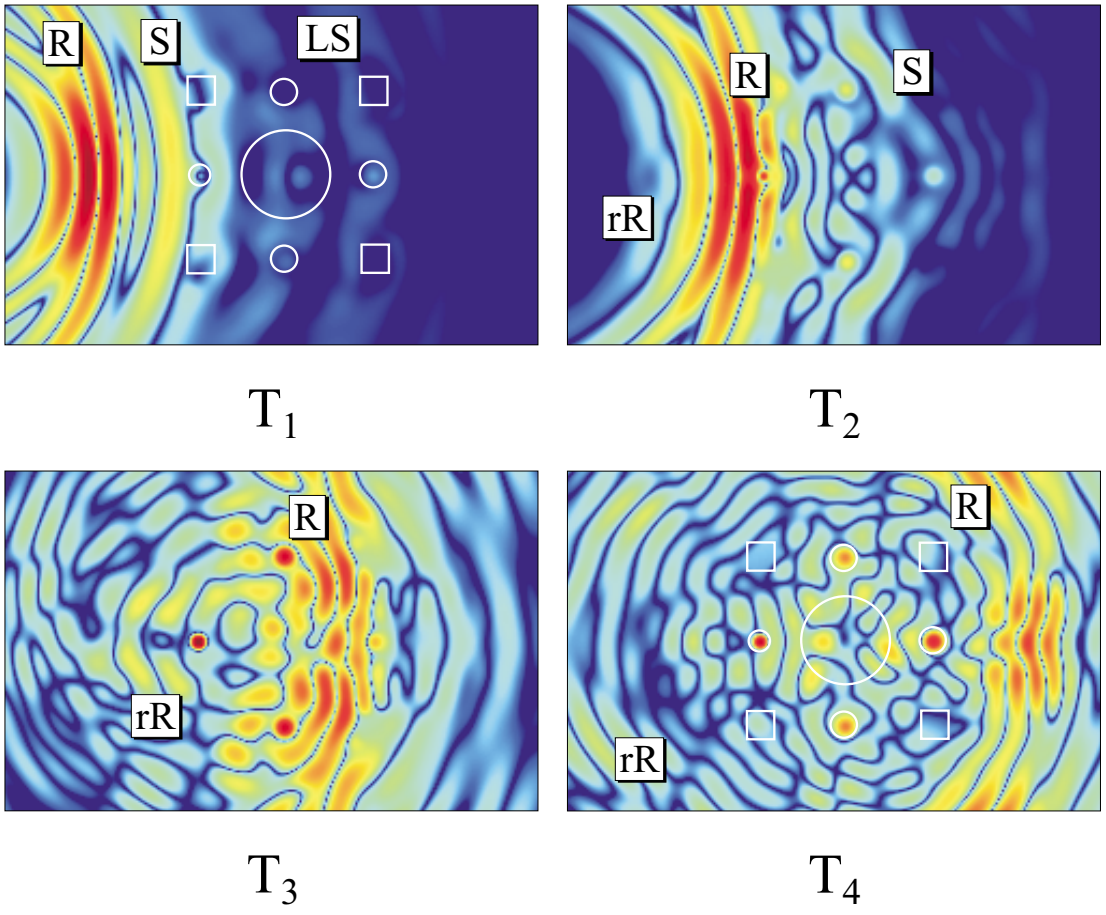


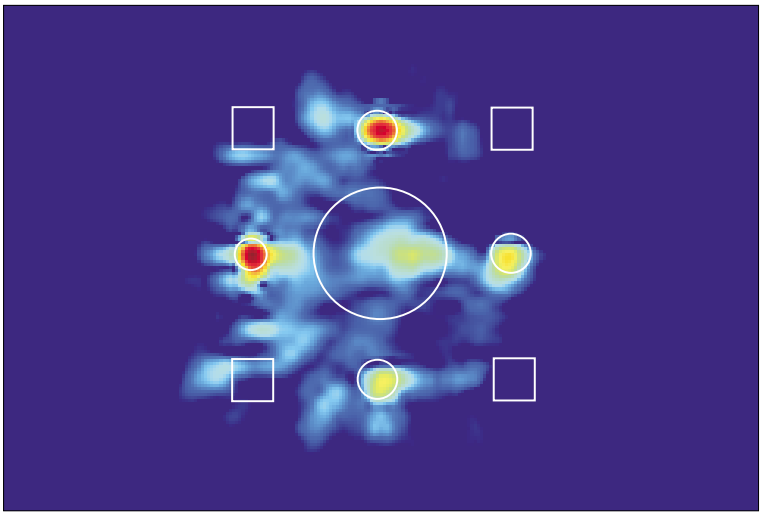
Figure 7.23: Interaction of elastic waves with four anti-personnel mines, one anti-tank mine and four rocks buried in the ground; pseudo color plots of the normal particle displacement on the surface at four instants in time.

the anti-tank mine is strong. But, as explained earlier, due to the substantially larger size of the anti-tank mine compared to the anti-personnel mines, no resonant oscillations occur at the mine location, but rather a pattern of flexural waves propagating back and forth above the mine. To facilitate the localization of the buried mines, an imaging algorithm has been developed by A. Behboodian *et al.* which forms an image by, in simple words, summing the energy of the back-scattered waves at each spatial position [77]. Fig. 7.24 shows the image that results when the algorithm is applied to the numerically computed data and compares it to experimental results for a similar set-up [78]. A dynamic range of 30 dB is used. For the experimental results, an M-14, a TS-50, a VS-50 and a PFM-1 anti-personnel mine were buried around a VS 1.6 anti-tank mine. In the image formed from the numerical data, the locations of the five mines are accurately predicted. The response of three of the five mines in the numerical data, however, is weaker than in the experimental data. For the anti-tank mine, the image shows the strongest response in both numerical and experimental data at the back-end of the mine. This is due to the reflections of the incident waves being strongest at the back of the mine, an effect that has also been observed earlier in Fig. 7.21.

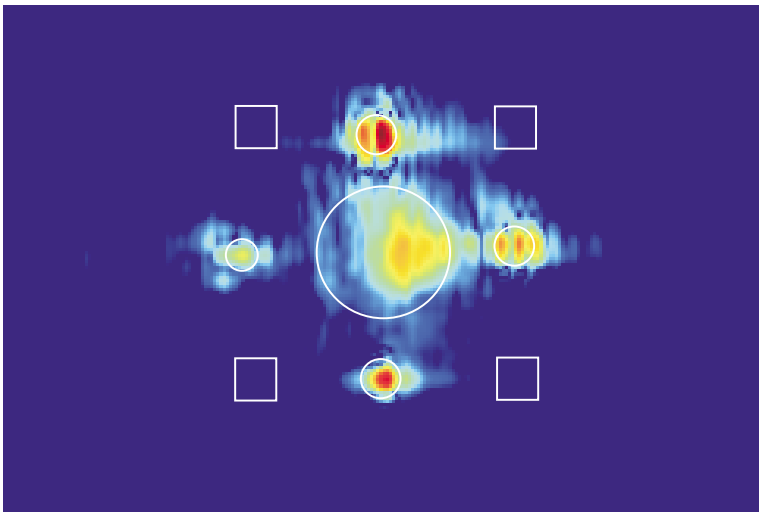
7.4 A Moving-Source/Moving-Receiver System

In the previous sections, a detection system has been investigated in which the source was at rest and either a stationary receiving array or a synthetic array with a moving receiver was used. Ultimately, this system will be mobile, scanning step by step across the potential mine field. For a mobile sensor, other configurations, however, are possible. Fig. 7.25 shows schematically a mine detection system in which both the source and the receiver are moving. In this system, a constant distance between source and receiver is maintained, while the source-receiver unit scans across the surface of the ground.

To investigate such a moving-source/moving-receiver system, the system is sim-



(a)



(b)

Figure 7.24: Image formed from the interaction of elastic waves with the buried mines using an imaging algorithm; (a) numerical data, (b) experimental data. The image describes a surface area of 120 cm by 80 cm.

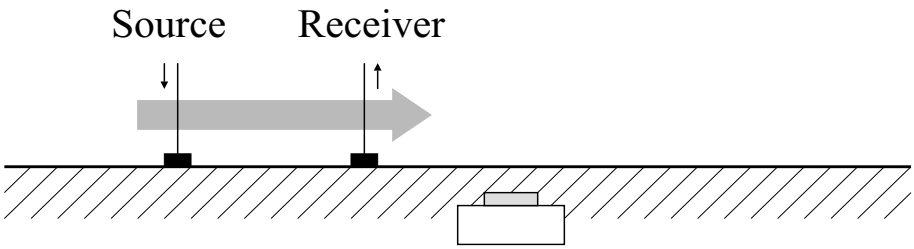


Figure 7.25: Schematic drawing of the moving-source/moving-receiver system.

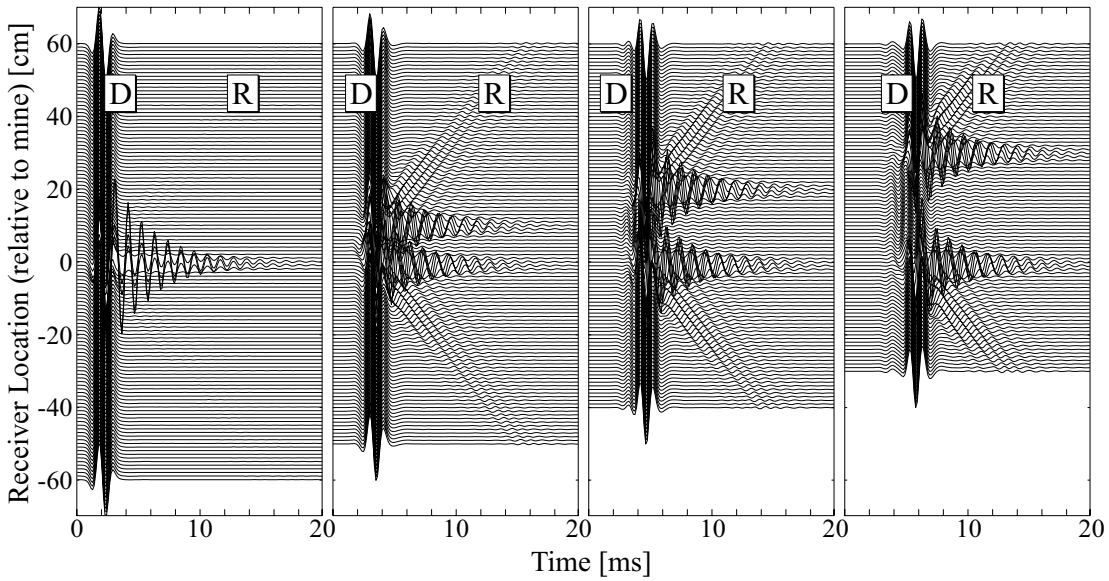
ulated with the numerical model. For this, a point source is placed on the surface and excites elastic waves. The normal particle displacement is excited using a differentiated Gaussian pulse with a center frequency of 450 Hz. At some distance from the source, the vertical particle displacement is recorded at one point, i. e., at the receiver location, as a function of time. Then, the source is moved by an incremental distance, and so is the receiver location. Each time trace is plotted and vertically displaced by the receiver location relative to the buried land mine. A waterfall graph arises, in which the vertical displacement is plotted as a function of time for all source/receiver positions.

Fig. 7.26 shows waterfall graphs for a moving-source/moving-receiver system with different distances between the source point and the receiver location. The distances between source and receiver are 0 cm, 10 cm, 20 cm, and 30 cm (from left to right). In the graphs, the time traces of the received signal are vertically displaced by the distance from the receiver to the center of the mine. Thus, a value of 0 cm on the vertical axis indicates that the receiver is directly above mine. Two sets of waterfall graphs are shown: in the first the results are scaled so that the maximum peak value of the first time trace is equal in all graphs, whereas in the second the results are not scaled. In the second set, the peaks for the 0 cm and 10 cm cases are clipped, because the maximum would otherwise exceed the range of the plot. Each waterfall graph exhibits three different kinds of waves, corresponding to different signal paths from the source to the receiver. One path is the direct path (D) and corresponds to the elastic waves that are excited by the source and propagate directly to the receiver.

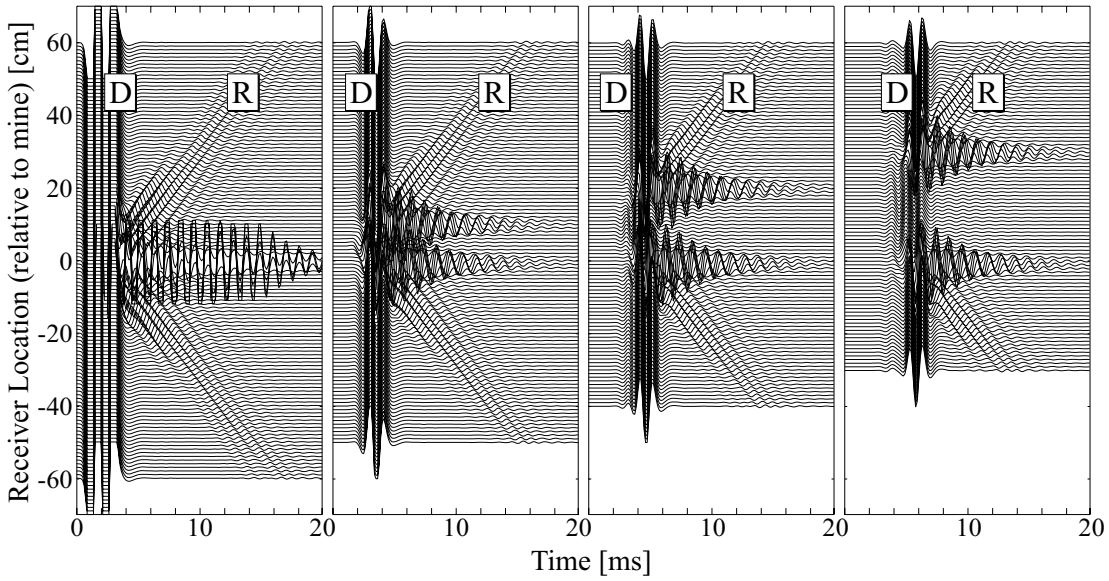
In general, this is the strongest signal picked up by the receiver. The direct path is a straight line in the waterfall graphs, because the distance from source to receiver is constant and, thus, the propagation time from source to receiver is independent of the receiver location. Another path is the reflected path (R). The incident waves propagate towards the buried mine, are reflected by the mine and are then recorded by the receiver. Because the receiver detects the surface motion, the reflected waves will primarily be Rayleigh surface waves. The reflected signal describes diagonal lines in the waterfall graphs, because, dependent on the receiver location with respect to the mine, the time the signal takes to travel between mine and receiver varies. The signal that is received via the reflected path contains the reflection from the mine as well as the waves that are radiated due to resonant oscillations excited at the mine location. From the slope of the reflected signal in Fig. 7.26, the wave speed of the reflected waves is determined to be approximately 80 m/s, and, thus, agrees with the Rayleigh wave speed. When the receiver location or the source location coincide with the position of the buried mine, strong resonant oscillations are recorded by the receiver. If the receiver is at the mine location, it will pick up the oscillations excited at the mine and thus will receive a strong signal, stronger than any reflections or radiated waves that the receiver records apart from the mine. If the source is at the mine location, it primarily will excite the resonant oscillations at the mine, yielding an enhanced resonance. Due to the enhanced resonance, the radiation from the mine will be increased, which again leads to a strong signal that is recorded at the receiver location. As visible in Fig. 7.26, the system is reciprocal: the location of source and receiver are interchangeable, and it does not matter, whether the source or the receiver are located at the mine location. The received signal will in both cases be the same. If the distance between source and receiver is zero (left in Fig. 7.26), strong resonant oscillations are visible at the mine location. If the distance between source and receiver is increased, the resonant oscillations appear at two distinct receiver locations: once when the receiver is above the mine, and once when the source is above the mine. The reflected signal is not visible in between these two resonant

signals. When the results are scaled according to their maximum (as in (a)), the reflected waves disappear if the source-receiver distance is zero. This is due to the enhanced resonance picked up by the receiver at the mine location. If the results are not scaled, it is evident that the reflected waves in each case have the same strength. Note that, due to the depth variation in the ground, the pulse disperses as it propagates along the surface, and thus the received pulse changes in shape as the distance from the source to the receiver is increased.

Fig. 7.27 explains schematically the paths the signal takes. If source and receiver are both on the same side of the mine (as in 1 and 3), the receiver will pick up the signal via the direct path (D) and via the reflected path (R). If source and receiver are on opposite sides of the mine (as in 2), the receiver will record the direct path, but the signal induced by the reflections and the radiation from the mine will be weak. This is clearly visible in Fig. 7.26.



(a)



(b)

Figure 7.26: Moving-source/moving-receiver system; water fall graphs of the vertical particle displacement at the receiver location; for four distances between source and receiver: 0 cm, 10 cm, 20 cm, 30 cm (from left to right). (a) Scaled to the maximum value of each graph; (b) not scaled. The peaks in (b) are clipped for the 0 cm- and the 10 cm-case.

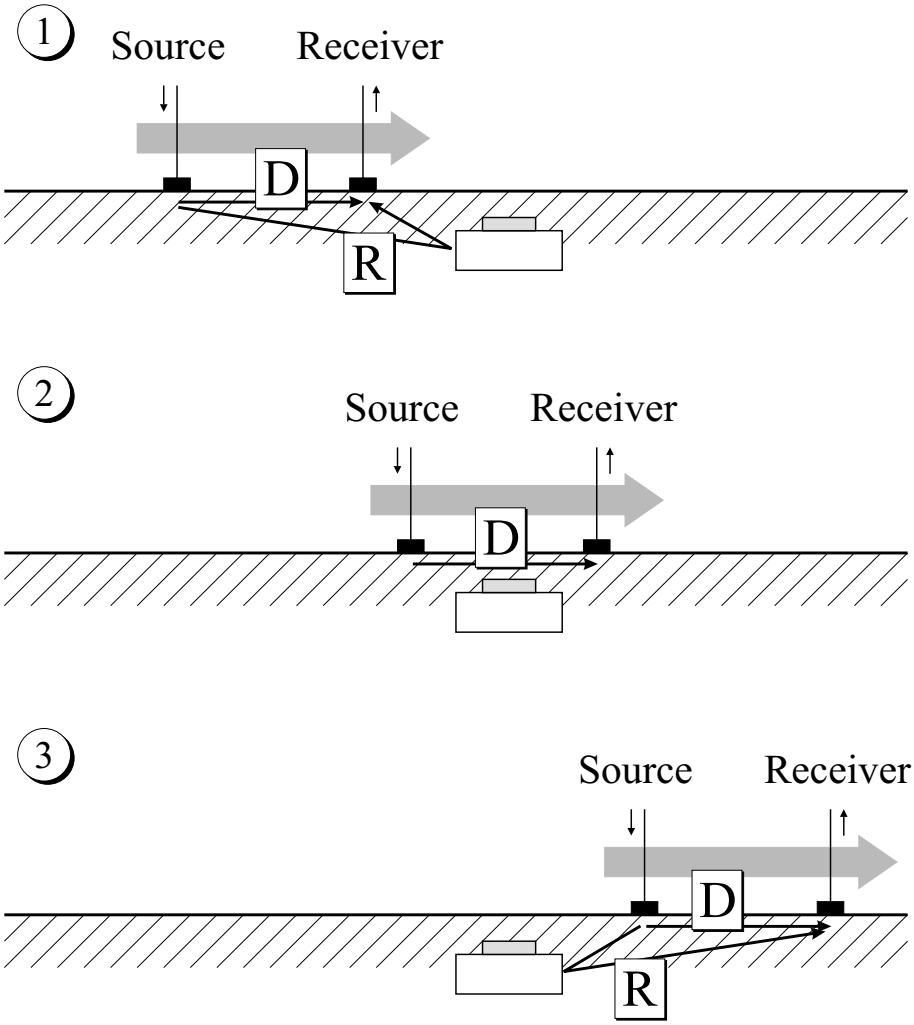


Figure 7.27: Signal paths for three positions of the source-receiver system.

CHAPTER 8

Modeling of Resonant Structures: the Resonance Behavior of a Buried Land Mine

In experimental and numerical studies, resonant oscillations have been observed at the location of a buried land mine, when the mine is excited by elastic waves. These resonant effects have occurred for mines of different types, shapes and sizes, including anti-personnel and anti-tank mines. The resonance significantly enhances the signature of a buried mine and clearly distinguishes it from surrounding clutter objects.

In the previous chapter, a very simple model has been used to compute the interaction of elastic waves with a buried anti-personnel mine on a large scale. Although the mine model is very simple, it predicts resonant oscillations at the location of the buried land mine that are qualitatively very similar to the oscillations observed in the experiments. It has been shown that the resonant oscillations in the numerical model are due to flexural waves being trapped in the thin soil layer above the mine. While this explains the resonance for the numerical model, it gives only one possible cause for the resonance of an actual mine. An actual mine contains explosives, various air chambers, a triggering mechanism, and a flexible case. All these components of the mine might have a significant impact on the mine-wave interaction. Therefore, to investigate the resonant behavior in more detail, a refined mine model must be developed that more closely approximates the structure of an actual mine.

This chapter is divided into two sections. In the first section, the results of a preliminary study are presented. In this study, the necessary spatial discretization is determined that is required to simulate the resonant behavior of mechanical structures accurately with the numerical model. In the second section, a refined mine model is introduced to study the resonance at the location of a buried mine as a function of burial depth and material properties.

8.1 Discretization of Elastic Objects

A refined mine model will have to incorporate the details of the mine. Thus, the new model will require a much finer discretization than used previously in the numerical model. To investigate how fine the structural details of the mine must be discretized to model their elastic behavior accurately, the resonant frequencies of various mechanical objects are computed as a function of the spatial discretization in the numerical model. Results are presented for a long thin bar, a thin circular plate and a tuning fork.

8.1.1 Long Thin Bar

The long thin bar is shown in Fig. 8.1 as discretized within the numerical model. It has a length of 20 cm and a cross section of 1 cm by 1 cm. Its material properties are assumed to be those of plastic: $\rho = 1200$ kg/m^3, $c_P = 2700$ m/s, and $c_S = 1100$ m/s. In Fig. 8.1, a grid spacing of 0.5 cm is assumed, yielding a discretization of 2 cells across the width of the bar. Using the numerical model, the resonant frequencies of the bar in air are determined in the range up to 1.5 kHz and are compared to results from a simple analytical theory.

The resonant frequencies for the flexural modes of a long thin bar can be calculated analytically using the simple Bernoulli-Euler theory for long thin rods [58, 79]. The Bernoulli-Euler theory is an approximate theory, because it neglects the effects of shear deformations and the rotary inertia of the bar. It will predict the resonant

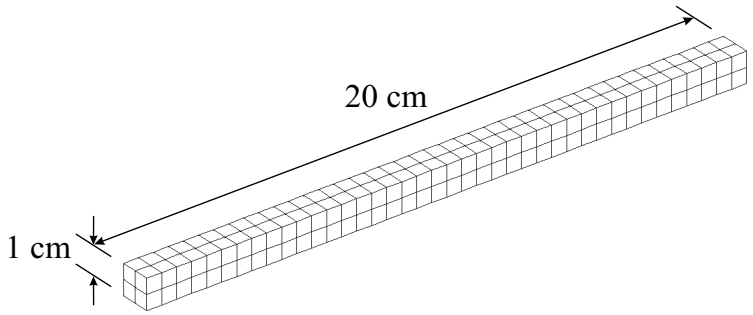


Figure 8.1: Long thin bar.

frequencies accurately only for the low-frequency harmonics, i. e., if the harmonic wave length is much larger compared to the width of the bar. For higher harmonics, the Bernoulli-Euler theory will significantly overestimate the resonant frequencies.

For a bar with free ends, the Bernoulli-Euler theory predicts the resonant frequency of the i-th harmonic to be [79]

$$f_i = \frac{K}{2\pi l^2}\sqrt{\frac{E}{\rho}}(\beta_i l)^2, \tag{8.1}$$

where K is the *radius of gyration* (equal to the height of the bar divided by $\sqrt{12}$ for a bar with a rectangular cross section), l is the length of the bar, and E is Young's Modulus (see Appendix D). The transverse wave number β_i is the solution to a transcendental equation and is given in Table 8.1 for the first two harmonics.

Table 8.1: Wave numbers and wave length of the first two harmonics of a bar with free ends according to the Bernoulli-Euler theory.

Harmonic	Wave Number β_i	Wavelength $\lambda = \frac{2\pi}{\beta_i}$
1	$4.730/l$	$1.330l$
2	$7.853/l$	$0.800l$

Using Eq. (8.1), the first and second harmonic of the bar are estimated to have resonant frequencies of 473 Hz and 1303 Hz, respectively. The resonant frequencies as

computed with the numerical model are summarized in Table 8.2. With the numerical model, the resonant frequencies are computed as a function of the discretization, i. e., for a varying spacing between the nodes in the numerical grid. In Fig. 8.2, the resonant frequencies and the error relative to the theoretical prediction are plotted as a function of the number of nodes that are placed across the bar. With two nodes across the bar, the numerical model predicts resonant frequencies of 405 Hz and 1100 Hz for the first two modes, corresponding to an error of about 15 %. With four nodes, the error has reduced to about 5 %, and for ten nodes it is about 3 %. The relative error for the second harmonic always exceeds the error for the first harmonic. This is due to the simple analytical theory overestimating the resonant frequency of the second harmonic. In general, the true error of the numerical results will be smaller than predicted here, because the actual resonant frequencies will be slightly smaller than the theoretical predictions.

Table 8.2: Resonant frequencies of a long thin bar.

# nodes across bar	1^{st} Mode [Hz]	2^{nd} Mode [Hz]
2	405	1100
3	430	1170
4	455	1230
5	460	1245
6	460	1255
8	465	1260
10	465	1265
Theoretical	473	1303

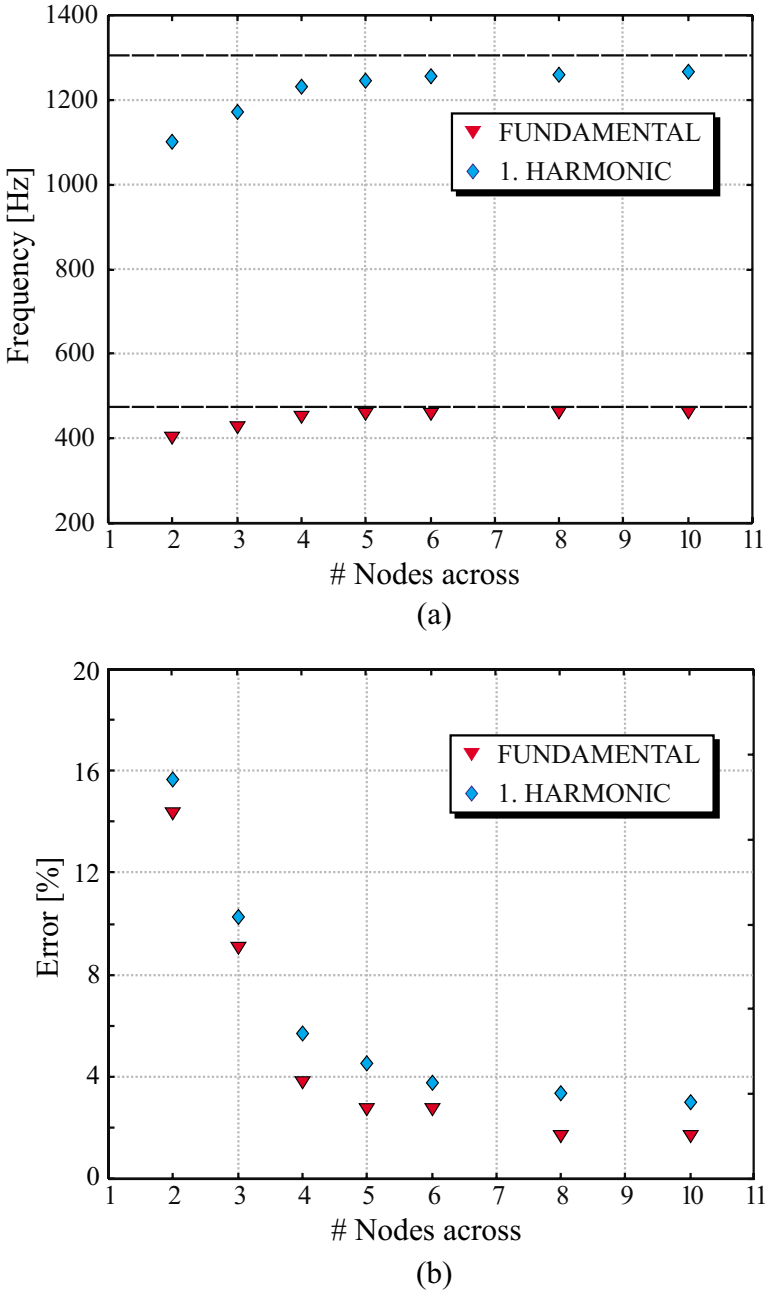


Figure 8.2: (a) Resonant frequencies of the bar and (b) error relative to the theoretical prediction as a function of the number of nodes across the width of the bar.

8.1.2 Thin Circular Plate

Next, a thin circular plate is considered. The plate is shown in Fig. 8.3. It has a diameter of 20 cm and a thickness of 1 cm. As before, the material properties of plastic are assumed for the plate. The plate is placed in air and is free along its edge.

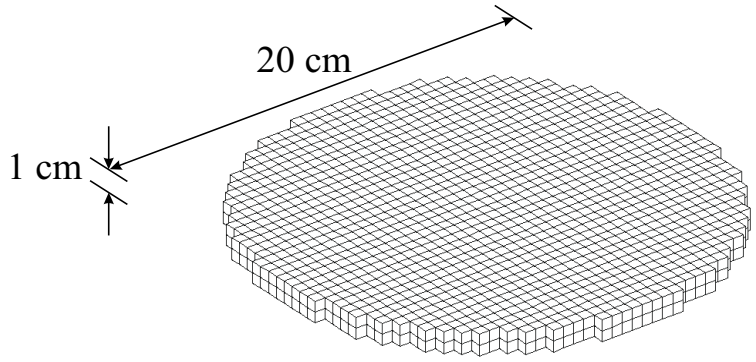


Figure 8.3: Thin circular plate.

The shape of the first three modes of a circular plate with a free edge is schematically shown in Fig. 8.4. The modes are labeled by the number of their nodal diameters and their nodal circles. For example, the (20)-mode describes a mode with two nodal diameters and zero nodal circles. Along a nodal diameter or a nodal circle, the displacement of the plate is zero. The fundamental mode of a thin circular plate that is free along its edge is the (20)-mode, the next higher mode in frequency is the (01)-mode, followed by the (30)-mode.

Using an approximate theory, the resonant frequencies of a thin circular plate can be obtained as solutions to a rather complicated transcendental equation [58]. The analytically determined resonant frequencies of the first three modes of the thin circular plate are summarized in Table 8.3 in terms of the frequency of the fundamental (20)-mode.

The resonant frequencies of the first three modes as computed with the numerical model are summarized in Table 8.4. Results are listed for three discretizations, namely

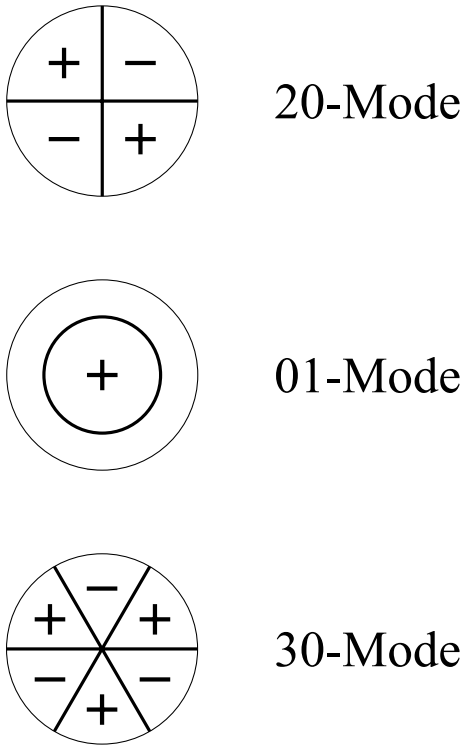


Figure 8.4: Modes of a circular plate with free edge. The bold lines indicate the nodal diameters and nodal circles.

Table 8.3: Resonant frequencies of the first three modes of a thin circular disc that is free along its edge.

Mode	Resonant Frequency [Hz]
(20)	$f_{(20)} = 0.2413 c_L h/a^2$
(01)	$f_{(01)} = 1.73 f_{(20)}$
(30)	$f_{(30)} = 2.33 f_{(20)}$

where $c_L = \sqrt{E/(\rho(1-\nu^2))}$ is the longitudinal wave speed in an infinite plate, h is the thickness of the plate, and a is its radius.

for two, four and eight cells across the thickness of the plate. The error relative to the theoretical values is given in brackets. For all three modes, the theoretical predictions are larger than the numerical values. The relative error reduces as the number of nodes across the plate is increased, from about 20 % for two nodes to below 10 % for eight nodes. Again, the simple theory overestimates the resonant frequencies, and, thus, the actual error of the numerical results will be smaller than stated here.

Table 8.4: Resonant frequencies of the thin circular plate as computed with the numerical model.

# nodes across plate	$f_{(20)}$ [Hz]	$f_{(01)}$ [Hz]	$f_{(30)}$ [Hz]
2	380 (21%)	715 (15%)	870 (23%)
4	440 (9%)	805 (4%)	1000 (11%)
8	455 (6%)	825 (1.5%)	1030 (9%)
Theoretical	484	839	1129

For both the bar and the plate it is apparent that, in general, the accuracy of the numerical model will be improved, if the spacing between the nodes of the

numerical grid is decreased. Considering the high uncertainty in, for example, the material parameters in the numerical model, one can conclude from the above results that the resonant behavior of an object will be modeled with a reasonable accuracy, if four cells are placed across its smallest structural width. Even a discretization with only two or three cells might suffice, if a slightly increased error is acceptable in the numerical model.

8.1.3 Tuning Fork

As a more complex mechanical structure, the resonant behavior of a tuning fork is explored. The tuning fork, as discretized in the numerical model, is shown in Fig. 8.5. Its measures are identical to an actual tuning fork with a fundamental frequency of 329.6 Hz, corresponding to the musical tone E. The tuning fork is made out of steel, approximately described by $\rho = 4100$ kg/m^3, $c_P = 5000$ m/s, and $c_S = 3500$ m/s. With the numerical model, the resonant frequencies of the fork in the range up to 1.5 kHz are determined.

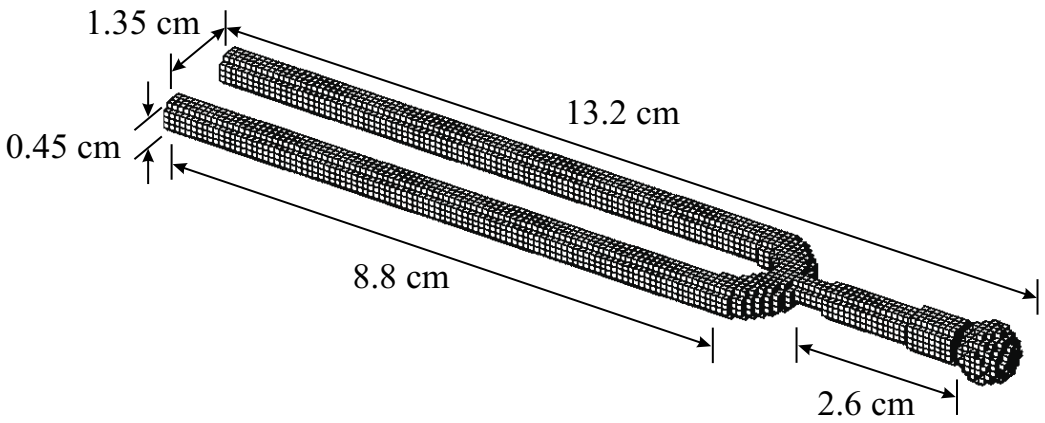


Figure 8.5: Tuning fork.

The modes of a tuning fork can be classified into four groups: symmetrical

modes in the plane of the fork's arms (caused by the symmetric motion of the two arms), antisymmetrical modes in the plane, symmetrical modes normal to the plane and antisymmetrical modes normal to the plane [80]. In the range up to 1.5 kHz, the tuning fork will have four modes, one of each of the four groups. The four different modes are schematically shown in Fig. 8.6. The fundamental mode is the symmetric in-plane mode.

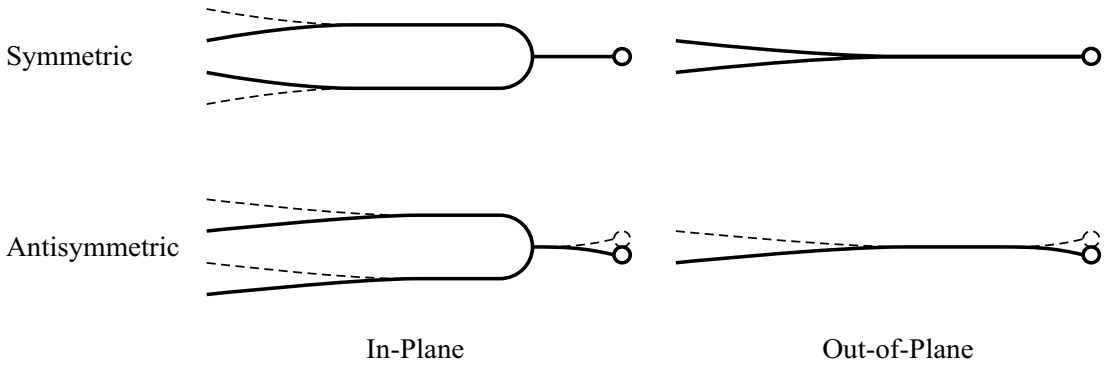


Figure 8.6: Schematic drawing of the first four modes of a tuning fork.

Using an approximate theory, the resonant frequencies of the four modes can be estimated [79, 80]. For the symmetrical in-plane and out-of-plane modes, the motion of the tines of the tuning fork resembles that of a cantilever beam, clamped at the forking point. For a cantilever beam with dimensions similar to the arms of the fork in Fig. 8.5, the resonant frequency is predicted to be approximately 363 Hz. For the antisymmetrical modes, the entire fork bends in the manner of a beam with free ends. A beam with a uniform cross section equivalent to the cross section of one arm of the tuning fork has a resonant frequency of about 1051 Hz. Of course, the fork does not have a uniform cross section along its length, and, thus, this value is only a very crude estimate for the resonant frequency of the antisymmtrical modes of the tuning fork. It must be noted that this theory does not distinguish between in-plane

and out-of-plane modes. However, the in-plane and out-of-plane of a tuning-fork will experience different bending stiffnesses, and, therefore, their resonant frequencies will differ significantly.

Using the numerical model, the resonant frequencies of the tuning fork are computed. Results are obtained for discretizations with five and nine nodes across each arm, corresponding to a grid spacing of 0.09 cm and 0.05 cm, respectively. Table 8.5 summarizes the results. The resonant frequency of the fundamental mode comes out to be 320 hz when discretizing the tuning fork with five nodes and 330 Hz for nine nodes. Although this differs from the theoretical result, it agrees well with the resonant frequency of the actual tuning fork (329.6 Hz). The numerical model predicts the resonant frequency of the symmetric out-of-plane mode to be roughly 515 Hz (for nine cells across each arm), which is substantially larger than for the in-plane mode. The anti-symmetric modes have, according to the numerical model, only slightly different resonant frequencies.

Table 8.5: Resonant frequencies of a tuning fork.

# nodes across arm	Symmetric modes		Antiymmetric modes	
	in-plane [Hz]	out-of-plane [Hz]	in-plane [Hz]	out-of-plane [Hz]
5	320	500	1000	1010
9	330	515	1105	1140
Theoretical	363		1051	

The motion of the fork, strongly amplified, is visualized in Fig. 8.7 (as computed with a discretization of five cells across each arm). The in-plane motion is plotted on the left and the out-of-plane motion on the right at four moments in time. Both in-plane and out-of-plane, the superposition of symmetrical and antisymmetrical modes is evident. Note that the symmetrical out-of-plane motion of the tines causes the

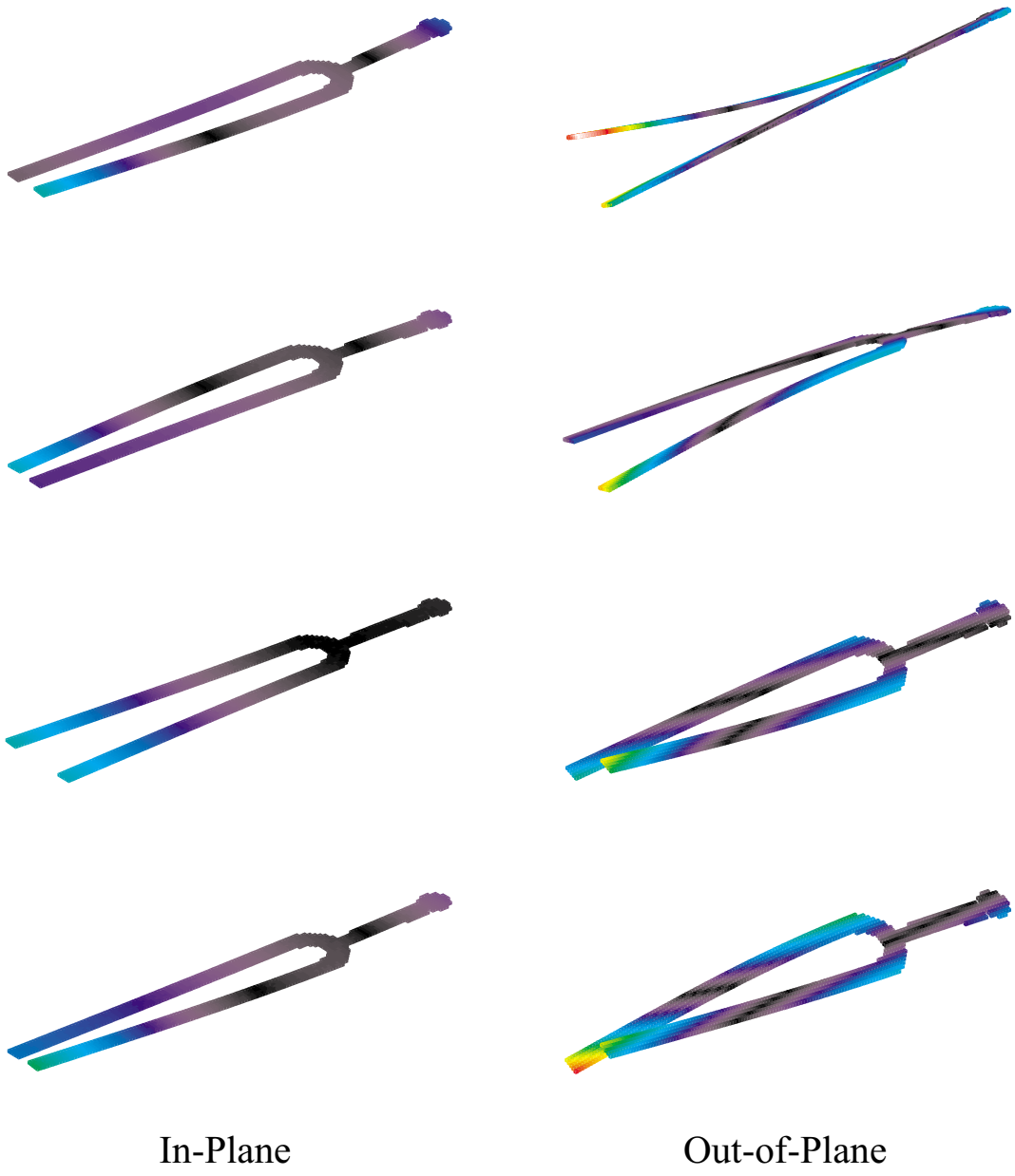


Figure 8.7: The motion of a tuning fork.

tuning fork to rotate back and forth about its axis, due to the fact that the rotational momentum must be conserved.

8.2 Resonance Behavior of a Buried Land Mine

To investigate the resonant oscillations at the location of a buried land mine in more detail, a mine model has been developed that more closely resembles the structure of an actual mine. The refined mine model is shown in Fig. 8.8. The mine model includes the details of the mine that are believed to have an impact on the mine's resonant behavior: the explosives, the plastic case, the rubber pressure plate, and the air chambers. The model has similar dimensions as an actual TS-50 anti-personnel mine (see Fig. 7.9). The diameter of the mine is about 8 cm, and its overall height is 4 cm. The material properties that are assumed for the various parts of the mine are summarized in Table 8.6. The grid spacing for the finite-difference grid is chosen to be $\Delta x = \Delta y = \Delta z = 0.1$ cm. At least two cells are placed across each structural detail of the mine. The mine's case, at its thinnest spot, is discretized by three cells across.

Table 8.6: Material properties of the parts of the detailed mine model.

	Shear Wave Speed c_S [m/s]	Pressure Wave Speed c_P [m/s]	Material Density ρ [kg/m^3]
Explosives	1340	2480	1610
Air	0	330	1.3
Plastic Case	1100	2700	1200
Pressure Plate	500	800	1000

To incorporate the features of an actual mine, a much finer grid than used previously must be employed in the numerical model. With such a grid, the entire space including the transducer on the surface and the buried mine cannot be modeled, because the requirements in computer memory and computation time would be immense. Therefore, with the fine grid the elastic wave fields are computed only in the

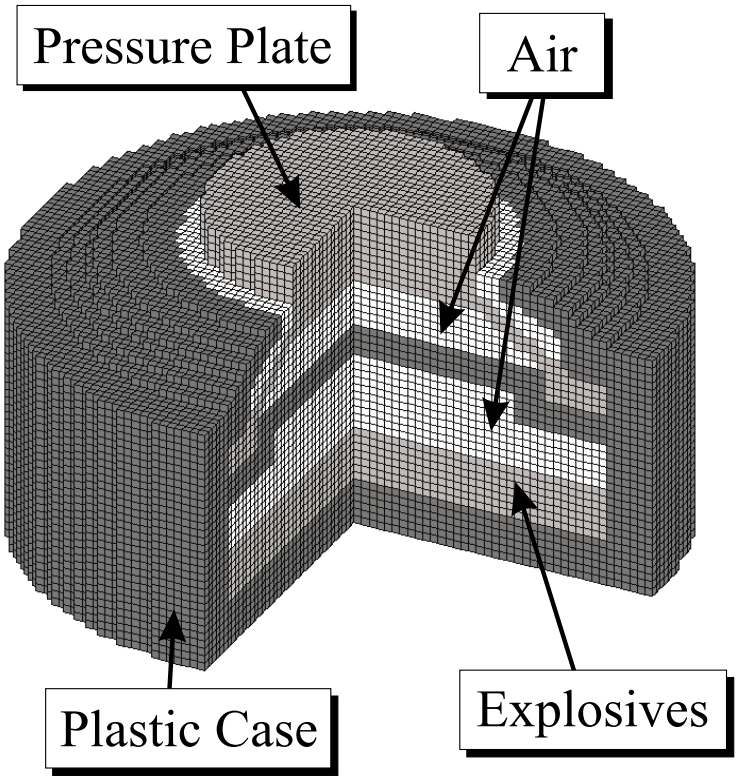


Figure 8.8: Refined mine model that includes the structural details of an actual TS-50 mine such as the explosives, the case, the rubber pressure plate, and the air chambers.

immediate surrounding of the mine. A grid with $120 \times 60 \times 120$ cells is used, yielding an effective solution space of 10 cm by 10 cm by 10 cm (because a PML of 10 cells is employed and the symmetry in the y-direction is exploited, reducing the size of the numerical grid by a factor of 2). The time step becomes, according to the Courant condition, $\Delta t = 2.138 \cdot 10^{-7}$ s. About 117,000 times steps are needed to compute the wave fields for 25 ms, requiring a computation time of about five hours and 23 minutes on the Beowulf PC cluster using 28 processors.

In the surrounding of the mine, the soil can be roughly approximated as homogeneous. Therefore, for all results presented in the following, the material properties in the ground are assumed constant throughout the medium. Instead of modeling the transducer on the surface, a vertically polarized plane shear wave is injected from the left. For this, the total-field/scattered-field formulation as described in Sec. 4.3 is used. The excitation is a differentiated Gaussian pulse with a center frequency of 1 kHz and, thus, has a wider bandwidth than the excitation used previously. A plane shear wave is chosen, because it resembles the incident guided shear wave and the Rayleigh surface wave that, as observed previously, are excited by the transducer on the surface. The scenario is schematically depicted in Fig. 8.9.

The interaction of an incident plane shear wave with a buried mine as computed with the detailed mine model is illustrated in Fig. 8.10. The vertical particle displacement on the surface plane (top) and on a cross section through the ground (bottom) is shown at four instants in time. A color scale ranging from black to white is employed. The mine is buried 2 cm beneath the surface. The material properties in the ground are assumed to be $\rho = 1400$ kg/m^3, $c_P = 250$ m/s, and $c_S = 80$ m/s.

In Fig. 8.10, resonant oscillations are seen to be excited at the buried mine. The oscillations appear to be mainly due to the motion of the thin soil layer above the mine and the rubber pressure plate of the mine. The oscillations decay as time progresses. The surface motion appears to be composed of at least two different modes. At T_1 and T_3, the motion is dominated by a mode describing an upward and downward motion of the surface above the mine. At T_4, a mode causing a rocking motion of the

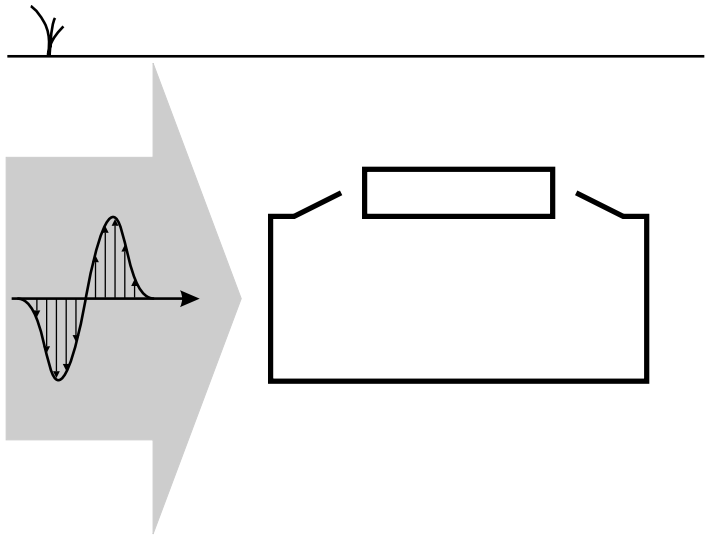


Figure 8.9: A vertically polarized shear wave is incident on the buried mine. The elastic wave fields are computed only in the surrounding of the mine.

surface is apparent. A null in the vertical particle displacement along the rotary axis is clearly visible on the surface and on the cross section. At T_2, the surface motion is composed of the superposition of at least these two modes.

From the results obtained with the detailed mine model, it becomes clear that the resonance is not only due to flexural modes of the thin soil layer, but rather must be attributed to the resonant behavior of the soil-mine system. The resonant oscillations are caused by the motion of both the soil layer and the rubber pressure plate. This motion will be influenced by various factors. On the one hand, it will strongly depend on the stiffness and thickness of the soil layer above the mine. On the other hand, it will be affected by the mine and its structure. For example, below the rubber pressure plate lies a closed air chamber (see Fig. 8.8). This air chamber acts like a spring and will have a significant impact on the motion of the pressure plate.

To analyze the resonant oscillations at the location of a buried mine in detail, a parametric study of the resonance as a function of the burial depth and the shear wave speed in the ground is conducted. For this, the vertical displacement at one point on

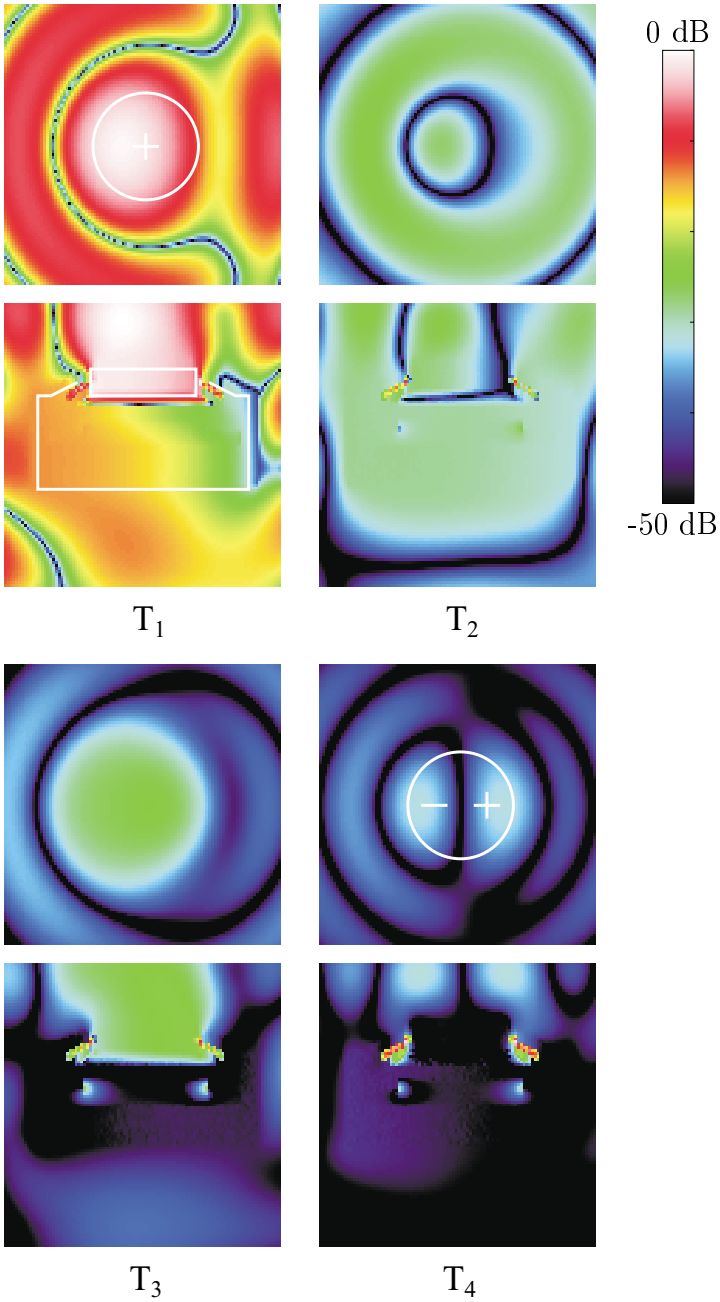


Figure 8.10: Resonant oscillations at the location of a buried land mine; mine buried 2 cm beneath the surface. The vertical particle displacement in the surrounding of the mine as computed with the detailed mine model is shown at four instants in time on the surface plane (top) and on a cross section through the ground (bottom).

the surface above the mine is computed in the time domain. A point is picked that lies 1 cm off the vertical symmetry axis of the mine. The motion is Fourier transformed from the time domain into the frequency domain, and, by dividing by the Fourier transform of the incident pulse, the transfer function is determined. From the transfer function, the resonant frequencies of the vertical particle motion can be identified. For the parametric study, shear wave speeds in the range from 40 m/s up to 150 m/s are considered in order to determine the response for fairly loose as well as rather stiff soils. These values are representative for the shear wave speeds at the surface for many common soil types. The pressure wave speed and the material density will be kept constant throughout the study, at $c_P = 250$ m/s and $\rho = 1400$ kg/m^3.

Fig. 8.11 shows the transfer function for a mine buried 1 cm beneath the surface. The shear wave speed in the ground is assumed to be 40 m/s. In the range up to 2.1 kHz, two resonant peaks can be distinguished. The two peaks correspond to two modes of the surface motion. The first mode has a resonant frequency of about 691 Hz and corresponds to the up- and downward motion of the surface. The second mode, at about 1.373 kHz, describes the rocking motion. The surface motion associated with the modes is illustrated in Fig. 8.11. The two modes that are excited within the soil-mine system are similar in shape to the first two modes of a thin circular disc that is rigidly clamped along its edge [58, 79]. However, the analogy to the circular disc cannot be used to accurately predict the resonant frequencies of the soil-mine system, because along its edge the soil layer in the model is not rigidly fixed and the structure of the mine has a significant impact on the resonance. An additional peak at about 2 kHz is apparent in Fig. 8.11. The origin of this peak, however, is not clearly understood. It might be due to a resonance in the mine itself. Further studies, beyond the scope of this work, are needed to explore the resonant behavior at higher frequencies.

To explore the resonance for mines buried in different types of soil, the resonant behavior is determined as a function of the shear wave speed in the ground and the burial depth of the mine. Fig. 8.12 shows the transfer function as a function of the

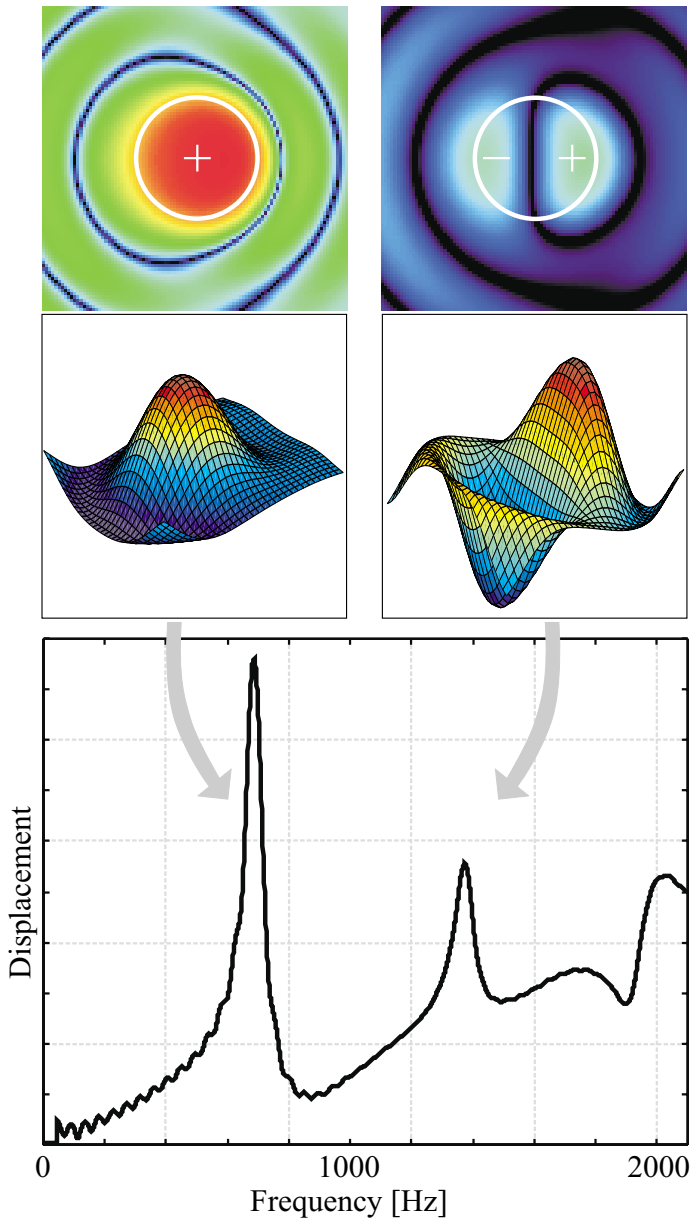
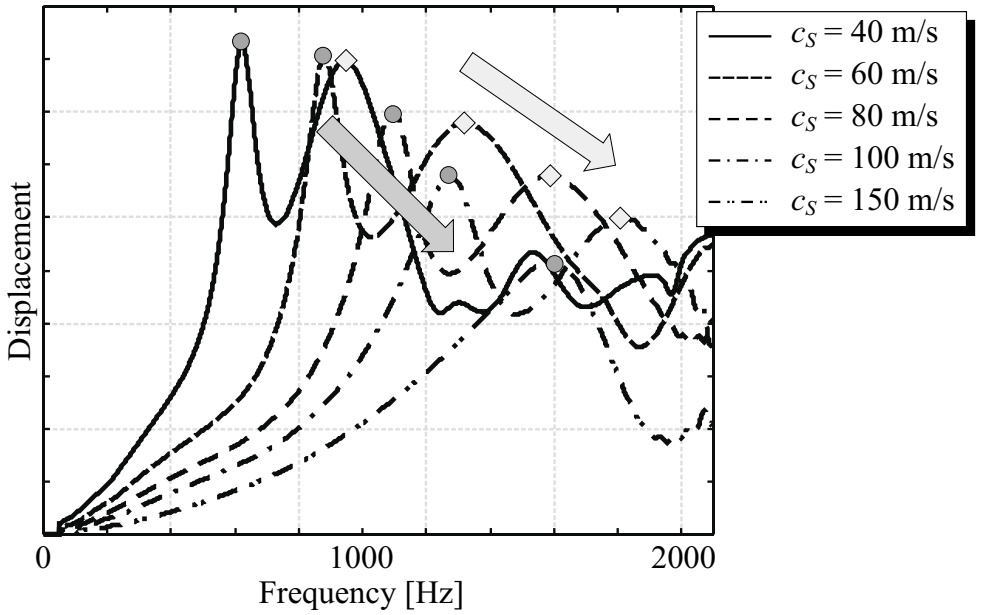


Figure 8.11: Transfer function of a point on the surface above the mine; mine buried at 1 cm; $c_S = 40$ m/s. The two resonant peaks correspond to two distinct modes of the vertical surface motion.

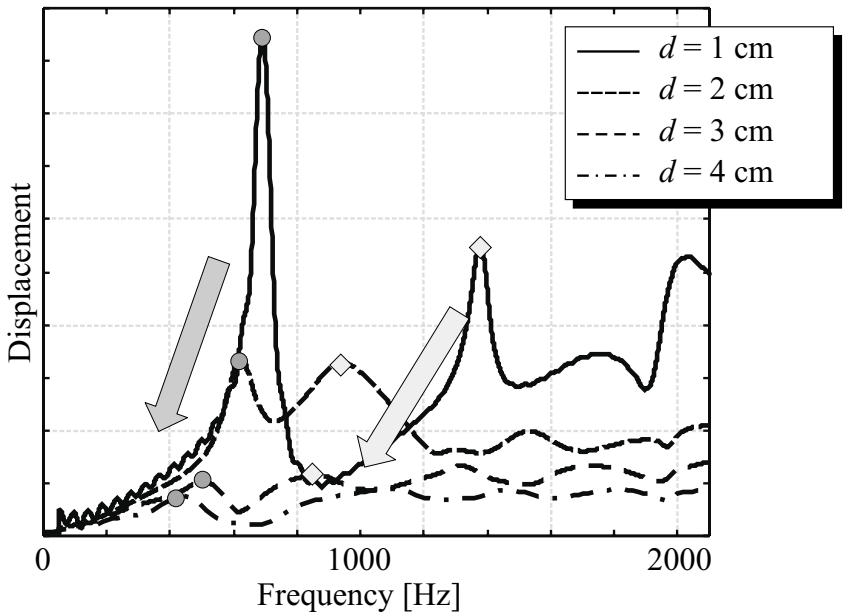
shear wave speed in the ground with the burial depth of the mine fixed at 2 cm (top) and as a function of the burial depth with the shear wave speed kept constant at 40 m/s (bottom). It is apparent that for increasing shear wave speed the resonant frequencies of the first and second peak are shifted upwards in frequency. At the same time, the amplitudes of the peaks drop, and the quality factor of the modes is degraded. For increasing burial depth, the two resonant peaks are shifted slightly downwards, and their amplitudes drop drastically. For the mine at 4 cm, the second resonant peak is not distinguishable any more.

To summarize the results, the resonant frequency and the quality factor of the first resonant peak are plotted in two parametric graphs as a function of the shear wave speed in the ground with the burial depth as a parameter. The results are shown in Fig. 8.13. Note that, by varying the shear wave speed in the ground while keeping the material density constant, effectively the stiffness in the ground is being varied. Thus, in these plots the resonant behavior is shown essentially as a function of the stiffness in the ground. The top line in each plot corresponds to the mine at a burial depth of 1 cm. Qualitatively, it can be seen that the resonant frequency increases with shear wave speed and decreases with burial depth. The quality factor of the resonance decreases with both shear wave speed and burial depth. The decrease in the quality factor of the resonance with both increasing stiffness and burial depth indicates that the resonance is dampened. The only damping mechanism affecting the resonance within the model is the radiation of energy into the surrounding soil. The coupling of the thin soil layer above the mine to the surrounding soil increases with the stiffness and the thickness of the layer. Thus, the energy that is radiated increases, and the damping of the resonance is enhanced.

Resonant oscillations at the location of a buried land mine have also been observed experimentally. Fig. 8.14 shows measured results for an actual TS-50 mine buried 1 cm beneath the surface. In the top graph the measured transfer function of the surface motion at a point above the mine is plotted, whereas in the bottom graph various numerically determined transfer functions are shown. The curves are seen



(a)



(b)

Figure 8.12: Surface motion at a point above the mine; transfer functions. (a) Mine buried 2 cm beneath the surface, variable shear wave speed; (b) shear wave speed $c_S = 40$ m/s, variable depth.

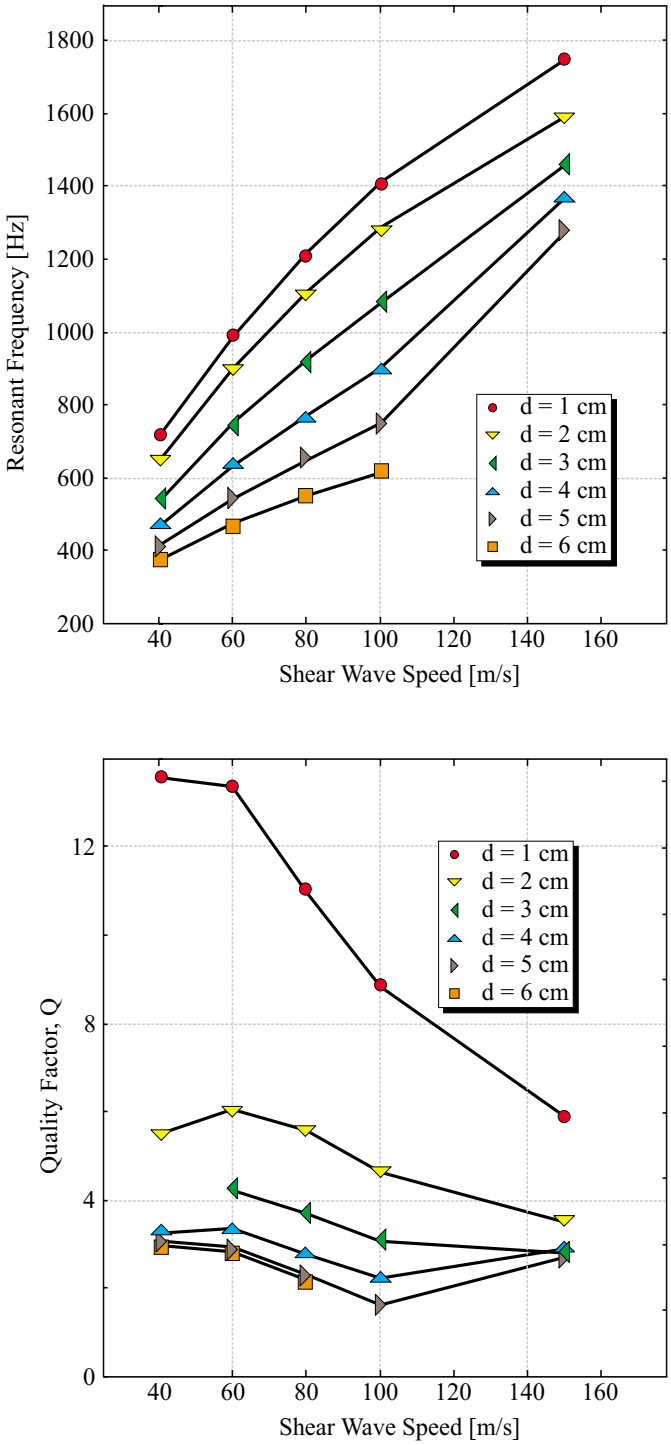


Figure 8.13: Parametric graphs of the resonant frequency and the quality factor of the first resonant peak as a function of the shear wave speed in the ground with the burial depth as a parameter.

to be quite different. The measured transfer function exhibits two resonant peaks at about 450 Hz and 700 Hz. The second peak is very weak, and the response is negligible for frequencies higher than about 1 kHz. The numerical transfer functions for four cases are plotted: for a mine at 1 cm assuming shear wave speeds in the ground of 20 m/s, 30 m/s and 40 m/s, and for a mine at 2 cm assuming a shear wave speed in the ground of 40 m/s. The best resemblance with the experimental data is obtained for the first of these cases. When assuming a burial depth of 1 cm and a shear wave speed of 20 m/s, two peaks occur, approximately at 400 Hz and at 750 Hz (bottom, solid curve). Contrary to the experimental result, however, the numerical results indicate a significant response at high frequencies. It is believed that the differences between experimental and numerical results are mainly due to uncertainties in the experimental model and to non-linear effects occurring in the sand. Earlier, an estimate of the wave speed variation with depth has been shown (see Fig. 7.1). This depth profile might be especially inaccurate at and close to the surface, predicting wave speeds at the surface that are far too large. Nominally, in the experiment the mine is buried at 1 cm. However, the actual burial depth of the mine might differ, because the level of accuracy that can be achieved when placing the mine in the ground is limited. Thus, the actual burial depth of the mine will lie roughly within ± 0.5 cm of the nominal burial depth. The sand in the experiments behaves in a strongly non-linear way. The non-linear effects dampen the high frequency components of the elastic waves and effectively suppress frequency components above about 500 Hz. Due to the resonant oscillations at a buried mine, the surface motion above the mine is amplified and driven strongly into the non-linear regime, thus enhancing the effects of loss and non-linearities. To obtain a better agreement of experiment and numerical simulation, the experiment must be repeated using smaller drive levels, thus exciting elastic waves with a smaller amplitude to achieve operation in the linear regime of the sand.

The numerically determined transfer functions exhibit ripples at low frequencies. These ripples are numerical artifacts and are due to the fact that the resonant response

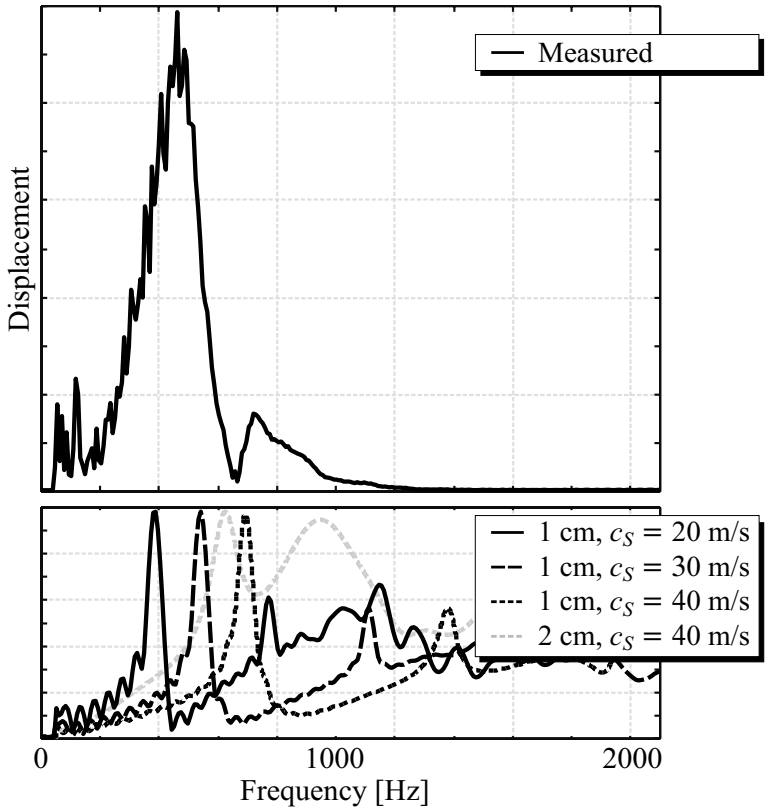


Figure 8.14: Experimental results for a mine buried at 1 cm; transfer function.

has not been computed long enough. Within the time window for which the wave fields are computed, the response of the mine has not completely decayed, and thus the numerical transfer function exhibits the true response of the mine convolved with the shape of the time window. In other words, the transfer function in the frequency domain is the product of the Fourier transforms of the true response and the rectangular time window (i. e., a sinc-function), causing the ripples visible in Fig. 8.14.

8.3 Excitation by a Pressure Wave

In the mine detection system that is investigated in this work, the elastic waves are excited by an electrodynamic transducer placed on the surface of the ground. Such a

transducer will launch predominantly shear and Rayleigh surface waves which propagate along the surface of the ground and interact with the buried mine. This scenario resembles the case that has been studied in the last section, where the buried mine was excited by a shear wave incident from the side. A plane shear wave incident onto the mine from the side will excite resonant oscillations at the mine similar to the ones excited by a Rayleigh wave or a guided shear wave as observed in Chapter 7. Different types of excitations are, however, possible in an acoustic mine detection system. In a different approach, elastic waves in the ground are excited by a loudspeaker mounted above the surface [3, 10]. The loudspeaker radiates acoustic pressure waves, which impinge upon the surface of the ground and couple primarily into pressure waves in the ground. The pressure waves do not propagate along the surface, but rather are directed vertically into the ground. To study this form of excitation, the response of a buried mine to pressure waves being incident from above the mine is analyzed using the detailed mine model.

Fig. 8.15 shows the configuration schematically. A pressure wave is vertically incident upon the buried mine. A differentiated Gaussian pulse with a center frequency of 1 kHz is used as excitation. As before, the material properties are assumed to be homogeneous around the mine.

Fig. 8.16 shows the transfer function of a point on the surface above the mine for the excitation with a shear wave incident from the side (solid line) and with a pressure wave incident from above (dashed line). The mine is buried 1 cm beneath the surface, and the shear wave speed in the ground is $c_S = 40$ m/s. The differences are apparent. As described previously, the shear wave excites two resonant peaks in the range up to 2.1 kHz. The pressure wave, however, excites only the first resonant mode. This is readily explained by noting the symmetry of the incident pressure wave. The plane pressure wave is symmetric with respect to the vertical axis of the mine. Thus, it will excite only modes that are equally symmetric. The second mode excited by the shear wave is asymmetric with respect to the mine's axis, as visible in Fig. 8.11, and consequently it is not excited by the pressure wave incident from

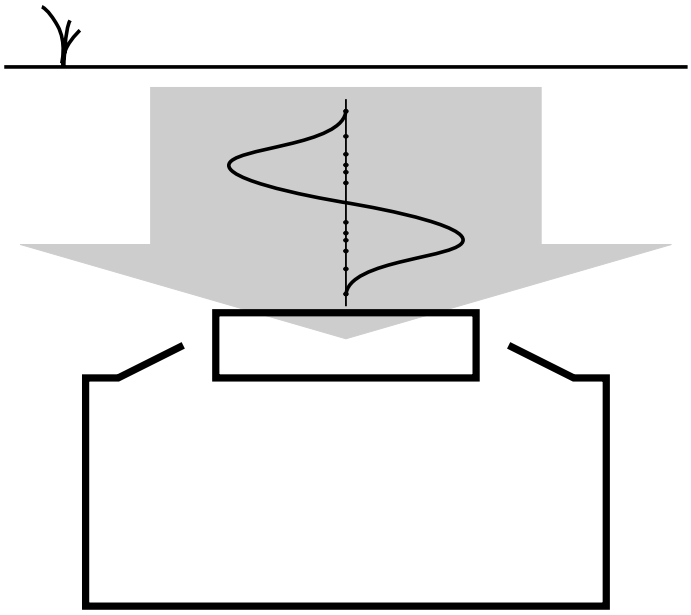


Figure 8.15: A pressure wave is vertically incident on the buried mine.

above.

Fig. 8.17 shows the transfer function due to the excitation with an incident pressure wave in (a) as a function of the shear wave speed in the ground for the mine buried at 2 cm and in (b) as a function of the burial depth for a shear wave speed of 40 m/s. For increasing shear wave speed and, thus, for increasing shear stiffness, the resonant frequency of the mode excited by the pressure wave is shifted upwards in frequency and the amplitude of the resonance decreases, causing the quality factor of the resonance to be degraded. When increasing the burial depth, the resonant frequency is shifted slightly downwards in frequency and the amplitude drops, indicating that the resonance is increasingly dampened. The resonant mode that is excited by the pressure wave behaves in a similar way as the first resonant mode excited by the shear wave incident from the side (see Fig. 8.12). The second resonant mode that is excited by the shear wave, however, is not apparent for the pressure wave excitation. Note that the resonant peak for the mine at 3 cm is barely distinguishable and for the mine at 4 cm has vanished or is superimposed by other

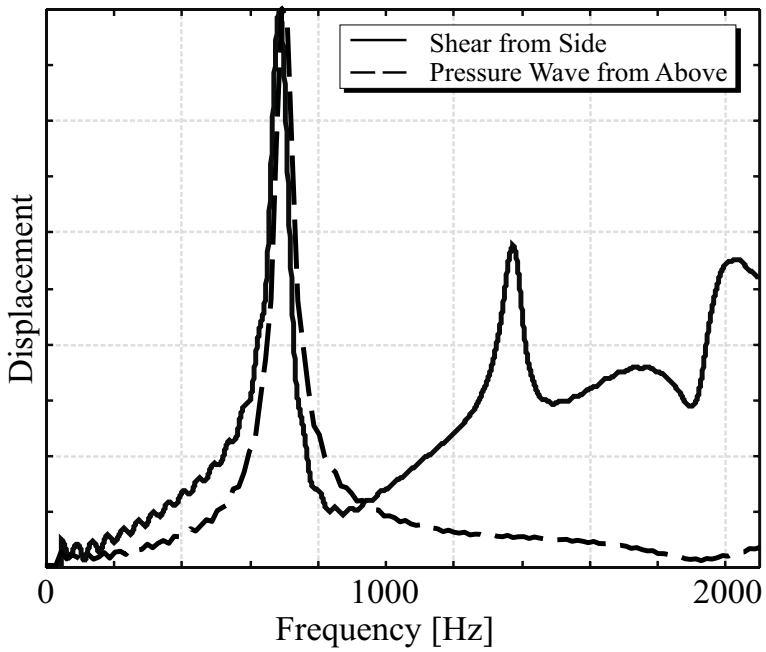
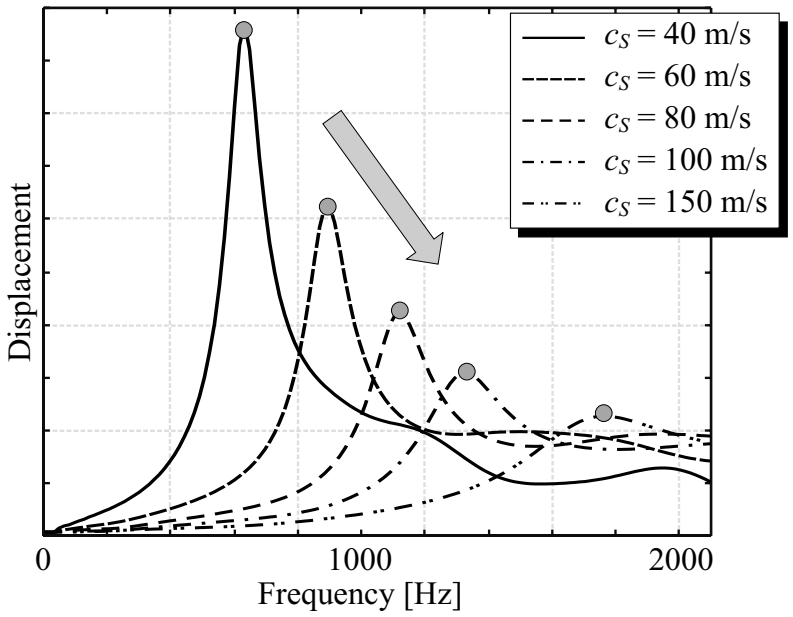


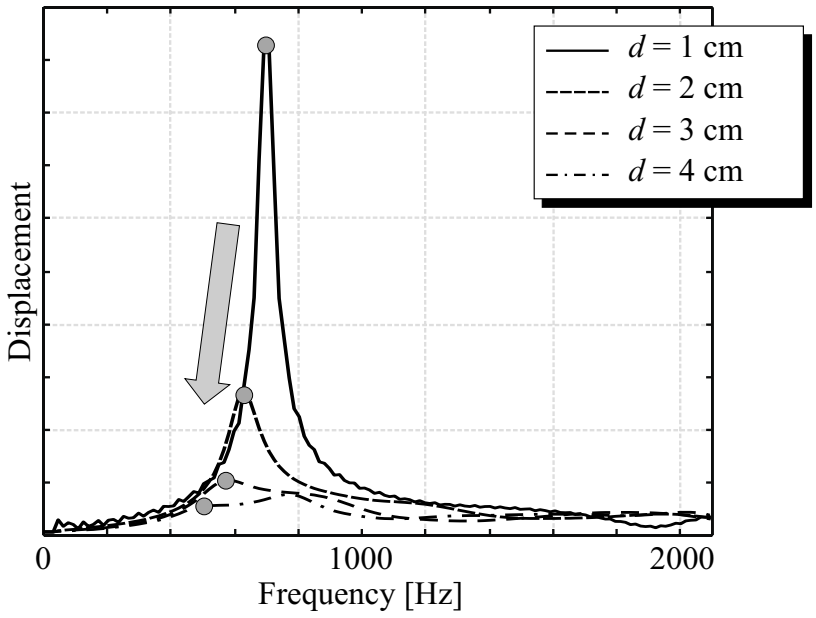
Figure 8.16: Transfer functions of the surface motion due to the excitation with shear wave from the side and a pressure wave from above.

effects.

The results are summarized in Fig. 8.18 in parametric graphs showing the resonant frequency and the quality factor of the first mode as a function of the shear wave speed in the ground with the burial depth as a parameter. Burial depths of 1 cm, 2 cm and 3 cm are considered. For larger burial depths, the first resonant mode is not clearly identifiable, and other resonant effects of the mine become dominant. The results are similar to the results for the shear wave excitation. The resonant frequency increases with the shear wave speed in the ground and decreases with the burial depth. The quality factor of the resonance decreases with both shear wave speed and burial depth.



(a)



(b)

Figure 8.17: Surface motion at a point above the mine; tranfer functions. Excitation with a pressure wave from above. (a) Mine buried 2 cm beneath the surface, variable shear wave speed; (b) shear wave speed $c_S = 40$ m/s, variable depth.

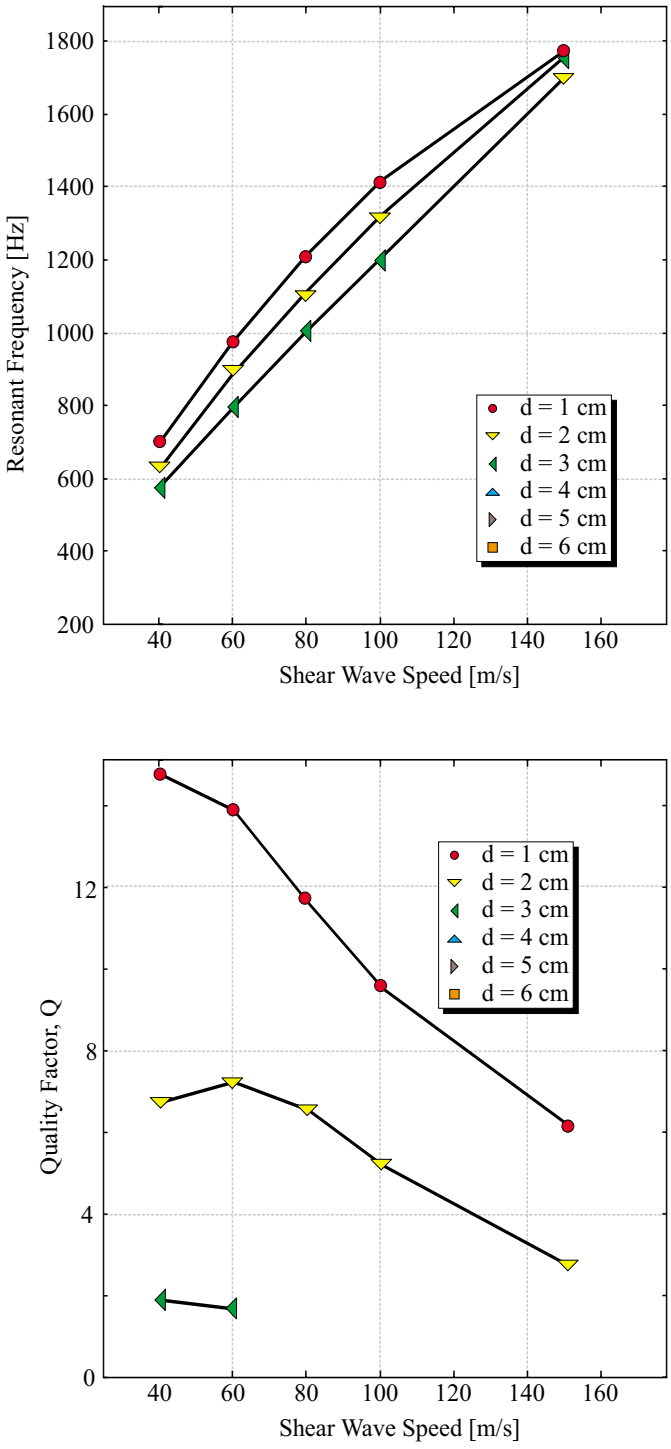


Figure 8.18: Parametric graphs of the resonant frequency and the quality factor of the first resonant mode as a function of the shear wave speed in the ground with the burial depth as a parameter. Pressure wave excitation.

CHAPTER 9

Experiments

In experimental measurements, elastic waves propagating along the surface of the ground have been observed to be dispersed. The dispersion is believed to be largely due to the material properties in the ground being inhomogeneous. As has been argued earlier (see Chapter 7), a Rayleigh surface wave will be dispersed if the wave speeds in the ground are a function of depth, because the reach of the Rayleigh wave into the ground depends on frequency, and, thus, different frequency components of the Rayleigh wave will experience different wave speeds. To explore the wave propagation in the ground, the elastic wave fields in the ground are measured.

The laboratory experiments are performed in a large sand-filled tank. If the elastic wave fields in sand are assumed to behave in a linear manner, the material properties of sand can be described by three independent constants: the material density ρ, the shear wave speed c_S, and the pressure wave speed c_P. In this chapter, procedures to measure these quantities as a function of depth are described, and measurement results are presented.

The measurement of soil properties has been treated extensively in the literature [81] – [83]. The signal processing techniques applied in this chapter are adapted from Oppenheim and Schafer [84] and Smith *et al.* [62].

9.1 Measuring the Water Content in the Ground

Before a measurement, the sand in the laboratory tank is usually uniformly wetted and packed. This is primarily done to minimize the loss effects at the surface and to improve the lateral uniformity of the material properties in the tank. A water table is maintained within the tank to approximate realistic conditions. Consequently, the water content within the sand is largely inhomogeneous with depth.

Changes in the water content with depth might have a significant impact on the wave propagation in the ground. If the water in the ground is assumed to primarily fill in the voids of the sand matrix (which is constituted by the grains of the sand), the water content will not influence the mass of the sand per volume, but it will affect the overall mass per volume, thus changing the material density in the ground. Additionally, the water content in the ground will affect the cohesion between the grains of the sand. A change in the cohesion will primarily influence the wave speeds in the ground, because an increase in the cohesion is synonymous with an increase in the stiffness.

To measure the water content within the sand, an easy and straightforward procedure has been applied. Various samples of the sand were taken at a range of depths. The samples were weighed using a conventional scale accurate to 0.01 g and then dried to remove the water. A conventional oven at 400 degrees Fahrenheit was used to dry the samples, baking the samples for about twelve hours. The samples were finally weighed again. The weight difference between the initial sample and the dry sample represents the water content. Table 9.1 lists the results of the water content measurement, and Fig. 9.1 shows the water content in percent of the mass of sand as a function of depth. The water content was calculated as

$$\text{Water Content [\% mass]} = \frac{W_1 - W_2}{W_2} \cdot 100, \tag{9.1}$$

where W_1 indicates the weight of the sample before drying, and W_2 is the weight after drying. In Eq. (9.1), the atmospheric pressure and the pressure within the sand were neglected. Note that the water table has been found to lie at a depth of about 52 cm.

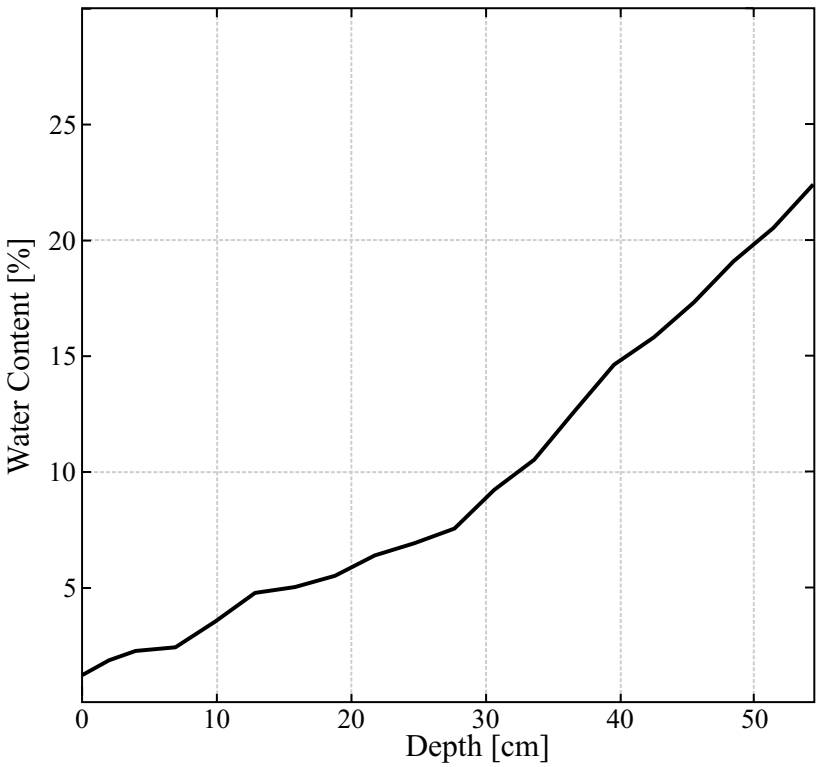


Figure 9.1: Water content (mass percent) as a function of depth.

In the water table, the mass of the water per volume is about 22 % of the mass of sand per volume.

As stated above, if the water fills in the matrix of the sand, the mass of the sand per volume will be constant, but the overall mass per volume will change due to the additional mass of the water. The change in the material density is then

$$\rho = \rho_{\text{sand}} + \frac{\text{Water Content [\% mass]}}{100}\rho_{\text{sand}}. \tag{9.2}$$

The density of dry sand is assumed to be $\rho_{\text{sand}} = 1400$ kg/m^3. The resulting material density as a function of depth is graphed in Fig. 9.2.

The depth dependence of the material density might have an effect on the other material properties. For example, if the stiffness in the ground is assumed to be constant, an increase in the material density will cause a decrease in the wave speeds.

Table 9.1: Measurement of the water content as a function of depth.

Depth [cm]	W_1 [g]	W_2 [g]	Water Content [% mass]
0	127.10	125.31	1.44
2	140.93	138.09	2.06
4	167.66	163.62	2.46
7	164.60	160.40	2.61
10	148.85	143.50	3.70
13	164.04	156.16	4.90
16	151.87	144.25	5.15
19	155.18	146.65	5.61
22	191.96	179.75	6.48
25	142.04	132.33	6.99
28	156.24	144.64	7.60
31	166.07	151.03	9.23
34	162.36	145.65	10.50
37	228.01	199.95	12.51
40	228.92	196.19	14.50
43	196.63	166.33	15.64
46	234.54	195.04	17.12
49	233.81	190.35	18.85
52	203.94	163.34	20.25
55	153.06	119.96	22.08

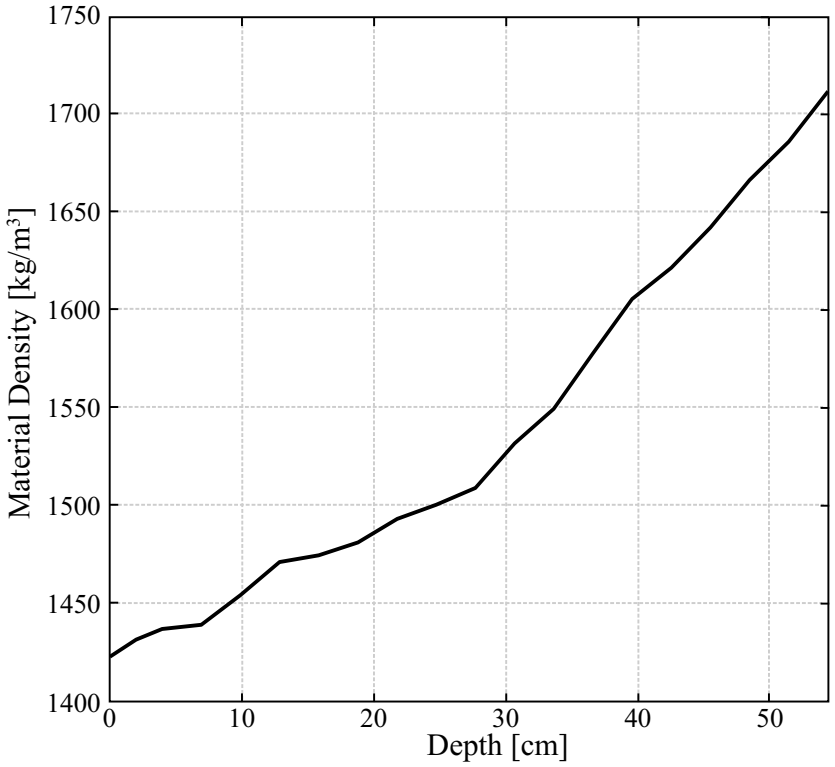


Figure 9.2: Material density as a function of depth.

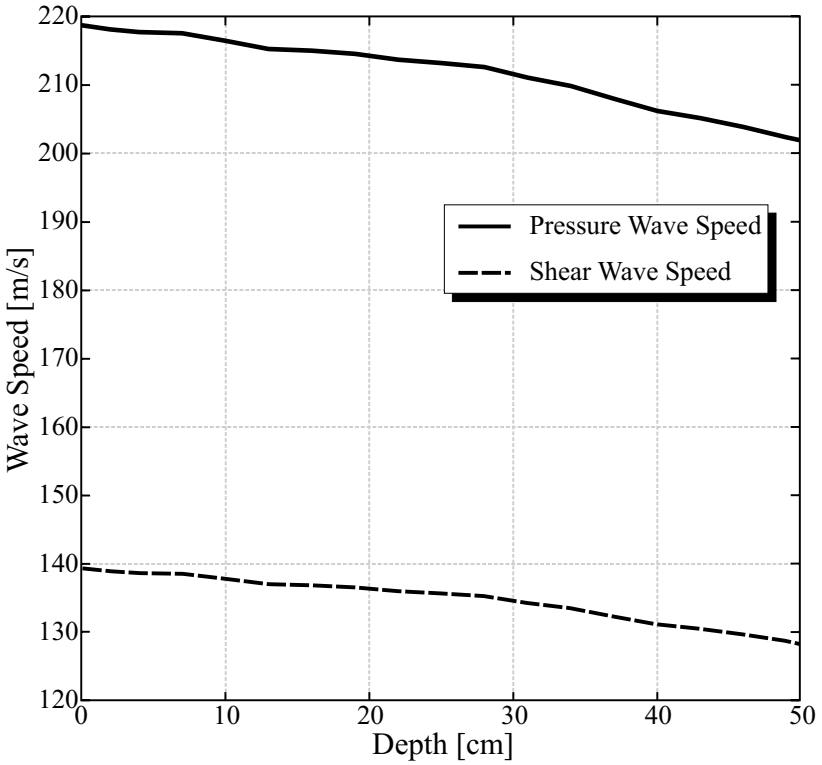


Figure 9.3: Pressure wave speed and shear wave speed as a function of depth, if the stiffness in the ground is assumed to be constant and the density changes as in Fig. 9.2.

This is depicted in Fig. 9.3. Here, the pressure wave speed and the shear wave speed in dry soil have been assumed to be 220 m/s and 140 m/s, respectively. Due to the increase of the material density with depth, the wave speeds decrease to values of about 202 m/s and 128 m/s. A qualitatively similar behavior of the wave speeds as a function of the water content in the soil has also been observed by Bachrach and Nur [83].

9.2 Measuring the Wave Speeds in the Ground

To measure the elastic wave fields in the ground, a vertical array of accelerometers is buried in the ground. With the accelerometers, the arrival times of the waves as a function of depth are estimated. From the arrival times, the wave speeds of the

different wave types can be determined.

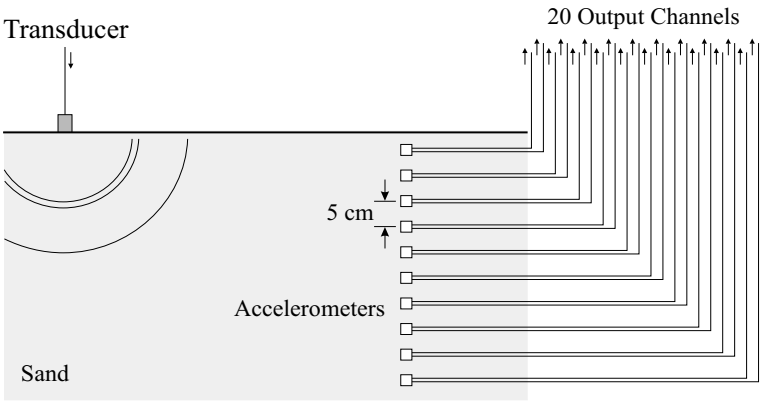
9.2.1 Experimental Set-Up

Fig. 9.4 gives a schematic overview of the measurement setup. The elastic waves are excited by a *transducer* placed on the surface of the ground. A vertical array of *accelerometers* is buried in the ground. Ten bi-axial accelerometer pairs measuring the horizontal and vertical particle acceleration are used, thus providing 20 signals. The shallowest accelerometer pair is placed at a depth of 5 cm, the deepest at 50 cm, and the spacing between the accelerometer pairs is 5 cm. The accelerometers are powered by a *power supply decoupling unit* supplying a DC voltage of 27 Volts. The 20 measured signals are amplified by a *preamplifier* with a constant gain of 320 and fed into a computer through a *data acquisition card*.

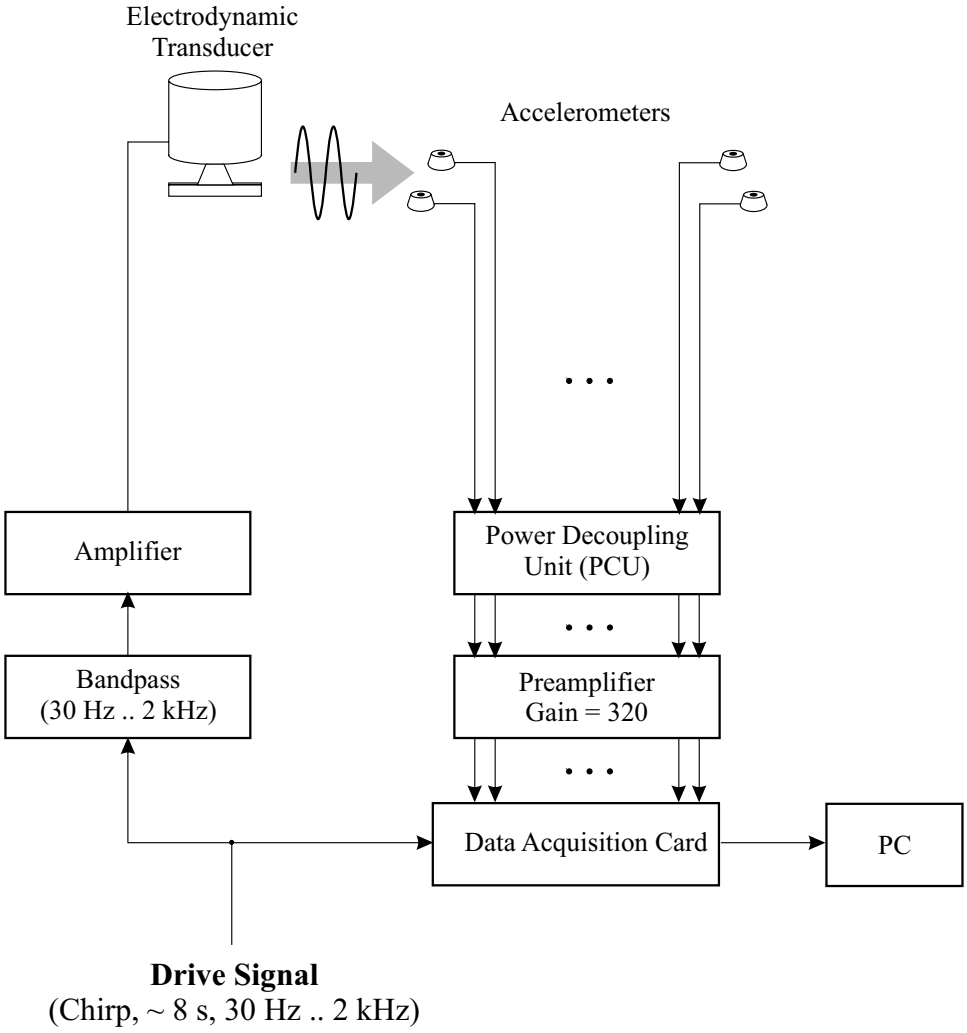
The accelerometers (manufactured by VIBRA-METRICS, Model 9002 A) are coated to improve their water resistance and enclosed in syntactic foam, making the accelerometers neutrally buoyant in water. Two accelerometers were orthogonally mounted on a piece of plexiglas, raising the material density of the formation to roughly the density of sand (≈ 1400 kg/m^3).

The drive signal fed to the transducer is a frequency modulated chirp with a length of 8 s and has a frequency range from 30 Hz to 2 kHz. Because the response to the chirp is rather complicated to interpret, the measured signal is processed. For each channel, the transfer function is determined and convolved with a differentiated Gaussian pulse.

Fig. 9.5 illustrates the signal processing. A chirp is supplied to the transducer (*Drive Signal*, top right). Ten accelerometer pairs measure the acceleration in the ground, providing 20 signals (*Accelerometer Signal*, top left). The accelerometer signal and the drive signal are Fourier transformed. By dividing the Fourier transform of the accelerometer signal by the Fourier transform of the drive signal, the transfer function for each channel is obtained (*Transfer Function*, third from top on the left).



(a)



(b)

Figure 9.4: Set-up for accelerometer measurements. (a) Transducer and accelerometer array; (b) measurement set-up.

The transfer function is multiplied by the Fourier transform of a *Reference Signal* (differentiated Gaussian pulse, bottom right) and inverse-Fourier transformed. In this way, the signal due to the reference signal is obtained, i. e., the signal that would have been measured if the reference signal had been fed to the transducer instead of the chirp.

9.2.2 Measurement Procedure

Fig. 9.6 schematically shows the arrangement of the transducer and the accelerometers in the laboratory tank. The accelerometer measurements were repeated numerous times. The accelerometers were first buried by digging a large hole ($\sim$ 50 cm in diameter) using a conventional shovel. The hole was refilled and repacked, as the accelerometers were positioned in the ground. The measured results were noisy, and reflections occurred which could be attributed primarily to the disturbance of the sand due to the hole that was dug in the process of burying the accelerometers. To improve the measurements, the accelerometers were removed from the hole, and the soil in the entire sand box was repacked by using a commercial packing machine (VIBCO plate compactor). According to its specifications, the packing machine packs the sand well down to a depth of about 30 cm. The accelerometers were then buried by drilling a bore-hole-like cylindrical hole with a diameter of about 5 cm. The accelerometers were carefully aligned in the hole and regularly spaced at depths between 5 cm and 50 cm. The water table was measured to be at about 52 cm, and thus the deepest accelerometer was placed just above the water table. The particle acceleration in the ground was measured for various transducer positions with respect to the burial hole. The transducer was placed at distances of 0.5 m, 1 m and 1.5 m from the location of the accelerometers and was displaced in two directions to investigate the effects of local inhomogeneities in the sand. The measurements were repeated several times, studying various conditions. The measurements were performed directly after burying the accelerometers and repeated throughout the following days, for a

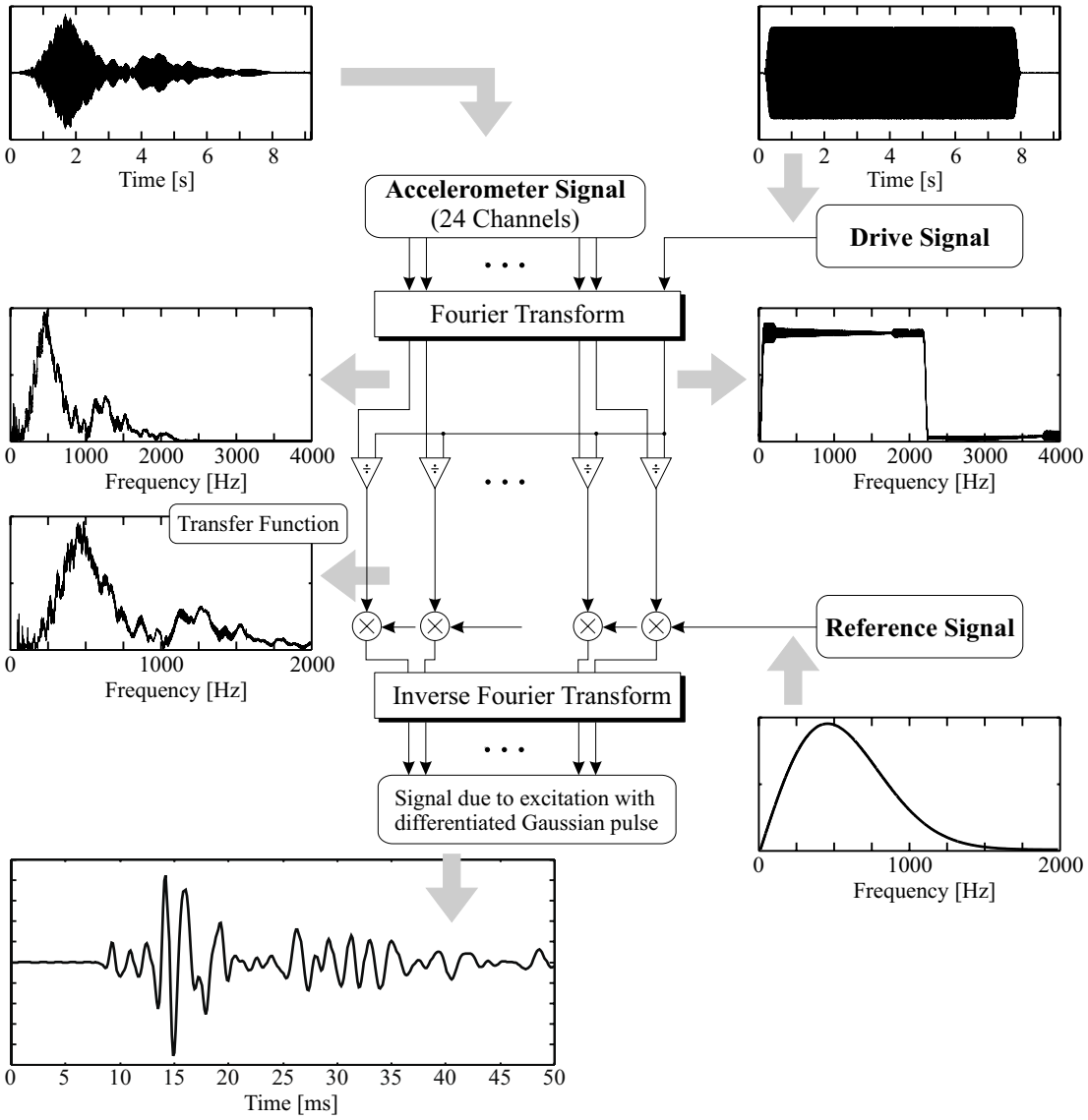


Figure 9.5: Signal processing: the accelerometer signal is Fourier transformed, divided by the Fourier transform of the drive signal (chirp), multiplied by the reference signal (differentiated Gaussian pulse), and then inverse-Fourier transformed.

wet surface as well as for a dry surface. To investigate the influence of the burial hole, the accelerometers were removed and reburied at a different location, and the measurements were repeated.

Before burying the accelerometers, the accelerometers were calibrated. The sensitivity of all accelerometers was found to lie within ± 10 % of the mean sensitivity. The differences in the accelerometer response were adjusted for during the signal processing of the measured signals.

Minimizing the disturbance of the sand by burying the accelerometers in a bore-hole-like hole clearly improved the measurement results, eliminating reflections from local inhomogeneities due to the accelerometer burial. The measurements with the bore hole arrangement were found to be largely independent of the time that had passed after the burial of the accelerometers, of the state of the surface (dry or wet), and of the burial location of the accelerometers. No reflections from the water table or the tank walls could be observed. It is believed that a residual content of air is present below the water table, causing the values of the wave speeds above and below the water table to be comparable in size [83]. In that case, the reflections from the water table are negligible. The reflections from the tank walls are believed to be largely attenuated by the lossy sand.

Various complications arose during the measurements. Several accelerometers became erratic during the initial measurements. This was believed to be due to damages in the sealed cases of the accelerometers, allowing water to seep into the accelerometers. To improve their sealing, the accelerometers were coated with multiple layers of polyurethane varnish. With the coating, the accelerometers were buried in the ground for an extended period of time without a noticeable deterioration in their performance.

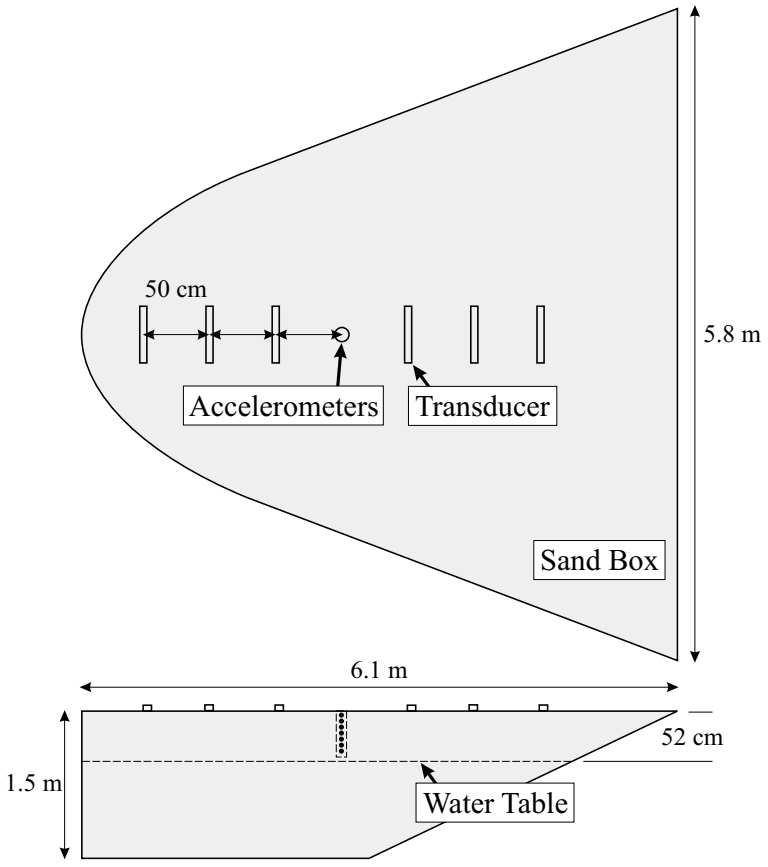


Figure 9.6: Accelerometer and transducer positions in the laboratory tank.

9.2.3 Results

Using the accelerometer array, the elastic wave fields in the ground are measured as a function of depth. Fig. 9.7 shows the wave fields as a function of depth and time for a lateral distance of 1.5 m between transducer and accelerometer location. At the top, the vertical and horizontal particle acceleration are graphed on a linear scale. The two plots below show the field components on a logarithmic scale using a dynamic range of 35 dB. The same color map as earlier is employed. The excitation is a differentiated Gaussian pulse with a center frequency of 450 Hz, launched by the transducer at 0 ms.

Five wave fronts are distinguishable. The fastest wave front (P) is the pressure wave. Due to the spherical shape of the propagating waves, the wave front recorded by the vertical array is curved. Note that at the surface the pressure wave arrives later than in the ground, indicating that the pressure wave speed at the surface is smaller than within the ground. The second wave front (LS), primarily visible in the horizontal field component, is the leaky surface wave. The leaky surface wave appears as a straight line that originates at the surface. The third wave front (S) represents the shear wave in the ground. The fourth wave (GS) corresponds to the guided shear wave and is strongest close to the surface. The slowest wave is the Rayleigh surface wave (R). Note that the Rayleigh surface wave at a depth of 5 cm (i. e., the top line in Fig. 9.7) has a prograde particle motion, indicated by the vertical component leading the horizontal component in phase, rather than a retrograde motion as was found earlier for a Rayleigh wave at the surface. This is due to the fact that the shallowest accelerometer is placed at a depth of 5 cm and not at the surface. As it was shown previously in Fig. 6.2, the particle motion of the Rayleigh wave reverses its direction at a depth of about 0.2 times the wavelength of the Rayleigh wave. At a frequency of 450 Hz, this equals roughly 3 cm. Thus, at a depth of 5 cm the Rayleigh wave indeed describes a prograde particle motion, and the measured results are in agreement with the theoretical prediction.

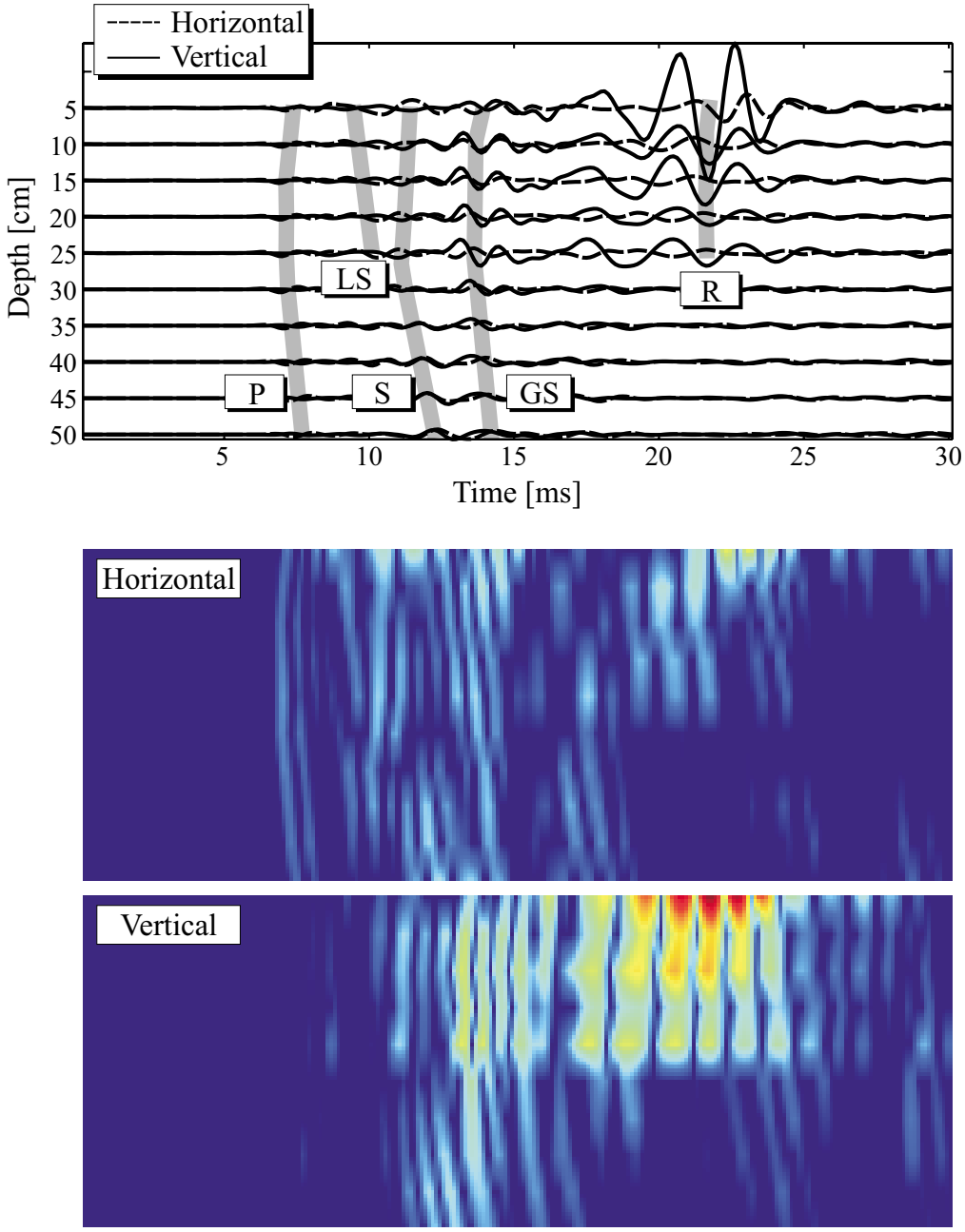


Figure 9.7: Horizontal and vertical particle acceleration as a function of time and depth; accelerometer measurements. The distance from transducer to accelerometers is 1.5 m. P: pressure wave, LS: leaky surface wave, S: shear wave, GS: guided shear wave, R: Rayleigh wave.

From Fig. 9.7, the arrival times of the separate wave fronts can be estimated as a function of depth. Table 9.2 lists the arrival times for all wave types and the corresponding wave speeds derived from the arrival times. The wave speeds are estimated by assuming a direct propagation path, thus neglecting the refraction of the waves due to the inhomogeneous material properties:

$$c \approx \frac{\sqrt{D^2 + d^2}}{t_{\text{arrival}}}, \tag{9.3}$$

where D is the lateral distance from the transducer to the accelerometers, and d is the burial depth of the accelerometers. From Table 9.2, the wave speeds of the shear wave

Table 9.2: Measured arrival times and wave speeds: pressure wave front.

	Pressure Wave		Leaky Wave		Shear Wave		Guided Shear Wave		Rayleigh Wave	
Depth [cm]	Arrival Time [ms]	Wave Speed [m/s]	Arrival Time [ms]	Wave Speed [m/s]	Arrival Time [ms]	Wave Speed [m/s]	Arrival Time [ms]	Wave Speed [m/s]	Arrival Time [ms]	Wave Speed [m/s]
5	7.32	205	9.71	154	10.86	138	14.09	107	21.59	69
10	7.14	211	9.68	155	10.74	140	13.59	111	21.69	69
15	7.07	213	9.65	156	10.67	141	13.59	111	21.59	69
20	7.07	214	9.77	155	10.64	142	13.62	111	21.56	70
25	7.11	214	9.77	156	10.64	143	13.59	112	21.53	70
30	7.07	216			10.58	145	13.78	111		
35	7.17	215			10.89	141	13.93	111		
40	7.23	215			11.05	141	13.99	111		
45	7.32	214			11.11	141	14.36	109		
50	7.48	211			11.39	139	14.64	108		

and the pressure wave as a function of depth can be deducted in an approximate sense. The pressure wave speed appears to be slightly smaller at the surface ($\approx$ 205 m/s at a depth of 5 cm) than within the ground ($\approx$ 215 m/s). The shear wave speed in the ground is roughly 140 m/s. The guided shear wave propagates at a speed of about 110 m/s, and the Rayleigh wave has a speed of about 70 m/s. The speeds of shear wave, guided shear wave and Rayleigh wave give indications about the shear wave speed variation in the ground. From the speed of the Rayleigh wave it can be

concluded that a layer with a rather slow shear wave speed exists close to the surface. Assuming that the Rayleigh wave speed equals about 0.9 times the shear wave speed at the surface, the shear wave speed of the surface layer can be estimated to be about 80 m/s. The shear wave front indicates that the shear wave speed beyond the surface layer is about 140 m/s. The guided shear wave propagates primarily in the region beyond the surface layer. The guided shear wave arises due to the rapid change in the shear wave speed in the transition zone between surface layer and the deeper medium. Because of the change in the shear wave speed, the guided shear wave is refracted towards the surface, is reflected by the surface, refracted back towards the surface and so forth. Due to the refraction, the path of the guided shear wave is curved, and the wave speed of the guided shear wave along the surface is smaller than the shear wave speed within the ground.

From the wave speeds within the ground, an approximate wave speed profile has been derived. This approximate profile was subsequently refined by trying to match numerical data to the measured results and by repeatedly tweaking the depth profile of the wave speeds within the numerical model. The estimated depth profile that yielded best agreement is shown in Fig. 9.8.

The depth profile shown in Fig. 9.8 is a crude approximation to to the actual depth variation of the material properties in the ground. To obtain a more accurate representation of the depth variation, more sophisticated techniques, such as numerical inversion algorithms, must be employed, which are, however, beyond the scope of this work.

Fig. 9.9 shows the wave fields as computed with the numerical model. The particle displacement is plotted on a linear scale for a transducer-accelerometer distance of 1.5 m. The arrival times and the curvature of the separate wave fronts agree fairly well with the experimental results (see Fig. 9.7). The measurements revealed that at a depth of 5 cm the Rayleigh wave describes a prograde particle motion. The numerical results predict the same. At the surface (top line), the particle motion for the Rayleigh wave is retrograde. At a depth of 5 cm, the horizontal motion reverses

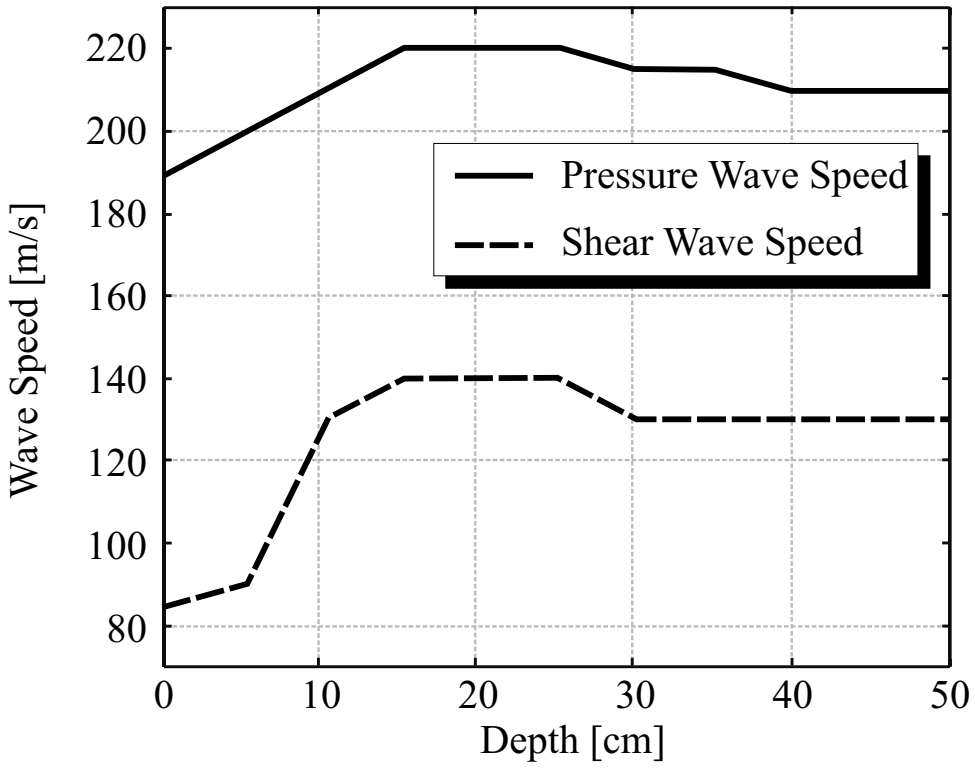


Figure 9.8: Pressure wave speed and shear wave speed as a function of depth.

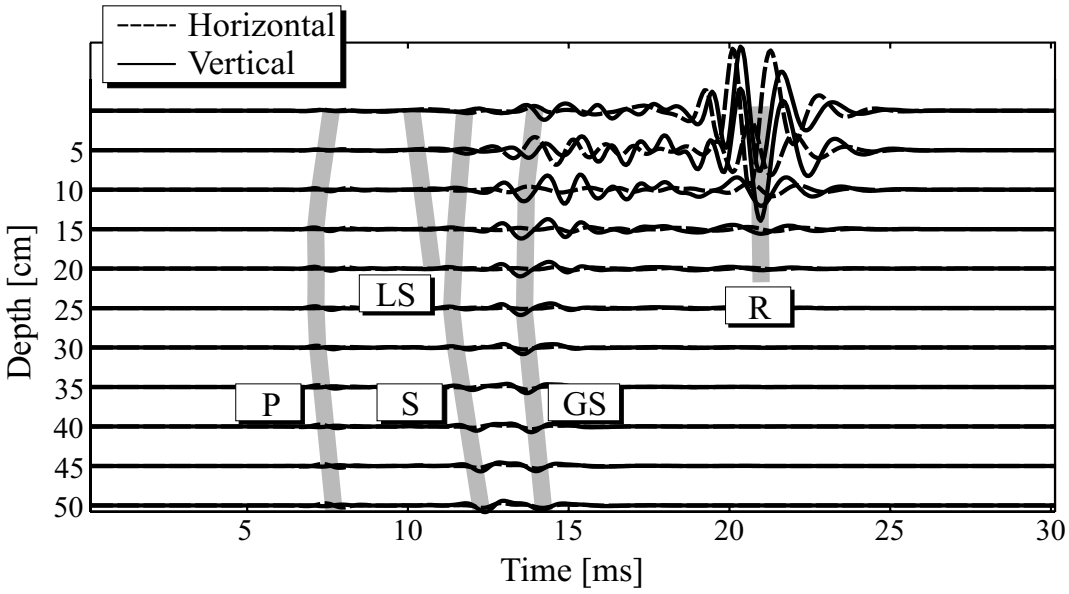


Figure 9.9: Horizontal and vertical particle acceleration as a function of time and depth; numerical simulation. The distance from transducer to accelerometers is 1.5 m. P: pressure wave, LS: leaky surface wave, S: shear wave, GS: shear wave, R: Rayleigh wave.

its sign, and thus the particle motion becomes prograde.

In Fig. 9.10, the wave fields in the ground as computed with the numerical model are shown and compared to the measured results for three transducer-accelerometer distances (0.5 m, 1 m, 1.5 m from top to bottom). The dynamic range is 35 dB for the experimental results (left column) and 50 dB for the numerical results (right column). Note that the measured results show the particle acceleration, whereas the numerical results show the particle displacement. The same qualitative behavior is observed for the experimental and the numerical results. The separate wave fronts have roughly the same shapes and the same wave speeds. Differences can mainly be observed for the leaky surface wave and the Rayleigh wave. The leaky surface wave is seen to be much weaker in the numerical results. The Rayleigh wave, in the measured results, reaches deeper into the ground than predicted by the numerical simulation. Furthermore, the horizontal component for the Rayleigh wave is significantly weaker in the measurement than in the numerical simulation.

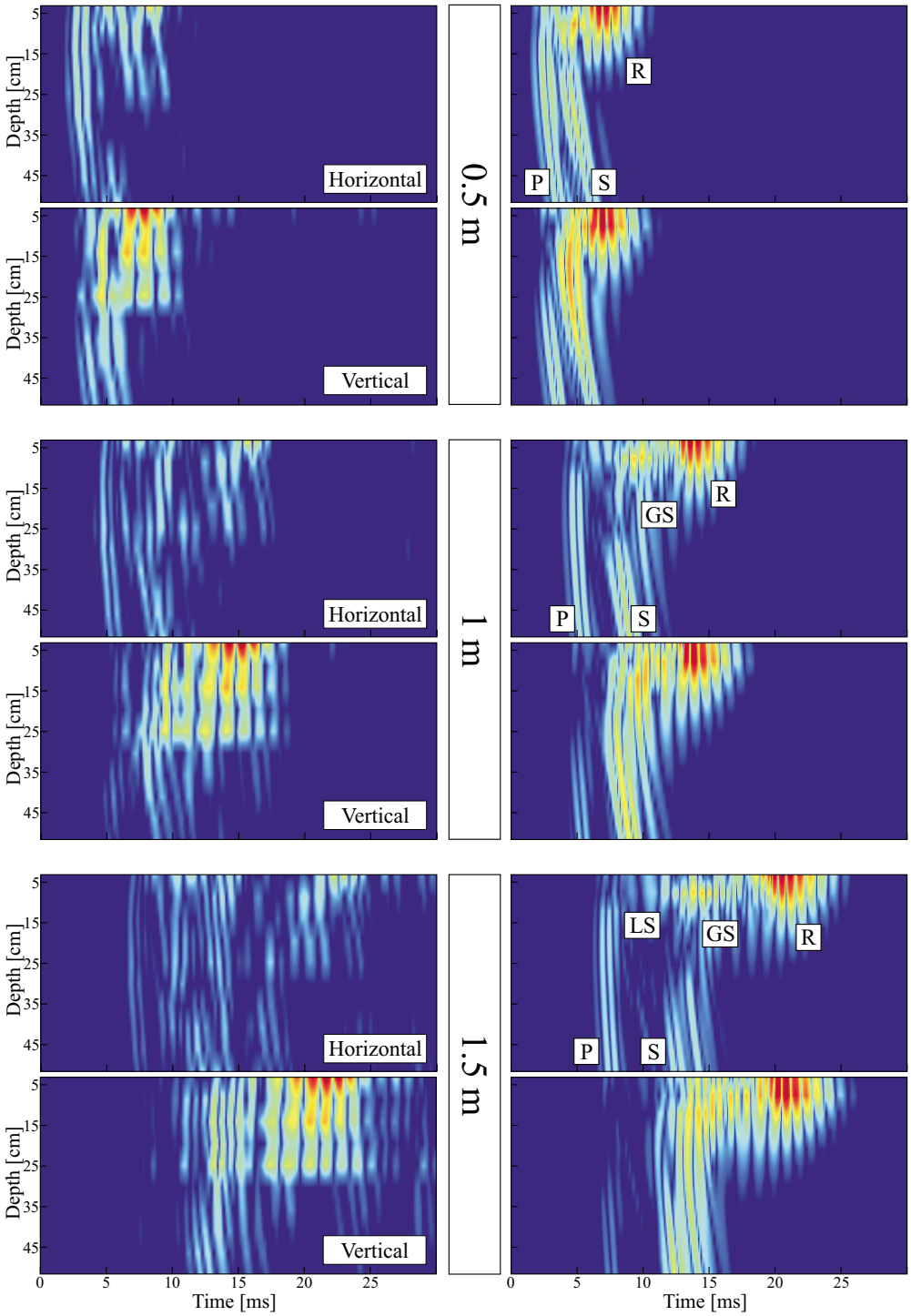


Figure 9.10: Horizontal and vertical particle acceleration as a function of time and depth; for three transducer distances: 0.5 m, 1 m, 1.5 m; left: accelerometer measurements, right: numerical simulation.

Frequency-Time Decomposition

The depth profiles for the material properties represent an approximation to the actual depth profile in the soil. The measured and computed results agree remarkably well, considering the uncertainties that still remain in the description of the material properties in the ground. For example, the material properties have been assumed to be laterally invariant, which certainly is not true in the actual soil. Furthermore, sand exhibits a strongly non-linear behavior, especially close to the surface. These non-linearities have been completely neglected in the numerical model and in the description of the material properties.

To identify and analyze the differences between experimental and numerical results in more detail, a *frequency-time decomposition* is applied. In this frequency-time decomposition, the particle motion at one point (1.5 m from the transducer, 5 cm beneath the surface, i. e., at the location of the shallowest accelerometer) is decomposed in its frequency components. For this, rather than convolving the transfer function with a differentiated Gaussian pulse, the transfer function is convolved with a Gaussian pulse shifted in frequency. The resulting signal represents the portion of the measured signal that is excited within the frequency band described by the Gaussian pulse. The shifted pulse can be expressed simply as a Gaussian pulse modulated by a carrier frequency:

$$\text{Reference Signal} = e^{-0.5(2\pi BWt)^2} \cdot \cos(2\pi f_C t), \tag{9.4}$$

where BW represents the bandwidth of the pulse and f_C is the carrier frequency. If, for example, the transfer function is convolved with a pulse having a bandwidth of 100 Hz and a carrier frequency of 800 Hz, then the signal excited in a 100 Hz band around a frequency of 800 Hz is obtained as a function of time.

Fig. 9.11 illustrates schematically the frequency-time decomposition. For the frequency-time decomposition, the frequency components between 50 Hz and 1.5 kHz are determined in 50 Hz increments. A Gaussian pulse with a bandwidth of 100 Hz is applied. The frequency-time decomposition as described here is similar to the

more commonly used decomposition by windowing along the time-axis and applying a Fourier transform to the windowed signal to obtain the frequency content within the time window [68, 84].

Fig. 9.12 shows the particle motion at a depth of 5 cm as a function of time and frequency for the measured and the computed results. The time-frequency decomposition as described above is applied. The distance from the transducer to the observation point is 1.5 m. The experimental results are plotted on a 35-dB scale, whereas the numerical results are plotted on a 50-dB scale. The horizontal axis indicates the time, the vertical axis shows frequency.

In the frequency-time decomposition plots, specific areas in the time-frequency plane can be attributed to specific wave types. The wave types are distinguished mainly by their arrival time and their frequency content. The fastest wave in Fig. 9.12 describes the pressure wave (P). Both in the measured and the numerical data the main frequency content of the pressure wave lies above 1 kHz. The second wave is the leaky surface wave (LS). The leaky surface wave has a much lower frequency content (below 0.5 kHz). Note that the leaky wave is much weaker in the numerical data. The shear wave (S) lies between 0.5 kHz and 1 kHz for the experimental results, but seems to extend to higher frequencies for the numerical data. The guided shear wave (GS), slower than the shear wave, is strongest in the range above 1 kHz. The Rayleigh surface wave (R), finally, has a frequency content below 1 kHz in the measurements, but extends to higher frequencies in the numerical results. Overall, in the experimental data the low frequency components appear to be stronger, whereas in the numerical data the high frequency components are enhanced. The differences are especially striking for the Rayleigh wave and the leaky surface wave. The differences between numerical and experimental data are believed to be due to the representation of the depth variation in the numerical model being especially inaccurate close to the surface and to non-linear effects occurring in the sand. The non-linear effects in the sand primarily dampen the high frequency components. The non-linearities are believed to be strongest at the surface, thus affecting most the Rayleigh surface wave

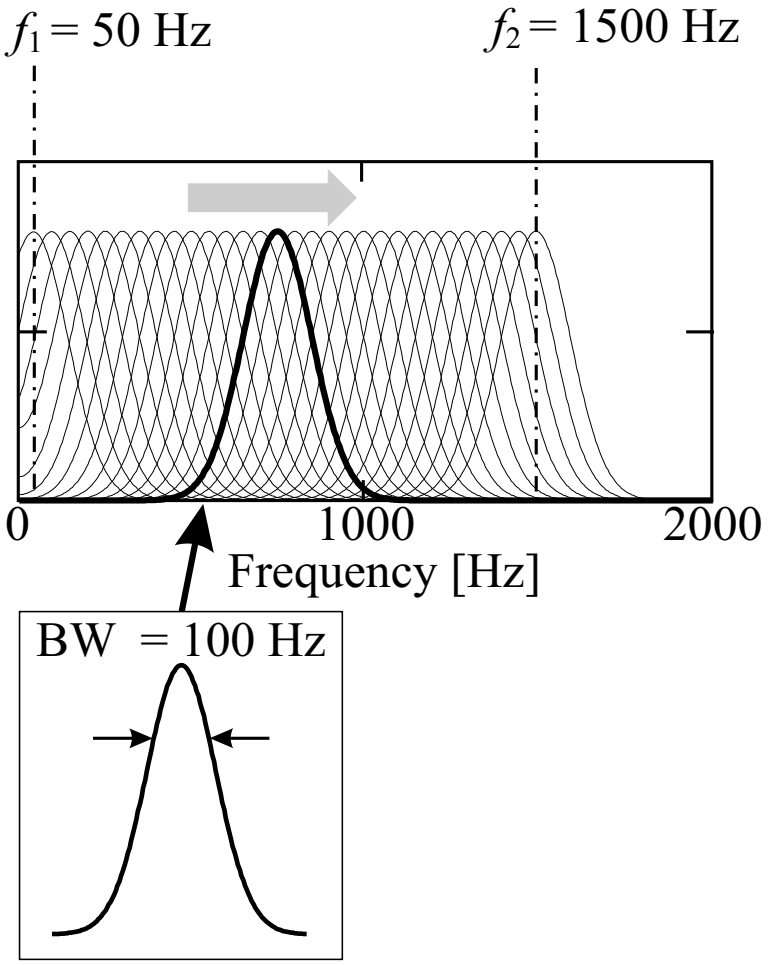


Figure 9.11: Frequency-time decomposition. The transfer function is convolved with a Gaussian pulse shifted in frequency, and thus the signal excited in a limited frequency band is determined.

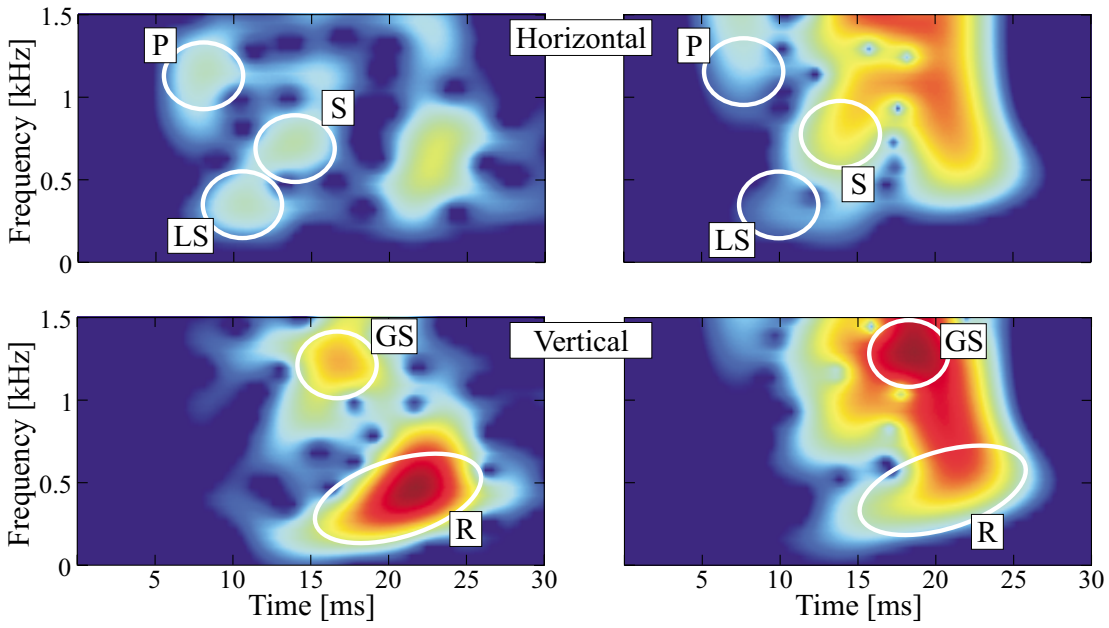


Figure 9.12: Frequency-time decomposition; particle motion 1.5 m from source, at a depth of 5 cm. Top: horizontal component; bottom: vertical component. Left: measurement (acceleration); right: numerical simulation (displacement).

and the leaky surface wave.

The difference in the frequency content is illustrated by the transfer functions of the experimental and the numerical signals. Fig. 9.13 shows the magnitude of the transfer functions for both the horizontal component and the vertical component as measured and as computed for the observation point 5 cm beneath the surface. The frequency content of the measured data seems to be strongest below 1 kHz. For the numerical data, however, the frequency content grows towards higher frequencies.

In the measurements the particle acceleration in the ground is determined, whereas in the numerical model the particle displacement is computed. In general, the particle displacement and the particle acceleration will exhibit a similar behavior when comparing the arrival times and the curvature of the wave fronts. Substantial differences arise, however, if the frequency content is analyzed. The acceleration is the second time derivative of the displacement. Thus, when computing the particle ac-

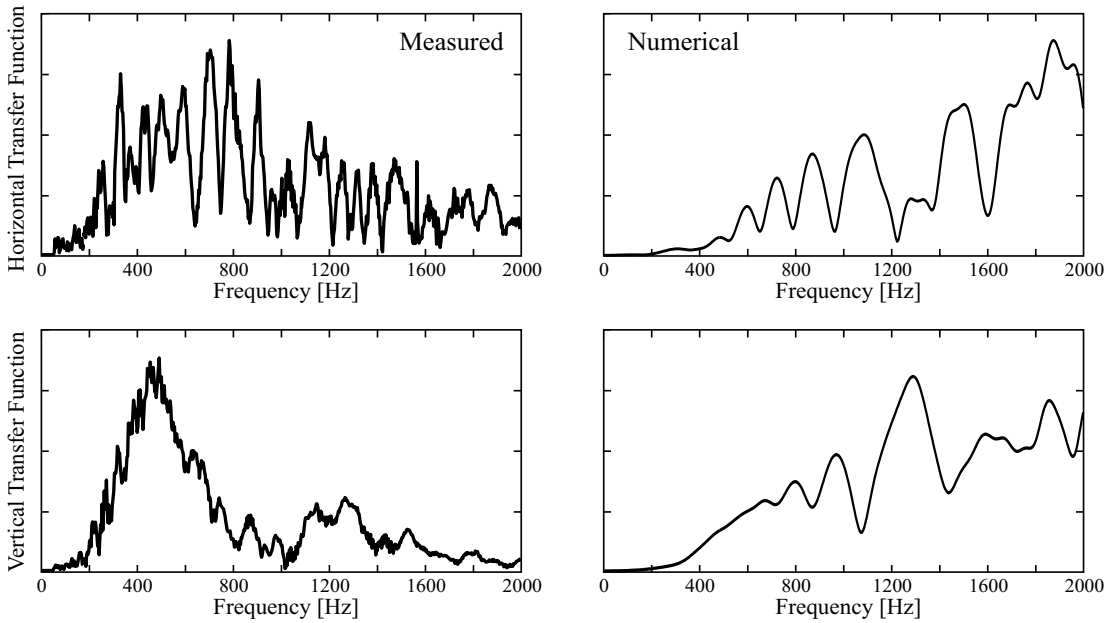


Figure 9.13: Transfer functions for the horizontal and the vertical field component; distance to source is 1.5 m, at a depth of 5 cm. Left: measurement (acceleration), right: numerical simulation (displacement).

celeration from the displacement, the higher frequency components will be enhanced in the particle acceleration. Throughout this section, the measured acceleration is compared to the numerical displacement, because primarily the arrival times of the wave fronts are to be analyzed to derive information about the wave speeds in the ground. However, Figs. (9.12) and (9.13) can be interpreted purely in a qualitative sense, because the frequency content of the experimentally measured acceleration and the numerically determined displacement are compared. If the particle accelerations were compared, however, the high frequency components would be strongly enhanced in the numerical data and a meaningful interpretation of the results would hardly be feasible.

9.3 Measurements of the Mine-Wave Interaction

In Chapter 7, numerical results describing the interaction of elastic waves with buried land mines are compared to experimental measurements. These measurements have been performed by Gregg D. Larson, James S. Martin, and Waymond R. Scott at the Georgia Institute of Technology [14, 12]. Fig. 9.14 illustrates the measurement procedure. A transducer is placed on the surface of the ground. The transducer is driven with a chirp having a uniform frequency spectrum from 30 Hz to 2 kHz. A radar is mounted above the surface and measures the displacement of the surface. The radar is highly optimized and has been developed specifically for this purpose. A stepwise procedure is employed to measure the interaction of the elastic waves with the buried mine. The transducer excites the elastic waves, and the radar records the displacement at one point on the surface. The radar is then moved, the waves are launched again, and the signal at the next point is recorded. In this way, the radar scans the surface, successively measuring the displacement along a line on the surface. The recorded signal is processed as described in Fig. 9.5 for the accelerometer measurements, and the waterfall graphs presented in Chapter 7 are obtained.

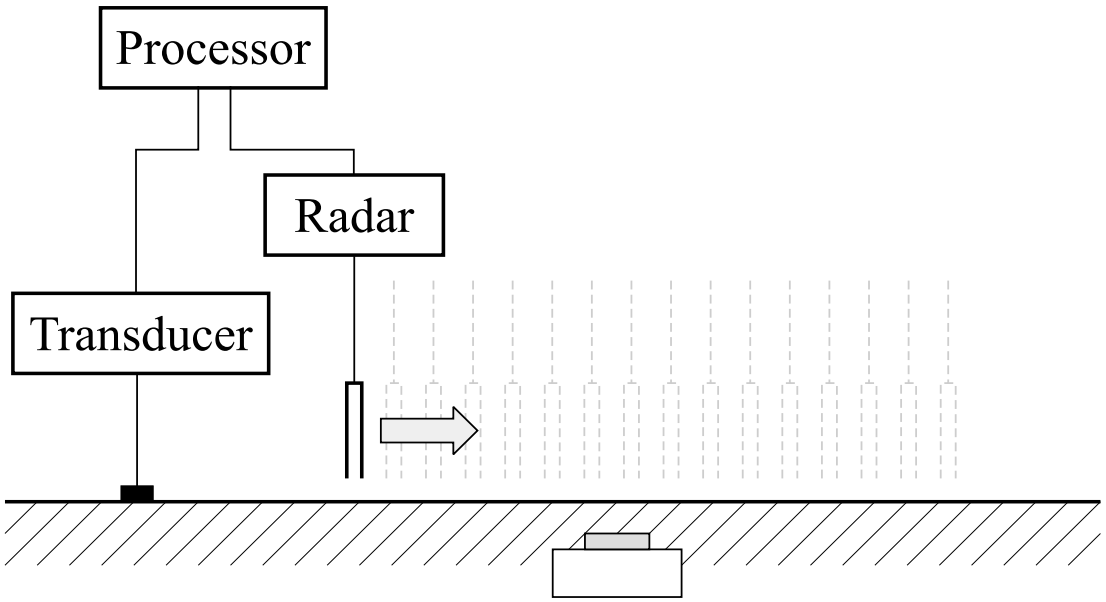


Figure 9.14: Schematic drawing of the radar measurement.

CHAPTER 10

Summary and Conclusions

This dissertation describes various aspects of elastic wave motion in solid media and explains the interaction of elastic waves with buried land mines. In this work, results of numerical, theoretical and experimental studies are presented, providing a comprehensive description of the fundamental functionality of a land mine detection system using elastic waves.

This text can be divided into two parts. Chapters 2 – 6 describe a numerical model, its theoretical background and its implementation. Chapters 7 – 9 depict the application of the numerical model.

To compute the elastic wave motion in the ground, a numerical model has been developed and implemented in two and three dimensions. The model uses the finite-difference time-domain modeling scheme and is implemented to run on a parallel computer with multiple processors. In the model, the ground is assumed to be an infinite, isotropic, lossless, heterogeneous half space. The air-ground interface is approximated by a stress-free boundary condition, and a perfectly matched layer absorbing boundary is implemented at the outer faces of the numerical grid.

In general, the numerical finite-difference algorithm will be stable, if the Courant condition (which links the time increment to the spacing between the nodes of the discrete grid) is satisfied. During the development of the numerical model, however, it was observed that a more restrictive stability condition arises at the interface between two different media. An analysis of the stability of the finite-difference algorithm at a

material interface is presented, and it is conclusively shown that the Courant condition will always be a sufficient condition for stability, if the material properties for the field components on the interface in the discrete numerical grid are averaged properly. The averaging procedure is derived, and its implementation is presented.

While investigating the elastic wave motion that is excited by a source on the surface of an isotropic, homogeneous medium, two distinct surface waves have been observed: the well-known Rayleigh surface wave and a leaky surface wave. Both surface waves are solutions to the governing partial-differential equations and arise from the roots of the Rayleigh equation. The Rayleigh surface wave and the leaky surface wave have distinct characteristics. The Rayleigh wave always exists, independent of the properties of the medium. It is slower than both the pressure and the shear wave in the medium. It propagates along the surface and decays exponentially into the ground. On the surface, it induces a retrograde particle motion. The leaky surface wave, on the other hand, exists only for materials with a high Poisson ratio ($\nu > 0.263$). Along the surface, it propagates with a wave speed smaller than the pressure wave, but larger than the shear wave and, therefore, couples into a plane shear wave in the ground. Due to the coupling, the leaky surface wave decays in its propagation direction. On the surface, it causes a prograde particle motion.

The numerical model is used to investigate the interaction of elastic waves with buried land mines. As a first approach, the mines are approximated in the numerical model by fairly simple models, neglecting the structural details and the complexity of actual mines. Various results are presented illustrating the mine-wave interaction for both anti-personnel and anti-tank mines. In general, when elastic waves impinge on a buried mine, resonant oscillations are excited at the mine location. Oscillations similar in size and shape are observed both in numerical results and in experimental measurements. The resonant oscillations enhance the response of the buried mine and, in that way, clearly distinguish it from surrounding clutter objects.

The resonant oscillations, in general, will depend on the type of mine, on the properties of the soil, and on the burial depth of the mine. To explore the resonant

oscillations in more detail, a refined mine model has been developed emulating a TS-50 anti-personnel mine. The new mine model includes the details of the actual mine and, thus, models its resonant behavior more accurately. With the detailed mine model a parametric study is conducted in which the response of the mine is investigated as a function of the burial depth and the stiffness of the ground. If a shear wave impinges on a buried mine, two resonant modes (in the range up to 2.1 kHz) are excited at the mine location. The resonant modes are primarily due to the flexural motion of the soil layer above the buried mine. In essence, it is concluded that the resonant frequencies of the modes will increase with the stiffness of the soil and decrease with the burial depth of the mine, whereas the quality factor of the resonance will decrease with both stiffness and burial depth.

To study the wave propagation in the ground, the wave fields have been measured experimentally. For this, a vertical array of accelerometers has been buried. With the accelerometers, the wave fields underground are determined as a function of depth, and conclusions are drawn about the depth variation of the soil properties.

The numerical model as presented in this work represents a compromise in accuracy and complexity. Considering the uncertainties in the material properties of the ground and the assumptions and simplifications that have been made for the numerical model, the agreement of the numerical and the experimental results that are presented throughout this text is remarkably good. Various paths can be followed to improve the agreement in future models. In the numerical model, the depth variation of the material properties is approximated by a fairly crude profile, and thorough measurements are necessary to obtain a better picture of the variation of the material properties in the ground. Furthermore, the numerical model as presented in this work does not incorporate any loss or non-linear effects in the soil. Especially the non-linear behavior of the soil has a significant impact on the elastic wave propagation in the ground. However, the implementation of the non-linearities will cause a substantial increase in algorithmic complexity. Moreover, the non-linear behavior of the soil has not yet been fully characterized and will differ between different soil

types, making the implementation of the non-linearities almost infeasible.

Parts of this work have been presented at conferences and meetings [8], [9], [78], [85] – [92] and have been published in or have been submitted for publication to various scientific journals [93] – [96].

APPENDIX A

3-D FDTD Equations

To implement the perfectly matched layer according to Chew and Liu [55], the field variables have to be split. For each particle velocity component, three split field quantities arise, which correspond to the spatial derivatives on the right hand side of Eq. (4.11) – Eq. (4.13). The stress components are split in a similar way. For each of the longitudinal stress components, τ_{xx}, τ_{yy} and τ_{zz}, three spatial derivatives appear on the right hand side of Eq. (4.14) – Eq. (4.16). Some of the derivative terms are equal for all longitudinal stresses and, thus, their split field components are also equal. For example, $\tau_{xx,y} = \lambda \frac{\partial u_y}{\partial y} = \tau_{zz,y}$. Consequently, only six split field components are introduced for the three longitudinal stresses. The equations for the shear stresses (Eq. (4.17) – (4.19)) have only two derivatives each on the right side. Thus, two split components for each shear stress component arise. Altogether, a set of 21 equations with 9 combining relations have to be solved for the 21 split field components and the 9 original field components.

In the following, the position of the grid cell in x, y, z is labeled with i, j, k and the time step is labeled with l. For example, $V_x^{l+0.5}|_{i,j,k}$ is the x-component of the numerical particle velocity at the (i, j, k)-th cell at the time $(l + 0.5) \cdot \Delta t$. The split finite-difference equations for the (i, j, k)-th cell are then:

$$V_{x,x}^{l+0.5}|_{i,j,k} = V_{x,x}^{l-0.5}|_{i,j,k} \frac{2 - \Delta t \Omega_x|_{i,j,k}}{2 + \Delta t \Omega_x|_{i,j,k}} \tag{A.1}$$

$$+ \quad \frac{2\Delta t}{\rho|_{i,j,k}\Delta x(2+\Delta t\Omega_x|_{i,j,k})}(T^l_{xx}|_{i,j,k} - T^l_{xx}|_{i-1,j,k})$$

$$V^{l+0.5}_{x,y}|_{i,j,k} = V^{l-0.5}_{x,y}|_{i,j,k}\frac{2-\Delta t\Omega_y|_{i,j,k}}{2+\Delta t\Omega_y|_{i,j,k}} \tag{A.2}$$

$$+ \quad \frac{2\Delta t}{\rho|_{i,j,k}\Delta y(2+\Delta t\Omega_y|_{i,j,k})}(T^l_{xy}|_{i,j,k} - T^l_{xy}|_{i,j-1,k})$$

$$V^{l+0.5}_{x,z}|_{i,j,k} = V^{l-0.5}_{x,z}|_{i,j,k}\frac{2-\Delta t\Omega_z|_{i,j,k}}{2+\Delta t\Omega_z|_{i,j,k}} \tag{A.3}$$

$$+ \quad \frac{2\Delta t}{\rho|_{i,j,k}\Delta z(2+\Delta t\Omega_z|_{i,j,k})}(T^l_{xz}|_{i,j,k} - T^l_{xz}|_{i,j,k-1})$$

$$V^{l+0.5}_{y,x}|_{i,j,k} = V^{l-0.5}_{y,x}|_{i,j,k}\frac{2-\Delta t\Omega_x|_{i,j,k}}{2+\Delta t\Omega_x|_{i,j,k}} \tag{A.4}$$

$$+ \quad \frac{2\Delta t}{\rho|_{i,j,k}\Delta x(2+\Delta t\Omega_x|_{i,j,k})}(T^l_{xy}|_{i+1,j,k} - T^l_{xy}|_{i,j,k})$$

$$V^{l+0.5}_{y,y}|_{i,j,k} = V^{l-0.5}_{y,y}|_{i,j,k}\frac{2-\Delta t\Omega_y|_{i,j,k}}{2+\Delta t\Omega_y|_{i,j,k}} \tag{A.5}$$

$$+ \quad \frac{2\Delta t}{\rho|_{i,j,k}\Delta y(2+\Delta t\Omega_y|_{i,j,k})}(T^l_{yy}|_{i,j+1,k} - T^l_{yy}|_{i,j,k})$$

$$V^{l+0.5}_{y,z}|_{i,j,k} = V^{l-0.5}_{y,z}|_{i,j,k}\frac{2-\Delta t\Omega_z|_{i,j,k}}{2+\Delta t\Omega_z|_{i,j,k}} \tag{A.6}$$

$$+ \quad \frac{2\Delta t}{\rho|_{i,j,k}\Delta z(2+\Delta t\Omega_z|_{i,j,k})}(T^l_{yz}|_{i,j,k} - T^l_{yz}|_{i,j,k-1})$$

$$V^{l+0.5}_{z,x}|_{i,j,k} = V^{l-0.5}_{z,x}|_{i,j,k}\frac{2-\Delta t\Omega_x|_{i,j,k}}{2+\Delta t\Omega_x|_{i,j,k}} \tag{A.7}$$

$$+ \quad \frac{2\Delta t}{\rho|_{i,j,k}\Delta x(2+\Delta t\Omega_x|_{i,j,k})}(T^l_{xz}|_{i+1,j,k} - T^l_{xz}|_{i,j,k})$$

$$V_{z,y}^{l+0.5}|_{i,j,k} = V_{z,y}^{l-0.5}|_{i,j,k}\frac{2-\Delta t\Omega_y|_{i,j,k}}{2+\Delta t\Omega_y|_{i,j,k}} \tag{A.8}$$

$$+ \frac{2\Delta t}{\rho|_{i,j,k}\Delta y(2+\Delta t\Omega_y|_{i,j,k})}(T_{yz}^{l}|_{i,j,k} - T_{yz}^{l}|_{i,j-1,k})$$

$$V_{z,z}^{l+0.5}|_{i,j,k} = V_{z,z}^{l-0.5}|_{i,j,k}\frac{2-\Delta t\Omega_z|_{i,j,k}}{2+\Delta t\Omega_z|_{i,j,k}} \tag{A.9}$$

$$+ \frac{2\Delta t}{\rho|_{i,j,k}\Delta z(2+\Delta t\Omega_z|_{i,j,k})}(T_{zz}^{l}|_{i,j,k+1} - T_{zz}^{l}|_{i,j,k})$$

$$T_{xx,x}^{l+1}|_{i,j,k} = T_{xx,x}^{l}|_{i,j,k}\frac{2-\Delta t\Omega_x|_{i,j,k}}{2+\Delta t\Omega_x|_{i,j,k}} \tag{A.10}$$

$$+ \frac{2(\lambda|_{i,j,k}+2\mu|_{i,j,k})\Delta t}{\Delta x(2+\Delta t\Omega_x|_{i,j,k})}(V_x^{l+0.5}|_{i+1,j,k} - V_x^{l+0.5}|_{i,j,k})$$

$$T_{xx,y}^{l+1}|_{i,j,k} = T_{zz,y}^{l+1}|_{i,j,k} = T_{xx,y}^{l}|_{i,j,k}\frac{2-\Delta t\Omega_y|_{i,j,k}}{2+\Delta t\Omega_y|_{i,j,k}} \tag{A.11}$$

$$+ \frac{2\lambda|_{i,j,k}\Delta t}{\Delta y(2+\Delta t\Omega_y|_{i,j,k})}(V_y^{l+0.5}|_{i,j,k} - V_y^{l+0.5}|_{i,j-1,k})$$

$$T_{xx,z}^{l+1}|_{i,j,k} = T_{yy,z}^{l+1}|_{i,j,k} = T_{xx,z}^{l}|_{i,j,k}\frac{2-\Delta t\Omega_z|_{i,j,k}}{2+\Delta t\Omega_z|_{i,j,k}} \tag{A.12}$$

$$+ \frac{2\lambda|_{i,j,k}\Delta t}{\Delta z(2+\Delta t\Omega_z|_{i,j,k})}(V_z^{l+0.5}|_{i,j,k} - V_z^{l+0.5}|_{i,j,k-1})$$

$$T_{yy,x}^{l+1}|_{i,j,k} = T_{zz,x}^{l+1}|_{i,j,k} = T_{yy,x}^{l}|_{i,j,k}\frac{2-\Delta t\Omega_x|_{i,j,k}}{2+\Delta t\Omega_x|_{i,j,k}} \tag{A.13}$$

$$+ \frac{2\lambda|_{i,j,k}\Delta t}{\Delta x(2+\Delta t\Omega_x|_{i,j,k})}(V_x^{l+0.5}|_{i+1,j,k} - V_x^{l+0.5}|_{i,j,k})$$

$$T_{yy,y}^{l+1}|_{i,j,k} = T_{yy,y}^{l}|_{i,j,k}\frac{2-\Delta t\Omega_y|_{i,j,k}}{2+\Delta t\Omega_y|_{i,j,k}} \tag{A.14}$$

$$+ \quad \frac{2(\lambda|_{i,j,k} + 2\mu|_{i,j,k})\Delta t}{\Delta y(2 + \Delta t\Omega_y|_{i,j,k})}(V_y^{l+0.5}|_{i,j,k} - V_y^{l+0.5}|_{i,j-1,k})$$

$$T_{zz,z}^{l+1}|_{i,j,k} = T_{zz,z}^{l}|_{i,j,k}\frac{2 - \Delta t\Omega_z|_{i,j,k}}{2 + \Delta t\Omega_z|_{i,j,k}} \tag{A.15}$$

$$+ \quad \frac{2(\lambda|_{i,j,k} + 2\mu|_{i,j,k})\Delta t}{\Delta z(2 + \Delta t\Omega_z|_{i,j,k})}(V_z^{l+0.5}|_{i,j,k} - V_z^{l+0.5}|_{i,j,k-1})$$

$$T_{yz,y}^{l+1}|_{i,j,k} = T_{yz,y}^{l}|_{i,j,k}\frac{2 - \Delta t\Omega_y|_{i,j,k}}{2 + \Delta t\Omega_y|_{i,j,k}} \tag{A.16}$$

$$+ \quad \frac{2\mu|_{i,j,k}\Delta t}{\Delta y(2 + \Delta t\Omega_y|_{i,j,k})}(V_z^{l+0.5}|_{i,j+1,k} - V_z^{l+0.5}|_{i,j,k})$$

$$T_{yz,z}^{l+1}|_{i,j,k} = T_{yz,z}^{l}|_{i,j,k}\frac{2 - \Delta t\Omega_z|_{i,j,k}}{2 + \Delta t\Omega_z|_{i,j,k}} \tag{A.17}$$

$$+ \quad \frac{2\mu|_{i,j,k}\Delta t}{\Delta z(2 + \Delta t\Omega_z|_{i,j,k})}(V_y^{l+0.5}|_{i,j,k+1} - V_y^{l+0.5}|_{i,j,k})$$

$$T_{xz,x}^{l+1}|_{i,j,k} = T_{xz,x}^{l}|_{i,j,k}\frac{2 - \Delta t\Omega_x|_{i,j,k}}{2 + \Delta t\Omega_x|_{i,j,k}} \tag{A.18}$$

$$+ \quad \frac{2\mu|_{i,j,k}\Delta t}{\Delta x(2 + \Delta t\Omega_x|_{i,j,k})}(V_z^{l+0.5}|_{i,j,k} - V_z^{l+0.5}|_{i-1,j,k})$$

$$T_{xz,z}^{l+1}|_{i,j,k} = T_{xz,z}^{l}|_{i,j,k}\frac{2 - \Delta t\Omega_z|_{i,j,k}}{2 + \Delta t\Omega_z|_{i,j,k}} \tag{A.19}$$

$$+ \quad \frac{2\mu|_{i,j,k}\Delta t}{\Delta z(2 + \Delta t\Omega_z|_{i,j,k})}(V_x^{l+0.5}|_{i,j,k+1} - V_x^{l+0.5}|_{i,j,k})$$

$$T_{xy,x}^{l+1}|_{i,j,k} = T_{xy,x}^{l}|_{i,j,k}\frac{2 - \Delta t\Omega_x|_{i,j,k}}{2 + \Delta t\Omega_x|_{i,j,k}} \tag{A.20}$$

$$+ \quad \frac{2\mu|_{i,j,k}\Delta t}{\Delta x(2 + \Delta t\Omega_x|_{i,j,k})}(V_y^{l+0.5}|_{i,j,k} - V_y^{l+0.5}|_{i-1,j,k})$$

$$T_{xy,y}^{l+1}|_{i,j,k} = T_{xy,y}^{l}|_{i,j,k}\frac{2-\Delta t\Omega_y|_{i,j,k}}{2+\Delta t\Omega_y|_{i,j,k}} \tag{A.21}$$

$$+ \frac{2\mu|_{i,j,k}\Delta t}{\Delta y(2+\Delta t\Omega_y|_{i,j,k})}(V_x^{l+0.5}|_{i,j+1,k} - V_x^{l+0.5}|_{i,j,k}).$$

The relations combining the split fields with the physical fields are in discretized form

$$V_x^{l+0.5}|_{i,j,k} = V_{x,x}^{l+0.5}|_{i,j,k} + V_{x,y}^{l+0.5}|_{i,j,k} + V_{x,z}^{l+0.5}|_{i,j,k} \tag{A.22}$$

$$V_y^{l+0.5}|_{i,j,k} = V_{y,x}^{l+0.5}|_{i,j,k} + V_{y,y}^{l+0.5}|_{i,j,k} + V_{y,z}^{l+0.5}|_{i,j,k} \tag{A.23}$$

$$V_z^{l+0.5}|_{i,j,k} = V_{z,x}^{l+0.5}|_{i,j,k} + V_{z,y}^{l+0.5}|_{i,j,k} + V_{z,z}^{l+0.5}|_{i,j,k} \tag{A.24}$$

$$T_{xx}^{l+1}|_{i,j,k} = T_{xx,x}^{l+1}|_{i,j,k} + T_{xx,y}^{l+1}|_{i,j,k} + T_{xx,z}^{l+1}|_{i,j,k} \tag{A.25}$$

$$T_{yy}^{l+1}|_{i,j,k} = T_{yy,x}^{l+1}|_{i,j,k} + T_{yy,y}^{l+1}|_{i,j,k} + T_{xx,z}^{l+1}|_{i,j,k} \tag{A.26}$$

$$T_{zz}^{l+1}|_{i,j,k} = T_{yy,x}^{l+1}|_{i,j,k} + T_{xx,y}^{l+1}|_{i,j,k} + T_{zz,z}^{l+1}|_{i,j,k} \tag{A.27}$$

$$T_{yz}^{l+1}|_{i,j,k} = T_{yz,y}^{l+1}|_{i,j,k} + T_{yz,z}^{l+1}|_{i,j,k} \tag{A.28}$$

$$T_{xz}^{l+1}|_{i,j,k} = T_{xz,x}^{l+1}|_{i,j,k} + T_{xz,z}^{l+1}|_{i,j,k} \tag{A.29}$$

$$T_{xy}^{l+1}|_{i,j,k} = T_{xy,x}^{l+1}|_{i,j,k} + T_{xy,y}^{l+1}|_{i,j,k}. \tag{A.30}$$

APPENDIX B

Total-Field/Scattered-Field Formulation: Correction Terms

In this section, the field components at the six artificial interfaces arising for the total-field/scattered-field formulation are identified. In six tables, all components are listed, and their correction terms are given. The numbering of the interfaces is indicated in Fig. B.1. The components at the interface are labeled by i, j, k, and the field components within the correction terms are labeled with respect to their corresponding interface component. Note that components at the *edges* between inner and outer region must be adjusted by the correction terms associated with both adjacent interfaces and the field components on the *corners* (i. e., the longitudinal stresses) must be corrected by adding three terms.

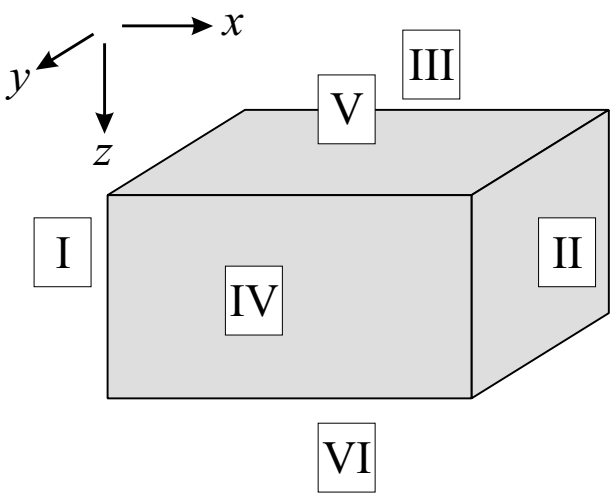


Figure B.1: Numbering of the interfaces.

Table B.1: Field components and correction terms at Interface I and II.

I		II	
Scattered Fields		Scattered Fields	
$V_x^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta x}\cdot T_{xx,\mathrm{inc}}^{l}\|_{i,j,k}$	$V_x^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta x}\cdot T_{xx,\mathrm{inc}}^{l}\|_{i-1,j,k}$
$T_{xy}^{l+1}\|_{i,j,k}$	$-\frac{\mu\Delta t}{\Delta x}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$	$T_{xy}^{l+1}\|_{i,j,k}$	$\frac{\mu\Delta t}{\Delta x}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i-1,j,k}$
$T_{xz}^{l+1}\|_{i,j,k}$	$-\frac{\mu\Delta t}{\Delta x}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$	$T_{xz}^{l+1}\|_{i,j,k}$	$\frac{\mu\Delta t}{\Delta x}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i-1,j,k}$
Total Fields		Total Fields	
$V_y^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta x}\cdot T_{xy,\mathrm{inc}}^{l}\|_{i,j,k}$	$V_y^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta x}\cdot T_{xy,\mathrm{inc}}^{l}\|_{i+1,j,k}$
$V_z^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta x}\cdot T_{xz,\mathrm{inc}}^{l}\|_{i,j,k}$	$V_z^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta x}\cdot T_{xz,\mathrm{inc}}^{l}\|_{i+1,j,k}$
$T_{xx}^{l+1}\|_{i,j,k}$	$-\frac{(\lambda+2\mu)\Delta t}{\Delta x}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$	$T_{xx}^{l+1}\|_{i,j,k}$	$\frac{(\lambda+2\mu)\Delta t}{\Delta x}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i+1,j,k}$
$T_{yy}^{l+1}\|_{i,j,k}$	$-\frac{\lambda\Delta t}{\Delta x}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$	$T_{yy}^{l+1}\|_{i,j,k}$	$\frac{\lambda\Delta t}{\Delta x}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i+1,j,k}$
$T_{zz}^{l+1}\|_{i,j,k}$	$-\frac{\lambda\Delta t}{\Delta x}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$	$T_{zz}^{l+1}\|_{i,j,k}$	$\frac{\lambda\Delta t}{\Delta x}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i+1,j,k}$

Table B.2: Field components and correction terms at Interface III and IV.

III		IV	
Scattered Fields		Scattered Fields	
$V_y^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta y}\cdot T_{yy,\mathrm{inc}}^{l}\|_{i,j+1,k}$	$V_y^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta y}\cdot T_{yy,\mathrm{inc}}^{l}\|_{i,j,k}$
$T_{yz}^{l+1}\|_{i,j,k}$	$-\frac{\mu\Delta t}{\Delta y}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j+1,k}$	$T_{yz}^{l+1}\|_{i,j,k}$	$\frac{\mu\Delta t}{\Delta y}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$
$T_{xy}^{l+1}\|_{i,j,k}$	$-\frac{\mu\Delta t}{\Delta y}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i,j+1,k}$	$T_{xy}^{l+1}\|_{i,j,k}$	$\frac{\mu\Delta t}{\Delta y}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$
Total Fields		Total Fields	
$V_x^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta y}\cdot T_{xy,\mathrm{inc}}^{l}\|_{i,j-1,k}$	$V_x^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta y}\cdot T_{xy,\mathrm{inc}}^{l}\|_{i,j,k}$
$V_z^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta y}\cdot T_{yz,\mathrm{inc}}^{l}\|_{i,j-1,k}$	$V_z^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta y}\cdot T_{yz,\mathrm{inc}}^{l}\|_{i,j,k}$
$T_{xx}^{l+1}\|_{i,j,k}$	$-\frac{(\lambda+2\mu)\Delta t}{\Delta y}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j-1,k}$	$T_{xx}^{l+1}\|_{i,j,k}$	$\frac{(\lambda+2\mu)\Delta t}{\Delta y}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$
$T_{yy}^{l+1}\|_{i,j,k}$	$-\frac{\lambda\Delta t}{\Delta y}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j-1,k}$	$T_{yy}^{l+1}\|_{i,j,k}$	$\frac{\lambda\Delta t}{\Delta y}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$
$T_{zz}^{l+1}\|_{i,j,k}$	$-\frac{\lambda\Delta t}{\Delta y}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j-1,k}$	$T_{zz}^{l+1}\|_{i,j,k}$	$\frac{\lambda\Delta t}{\Delta y}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$

Table B.3: Field components and correction terms at Interface V and VI .

V		VI	
Scattered Fields		Scattered Fields	
$V_z^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta z}\cdot T_{zz,\mathrm{inc}}^{l}\|_{i,j,k+1}$	$V_z^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta z}\cdot T_{zz,\mathrm{inc}}^{l}\|_{i,j,k}$
$T_{xz}^{l+1}\|_{i,j,k}$	$-\frac{\mu\Delta t}{\Delta z}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i,j,k+1}$	$T_{xz}^{l+1}\|_{i,j,k}$	$\frac{\mu\Delta t}{\Delta z}\cdot V_{x,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$
$T_{yz}^{l+1}\|_{i,j,k}$	$-\frac{\mu\Delta t}{\Delta z}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j,k+1}$	$T_{yz}^{l+1}\|_{i,j,k}$	$\frac{\mu\Delta t}{\Delta z}\cdot V_{y,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$
Total Fields		Total Fields	
$V_x^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta z}\cdot T_{xz,\mathrm{inc}}^{l}\|_{i,j,k-1}$	$V_x^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta z}\cdot T_{xz,\mathrm{inc}}^{l}\|_{i,j,k}$
$V_y^{l+0.5}\|_{i,j,k}$	$-\frac{\Delta t}{\rho\Delta z}\cdot T_{yz,\mathrm{inc}}^{l}\|_{i,j,k-1}$	$V_y^{l+0.5}\|_{i,j,k}$	$\frac{\Delta t}{\rho\Delta z}\cdot T_{yz,\mathrm{inc}}^{l}\|_{i,j,k}$
$T_{xx}^{l+1}\|_{i,j,k}$	$-\frac{\lambda\Delta t}{\Delta z}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j,k-1}$	$T_{xx}^{l+1}\|_{i,j,k}$	$\frac{\lambda\Delta t}{\Delta z}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$
$T_{yy}^{l+1}\|_{i,j,k}$	$-\frac{\lambda\Delta t}{\Delta z}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j,k-1}$	$T_{yy}^{l+1}\|_{i,j,k}$	$\frac{\lambda\Delta t}{\Delta z}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$
$T_{zz}^{l+1}\|_{i,j,k}$	$-\frac{(\lambda+2\mu)\Delta t}{\Delta z}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j,k-1}$	$T_{zz}^{l+1}\|_{i,j,k}$	$\frac{(\lambda+2\mu)\Delta t}{\Delta z}\cdot V_{z,\mathrm{inc}}^{l+0.5}\|_{i,j,k}$

APPENDIX C

Software Documentation

The numerical model has been implemented using FORTRAN90. The three-dimensional model is written in a fully parallel fashion using the *Message-Passing Interface* (MPI) to run on computers with multiple processors. It can be executed on any machine on which MPI is installed by using the default MPI execution command (mpirun) or a script file called execute_gui. The script file requires certain input parameters:

```
execute_gui <np> <machine file> <executable> <directory>
      <input file>
```

<np> specifies the number of processors that are to be used, <machine file> is a file containing a listing of all machines that are to be used (i. e., the names of the nodes of the parallel computer), <executable> is the program name to be executed, <directory> is the name (including the path) of the directory in which the output data is to be stored, and <input file> is the name of the file in which the program parameters are specified. For example,

```
execute_gui 28 machines 3d_parallel /nodes/data parameters.dat
```

executes the file 3d_parallel using 28 processors (whose names are specified in machines), reads in the program parameters from parameters.dat and stores the output in the directory /nodes/data.

C.1 Input File

The input parameters are specified in the input file. The name of the input file is arbitrary and must be specified during the program execution. The input file consists of five sections. In the first section, the general simulation parameters are specified. This section must appear first. In the four other sections, sources and objects in the solution space are specified, the background material is defined and the material properties of the objects are listed. Each section must be started with a characteristic letter:

- `G` for '*GENERAL*',
- `S` for '*SOURCE*',
- `O` for '*OBJECT*',
- `B` for '*BACKGROUND*', and
- `M` for '*MATERIAL*'.

Each section is ended with the statement `END`. Comments can be inserted anywhere and are indicated by either `C` followed by a space or `#` at the beginning of the line.

C.1.1 General Simulation Parameters

This section must appear first. The section beginning is indicated by `G`, and the end by `END`. The following parameters must be specified or errors will occur:

- `DIM =` $x\ y\ z$
  problem dimensions in x, y and x [cm]
- `DISCRETIZATION =` $dx\ dy\ dz$
  size of space steps Δx, Δy, Δz [cm]
- `TIME =` t
  time [ms]
- `FREESURFACE = <0/1>`
  0: no free surface, 1: free surface at $z = 0$

- SYMMETRIE_Y = <0/1>
  0: symmetry not exploited, 1: solution space symmetric to x-z-plane at $y = 0$.

Each statement begins with a characteristic phrase (e. g. TIME), followed by an equal sign. After the equal sign, the value(s) is (are) specified.

The time domain output data is written for entire planes. Up to 50 planes can be specified. Also, the frequency domain data for up to 1000 single points can be output.

- SAMPLERATEMOVIE = r
  rate for writing the output [ms]
- OUTPUT_PLANE = *Plane Position*
  plane (X,Y,Z, e. g., X for y-z-plane) and its position [cm].
  A maximum of 50 output planes can be specified. The data for the first specified plane is written in the files 'x1.dat', 'y1.dat', and 'z1.dat'; the data for the second plane in 'x2.dat', 'y2.dat', and 'z2.dat', and so on.
- OUTPUT_POINT = f_{min} f_{max} Δf x y z
  frequency range of the output data (min. freq., max. freq., step freq.) and coordinates of the point [cm];
  A maximum of 1000 output points can be specified. The frequency domain data for the first specified output point is written in 'freq_x1.dat', 'freq_y1.dat', and 'freq_z1.dat'; for the second point in 'freq_x2.dat', 'freq_y2.dat', and 'freq_z2.dat', and so on.

A sample input file might look like this:

```
DIM = 30.  30.  30.
DISCRETIZATION = 0.5 0.5 0.5
TIME = 40.
FREESURFACE = 1
SYMMETRIE_Y = 1
SAMPLERATEMOVIE = 0.125
OUTPUT_PLANE = Z 0.
OUTPUT_POINT = 0.  1500.  5.  15.  10.  10.
END
```

This input produces a solution space of size 30 cm by 30 cm by 30 cm (excluding PML) with a spacing between the nodes of 0.5 cm in x, y and z. The fields are

computed for 40 milliseconds, a free surface boundary is present at $z = 0$, the fields are symmetric to the x-z-plane at $y = 0$. The output is written for the x-y-plane at $z = 0$ (equivalent to the surface plane) with a rate of one frame per 0.125 milliseconds (equivalent to a sample rate of 8 kHz). The frequency domain data is written out for the point $(x, y, z) = (15\text{ cm}, 10\text{ cm}, 10\text{ cm})$ in the range from 0 Hz to 1500 Hz. Note that all length specifications are in centimeters, and all time specifications are in milliseconds.

C.1.2 Sources

Up to 100,000 source points can be specified. Additionally, 10 plane waves (shear or pressure) can be injected. Note that all length specifications are in centimeters, and all time specifications are in milliseconds.

- `CENTERFREQUENCY =` f_{center}
  center frequency of excitation [Hz]; must be specified!
- `POINT =` $x\ y\ z\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the point source [cm], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation
- `CIRCLE =` $x_{center}\ y_{center}\ z_{center}\ d\ \alpha_{az}\ \beta_{ele}\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the center of the source [cm], diameter [cm], azimuth and elevation angle [degrees], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation
- `SQUARE =` $x_{center}\ y_{center}\ z_{center}\ x_{dim}\ y_{dim}\ \alpha_{az}\ \beta_{ele}\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the center of the source [cm], dimensions in x and y [cm], azimuth and elevation angle [degrees], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation
- `RING =` $x_{center}\ y_{center}\ z_{center}\ d\ w\ \alpha_{az}\ \beta_{ele}\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the center of the source [cm], diameter and width of the ring [cm], azimuth and elevation angle [degrees], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation
- `CUBE =` $x_{center}\ y_{center}\ z_{center}\ x_{dim}\ y_{dim}\ z_{dim}\ \alpha_{az}\ \beta_{ele}\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the center of the source [cm], dimensions in x, y and z [cm], azimuth and elevation angle [degrees], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation

- `SPHERE =` $x_{center}\ y_{center}\ z_{center}\ d\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the center of the source [cm], diameter [cm], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation

- `CYLINDER =` $x_{center}\ y_{center}\ z_{center}\ d\ z_{dim}\ \alpha_{az}\ \beta_{ele}\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the center of the source [cm], diameter and height of cylinder [cm], azimuth and elevation angle [degrees], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation

- `SPHERICAL_SHELL =` $x_{center}\ y_{center}\ z_{center}\ d\ w\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the center of the source [cm], diameter and width of shell [cm], azimuth and elevation angle [degrees], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation

- `CYLINDRICAL_SHELL =` $x_{center}\ y_{center}\ z_{center}\ d\ w\ z_{dim}\ \alpha_{az}\ \beta_{ele}\ t_{delay}\ i_{exc}\ D\ A$
  coordinates of the center of the source [cm], diameter, width and height of cylinder [cm], azimuth and elevation angle [degrees], time delay [ms], type of excitation (0,1,2), direction (`X`,`Y`,`Z`) and amplitude of excitation

- `PLANEWAVE =` $Type\ Dir\ t_{delay}\ i_{exc}\ A$
  type of wave (`PRESSURE`,`SHEAR`), propagation direction and polarization (`X`,`Y`,`Z` for `PRESSURE`; `XZ`, ... for `SHEAR`, indicating a shear wave polarized in x and propagating in z), time delay [ms], type of excitation (0,1,2) and amplitude of excitation

- `PRESSURE_SOURCE =` $x_{center}\ y_{center}\ z_{center}\ d\ w\ t_{delay}\ i_{exc}\ A$
  Spherical source that emits mainly spherical pressure waves; coordinates of the center of the source [cm], diameter of source [cm], width of buffer layer surrounding the source to suppress shear waves [cm] (recommended: 2.), time delay [ms], type of excitation (0,1,2) and amplitude of excitation

The *azimuth angle* is the angle by which the source is to be rotated from the x-axis towards the y-axis in the x-y-plane.

The *elevation angle* is the angle by which the source is to be rotated from the x-axis towards the z-axis in the x-z-plane.

Three types of excitation are possible indicated by 0, 1, and 2:

0 = differentiated Gaussian pulse

1 = Gaussian pulse

2 = external data specified in the file *source.dat*

Format of *source.dat*: the first column specifies the time data, the second column the excitation values.

C.1.3 Objects

Objects with up to 1,000,000 elements can be specified.

- `ROUNDMINE =` $x_{ref}\ y_{ref}\ z_{ref}\ d_1\ z_{dim,1}\ d_2\ z_{dim,2}\ \alpha_{az}\ \beta_{ele}\ i_{Mat1}\ i_{Mat2}$
  reference location of mine (center in x and y, top of mine in z) [cm], diameter and height of large chamber, diameter and height of small chamber [cm], azimuth and elevation angle [degrees], material indices of large and small chamber, respectively

- `SQUAREMINE =` $x_{ref}\ y_{ref}\ z_{ref}\ x_{dim,1}\ y_{dim,1}\ z_{dim,1}\ x_{dim,2}\ y_{dim,2}\ z_{dim,2}\ \alpha_{az}\ \beta_{ele}\ i_{Mat1}\ i_{Mat2}$
  reference location of mine (center in x and y, top of mine in z) [cm], dimensions in x, y and z of large chamber, dimensions in x, y and z of small chamber [cm], azimuth and elevation angle [degrees], material indices of large and small chamber, respectively

- `DETAILED_TS50 =` $x_{ref}\ y_{ref}\ z_{ref}\ \alpha_{az}\ \beta_{ele}\ i_{Mat1}\ i_{Mat2}\ i_{Mat3}\ i_{Mat4}$
  reference location of mine (center in x and y, top of mine in z) [cm], azimuth and elevation angle [degrees], material indices of explosives, air, case and rubber plate, respectively

- `POINT =` $x_{ref}\ y_{ref}\ z_{ref}\ i_{Mat}$
  location in x, y and z [cm], material index

- `CUBE =` $x_{ref}\ y_{ref}\ z_{ref}\ x_{dim}\ y_{dim}\ z_{dim}\ \alpha_{az}\ \beta_{ele}\ i_{Mat}$
  reference location of cube (center in x and y, top of object in z) [cm], dimensions in x, y and z [cm], azimuth and elevation angle [degrees], material index

- `SPHERE =` $x_{ref}\ y_{ref}\ z_{ref}\ d\ i_{Mat}$
  reference location of sphere (center in x and y, top of object in z) [cm], diameter of sphere [cm], material index

- `SPHERICAL_SHELL =` $x_{ref}\ y_{ref}\ z_{ref}\ d\ w\ i_{Mat1}\ i_{Mat2}$
  reference location of shell (center in x and y, top of object in z) [cm], overall

diameter of shell [cm], width of shell wall [cm], material index of shell wall, material index of inner material

- `CYLINDER =` $x_{ref}\ y_{ref}\ z_{ref}\ d\ z_{dim}\ \alpha_{az}\ \beta_{ele}\ i_{Mat}$
  reference location of disk (center in x and y, top of object in z) [cm], diameter and height of disk [cm], azimuth and elevation angle [degrees], material index

- `CYLINDRICAL_SHELL =` $x_{ref}\ y_{ref}\ z_{ref}\ d\ z_{dim}\ w\ \alpha_{az}\ \beta_{ele}\ i_{Mat,1}\ i_{Mat,2}$
  reference location of shell (center in x and y, top of object in z) [cm], diameter, height and width of shell [cm], azimuth and elevation angle [degrees], material indices of the shell and the inside of the shell, respectively

- `DISK_LOOSESAND =` $x_{ref}\ y_{ref}\ z_{ref}\ d_{obj}\ z_{dim,obj}\ d_{sand}\ z_{dim,sand}\ \alpha_{az}\ \beta_{ele}\ i_{Mat1}\ i_{Mat2}$
  disk buried in medium with different properties than surrounding soil to simulate burying effects; reference location of disk (center in x and y, top of object in z) [cm], diameter and height of disk [cm], diameter and height of surrounding region, azimuth and elevation angle of disk [degrees], material indices of disk and surrounding region

- `STICK` $= x_{ref}\ y_{ref}\ z_{ref}\ d\ l\ \alpha_{az}\ \beta_{ele}\ i_{Mat}$
  emulates the random shape of a wooden stick, aligned with the x-axis; reference location of the stick is the center of the cross section at the front end; reference location [cm], diameter and length of stick [cm], azimuth and elevation angle [degrees], material index

- `RANDOMROCK =` $x_{ref}\ y_{ref}\ z_{ref}\ z_{dim}\ \alpha_{az}\ \beta_{ele}\ i_{Mat}$
  emulates the random shape of a rock; reference location is the center of the rock's upper face; reference location [cm], height [cm], azimuth and elevation angle [degrees], material index

- `TUNINGFORK =` $x_{ref}\ y_{ref}\ z_{ref}\ i_{Mat}$
  tuning fork; reference location is the center of the forking point; reference location [cm], material index

- `EXTERNAL =` *file* i_{Mat}
  reads in externally specified objects from *file* (*file* contains three columns specifying the coordinates of the elements in x, y, and z); material index of the elements

The coordinates of all source and object elements are written into the file *prop.dat*. Format of *prop.dat*: four columns containing the coordinates in x, y, and z, and the material index.

C.1.4 Background

Four different background conditions (i. e., properties of the soil) can be specified: a homogeneous background, a layered background, a linearly varying background, and a background in which the shear wave speed varies smoothly. The variation is assumed to be in the z-direction. The background must be specified only once!

- `HOMOGENEOUS =` $\rho\ c_P\ c_S$
  material density [kg/m^3], pressure wave speed and shear wave speed [m/s]

- `LAYERED =` n

  $$\begin{array}{l} d_1\ \rho_1\ c_{P_1}\ c_{S_1} \\ d_2\ \rho_2\ c_{P_2}\ c_{S_2} \\ \quad\vdots \\ d_n\ \rho_n\ c_{P_n}\ c_{S_n} \end{array}$$

  Specification of a layered background medium, in which the material properties are piecewise constant; n is the number of horizontal layers in the medium, and in subsequent lines the specification of these layers: thickness of layer [cm], material density [kg/m^3], pressure wave speed and shear wave speed [m/s]

- `LINEAR =` n

  $$\begin{array}{l} d_1\ \rho_{1,1}\ \rho_{1,2}\ c_{P_{1,1}}\ c_{P_{1,2}}\ c_{S_{1,1}}\ c_{S_{1,2}} \\ d_2\ \rho_{2,1}\ \rho_{2,2}\ c_{P_{2,1}}\ c_{P_{2,2}}\ c_{S_{2,1}}\ c_{S_{2,2}} \\ \quad\vdots \\ d_n\ \rho_{n,1}\ \rho_{n,2}\ c_{P_{n,1}}\ c_{P_{n,2}}\ c_{S_{n,1}}\ c_{S_{n,2}} \end{array}$$

  Specification of a layered background medium, in which the material properties vary linearly; n is the number of horizontal layers in the medium, and in subsequent lines the specification of these layers: thickness of layer [cm], material densities at upper and lower interface [kg/m^3], pressure wave speeds and shear wave speeds at upper and lower interface [m/s]

- `DEPTHVARYING =` $\rho\ c_P\ c_{S,min}\ c_{S,max}\ A\ B$
  Specification of a background medium in which the shear wave speed varies smoothly; material density [kg/m^3], pressure wave speed, minimum shear wave speed, maximum shear wave speed [m/s], and the constants A and B, indicating

the fraction of the maximum shear wave speed reached at a depth of one meter, and the growth rate, respectively.

The material properties in the soil are written out to the file *background.dat*. Format of *background.dat*: four columns containing the depth, the material density, the pressure wave speed and the shear wave speed in the model as a function of depth.

C.1.5 Materials

The materials that are used to characterize the objects must be specified. Only the materials that are actually used should be specified, because the size of the time step is determined from the maximum occurring wave speed in the list of the material specifications. The materials are identified by indices. Up to 100 different materials can be input.

Each material is defined by a material index i_{Mat}, the density ρ, the pressure wave speed c_P, and the shear wave speed c_S. All of these must be specified. A sample input file might look like this:

```
M
# air
1 1.3 330.  0.
# plastic
2 1200.  2700.  1100.  END
```

Here, the parameters for air and plastic are defined. Air has the index 1, a density of 1.3 kg/m^3, a pressure wave speed of 330 m/s and a shear wave speed of 0 m/s. Plastic has the index 2, a density of 1200 kg/m^3, a pressure wave speed of 2700 m/s and a shear wave speed of 1100 m/s.

C.1.6 Sample Input File

```
# Sample Input File
# input general problem parameters
G
DIM = 30.0 30.0 30.0
DISCRETIZATION = 0.5 0.5 0.5
TIME = 30.0
FREESURFACE = 1
SYMMETRIE_Y = 1
SAMPLERATEMOVIE = 0.1
OUTPUT_PLANE = Z 0.0
OUTPUT_PLANE = X 10.0
OUTPUT_PLANE = Y 0.0
OUTPUT_POINT = 0.0 1500.0 5.0 10.0 10.0 10.0
OUTPUT_POINT = 0.0 1500.0 5.0 15.0 10.0 10.0
END

# input sources
S
CENTERFREQUENCY = 450.0
SQUARE = 10.0 0.0 10.0 1.5 10.0 0.0 0.0 5.0 1 X 1.0
END

# input objects
O
ROUNDMINE = 15.0 0.0 3.0 8.0 3.0 4.0 1.0 0.0 0.0 2 1
EXTERNAL = ext.dat 1
END

# input background
B
HOMOGENEOUS = 1400.0 250.0 87.0
END

# define materials
M
# air
1 1.3 330.0 0.0
# plastic
2 1200.0 2700.0 1100.0
END
```

APPENDIX D

Elastic Constants for Isotropic Solids

Pressure Wave Speed in a Medium:

$$c_P = \sqrt{\frac{\lambda + 2\mu}{\rho}}$$

Shear Wave Speed in a Medium:

$$c_S = \sqrt{\frac{\mu}{\rho}}$$

Poisson's Ratio:

$$\nu = \frac{0.5\lambda}{\lambda + \mu}$$

Young's Modulus:

$$E = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu}$$

Longitudinal Wave Speed in a Plate:

$$c_L = \sqrt{\frac{E}{\rho(1 - \nu^2)}}$$

Transverse Flexural Wave Speed in a Plate:

$$c_{TF} = \sqrt{h\omega \left(\frac{E}{12\rho(1 - \nu^2)}\right)^{1/2}}$$

Bibliography

[1] C. Stewart, "Proposed massless remote vibration pickup," *J. Acoust. Soc. Am.*, vol. 30, pp. 644–645, 1958.

[2] J. M. Sabatier, H. E. Bass, L. N. Bolen, K. Attenborough, and V. S. Sastry, "The interaction of airborne sound with the porous ground: The theoretical formulation," *J. Acoust. Soc. Am.*, vol. 79, no. 5, pp. 1345–1352, May 1986.

[3] J. M. Sabatier, H. E. Bass, L. N. Bolen, and K. Attenborough, "Acoustically induced seismic waves," *J. Acoust. Soc. Am.*, vol. 80, no. 2, pp. 646–649, Aug. 1986.

[4] W. P. Arnott and J. M. Sabatier, "Laser-doppler vibrometer measurements of acoustic to seismic coupling," *Applied Acoustics*, vol. 30, pp. 279–291, 1990.

[5] "Feasibility of acoustic landmine detection: Final technical report," Tech. Rep., BBN Systems and Technologies, Cambridge, MA, 1992.

[6] K. Attenborough, J. M. Sabatier, and C. Frederickson, "The sound field within a rigid porous layer," *Applied Acoustics*, vol. 45, pp. 283–296, 1995.

[7] D. Donskoy, "Nonlinear vibro-acoustic technique for landmine detection," in *Detection and Remediation Technologies for Mines and Minelike Targets III, Proc. SPIE*, Apr. 1998, vol. 3392, pp. 211–217.

[8] W. R. Scott, Jr., C. T. Schröder, and J. S. Martin, "An acousto-electromagnetic sensor for locating land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets III, Proc. SPIE*, Apr. 1998, vol. 3392, pp. 176–186.

[9] W. R. Scott, Jr., C. T. Schröder, and J. S. Martin, "A hybrid acousto/electromagnetic technique for locating land mines," in *Proc. International Geoscience and Remote Sensing Symposium, Seattle, Wash.*, Apr. 1998, pp. 216–218.

[10] J. M. Sabatier and N. Xiang, "Acoustic-to-seismic coupling and detection of land mines," in *Proc. Geoscience and Remote Sensing Symposium*, Jul. 2000, vol. 4, pp. 1646–1648.

[11] J. M. Sabatier and N. Xiang, "An investigation of acoustic-to-seismic coupling to detect buried antitank landmines," *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, no. 6, pp. 1146–1154, Jun. 2001.

[12] W. R. Scott, Jr., J. S. Martin, and G. D. Larson, "Experimental model for a seismic landmine detection system," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 39, no. 6, pp. 1155–1164, Jun. 2001.

[13] W. R. Scott, Jr. and J. S. Martin, "Experimental investigation of the acousto-electromagnetic sensor for locating land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proc. SPIE*, Apr. 1999, vol. 3710, pp. 204–214.

[14] W. R. Scott, Jr., G. D. Larson, and J. S. Martin, "Simultaneous use of elastic and electromagnetic waves for the detection of buried land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets V, Proc. SPIE*, Apr. 2000, vol. 4038, pp. 667–678.

[15] W. R. Scott, Jr., J. S. Martin, and G. D. Larson, "Investigation of a technique that uses elastic waves to detect buried land mines," in *Proc. Geoscience and Remote Sensing Symposium*, Jul. 2000, vol. 4, pp. 1640–1642.

[16] E. E. Charette, "3D acoustic finite difference on a nCUBE massively parallel computer," Tech. Rep., ERL/nCUBE Geophysical Center for Parallel Processing, Sep. 1992.

[17] A. Tavlove, *Computational Electromagnetics: The Finite-Difference Time-Domain Method*, Artech House, Boston, 1995.

[18] K. J. Marfurt, "Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations," *Geophysics*, vol. 49, no. 5, pp. 533–549, May 1984.

[19] Z. Alterman and F. C. Karal, Jr., "Propagation of elastic waves in layered media by finite-difference methods," *Bulletin Seismological Society of America*, vol. 58, no. 1, pp. 367–398, Feb. 1968.

[20] M. Ottaviani, "Elastic wave propagation in two evenly-welded quarter-spaces," *Bulletin Seismological Society of America*, vol. 61, no. 2, pp. 1119–1152, 1971.

[21] K. R. Kelly, R. W. Ward, S. Treitel, and R. M. Alford, "Synthetic seismograms: a finite-difference approach," *Geophysics*, vol. 41, no. 1, pp. 2–27, Feb. 1976.

[22] A. Ilan and D. Loewenthal, "Instability of finite difference schemes due to boundary conditions in elastic media," *Geoph. Prosp.*, vol. 24, no. 3, pp. 431–453, Sep. 1976.

[23] B. Kummer and A. Behle, "Second-order finite-difference modeling of SH-wave propagation in laterally inhomogeneous media," *Bulletin Seismological Society of America*, vol. 72, no. 3, pp. 793–808, Jun. 1982.

[24] S. H. Emerman, W. Schmidt, and R. A. Stephen, "An implicit finite-difference formulation of the elastic wave equation," *Geophysics*, vol. 47, no. 11, pp. 1521–1526, Nov. 1982.

[25] R. Madariaga, "Dynamics of an expanding circular fault," *Bulletin Seismological Society of America*, vol. 66, pp. 639–666, 1976.

[26] J. Virieux, "SH-wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophysics*, vol. 49, no. 11, pp. 1933–1957, Nov. 1984.

[27] J. Virieux, "P-SV wave propagation in heterogenous media: Velocity-stress finite-difference method," *Geophysics*, vol. 51, no. 4, pp. 889–901, Apr. 1986.

[28] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propagat.*, vol. 14, pp. 302–307, 1966.

[29] A. R. Levander, "Fourth-order finite-difference P-SV seismograms," *Geophysics*, vol. 51, no. 11, pp. 1424–1436, Nov. 1988.

[30] N. Dai, A. Vaifidis, and E. R. Kanasewich, "Wave propagation in heterogeneous, porous media: a velocity-stress, finite-difference method," *Geophysics*, vol. 60, no. 2, pp. 327–340, Apr. 1995.

[31] Y. Q. Zeng and Q.-H. Liu, "Acoustic detection of buried objects in 3-D fluid saturated porous media: Numerical modeling," *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, no. 6, pp. 1165–1173, Jun. 2001.

[32] B. K. Sinha, Q.-H. Liu, and S. Kostek, "Acoustic waves in pressurized boreholes: a finite-difference formulation," *J. Geophysical Research*, vol. 101, no. B11, pp. 25,173–25,180, Nov. 1996.

[33] T. D. Mast, L. M. Hinkelman, L. A. Metlay, M. J. Orr, and R. C. Waag, "Simulation of ultrasonic pulse propagtion, distortion, and attenuation in the human chest wall," *J. Acoust. Soc. Am.*, vol. 106, no. 6, pp. 3665–3677, Jun. 2001.

[34] E. S. Krebes and G. Quiroga-Goode, "A standard finite-difference scheme for the time-domain computation of anelastic wavefields," *Geophysics*, vol. 59, no. 2, pp. 290–296, Feb. 1994.

[35] J. O. A. Robertson, J. O. Blanch, and W. W. Symes, "Viscoelastic finite-difference modeling," *Geophysics*, vol. 59, no. 9, pp. 1444–1456, Sep. 1994.

[36] E. L. Faria and P. L. Stoffa, "Finite-difference modeling in transversely isotropic media," *Geophysics*, vol. 59, no. 2, pp. 282–289, Feb. 1994.

[37] M. A. Biot, "Theory of propagation of elastic waves in a fluid-saturated porous solid. I. low-frequency range," *J. Acoust. Soc. Am.*, vol. 28, no. 2, pp. 168–178, Mar. 1956.

[38] J. E. Vidale and R. W. Clayton, "A stable free-surface boundary condition for two-dimensional elastic finite-difference wave simulation," *Geophysics*, vol. 51, no. 12, pp. 2247–2249, Dec. 1986.

[39] R. Clayton and B. Engquist, "Absorbing boundary conditions for acoustic and elastic wave equations," *Bulletin Seismological Society of America*, vol. 67, no. 6, pp. 1529–1540, Dec. 1977.

[40] S. H. Emerman and R. A. Stephen, "Comment on 'Absorbing boundary conditions for acoustic and elastic wave equations'," *Bulletin Seismological Society of America*, vol. 73, no. 2, pp. 661–665, Apr. 1983.

[41] K. D. Mahrer, "Numerical time step instability and Stacey's and Clayton-Engquist's absorbing boundary conditions," *Bulletin Seismological Society of America*, vol. 80, no. 1, pp. 213–217, Feb. 1990.

[42] R. G. Keys, "Absorbing boundary conditions for acoustic media," *Geophysics*, vol. 50, no. 6, pp. 892–902, Jun. 1985.

[43] C. J. Randall, "Absorbing boundary condition for the elastic wave equation," *Geophysics*, vol. 53, no. 5, pp. 611–624, May 1988.

[44] W.-F. Chang and G. A. McMechan, "Absorbing boundary conditions for 3-D acoustic and elastic finite-difference calculations," *Bulletin Seismological Society of America*, vol. 79, no. 1, pp. 211–218, Feb. 1989.

[45] R. A. Renault and J. Petersen, "Stability of wide-angle absorbing boundary conditions for the wave equation," *Geophysics*, vol. 54, no. 9, pp. 1153–1163, Sep. 1989.

[46] C. J. Randall, "Absorbing boundary condition for the elastic wave equation: Velocity-stress formulation," *Geophysics*, vol. 54, no. 9, pp. 1141–1152, Sep. 1989.

[47] L. T. Long and J. S. Liow, "A transparent boundary for finite-difference wave simulation," *Geophysics*, vol. 55, no. 2, pp. 201–208, Feb. 1990.

[48] C. Peng and M. N. Toksoez, "An optimal absorbing boundary condition for finite-difference modeling of acoustic and elastic wave propagation," *J. Acoust. Soc. Am.*, vol. 95, no. 2, pp. 733–745, Feb. 1994.

[49] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Physics*, vol. 114, pp. 185–200, 1994.

[50] Z. S. Sacks, D. M. Kingsland, R. Lee, and J. F. Lee, "A perfectly matched anisotropic absorber for use as an absorbing boundary condition," *IEEE Trans. Antennas Propagat.*, vol. 43, pp. 1460–1463, Dec. 1995.

[51] S. D. Gedney, "An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD-Lattices," *IEEE Trans. Antennas Propagat.*, vol. 44, pp. 1630–1639, Dec. 1996.

[52] L. Zhao and A. C. Cangellaris, "GT-PML: generalized theory of perfectly matched layers and its application to the reflectionless truncation of finite-difference time-domain grids," *IEEE Trans. Microwave Theory and Techniques*, vol. 44, no. 12, pp. 2555–2563, Dec. 1996.

[53] J.-P. Berenger, "Improved PML for the FDTD solution of wave-structure interaction problems," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 466–473, Mar. 1997.

[54] F. D. Hastings, J. B. Schneider, and S. L. Broschat, "Application of the perfectly matched layer (PML) absorbing boundary condition to elastic wave propagation," *J. Acoust. Soc. Am.*, vol. 100, no. 5, pp. 3061–3068, Nov. 1996.

[55] W. C. Chew and Q. H. Liu, "Perfectly matched layer for elastodynamics; a new absorbing boundary condition," *J. Comput. Acoustics*, vol. 4, pp. 341–359, 1996.

[56] X. Yuan, D. Borup, J. W. Wiskin, M. Berggren, R. Eidens, and S. A. Johnson, "Formulation and validation of Berenger's PML absorbing boundary for the FDTD simulation of acoustic scattering," *IEEE Trans. Ultrasonics, Ferroelectrics, Frequency Control*, vol. 44, no. 4, pp. 816–822, Jul. 1997.

[57] B. A. Auld, *Acoustic Fields and Waves in Solids*, vol. 1 and 2, Robert E. Krieger Publishing Company, 2nd edition, 1990.

[58] K. F. Graff, *Wave Motion in Elastic Solids*, Dover Publications, 1975.

[59] A. Tavlove, *Advances in Computational Electromagnetics: The Finite-Difference Time-Domain Method*, Artech House, Boston, 1998.

[60] *http://www.beowulf.org.*

[61] R. D. Richtmyer, *Difference Methods for Initial Value Problems*, John Wiley and Sons, New York, 2nd edition, 1967.

[62] G. D. Smith, *Numerical Solution of Partial Differential Equations; Finite Difference Methods*, Clarendon Press, 2nd edition, 1978.

[63] A. R. Mitchell and D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations*, John Wiley and Sons, New York, 1980.

[64] P. Fox, "The solution of hyperbolic partial differential equations by difference methods," in *Mathematical Methods for Digital Computers*. A. Ralston and H.S. Wilf, Editors, Ch. 16, Wiley, New York, 1964.

[65] D. S. Watkins, *Fundamentals of Matrix Computations*, John Wiley and Sons, New York, 1991.

[66] J. W. S. Rayleigh, "On waves propagated along the plane surface of an elastic solid," *Proc. Lond. math. Soc.*, vol. 17, no. 2, pp. 4–11, 1887.

[67] M. Roth, R. Spitzer, and F. Nitsche, "Seismic survey across an environment with very high Poisson's ratio," in *Proc.* 12$^{\text{th}}$ *Annu. Int. Meeting, Symp. Application of Geophysics to Engineering and Environmental Problems*, 1999, pp. 49–55.

[68] E. Smith, P. S. Wilson, F. W. Bacon, J. F. Manning, J. A. Behrens, and T. G. Muir, "Measurements and localization of interface wave reflections from a buried target," *J. Acoust. Soc. Am.*, vol. 103, no. 5, pp. 2333–2343, May 1998.

[69] N. E. Glass and A. A. Maradudin, "Leaky surface-elastic waves on both flat and strongly corrugated surfaces for isotropic, nondissipative media," *J. Appl. Phys.*, vol. 54, no. 2, pp. 796–805, Feb. 1983.

[70] G. F. Miller and H. Pursey, "The field and radiation impedance of mechanical radiators on the free surface of a semi-infinite isotropic solid," *Proc. R. Soc.*, vol. A 223, pp. 521–541, 1954.

[71] J. Miklowitz, *The Theory of Elastic Waves and Wave Guides*, North-Holland Publishing Company, 1978.

[72] J. D. Achenbach, *Wave Propagation in Elastic Solids*, North-Holland Publishing Company, 1973.

[73] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*, IEEE Press, 1994.

[74] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, 9$^{\text{th}}$ edition, 1970.

[75] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, Academic Press, 4$^{\text{th}}$ edition, 1995.

[76] G. S. Smith, *Classical Electromagnetic Radiation*, Cambridge University Press, 1997.

[77] A. Behboodian, W. R. Scott, Jr., and J. H. McLellan, "Signal processing of elastic surface waves for localizing buried land mines," in *Proc. 33rd Assilomar Conference on Signals, Systems, and Computers*, Oct. 1999.

[78] W. R. Scott, Jr., C. T. Schröder, G. D. Larson, and J. S. Martin, "Use of elastic waves for the detection of buried land mines," in *Proc. IGARSS*, Jul. 2001.

[79] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, Springer-Verlag, New York, 1st edition, 1991.

[80] T. D. Rossing, D. A. Russell, and David E. Brown, "On the acoustics of tuning forks," *Am. J. Phys.*, vol. 60, no. 6, pp. 620–626, Jul. 1992.

[81] M. W. Sprague, R. Raspet, H. E. Bass, and J. M. Sabatier, "Low frequency acoustic ground impedance measurement technique," *Applied Acoustics*, vol. 39, pp. 307–325, 1993.

[82] B. M. Das, *Principles of Soil Dynamics*, PWS-Kent Publishing Company, Boston, 1993.

[83] R. Bachrach and A. Nur, "High-resolution shallow seismic experiments in sand," *Geophysics*, vol. 63, no. 4, pp. 1225–1240, Jul. 1998.

[84] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, 1989.

[85] C. T. Schröder and W. R. Scott, Jr., "Finite-difference time-domain model for elastic waves in the ground," in *Detection and Remediation Technologies for Mines and Minelike Targets IV, Proc. SPIE*, Apr. 1999, vol. 3710, pp. 1361–1372.

[86] C. T. Schröder, *A Finite-Difference Model for Elastic Waves in the Ground*, Diploma Thesis, Technische Universität Braunschweig, Jun. 1999.

[87] C. T. Schröder and W. R. Scott, Jr., "Three-dimensional FDTD model to study the elastic-wave interaction with buried land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets V, Proc. SPIE*, Apr. 2000, vol. 4038, pp. 680–690.

[88] C. T. Schröder, K. Kim, and W. R. Scott, Jr., "A FDTD model for the interaction of elastic waves with buried land mines," *J. Acoust. Soc. Am.*, vol. 107, no. 5, Pt. II, pp. 2897, May 2000.

[89] C. T. Schröder and W. R. Scott, Jr., "A three-dimensional finite-difference model for the interaction of elastic waves with buried land mines," in *Proc. Geoscience and Remote Sensing Symposium*, Jul. 2000, vol. 1, pp. 26–28.

[90] C. T. Schröder and W. R. Scott, Jr., "Resonance behavior of buried land mines," in *Detection and Remediation Technologies for Mines and Minelike Targets VI, Proc. SPIE*, Apr. 2001, vol. 4394.

[91] W. R. Scott, Jr., C. T. Schröder, and J. S. Martin, "An acousto-electromagnetic method for detecting buried objects," in *Proc.XXVIth General Assembly of the International Union of Radio Science, Toronto, Canada*, Aug. 1999, vol. 3710, p. 724.

[92] W. R. Scott, Jr., C. T. Schröder, J. S. Martin, and G. D. Larson, "Investigation of a technique that uses both elastic and electromagnetic waves to detect buried land mines," in *Proc. AP2000 - Millennium Conference an Antennas and Propagation, Davos, Switzerland*, Apr. 2000.

[93] C. T. Schröder and W. R. Scott, Jr., "A finite-difference model to study the elastic-wave interactions with buried land mines," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 38, pp. 1505–1512, Jul. 2000.

[94] C. T. Schröder and W. R. Scott, Jr., "Elastic waves interacting with buried land mines: a study using the FDTD method," *submitted to the IEEE Trans. on Geoscience and Remote Sensing.*

[95] C. T. Schröder and W. R. Scott, Jr., "On the stability of the FDTD algorithm for elastic media at a material interface," *submitted to the IEEE Trans. on Geoscience and Remote Sensing.*

[96] C. T. Schröder and W. R. Scott, Jr., "On the complex conjugate roots of the rayleigh equation: the leaky surface wave," *submitted to the J. Acoust. Soc. Am.*

VITA

Christoph T. Schröder [redacted] He received the M.S. degree in electrical and computer engineering at the Georgia Institute of Technology, Atlanta, GA (USA), in 1997, the Diplom der Elektrotechnik from the Technical University of Braunschweig (Germany) in 1999, and the Ph.D. degree in electrical and computer engineering at the Georgia Institute of Technology in 2001. His research interests include general wave propagation, remote sensing, numerical modeling in elasticity and electromagnetics, and the finite-difference time-domain technique.



DETECTION OF SURFACE WAVES IN THE GROUND USING AN ACOUSTIC METHOD

A Thesis
Presented to
The Academic Faculty

By

Fabien Codron

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in Mechanical Engineering

Georgia Institute of Technology
July 2000

DETECTION OF SURFACE WAVES IN THE GROUND USING AN ACOUSTIC METHOD

Approved:

_____ Peter H. Rogers

_____ Waymond R. Scott

_____ Yves Berthelot

Date Approved

_____

ACKNOWLEDGEMENT

This work was accomplished in an environment very new to me. I had to adapt to the language, the culture, and discover the campus, the city and the people. Many people helped me go through this experience. Their knowledge, their advice and their support made my research easier. I would like to thank them all.

First, my advisor P. Rogers for giving me the opportunity to contribute to this research.

W. Scott for managing the landmine detection project at Georgia Tech, and also for his very useful debugging visits, and advice on electronics

Y. Berthelot for teaching me the basis on acoustic transducers and supporting the exchange with Georgia Tech Lorraine.

J. Martin for his availability, his everyday advice and teaching. His contribution to the good running of the acoustical laboratory was essential.

I would also like to thank the graduate students of the acousto-dynamic group for their availability. They maintained a nice work environment.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

SUMMARY ix

CHAPTER

I BACKGROUND 1

A. General 1

B. An acoustic method to detect the surface waves in the ground 10

C. Litterature Review 12

II THE TRANSDUCER SYSTEM 19

A. Presentation of the Transducers 20

B. Reflection of the Ultrasonic Waves from Soil 23

C. Focusing the Transducer Sound 26

III PHASE DEMODULATION 37

A. Investigation of Digital Demodulation 42

B. Analog Demodulation 52

IV NOISE AND FILTERING 65

A. Noise Measurement 65

B. Acoustic Noise 66

C. Electromagnetic Cross Talk 67

D. Source Noise 68

E. Filtering 73

F. Digital Signal Processing 76

V CONCLUSIONS 79

VI RECOMMENDATIONS 82

VIII APPENDICES 85

VII REFERENCES 93

LIST OF FIGURES

Page

Figure 1.1 - Schematic diagram of acousto-electromagnetic experimental system 3

Figure 1.2 - Top view of the experimental system 4

Figure 1.3 - Two dimensional finite difference model 7

Figure 1.4. - Numerical model results for the mine interaction with surface wave 8

Figure 1.5 -experimental and numerical results for the surface wave propagation 9

Figure 1.6 - spatial resolution 12

Figure 1.7 - Spectrum of a pure tone modulated by a sine wave 15

Figure 1.8 - schematic of the waves phases 15

Figure 1.9 -schematic of the phase demodulation signal processing 17

Figure 1.10 - Phase demodulation for a small amplitude modulation and signal in quadrature 18

Figure 2.1 - Schematic of the transducer assembly 20

Figure 2.2 - Cross section of a transducer 21

Figure 2.3 - Generation of the transducer signal 22

Figure 2.4 - Transmit response of the capacitance transducer 22

Figure 2.5 - Picture of transducer facing a sand sample 24

Figure 2.6 - Graph of the pressure on axis, in the nearfield 27

Figure 2.7 - sound field generated by piston at 50kHz 29

Figure 2.8 - on axis pressure generated by a 50kHz spherically focused transducer 31

Figure 2.9 - pressure field generated by a 50kHz spherically focused transducer 32

Figure 2.10 On axis pressure generated by a 200kHz spherically focused transducer 33

Figure 2.11 - Normalized pressure maximums versus true focusing distance for 50, 100 and 200 kHz focused transducers 33

Figure 2.12 - Cross-section of a spherically focused transducer 34

Figure 3.1 - influence of the transducer axis angle on sensitivity 38

Figure 3.2 - Square wave with jittered edges 42

Figure 3.3 - Schematic of bit coding 43

Figure3.4 - Fourier transform of the 50kHz pure tone acquired 44

Figure 3.5 - Fourier transform of the digitized signal coded on 32 bits 47

Figure 3.6 - Fourier transform of an ideally sampled signal coded on 12bits 47

Figure 3.7 - Fourier transform of a ideally sampled signal coded on 16bits 48

Figure 3.8 - Fourier transform of signal sampled with 5.10^{-8}s jitter 49

Figure 3.9 - Fourier transform of signal sampled with 5.10^{-9}s jitter 50

Figure 3.10 Measurement setup 53

Figure 3.11- Picture of the transducer facing the sound projector 54

Figure 3.12 Schematic of the signal processing 62

Figure 3.13 - Operational amplifier circuit for quadrature 63

Figure 4.1 - Spectral density of a pure tone 69

Figure - 4.2 three stages passive low pass filter 71

Figure - 4.3 One stage of the Chebyshev active low pass filter 74

Figure - 4.4 Amplitude response of the active filter implemented on board 75

Figure 4.5 LabView diagram of the digital signal processing 77

LIST OF TABLES

Page

Table 1.1 - acoustic waves velocities in sand and mines 6

Table 2.1 - received signal level and signal to noise ration 25

Table 3.1 - results of the mixer test 57

Table 3.2 - Transfer function of the K-H 3 poles Bessel filter for fc=3000Hz 58

Table 3.3 - Calibration results for the TM3 configuration 60

Table 3.4 - Calibration results for the AD534 configuration 60

Table 3.5 - Results for the AD630 configuration 60

Table 4.1 - Noise floor levels in dBm after filtering of the multiplier output 72

Table 4.2 - values of the resistors and capacitors for the filter 74

SUMMARY

Land mine detection techniques currently in use are not reliable for modern plastic mines. An acousto-electromagnetic technique that has the potential to detect such mines is being investigated at Georgia Tech. It uses an acoustic source for generating waves in the ground, which are detected with a radar, which scans the surface to be cleared of mines. The radar system visualizes the surface wave and its interaction with the mine by measuring the surface vibration. This radar ground vibration measuring system is expensive and may not be effective in all environments.

The purpose of this thesis is to investigate an ultrasonic vibrometer that could be used to supplement the radar or replace it. An ultrasonic system was implemented and tested with several different demodulation techniques. Emphasis was laid on getting a sensitivity of 1-nanometer, equal to that of the radar sensor. In order to obtain such sensitivity , design and optimization of the source, the transducer signal, the electronic filtering and the demodulation were conducted. The focusing of the ultrasound and the effects of spot size were also considered.

The system presented in this thesis has good potential characteristics for surface waves detection at a low cost. It achieves the required resolution with transducers running at 50kHz for a vibration of the soil in the frequency range 400-1200Hz. Placing the transducers a couple inches away from the vibrating surface produces a satisfactory spot size

CHAPTER I

BACKGROUND

General

Landmines are responsible for over 20,000 injuries or deaths per year. The recent Ottawa convention banning land mines has not been signed by all the major parties. Moreover, it does little to clean up the existing worldwide scourge of buried land mines. According to the United Nations, more than 100 million mines are buried over the planet. At the current rate of detection and removal, clearing the world's land mines could take hundreds of years.

The main problem for mine detection lies in the design of modern mines. A mine's structure includes a casing, explosive materials and a firing mechanism. Most modern mines are manufactured with plastic casings. The only metallic part is the small firing mechanism. Unfortunately, conventional metal detectors cannot discriminate between tin cans, bullets, and scrap metal, and the firing mechanism of a landmine. As a result, metal detectors have a considerable rate of false alarms. What is needed is a safe, reliable and cost effective technology for finding land mines.

As the metal part of the mine is very small, many researchers have turned to the development of techniques detecting other properties of landmines. Nuclear quadrupole resonance [1] has been used successfully to detect explosive materials. By probing the earth with radio-frequency signals, this technique can generate a coherent signal unique to certain compounds- including explosives such as RDX or TNT. Others tried to detect the plastic casing. Electro-quasistatic

[2] was developed and has the potential to detect the size and the shape of plastic objects. Many researchers have investigated acoustic techniques. The acoustic properties of the mines are very different from those of the surrounding soil regardless of the material used for the casing. Furthermore, the air trapped inside the complex structure of a mine creates a cavity. This cavity is likely to resonate at some frequencies when an external force is applied. Some effort has been directed at the development of pulse-echo techniques [3]. Generally these techniques did not solve the false alarm issue. Clutter, debris, rocks can also have acoustical properties very different from the soil too. They can reflect the pressure wave and have a signature similar to a mine's [4].

A new land mine detection that simultaneously uses both electromagnetic and acoustic waves in a synergistic manner is currently being investigated at Georgia Tech [5]. This combined technique has the potential to enhance the signature of the mine with respect to the clutter and make it possible to detect a mine when other methods fail. This technique is presented in detail as it motivated this thesis work. It has detected mines buried as deep as 12 inches. It has also performed successfully in the presence of clutter and when the ground was covered by vegetation (pine straw)

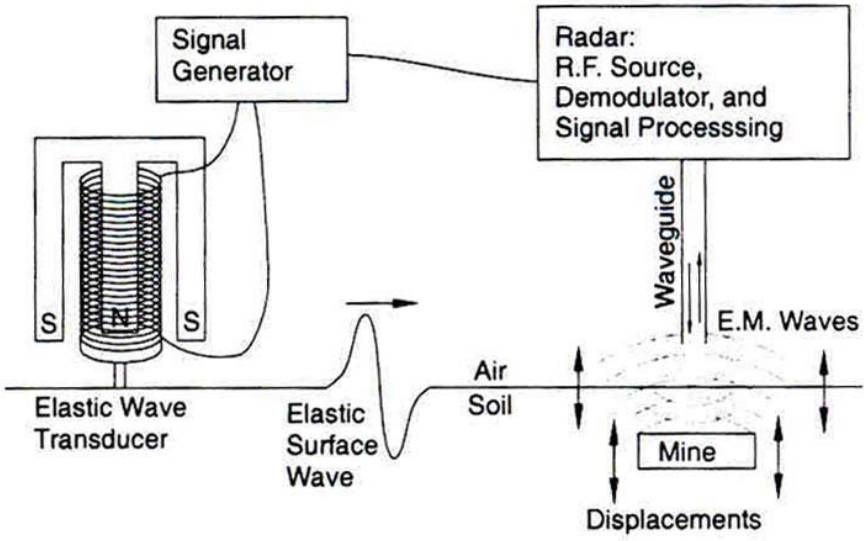


Figure 1.1 Schematic diagram of acousto-electromagnetic experimental system [7]

The configuration of the system consists of a radar and an seismic-acoustic source (electromagnetic shaker). The acoustic source induces an elastic seismic wave into the earth, which propagates along the surface of the soil. The elastic wave causes both the mine and the surface to be displaced. The displacement of the mine is different from that of earth, because the acoustic properties of the mine are quite different from those of the earth. Hence the displacement of the surface of the soil is affected by the presence of the mine. Resonance, scattering, and distortion of the surface wave corresponding to the presence of a mine can be observed by visualization of the surface wave propagation. In the current system, the electromagnetic radar is used to detect displacements of the surface and, hence visualize the surface waves.

The experimental set-up [6] consists of a concrete tank filled with sand (figure 1.2). Acoustic waves are generated using an electro-dynamic shaker mounted upside down and equipped with a foot contacting the sand. The radar is attached to an x-y positioner, located on a frame 50cm above the sand. The positioner scans the radar mechanically over a 120cm by 80cm surface of sand located in front of the shaker. Each point on the sand surface interrogated by the radar, is exited by an identical acoustic wave and the displacement of the soil is recorded. From this data, color animations of the surface wave propagating can be created, which display the wave interaction with the mine.

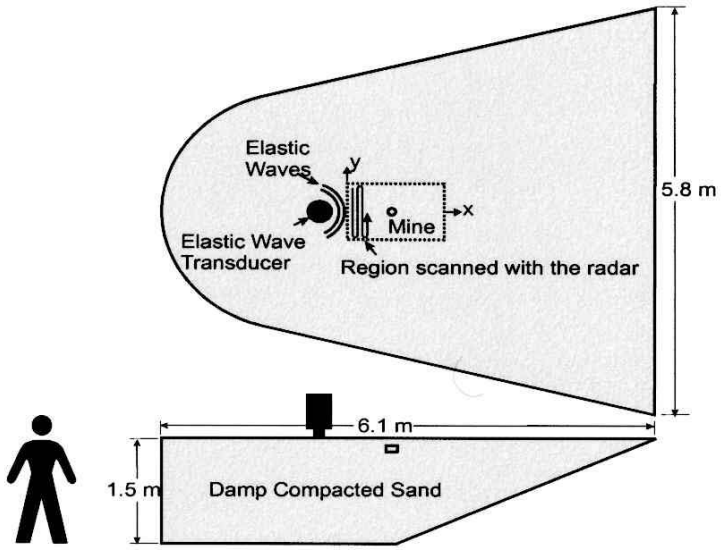


Figure 1.2 the experimental system[8]

The tank is filled with 50 tons of damp sand with a relatively uniform density and cohesion. The source is located at one end and radiates into a nearly free field. The tank dimensions (figure 1.2) are large enough so that the waves that are incident on the sidewalls do not cross the measurement region until after

the relevant signal have been recorded. Hence reflection on the walls does not appear on the data recorded. Due to the large dimensions, cavity resonances are at low frequencies in the tank.

The x-y positioner used to move the radar over the surface of the sand is under computer control. The mechanical scanning and the recording of the data are performed automatically. The scan is performed on a rectangular grid of discrete positions. The points are spaced every 1cm in the x direction and every 2 centimeter in the y direction. It currently takes 24 to 48 hours to perform a complete scan. The reason is that the measurement setup is designed to get the maximum data quality without concern about the scan time. This time can be greatly reduced by radar so that it can scan several points simultaneously.

There were two main challenges for the design of such a radar vibrometer. First, make it sufficiently sensitive to be able to detect small vibrations. Secondly, make the spot size sufficiently small. Measured vibration of the sand were usually smaller than a micrometer. Currently the radar configuration is able to detect 1-nanometer vibration amplitude. Such sensitivity has given a satisfactory vertical resolution. The spot size i.e. the area illuminated by the radar must be smaller than one half of the wavelength of the highest frequency ground wave. Currently a small spot size is obtained by using an open-ended wave-guide as the antenna for the radar. This antenna gives satisfactory results when its open end is placed within a few centimeters of the surface.

Validity of the media:

Sand was chosen as the soil medium for several different reasons. It is much easier to repeatably bury and dig up mines and clutter in sand. It also has

similar mechanical properties (wave velocities and displacement amplitude) to typical unconsolidated soils when used wet and compacted. As a result, a special device had to be installed to compensate for the evaporation of the water in the sand close to the surface. The table 1.1 gives acoustic properties representative of the damped compacted sand and compares them with those of a plastic mine. These figures were used in the numerical model exposed in the next section.

	Sand	Typical plastic mine
Pressure wave velocities (m/s)	250	2700
Shear wave velocities (m/s)	87	1100

Table 1.1 acoustic waves velocities in sand and mines [9]

Elastic waves – land mine interaction

In order to get a better understanding of the interaction of the plastic mines and elastic wave propagating in the ground, a numerical model was built [9]. It consists of a finite-difference method and uses equation of motion and the stress strain-relation. From these equations, a first order stress-velocity formulation is obtained. A free boundary limits the top of the discretized volume and a perfectly matched layer absorbs the outward going wave in all other directions.

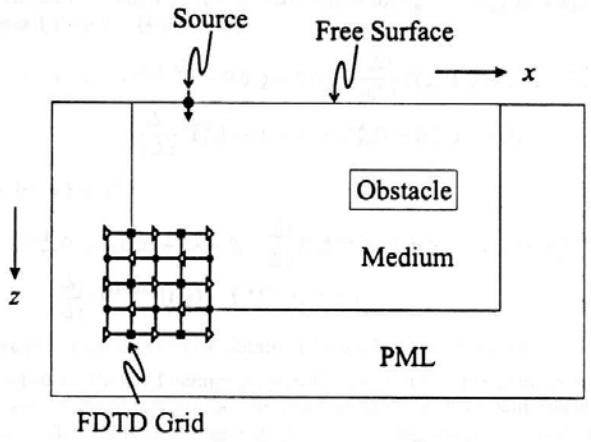


Figure 1.3 Two dimensional finite difference model [9]

Two simple models for antipersonnel mines were investigated: one containing an air filled chamber and one without an air filled chamber. The excitation producing the ground waves is a differentiated Gaussian pulse with a center frequency of 800Hz. The figure next page shows the results of a simulation for a mine with air cavity. The source is located at x=0 cm and the mine at 45cm, 2cm underneath the surface of the ground. They show that the waves hitting the mine are partially transmitted and partially reflected. This reflection is clearly seen for the mine without air filling. For the other mine, the reflected waves are only weakly dispersed. However, in this case, a resonance of the mine occurs. Both of these phenomena give the mine signature.

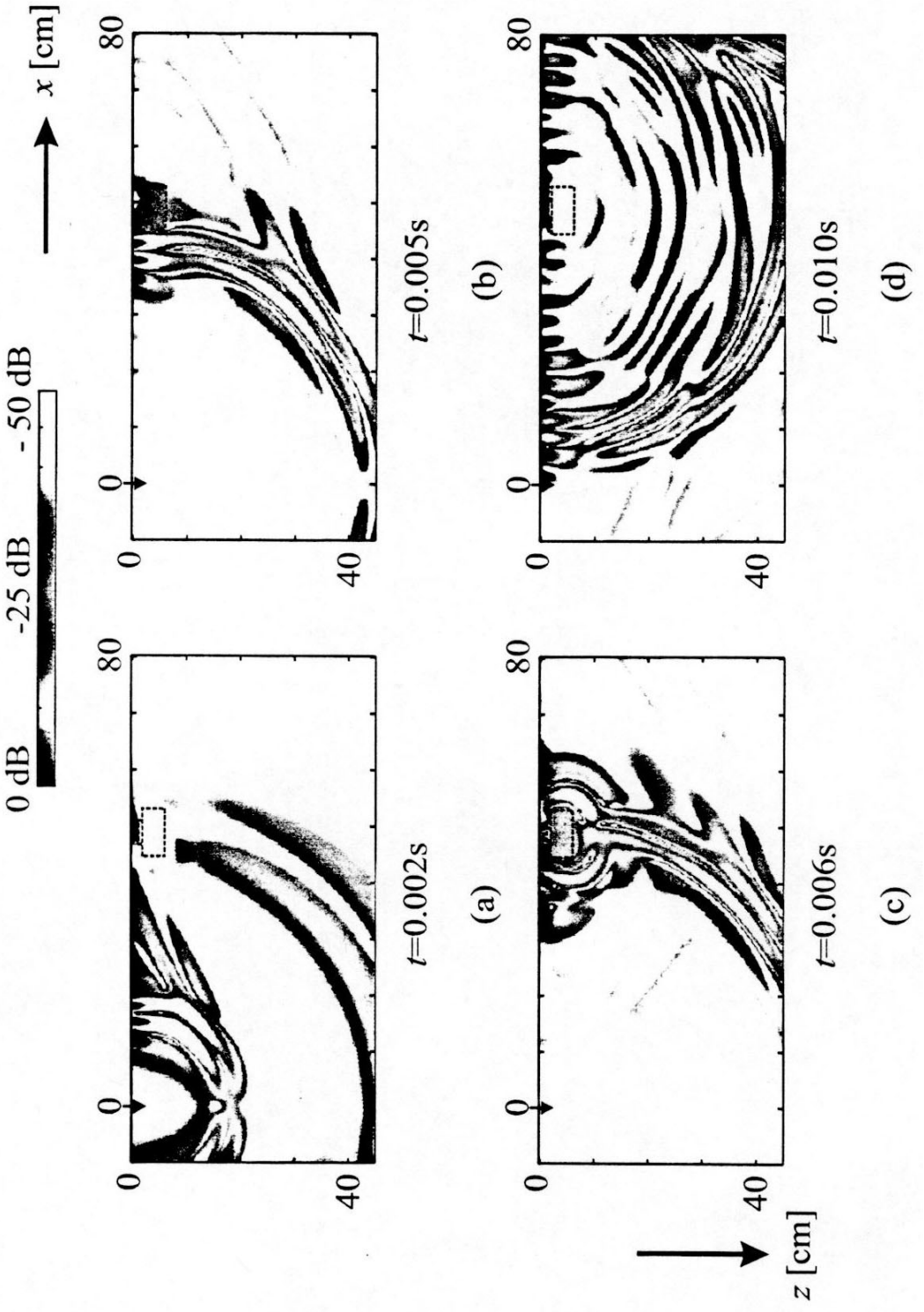


Figure1.4. Numerical model results for the mine interaction with surface waves[9]

Comparison with experimental model

The results computed with the finite difference model were compared to the experimental model. They are in fairly good agreement[6] (see figure 1.5).

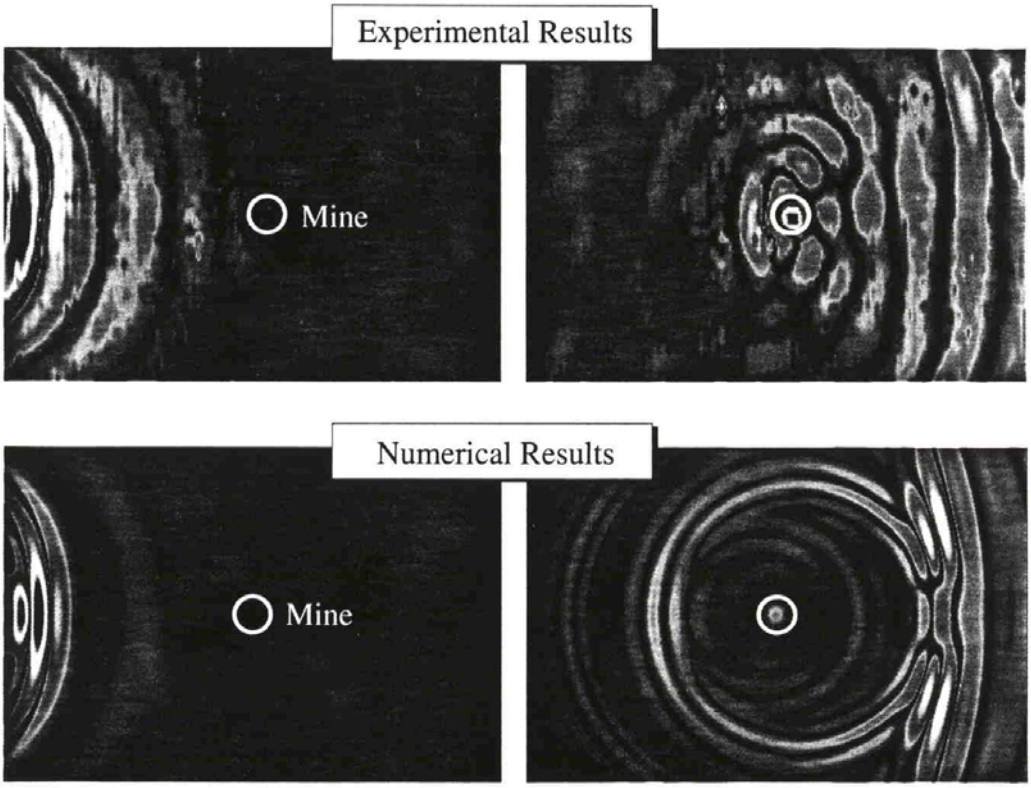


Figure 1.5 experimental and numerical results for the surface wave propagation [7]

AN ACOUSTIC METHOD TO DETECT THE SURFACE WAVES IN THE GROUND

A few aspects of the radar system are impractical. The radar has to be very close to the ground in order to get a high signal level and a small spot size. This is not practical for a use in fields with vegetation and soil irregularities. It is also heavy and big and so, difficult to scan mechanically over the soil.

Cost is also an issue. Most of the countries affected by the land mine problem are poor. Angola, Mozambique, Croatia, Afghanistan, and Cambodia cannot afford expensive techniques. Unfortunately it seems that a radar scan is bound to be expensive: Due to the cost of a single radar unit, it will not be possible to use a large array of them. Therefore, it will be difficult to get a scan time under 5 minutes for a square meter surface. Moreover, the electronics processing the demodulation are expensive due to the high performance required and the high frequencies of the signal.

On the other hand an acoustic device using ultrasonic waves between 50kHz and 200kHz could replace the radar sensor. The wavelength of airborne sound waves at these frequencies would be smaller than the radar's because of the difference of propagation speed of electromagnetic wave and sound waves. The transducers could be located at a reasonable distance from the ground and easily scanned due to their small weight and size. Because they are running at much lower frequencies, the electronics and the transducers are commercially available and a lot cheaper.

The basic idea for such an acoustic system that detects surface waves propagating in the ground is simple. An acoustic transducer sends sound to the ground, another one receives the sound reflected and modulated by the soil vibration. The signal is then processed to compute the ground vertical position (or height) at the spot interrogated by the transducers. This operation is a phase demodulation process. Using arrays of transducers would allow measuring the vibration at several spots simultaneously. Similarly as with the original system, by scanning this array over the soil, and repeating the exact same seismic wave, a map of the vibration at discrete locations spaced over the surface can be computed.

Objectives

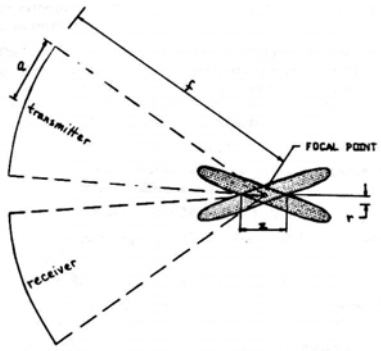
In the same manner as for the radar system, a few characteristics are critical for such a detection device: First, it must have a diplacement sensitivity of 1-nanometer in the frequency range of 100Hz to 2000Hz. The vibration amplitude of the surface waves do not exceed 1 micron. Detecting very small displacement is necessary to display the weak reflection of the surface waves on the mine. The ground source sends a Gaussian pulse centered at 800Hz, however the frequencies over 800Hz are heavily attenuated and create very small displacements. As a result the frequencies in the range 100-800Hz correspond to higher displacements of the ground more likely to show the mine signature.

Secondly, the spot size of the system has to be smaller than half a seismic wavelength. For a shear wave velocity of 80m/s and a frequency of 800Hz, this smallest wavelength is 10cm. Hence a spot size of the order of a few centimeter would be satisfactory.

LITERATURE REVIEW

Other active acoustic devices for detecting vibrations have been developed in the past. In particular M. Cox and P. Rogers implemented an underwater ultrasonic system measuring vibrations of fish hearing organs [10]. Though running underwater at very high frequencies, this system has a lot in common with an aeroacoustic soil vibration detector. The system used a 10Mhz source as a carrier and measured the sound reflected by the organ of the fish. The spectral analysis of these echoes provides the frequency and the amplitude of the organ vibration. The system was capable of measuring displacements of 1.2-nanometer with a spatial resolution of 0.28mm. This extremely small spot size was achieved by using high frequency focused transducers.

By crossing the transducer axis with an angle of 20 degrees, vertical axial resolution z of 0.8mm and lateral resolution r of 0.14mm were achieved.



$$r \propto (f/a)\,\lambda$$
$$z \propto (f/a)^2\,\lambda$$

Figure1.6 spatial resolution

Only harmonic vibrations were considered and the amplitude was determined with a spectrum analyzer. The receive signal spectrum features side lobes on both sides of the carrier tone at frequencies ω_c-ω_L and ω_c+ω_L, where $\omega_c$is the carrier frequency and ω_L the frequency of the vibration. The side band amplitude is down from the carrier tone amplitude by a factor of kx_0 [10]. Where k is the wave number of the ultrasound in the water and x_0 the peak amplitude of the vibration. Hence measurement of this factor on a spectrum analyzer provides the vibration amplitude.

This system used very high frequencies. The resulting wavelength of 0.15mm gives a good sensitivity and focusing of the sound. Such a wavelength in the air would correspond to a frequency of 2.25MHz. It would be unrealistic because of the large attenuation of such high frequencies in the air. The technique using a spectrum analyzer is valid only for single frequency measurement with constant amplitude in time. It is not the case for transient surface waves in the ground. A more elaborated demodulation process is required in this case.

Frequency demodulation techniques

Frequency modulation techniques find their origin in the last decade of the 19th century. They were developed intensively after 1930 for communication systems. The fundamentals of frequency demodulation are summarized in "High performance frequency demodulation" [11] This section describes the concepts of any phase demodulation process.

Frequency modulated signals can be represented by the following expression:

$$s(t) = A\cos(\omega_c t + \varphi(t))$$

A denotes the carrier amplitude and ω_c the carrier frequency (in rad/s). The information carried by the signal lies in the phase $\varphi(t)$. Frequency demodulators process the signal in order to obtain $d\varphi(t)/dt$.

Spectrum for sinusoidal modulation

Assume that the signal is given by

$$\varphi(t) = B \sin(\omega_L t)$$

then $s(t) = A\cos(\omega_c t + B \sin(\omega_l t))$

The spectrum is obtained by Fourier transformation of s(t). This spectrum correspond to the Fourier series expansion:

$$s(t) = A \sum_n BJn \cos(\omega_c + n\omega_L)t$$

where Jn denote the Bessel function of the first kind and order n.

For small modulation amplitudes (modulation indices), one can consider only the first few terms. Then the spectrum basically consists of a component at

the fundamental frequency ω_c, equal to $AJ_0(A)$ and the first harmonic components at ω_l equal to

$$AJ_{+/-1}(B) \approx .5AB$$

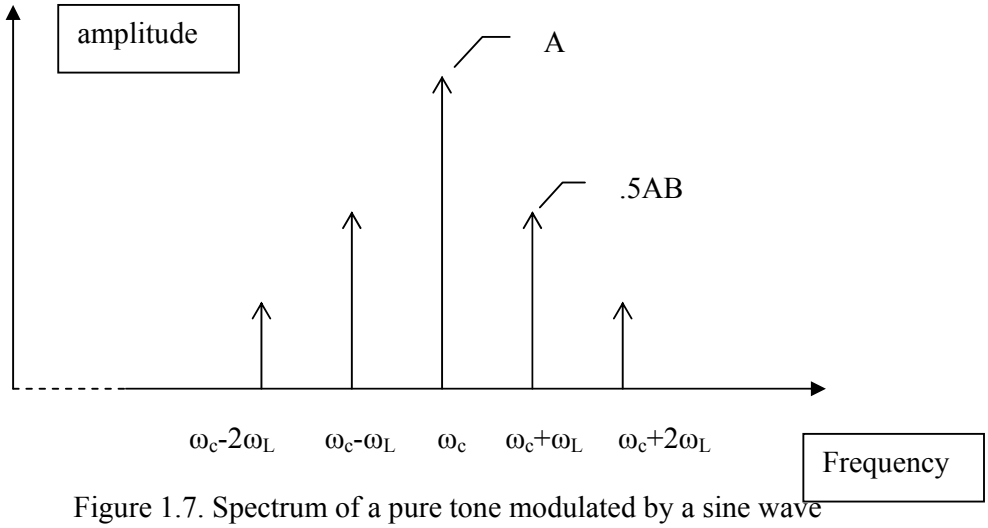


Figure 1.7. Spectrum of a pure tone modulated by a sine wave

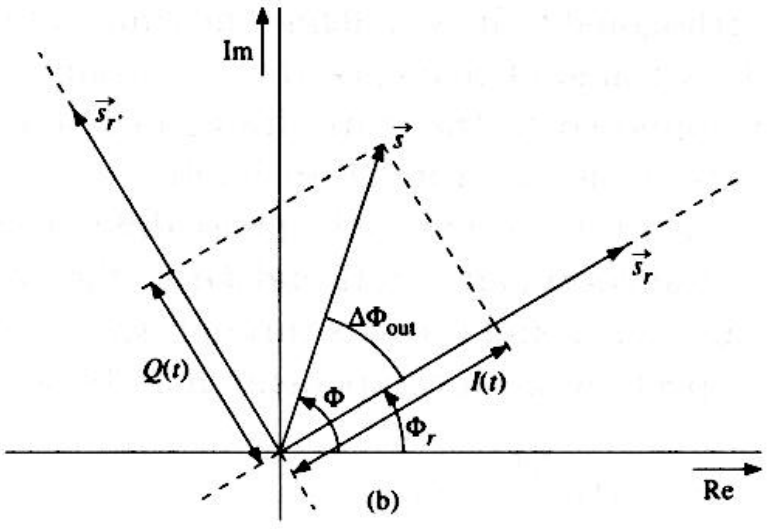


Figure 1.8 schematic of the waves phases [11]

What demodulation technique is appropriate depends on the carrier frequency, the amplitude and the bandwidth of the modulation, and the noise level tolerable.

General algorithm for a phase demodulation:

A phase demodulator determines the phase difference that exists between the two waves applied to its both inputs; the wave subjected to demodulation s(t) and a reference wave $s_r(t)$. These waves can be noted

$s(t) = A \cos\phi(t)$

$s_r(t) = A_r \cos\phi_r(t)$

The output of the phase demodulator is:

$$y_{out} = K\Delta\Phi_{out} = \Phi(t) - \Phi_r(t)$$

If s_r' is the wave in quadrature with the reference wave, s_r and s_r' define a frame of reference rotating around the origin a with angular velocity $\partial\Phi_r(t)/\partial t$ .The phase difference can be expressed in terms of the coordinates I(t) and Q(t), the coordinates of s(t) with respect to the rotating axis through s_r and s_r'. (see picture above).

The components I(t) and Q(t) may be written as

$I(t) = \mathbf{s} \cdot \mathbf{s_r}$

$= A\cos\Phi(t)\, A_r\cos\Phi_r(t) + A\sin\Phi(t)\, A_r\sin\Phi_r(t)$

$= AAr \cos(\Phi(t)-\Phi_r(t))$

$Q(t) = \mathbf{s} \cdot \mathbf{s_r}$,

$= A\sin\Phi(t)\, A_r\cos\Phi_r(t) - A\cos\Phi(t)\, A_r\sin\Phi_r(t)$

$= AAr \sin(\Phi(t)-\Phi_r(t))$

The phase difference between **s** and the reference wave $\mathbf{s_r}$ can therefore be expressed as

$\text{Arctan}(Q(t)/I(t))$ when $I(t) > 0$

$\text{Arctan}(Q(t)/I(t)) + \pi$ when $I(t) < 0$, $Q(t) > 0$

$\text{Arctan}(Q(t)/I(t)) - \pi$ when $I(t) < 0$, $Q(t) < 0$

Elimination of the Addition

When the bandwidth of s(t) and $s_r(t)$ is significantly smaller than their carrier frequency, I(t) and Q(t) can be obtained by one multiplication followed by low pass filtering, instead of two multiplication and a summation.[9]

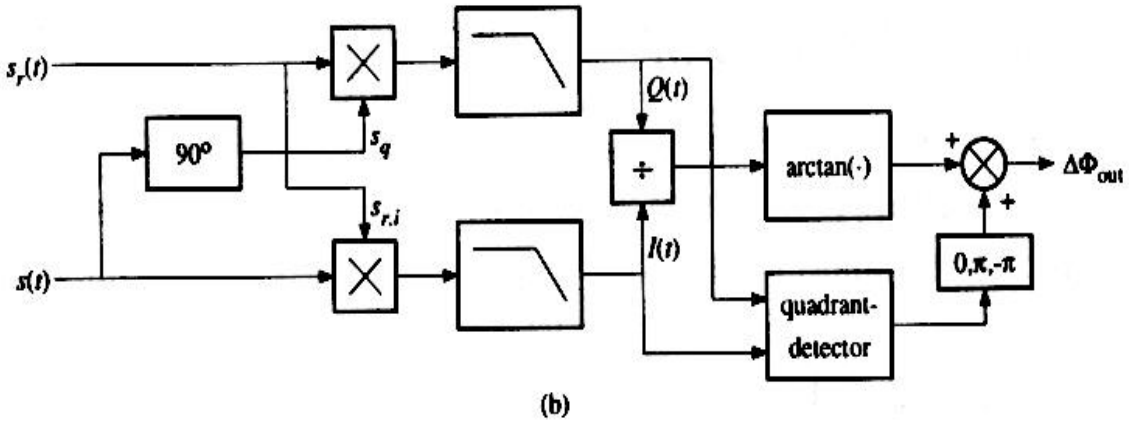


Figure 1.9 schematic of the phase demodulation signal processing [11]

Elimination of the Arctan and Division

In many cases, the phase difference, $\Phi(t)-\Phi_r(t)$, is relatively small, ie considerably smaller than one radian. In that case, I(t) and Q(t) may be approximated as

$I(t) \approx AAr$

$Q(t) \approx AAr[\Phi(t)-\Phi_r(t)]$

Under the same conditions, the arctan-function may be approximated by its first order Taylor term x. Therefore, the entire phase demodulation algorithm can in this case be reduced to the determination of Q(t). The demodulator is depicted below.

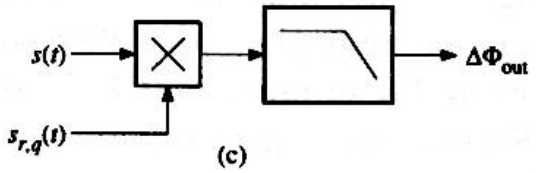


Figure 1.10 Phase demodulation for a small amplitude modulation and signal in quadrature [11]

The amplitude of the signals A and Ar are known, so the phase difference can be computed easily from the output of the demodulator:

$$\Delta\Phi = \frac{\Delta\Phi_{out}}{AA_r}$$

CHAPTER II

THE TRANSDUCER SYSTEM

There are several parameters to consider in choosing of the ultrasonic transducers:

The frequency: The smaller the wavelength of the sound, the bigger the modulation amplitude of the received signal, hence the more sensitive the system as will be shown in chapter III. To vibration of the ground, a smaller wavelength will give a bigger modulation of the carrier and hence, the sensitivity is increased. However in order to get a good sound reflection from an irregular soil surface and get a good penetration through vegetation, a longer wavelength is preferable. The attenuation of sound in air is higher for higher frequencies. This could cause a low signal level at the receiver, especially if the transducers are located relatively far from the ground. At 200kHz, attenuation is a few dB per meter, depending on the temperature and humidity of the air. Since the attenuation increases with the square of the frequency, higher frequencies should be avoided. At 200kHz the wavelength is 1.7mm. The sound beam will reflect correctly for surfaces with irregularities smaller than the wavelength.

The transducer type: At ultrasonic frequencies, piezoelectric, electrostatic and moving coil transducers can be used. All of them can create a loud sound if designed properly. The external shape of a piezoelectric material can be design to

create a focused transducer. This characteristic is desirable to minimize the spot size of the system. Piezoelectric transducers have usually a lower receive sensitivity than electrostatic transducers. However, the better focusing would probably compensate the loss in signal level resulting from this lower sensitivity.

Presentation of the transducers

The transducers used in the acoustic system are electrostatic (capacitance) devices which resonate at 50kHz. They were manufactured by B&K and are normally used for echo ranging in photography. They were chosen for their high transmitting source level and very high receiving sensitivity (receiving sensitivity of –45dB re 1V/Pa at 50kHz for a bias voltage of 150V). At 50kHz, the wavelength of airborne waves is 6.8mm, i.e. four times smaller than the 8GHz electromagnetic waves of the radar sensor. This wavelength difference insures a gain in sensitivity of 12dB relative to the radar system and for the same signal processing resolution. The transducers can give a spot size of a square centimeter when located a couple inches away from the soil.

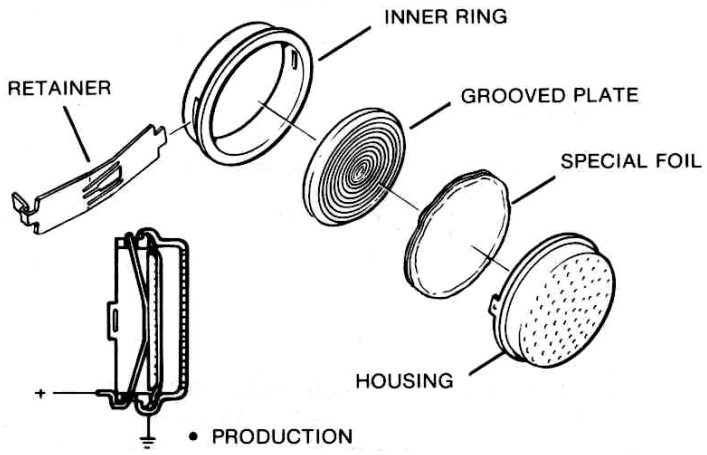


Figure 2.1 Schematic of the transducer assembly [12]

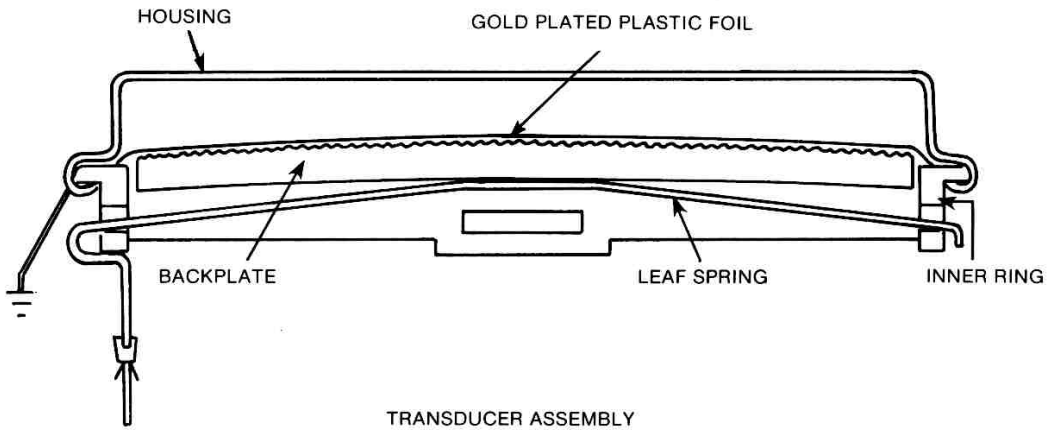


Figure 2.2 Cross section of a transducer [12]

The transducers require a DC bias voltage (up to 200V) and an additional AC voltage is creates the sound (peak amplitude up to the DC bias voltage). A thin foil is the moving surface that transforms electrical energy into acoustic energy, and, conversely, when operated as a receiver transforms the sound wave into electrical energy. The foil is plastic (Kapton) with a conductive coating (gold) on the front side. It is stretched over an aluminum backplate. The backplate and the foil constitute an electrical capacitor. When charged, an electrostatic force is exerted on the foil. This creates a displacement of the foil that is suspended over concentric grooves on the backplate. The AC voltage forces the foil to move at the same frequency and to emit sound waves.

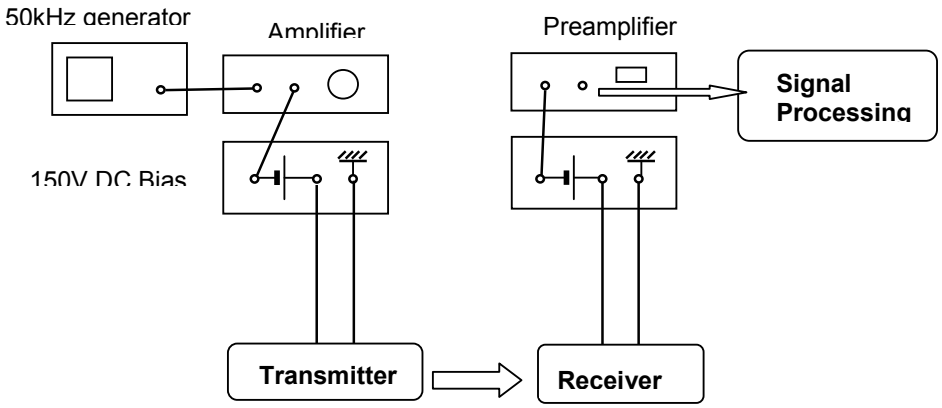


Figure 2.3 Generation of the transducer signal

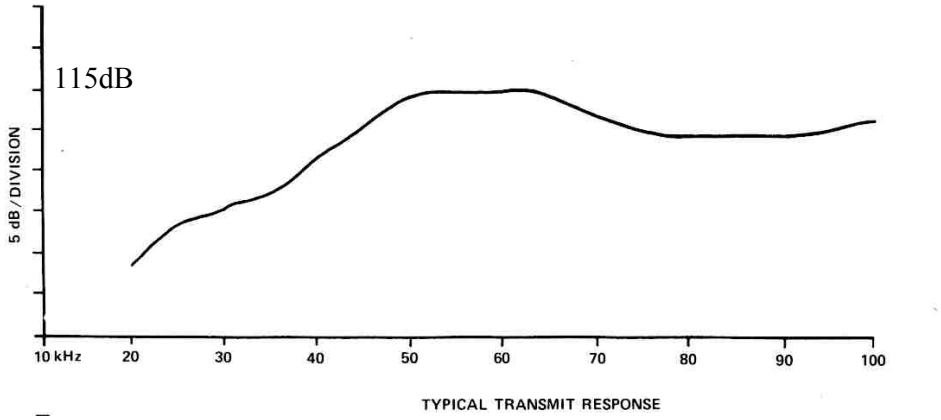


Figure 2.4 .Transmit response of the capacitance transducer at 1meter [12]

The transducers can produce a reasonably loud sound in the frequency range 20kHz – 100kHz with a resonance at 50kHz. Driven at 50kHz at 300Volts peak to peak and 150Volts DC bias, they produce a sound pressure level of approximately 118dB re 20μPa at 1meter on the axis. Measurements of the spectrum of the signal at the receiver between 48kHz and 52Kz did not show any noise above -130dBm when the transmitter was off. The level of the 50kHz signal when the source is on was about 500mV. Hence this signal level was higher than any acoustic noise in the room by over120dB.

Reflection of the ultrasonic waves from soil

In the vibration experiments presented in the following chapters, the transducers were aimed at an underwater sound projector piston. The metallic surface of the piston is covered by a seal for underwater use. Acoustically, (in air) it is a rigid surface. For these results to be applicable in an in situ use, the soil must be a good reflector. In order to determine whether this is a problem, a short experiment was conducted. It consisted in analyzing the received signal for different reflective surfaces.

Three different surfaces were tested: A rigid surface, a sample of sand and, a sample of Georgia's clay soil. The rigid surface used for the test was a wooden panel, 5mm thick, with a smooth and flat surface. It intended to set a reference for signal level and purity. The sand sample was taken from the sand filling the tank of the experimental mine detection set up. It was used in the same conditions as in the tank, wet and compacted. The sample of clay was taken from the campus grounds at a depth of a few centimeters. Both samples were 10mm thick and had a flat top surface after being compacted with a wooden board.

Experiment configuration:

The transmitter was run with a DC bias of 150Volts and an AC voltage of 150Vpeak. The transducers were located 50mm away from the sample, distance at which they focus naturally. The surfaces to be tested were placed at the same height without moving the transducer set-up. For each one of the tests, the received signal level before amplification was measured. The signal to noise ratio of the signal was also observed on a spectrum analyzer. The 49kHz-51kHz frequency range was explored with this apparatus with a resolution bandwidth of 3Hz.

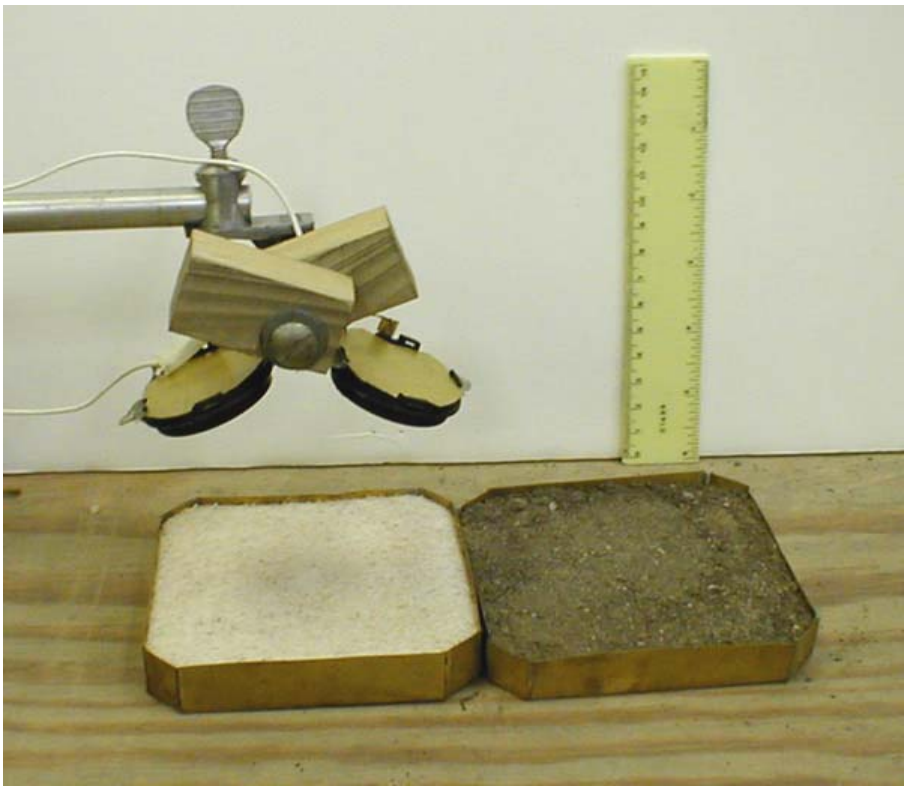


Figure 2.5 Picture of transducer facing a sand sample

Surface	Wooden board	Compacted sand	Clay
Signal level (mV)	600	500	500
Signal/noise (dB)	>100	>100	>100

Table 2.1 received signal level and signal to noise ratio

The results show a very high level for signal reflected by the sand and clay samples; only 20% less than the signal reflected by the flat and rigid wooden board. Sand and clay seem to reflect the 50kHz sound beam in the same manner. For the three tests, the received signal was very pure. Every time the signal to noise ratio exceeded the range that the spectrum analyzer can measure (100dB). Other tests performed with different samples of soil seemed to indicate that the signal level is much more sensitive to the geometrical irregularities of the surface than its nature. This test shows that sound at 50kHz reflects very well on flat surfaces of sand or clay. The signal loss compared to a reflection by ideally flat and rigid surface is of the order of 20%.

Focusing the transducer sound

Focusing the sound consists in generating a converging sound beam focused on the soil surface. A similar design can be adopted for the receive beam directivity, which will enhance the receiver sensitivity to sound coming from the collocated focal spot of the transmitter on the surface of the soil. This configuration would improve two characteristics of the system. First, the signal level at the receiver would be increased. The sound beam converges, so the sound pressure rises to a maximum at the soil's surface. The reflected sound field generates a higher electrical signal when hitting the receiver. The various sources of noise: ultrasonic noise in the air, thermal noise of the electronics, electromagnetic cross-talk etc. are not affected by the focusing. As a result, the signal to noise ratio is increased, which is crucial for the sensitivity of the system. The second characteristic of the system, improved by this technique, is the spot size, which in principle can be reduced considerably. The sound converges towards a particular point and thus the surface hit by the sound can have a much smaller area than the transducer. This would enable the apparatus to detect higher frequency waves in the ground and will give the system a better spatial resolution. Some focusing techniques were studied in order to determine their potential for the signal increase and spot size performance.

- Natural focusing
- spherically focused transducers
- mirrors focusing the sound generated by flat rigid pistons

Natural focusing of the electrostatic transducers

The transducers used in the experimental apparatus can be modeled as rigid flat pistons. At 50kHz, the sound generated by this type of transducer in the far field has a (–3dB) beam of angle 10 degrees with side lobes. In the far field, The further from the source, the lower the sound pressure level. However, closer to the transducer, in the near field, the sound pattern is more complex. The field on the symmetry axis is well known. The following expression for the axial amplitude p can be derived from the Raleigh integral:

$$p = -2\rho c v \sin((\, k(z^2+a^2)^{1/2}-kz)/2)$$

Where k is the wave number of the ultrasound, ρ the density of air, c the speed of sound in the air, a the radius of the transducer, an z the distance on the transducer axis.

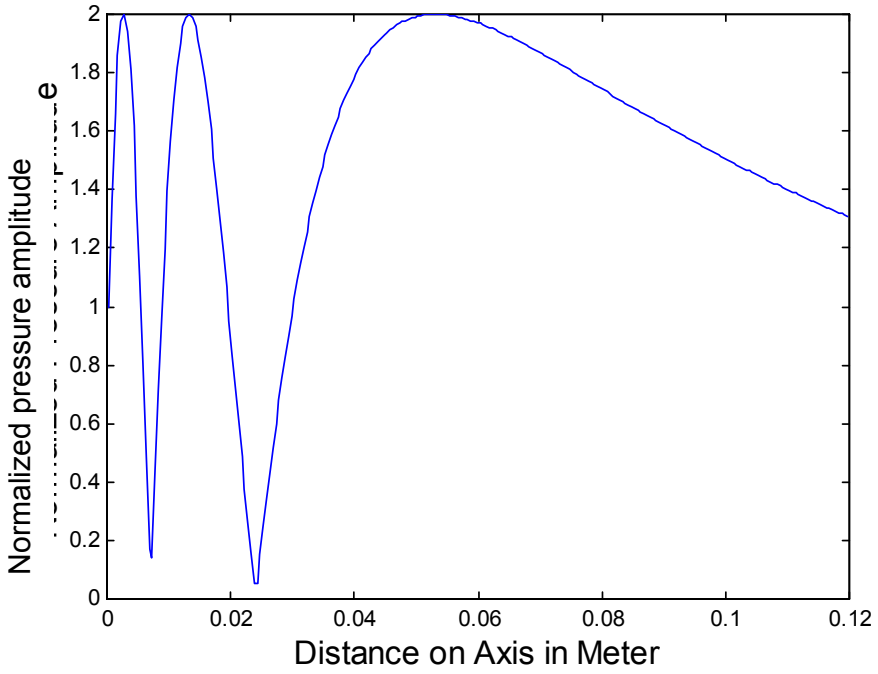


Figure 2.6 Graph of the pressure on axis, in the nearfield

The graph above shows the on axis, normalized, sound pressure level for a rigid piston vibrating at 50kHz. The pressure amplitude reaches equal amplitude maxima before attaining a 1/r decay. The last maximum, located at 53mm, has a sound pressure twice as high as the mean pressure of the nearfield beam. At this point, the peak particle velocity is twice the peak velocity of the piston. The beam narrows at this point in order to conserve energy. This effect is, in effect, a natural focusing of the sound for this configuration. It is difficult to study the sound field experimentally. However it is simple to compute the sound field in the case of a planar piston. A numerical model capable computing the sound field generated by a rigid piston was built to visualize the natural focusing and determine the spot size achieved. As we have the case of a planar vibrating surface, the Raleigh integral was used to calculate the sound pressure. The surface of the piston was discretized in 300 small elementary surfaces ds_n of similar areas. The pressure level generated is then:

$$p(x,y,z) = i\rho cvk\sum_n \frac{e^{ikr_n}}{r_n} ds_n$$

v is the peak velocity of the piston surface and r_n the distance between the elementary surface ds_n and the point of coordinate (x,y,z).The computation was performed using Matlab and the code is attached.

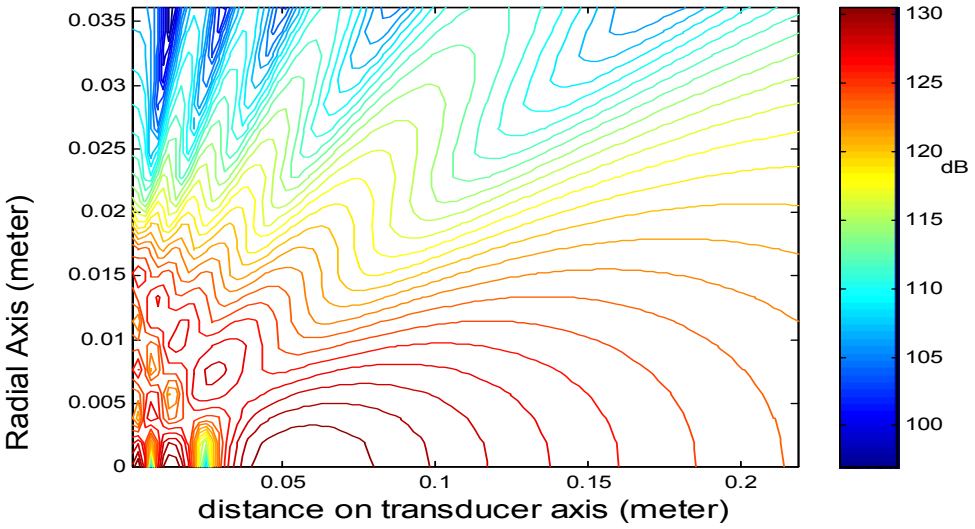


Figure 2.7 sound field generated by piston at 50kHz

As predicted, the beam narrows in the near field. It reaches it smallest dimension at the distance 53mm, corresponding to the last axial maximum. At this point, the 3dB spotwidth is of the order of a centimeter. This dimension is much smaller than the 3.8 cm diameter of the transducer. When placing the transducers at this distance from the soil's surface the spot size would thus be about a centimeter diameter. This is satisfactory for seismic waves of frequency up to 4kHz. This technique does not require any investment. It was successfully implemented on the system tested: a large increase in the signal level was achieved. However the spot size obtained was not checked experimentally.

Note: At distances close to 53mm, interference or acoustic cross talk was observed between the receiver and the transmitter, which led to deterioration of the signal. A higher noise floor appears giving a signal to noise ratio inferior to 100dB.

Placing the transducer at its natural focusing distance improves the signal level and reduces the spot size to a satisfactory value. However it constrains the transducers to be at a small distance from the ground that is not very practical for use in the fields. Therefore, investigation of more elaborated focusing techniques was conducted. This study concentrated on getting further from the ground, and determining the influence of frequency on the focusing.

Spherically focused transducers

In order to form an ideal focus, all the points on the transducer's active surface should vibrate in phase and be at the same distance of the focal point. Such transducers are called spherically focused transducers. The vibrating surface of a spherically focused transducer is a concave portion of a sphere centered on the focal point. Piezoelectric transducers can be designed to exhibit this characteristic. Even if the geometry of the transducer is perfect, the sound focusing is limited by the wavelength of the sound waves. Sound will not focus very well at low frequencies. In order to determine the focusing ability of such transducers, the on-axis pressure was calculated and a numerical model was built to display the sound field. The influences of frequency and focusing distance on the pressure level were investigated with these tools.

The numerical model was based on the same equations as for the piston's sound. In this case, the vibrating surface is not planar. However, the minimum radius of curvature of the surface considered was 50mm, which is much larger than the maximum wavelength, used in the calculations (6.8mm). Under these conditions, the Raleigh integral should give acceptable results.

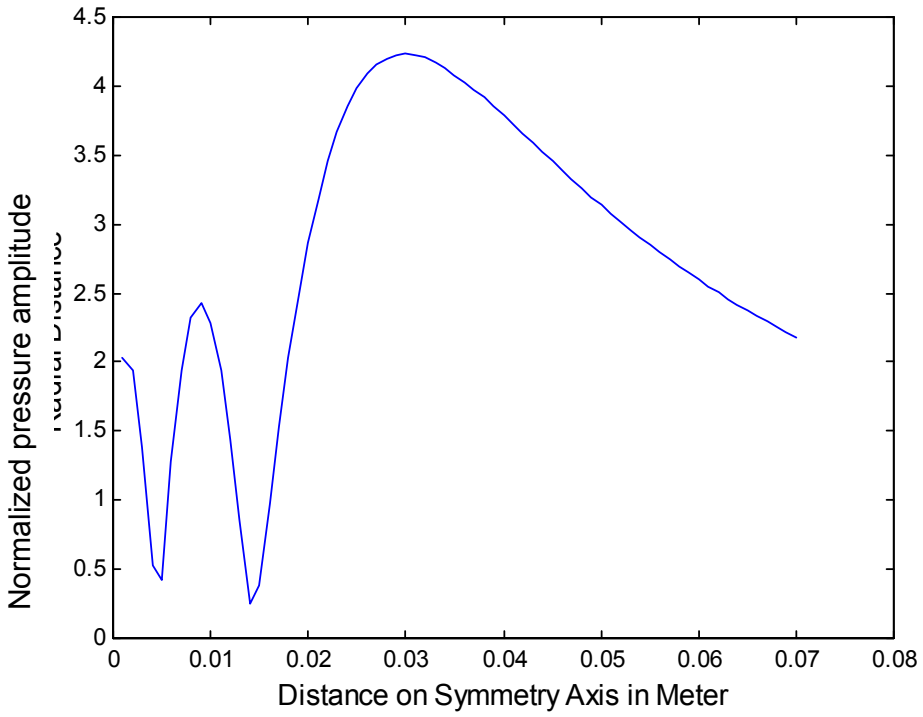


Figure 2.8 on axis pressure generated by a 50kHz spherically focused transducer

The graph above shows the normalized on-axis pressure amplitude for a 50kHz transducer with a diameter of 38mm and focusing at 50mm. The maximum pressure generated is at a shorter distance than the geometrical focal point. With this configuration, the maximum pressure is four times the 'normal pressure' or average pressure at the transducer surface. It is twice as good as the natural focusing.

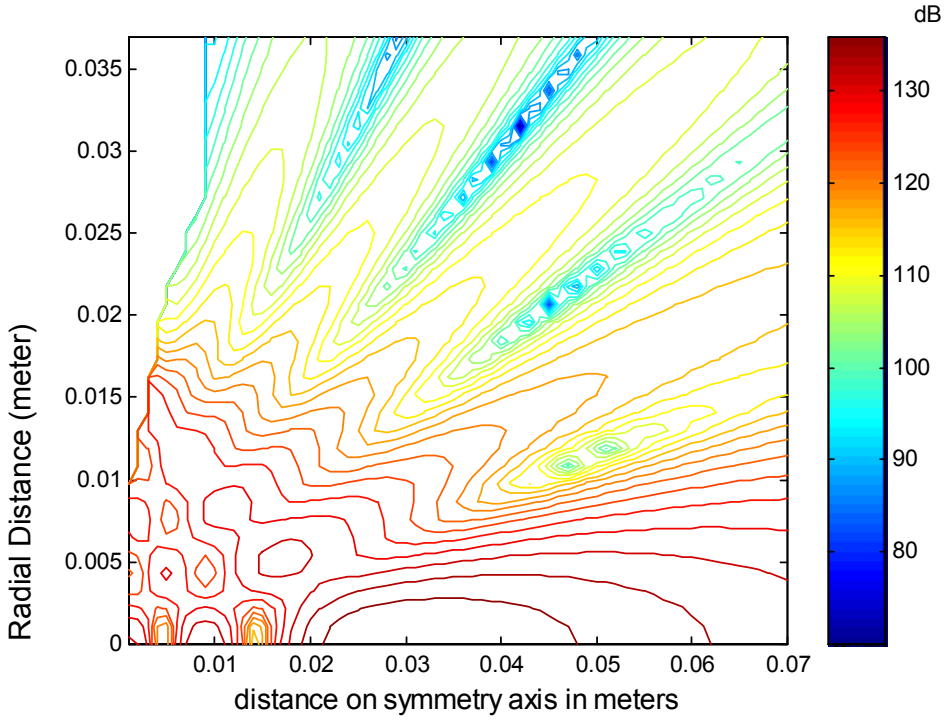


Figure 2.9 pressure field generated by a 50kHz spherically focused transducer

The beamwidth is not much smaller than with natural focusing. The acoustic field features several side lobes dissipating energy away from the focal point. The focusing distance of 30mm is shorter than for the plane transducer. These results indicate that focusing the sound at 50kHz with a transducer of 38mm diameter is not promising. In order to obtain focusing of the sound further from the transducer, a larger aperture (bigger transducer or array of them) and/or higher frequencies have to be used. The following of this section investigate the effects of frequency on the focusing of the sound. The transducers considered have the same diameter of 38mm.

The graphs below show the same transducer running at 200kHz. At this frequency, the normalized pressure maximum is 13. This performance deteriorates as try focusing further from the transducer.

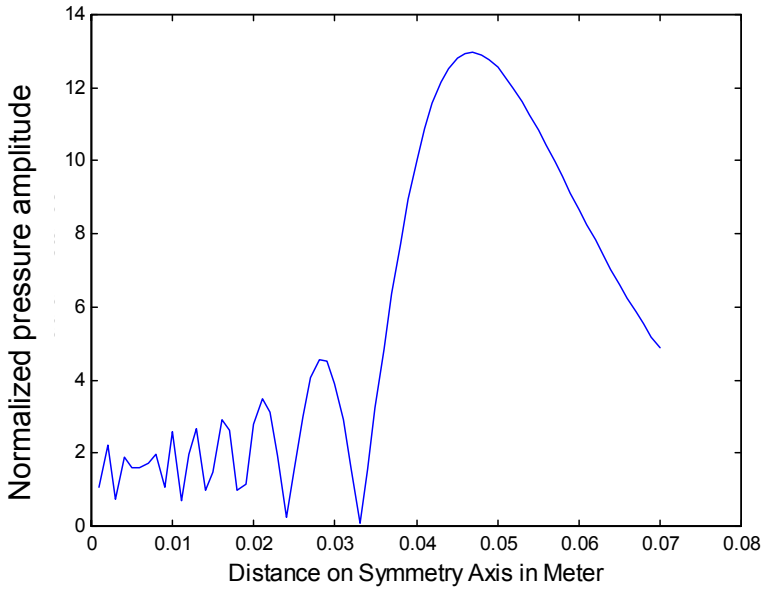


Figure 2.10 On axis pressure generated by a 200kHz spherically focused transducer

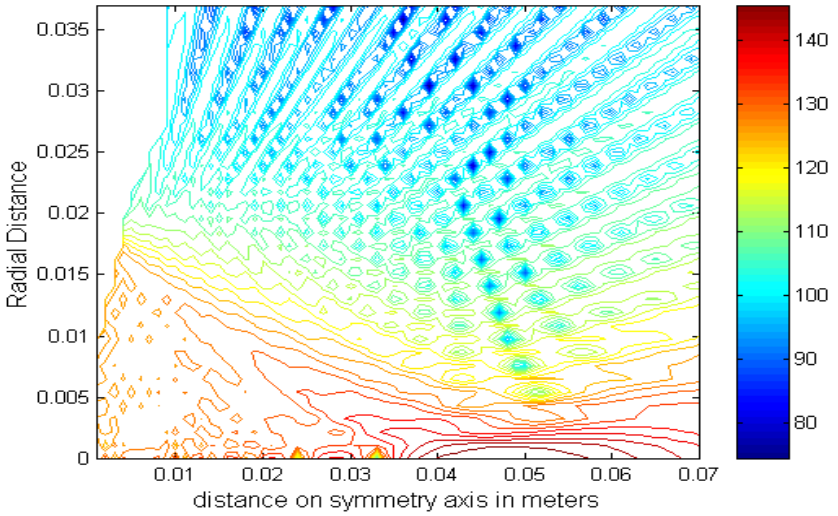


Figure 2.11 Pressure field generated by a 200kHz spherically focused transducer

The previous graph indicates as well that the focusing is a lot better at 200kHz. The sound beam concentrates very well. The side lobes are very weak and dissipate very little energy. Higher Pressure level are reached at the focal point, which is closer to the geometrical focal point. In order to quantify the effect of frequency and focal distance on the focusing, the same code was modified to compute the maximum pressure generated versus the true focusing distance. Geometrical focusing distances where considered from 50 to 200mm with a 5mm increment. For each focusing distance, the program calculated the on-axis pressure determined the pressure maximum and its location and added them to the graph. The results for three different frequencies (100,150, and 200kHz) are displayed below

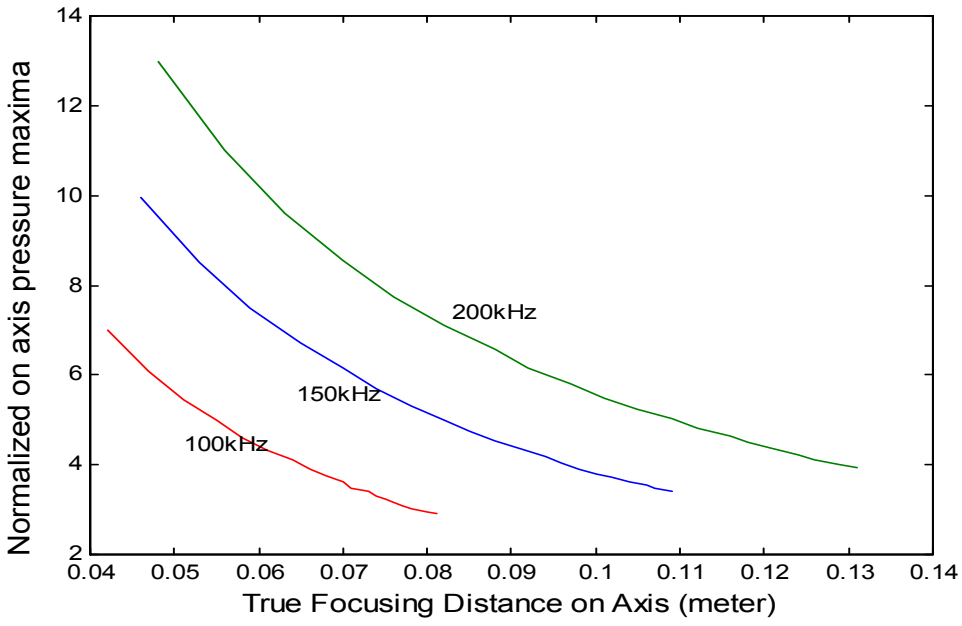


Figure 2.12 Normalized pressure max vs true focusing distance at 50, 100 and 200 kHz

It is clearly from the graph that the focusing is dependent on the frequency and the focusing distance. The curves exhibit a 1/r dependence on the focusing distance. For a given focusing distance (in the frequency range explored), the normalized pressure maximum is roughly proportional to the frequency. It seems that to achieve a focusing 10cm or further away from the soil surface, it would be necessary to use a carrier frequency of at least 100kHz. This would permit a focusing factor of four. Such a configuration produces a small spot size and a high signal level. Such an elaborated focusing technique produces a high signal level and a small spot size. However, an increase of the running frequency of a regular transducer will produce an increase of the natural focusing distance and spherical focusing should only be investigated if obtaining a high signal level is difficult.

Conclusion

The calculation of the on axis sound of the transducers allowed determination their natural focusing. Locating the transducer at 55mm from the reflective surface uses this natural focusing and leads to an increase of the signal level and a reduction of the spot size. The sound pressure level generated at this point is twice as high as the average pressure within the beam. The beamwidth of the sound beam at this distance is of the order of a centimeter.

A numerical model based on the Raleigh integral was built. It calculates the sound field generated by spherically focused transducers. The results of the computation indicates that sound does not focus well at 50kHz. A lot of energy is dispersed and it is not possible to focus far away from the transducers. On the other hand at higher frequencies the results are promising. At 200kHz sound can be focused at 130mm with a signal level gain of four. In the field, the soil surface will certainly have irregularities, obstacles, surface debris and vegetation. Under these conditions the transducer will have to be located at least a few inches from the soil surface. Sound waves of frequencies of 100kHz or higher would focus a lot better than 50kHz sound. The gain in signal level due to the focusing should not be significantly altered by the attenuation in the air. For such small distances at 200kHz, attenuation should be less than one dB.

CHAPTER III

PHASE DEMODULATION:

As discussed in Chapter I, the electrical signal generated by the receiver contains the displacement of the reflecting surface in the form of a phase modulation of the carrier signal. Computing the soil position from this signal thus requires a phase demodulation process. The well known demodulation processes used for communication, are not appropriate since the phase modulation of this signal has a very small amplitude and is therefore difficult to detect.

For simplification of the notation, t denotes the time when wave reaches the receiver. Therefore, the wave is reflected by the soil at $t-(d+\delta)/c$ and was emitted by the transmitter at $t-2(d+\delta)/c$. Where d is half of the average total path length and $\delta(t)$ is the change in acoustic path length due to the displacement of the soil. Hence δ is the change of acoustic length on one leg, at time $t_2=t-d/c$. The carrier and the received signals at time t can be respectively represented by the following expressions:

$$A_1\cos(\omega_c t)$$

$$A_2\cos(\omega_c t - 2k(d+\delta(t_2))$$

Where ω_c denotes the carrier frequency and k the corresponding wave number.

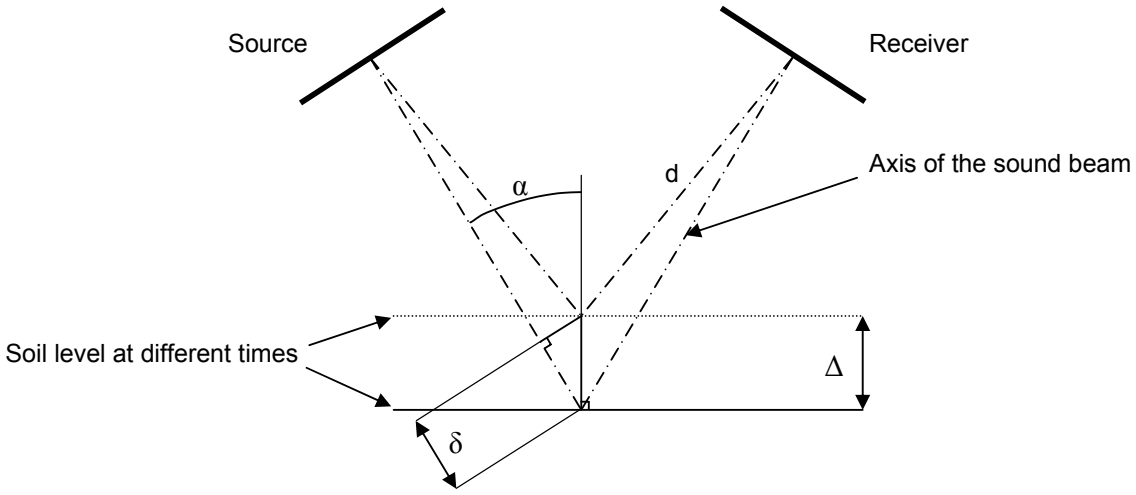


Figure 3.1 influence of the transducer axis angle on sensitivity

The change in length of the acoustic path, $2\delta(t_2)$, is proportional to the vertical displacement $\Delta(t)$ of the ground due to the surface wave. The figure 3.1 shows the geometry of the transducer configuration. The angle α between the sound beam axis and the normal to the soil is constant. The relation between the change in length of the acoustic path and the soil displacement is then:

$$\Delta(t_2) = \delta(t_2)/\cos(\alpha)$$

The central operation necessary for a phase demodulation is a simple multiplication between the received signal and the carrier:

$$F1 = A_1\cos(\omega_c t)*A_2\cos(\omega_c t - 2k(d+\delta(t_2))$$

$$F1 = 0.5*A_1A_2[\cos(2k(d+\delta(t_2))) + \cos(2\omega_c t - 2k(d+\delta(t_2)))]$$

The second term corresponds to a high frequency component. It can be filtered by a low pass filter. For the application considered, the vibration amplitude is much smaller than a wavelength. At 50kHz the wavelength in the air is about 7mm and the highest vibration amplitude recorded on the experimental

model was less than a micrometer. The maximum value for $k\delta(t)$ is $9x10^{-4}$. This is very small compared to 1 radian, and a first order Taylor series is a good approximation for the signal. After low pass filtering and this approximation, the product of the carrier and the received signal can be written:

$$LF1 = 0.5*A_1A_2(\underbrace{\cos(2kd)k)}_{DC1} \underbrace{- 2\sin(2kd)k\delta(t_2)}_{AC1})$$

Hence we get a signal composed of a constant term dependent of the transducer location and a term proportional to the displacement of the soil. d is unknown, as a result, this signal by itself cannot provide the displacement of the ground.

A second low frequency signal can be obtain by computing a similar multiplication of a quadrature of the carrier and the received signal :

$$F2= A_1\sin(\omega_c t)*A_2\cos(\omega_c t - 2k(d+\delta(t_2))$$

$$F2 = 0.5*A_1A_2[\sin(2k(d+\delta(t_2))) + \sin(2\omega_c t - 2k(d+\delta(t_2)))]$$

After low pass filtering,

$$LF2 = 0.5*A_1A_2(\underbrace{\sin(2kd)}_{DC2} + \underbrace{2\cos(kd)k\delta(t_2)}_{AC2})$$

The four terms AC1, AC2, DC1 and DC2 can be used to compute the displacement of the ground:

$$DC1^2+DC2^2= 0.25A_1^2A_2^2(\cos(2kd)^2+\sin(2kd)^2) = 0.25\ A_1^2A_2^2$$

$$AC1xDC2 = -0.5kA_1^2A_2^2\sin(2kd)^2x\delta(t_2)$$

$$AC2xDC1 = 0.5kA_1^2A_2^2\cos(2kd)^2x\delta(t_2)$$

And

$$AC2xDC1 - AC1xDC2 = 0.5kA_1^2A_2^2\delta(t_2)$$

There fore,the displacement of the ground can be computed by the following operation:

$$\Delta(t_2) = \frac{AC2 \times DC1 - AC1 \times DC2}{(DC1^2 + DC2^2)2k\cos(\alpha)}$$

DC1 designates the constant component of the signal, i.e. its average over time. AC1 designates the other component of the signal.

$$LF1= DC(LF1) + HP(LF1)$$

Resolution of the demodulation:

When corresponding to the minimum amplitude harmonic signal detectable of frequency ω_L, $\delta_m(t)$ can be noted:

$$\delta(t) = x_0\cos(\omega_L t)$$

The signal on the receiver is then

$$A_2\cos(\omega_c t - 2kd - 2kx_0\cos(\omega_L t))$$

The spectrum of this signal shows the carrier frequency ω_c and peaks at $\omega_c+\omega_L$ and $\omega_c-\omega_L$. The side lobe amplitude is down to the carrier by factor of kx_0.

For a 1-nanometer vibration amplitude and a carrier frequency of 50kHz, this factor corresponds to the sensitivity wanted for the demodulation device.

For a 1-nanometer displacement, $x_0=10^{-9}\cos(\alpha)$ and

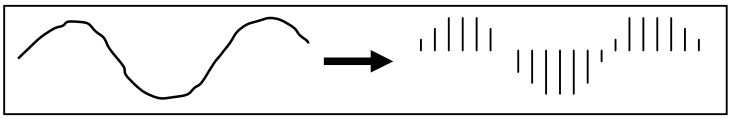
$$kx_0 = 9.24E\text{-}7$$

$$20*\text{Log10}(kx_0) = -120 \text{ dB}$$

Computation of the signal processing

The operations necessary for a phase demodulation are: quadrature, two multiplications, low pass filtering, and additions and multiplications of low frequency signals. These operations can be performed by different means. For the radar system, a mixer is used to perform the multiplication of an 8GHz signals. It is a passive device that multiplies two signals. Some of its advantages are a very low noise level, and very high frequency operation. At low-ultrasonic frequencies another alternative is to use an active multiplier. This device exhibits a better linearity than a mixer and the advantages of an active chip: low output impedance, high output level, and high input impedances. However, four quadrant multiplier chips have an internal noise that is dependant on the quality of the manufacturing and the design.

These operations can also be performed digitally. Many communication systems, particularly demodulation devices, use digital technology. It provides a very versatile tool that can perform many different operations with the same system. The performance of digital systems is improving rapidly with computer technology. More over, a computer has to be part of the total system anyway for data acquisition and storage.



As with any other demodulation technique, the problem is that the modulation amplitude is very small. Independent of the quality of the program performing operations on a signal, the performance of a digital demodulation system is limited by the data acquisition card (or component). The acquisition card converts an analog signal into a digital signal. It performs two operations that add noise to the input signal: sampling and coding. First it samples the signal at regular time intervals. This operation cannot be performed without some error. The time intervals vary slightly and the value of the signal is not taken at perfectly regular pace. This is called the clock jitter.

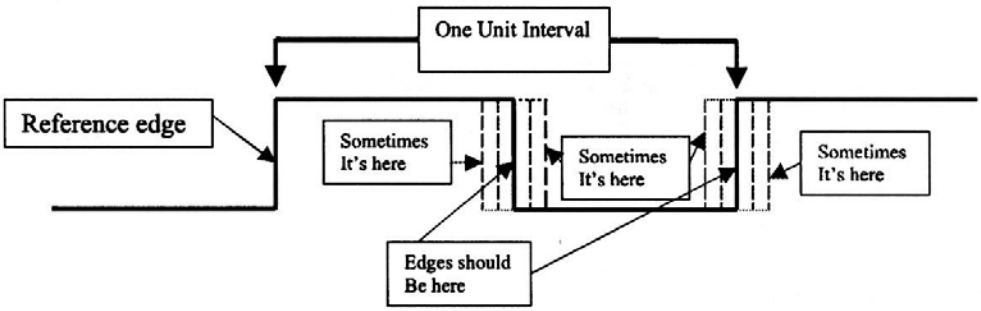


Figure 3.2 Square wave with jittered edges [14]

This jitter generates noise that has a spectral content generally not uniform over all the frequencies. The noise floor corresponding to this phase noise is usually small compared to the signal and it can be neglected for most applications. However, when dealing with large dynamic range, it has to be

considered. The noise floor level has to be lower than the smallest signal level to be acquired by the card.

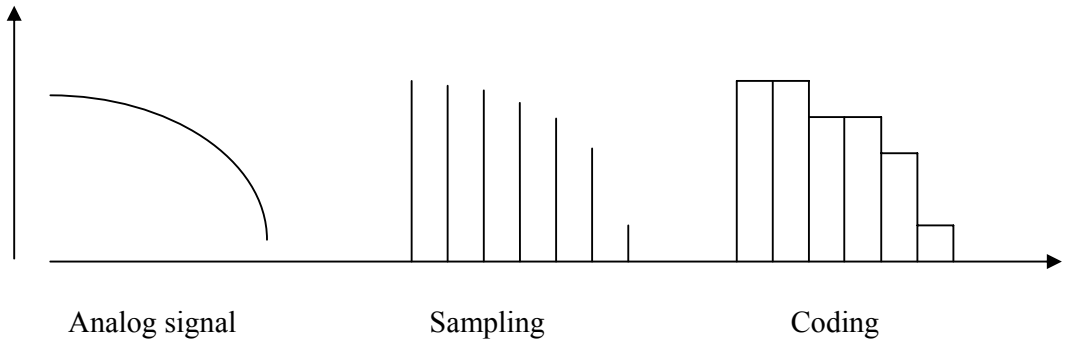


Figure 3.3 Schematic of bit coding

After sampling, the analog to digital converter card codes the amplitude of each sample over a limited number of values. An approximation is then made: the nearest value is taken. The user specifies a voltage range R for the input level, the resolution of the coding is then $R/2^n$ where n is the number of bits used to code the sample. This stage is the n-bit coding. Common A/D cards use 12 or 16 bits coding which correspond respectively to a maximum error of 0.02% and 0.0015%. The rounding process induces an additional error. In the same way as the clock jitter, this error generates a noise that has its own spectral component. Again, noise level due to the coding has to be lower than the smallest relevant signal level.

Test of an Analog to Digital card

To assess the capabilities of an A/D card in terms of noise floor level, a test was performed on a recently acquired card. This card was designed and manufactured by National Instrument and features: a sampling frequency of up to 1MHz, a 20MHz clock, 8-channel acquisition, and 12-bit coding.

The noise floor level was evaluated by computing a Fourier transform of a signal acquired on the card. A 50kHz pure tone was generated externally with a function generator and input to the card. The A/D card digitized the signal over a period of 1s, at a sampling frequency of 250kHz. This data was imported to Matlab to compute the Fourier transform. These settings give the spectral content of the discrete signal up to 125kHz, and with a resolution of 1Hz. The results are displayed below

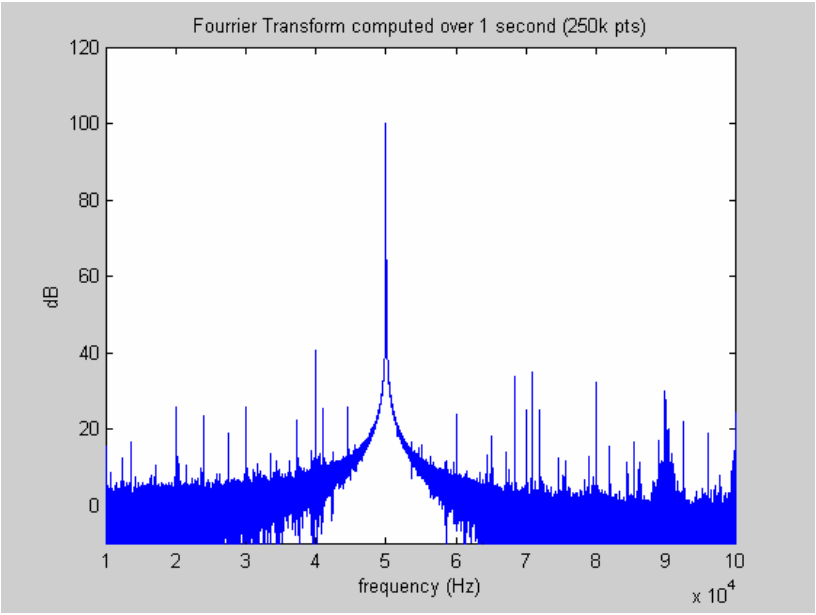


Figure 3.4 Fourier transform of the 50kHz pure tone acquired

This graph shows:

- A sharp peak corresponding to the pure input tone
- A broadband noise level about 95dB down to the 50kHz tone level
- Other noise peaks with amplitude level up to 30 dB above the broadband noise floor

When input into a spectrum analyzer with a resolution of 3Hz, the signal generated by the function generator shows a signal to noise ratio over 100dB. Hence, we have to conclude that the origin of noise shown by the Fourier transform is the card and the digitizing process. The desired sensitivity corresponds to a noise floor 120 dB below the carrier. A very generous signal to noise ratio for this card would be 95dB. This is not satisfactory. This particular card does not have a performance suitable for the surface wave detection application.

In order to determine the characteristics of an A/D card suitable for this application, a further investigation was done to determine appropriate characteristics for this device. Two programs modeling the jitter and bit coding noise sources discussed previously were designed.

Bit coding model:

The noise generated by the bit coding operation was evaluated by a Fourier transform of a digital signal. An ideal digital signal was created over a period of one second at a sampling frequency of 250kHz and machine precision (32 bit coding). It was then coded in 12 bits and 16 bits precision. The Fourier transform of these 3 data vectors shows the noise floor resulting from the bit coding. There was no clock Jitter modeled, hence the only noise results from the 32 bit coding for Matlab real numbers. This has very small effect compared to the errors modeled by the program.

For this test a 50kHz signal modulated at 200Hz with an amplitude B of 10^{-6} radian was used. This modulation corresponds to a signal 15dB above the resolution desired for the detection device.

$$S = \cos(\omega t + B\sin(400\pi t))$$

The following expression was used to model the bit coding of a signal

$$S_{nBits} = \frac{\text{round}(2^{n-1} S)}{2^{n-1}}$$

.

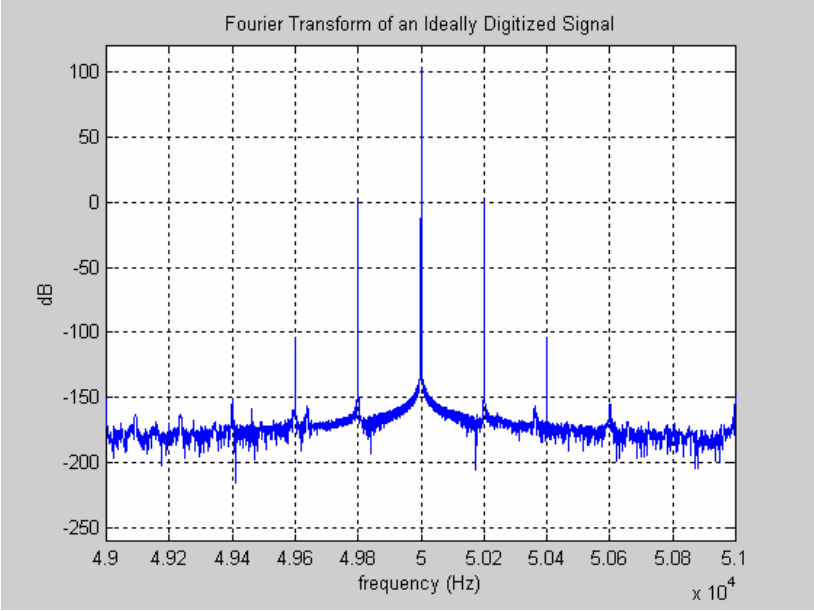


Figure 3.5 Fourier transform of the digitized signal coded on 32 bits

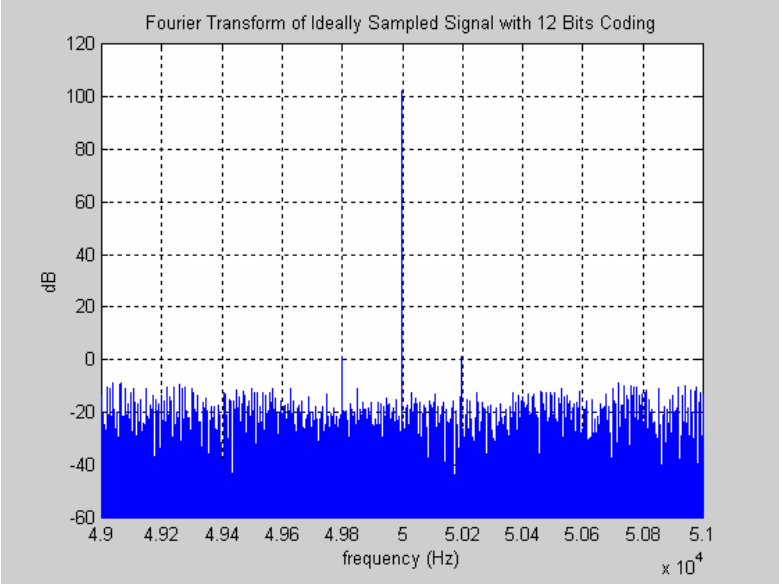


Figure 3.6 Fourier transform of an ideally sampled signal coded on 12bits

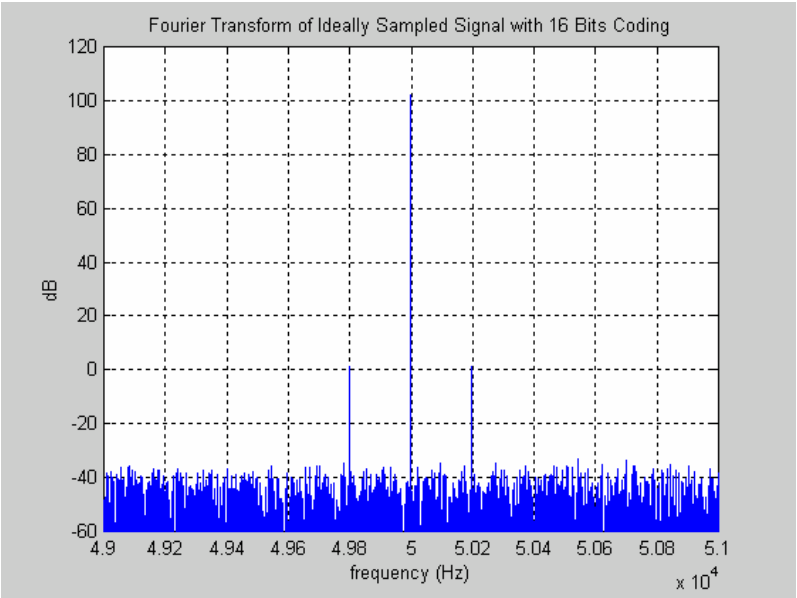


Figure 3.7 Fourier transform of a ideally sampled signal coded on 16bits

The 3 graphs above show the noise floor level in the frequency range 49kHz-51kHz, which is the relevant range for modulations up to 1kHz. The noise floor rises as the bit coding precision decreases. The 12 and 16 bits coded signal have respective signal to noise ratio of 115dB and 140 dB. This performance is satisfactory for the ground wave detection device: the side lobes appear clearly. Hence the 12 bit coding does not seem responsible for the broadband noise that appears on the A/D card test.

Clock Jitter noise

In the same manner as for the bit noise, the clock jitter noise was modeled and its effect on the broadband noise analyzed. Little information was found on the clock jitter for the card tested and for acquisition cards in general. A model was used to find the maximum clock jitter compatible with a good performance for the device.

A very simple model was used for the clock jitter. Each sampling time t_i was given an error Δt_i, normally distributed and centered on the exact time. The variance of the normal distribution was chosen to be one third of the peak jitter amplitude A_{jitter}. This figure insures that 99.75% of random time errors were in the interval [- $A_{jitter;}$ A_{jitter}].To each sampling time t_i, was added random error Δt_i:

$$\Delta t_i = A_{jitter} N(0,1)/3$$

Where N(0,1) a normally distributed random number with variance 1.

7

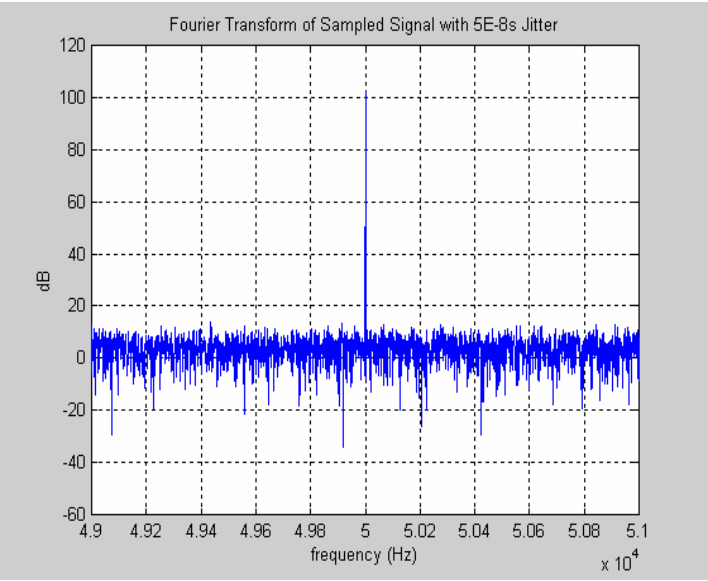


Figure 3.8 Fourier transform of a signal sampled with 5.10^{-8}s jitter

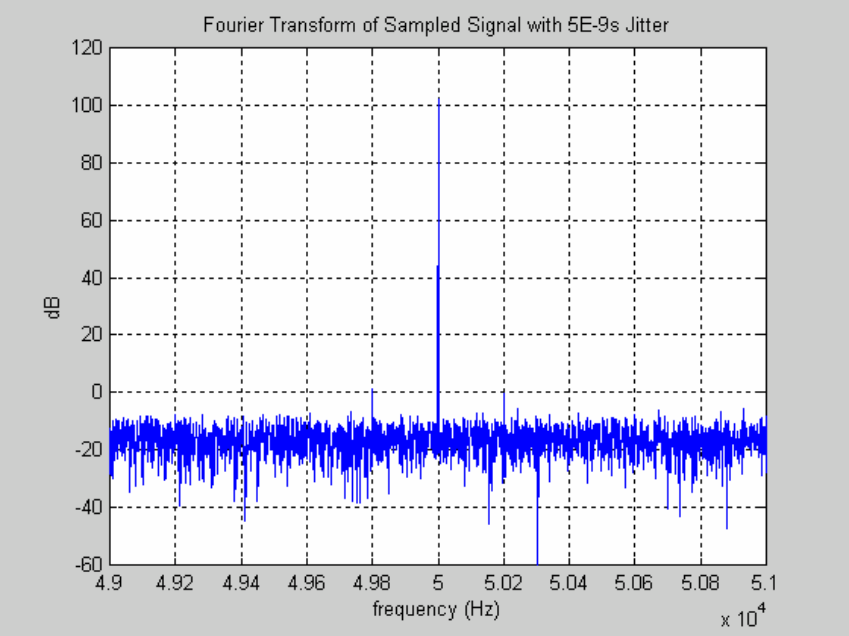


Figure 3.9 Fourier transform of a signal sampled with 5.10^{-9}s jitter

The results show a rise of the broadband noise as the jitter increases. A jitter amplitude of $5x10^{-8}$ second generates a noise level similar to the card tested. This jitter would have to be reduced by a factor of more than ten to give a performance suitable for the device (second graph). Little information was found on a possible model for the clock jitter, so several models were tried with different error distributions and means as well as accumulation of the error. The results varied only slightly. Hence, it is believed that the figures above give a good idea of the jitter amplitude. It seems a reasonable assumption that the sampling jitter is inversely proportional to the clock speed of the acquisition card. In this Case, a clock 50 times faster than the one on the card tested would be satisfactory, i.e. a 1GigaHertz clock.

Conclusion:

The simulations seem to indicate that the factor limiting the performance of the acquisition card tested is the clock jitter, not bit coding. Getting a signal to noise ratio of at least 120dB requires 16bits coding and a Jitter amplitude of 10^{-9} second. Little information was found on the jitter of commercially available A/D cards. However, if it is inversely proportional to the clock speed, a 1GHz clock would be satisfactory for the surface wave detection application.

Commercially available fast A/D products with a 16-bits coding and several acquisition channels are relatively expensive. Since digital demodulation seems to require electronic boards approaching the limits of current digital technology, it was decided to use an analog technique for the experimental system of this thesis. Analog technology is more mature and specialized. Analog chips and devices can perform demodulation operations at a low cost.

ANALOG DEMODULATION

An analog phase demodulation system was implemented. The system involves a quadrature circuit, multiplication devices, active filtering and amplification. This section discusses the processing and the tests of different configurations.

The performance of the process dictated by the dynamic range of the multiplication device. It has to be able to multiply signals at very different levels (The side lobe level is much smaller than the pure tone level). In order to determine the device most appropriate for the multiplication, a simplified demodulation configuration was used. Quadrature of the carrier, computer acquisition and digital processing were not implemented for these tests. Quadrature between the received signal and the carrier was realized manually. The transducers were moved slightly so that the length of the acoustic path created quadrature. This was feasible thanks to the relatively large wavelength of the ultrasonic wave. In this situation, the derivation of the multiplication becomes

$$F1 = A_1\cos(\omega_c t) \times A_2\cos(\omega_c t - \pi/2 - 2k\delta(t_2))$$

$$F1 = 0.5 * A_1A_2[\sin(-2k\delta(t_2)) + \sin(2\omega_c t - 2k\delta(t_2))]$$

After Low pass filtering and a Taylor approximation, the signal can be expressed by:

$$LF1 = A_1A_2\, k\delta(t_2)$$

. As the transducers are immobile, the amplitudes of the signal $A_1$and A_2 are constant. Hence the low frequency signal is directly proportional to change in length of the acoustic path δ(t). This configuration enables us to easily investigate individual analog system components and remove possible noise sources and implementation problems related to the quadrature process, the computer acquisition and the digital processing. This configuration was used to test both mixers and a multipliers.

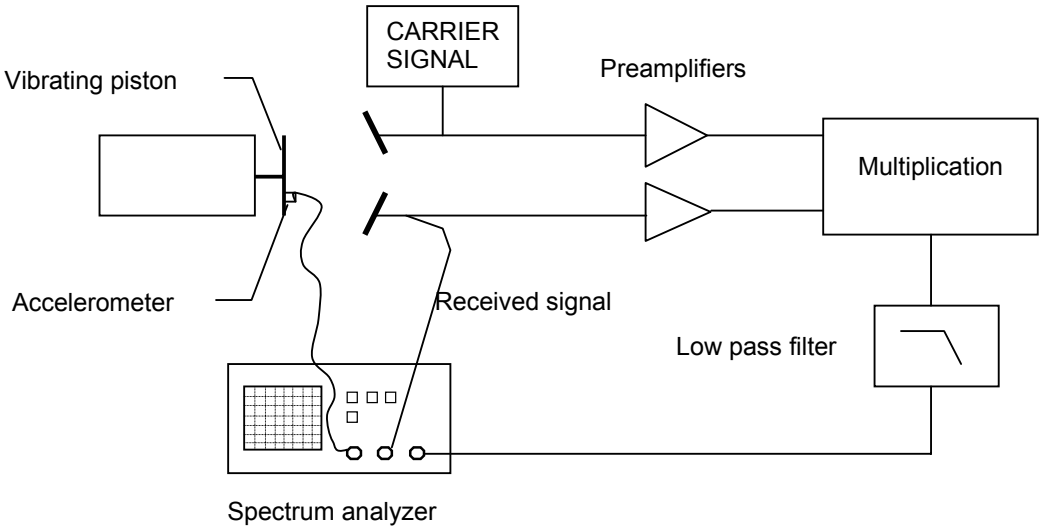


Figure 3.10 Measurement setup

Test of a mixer demodulation configuration

A mixer multiplication was implemented in the system and tests were performed in order to determine its performance. The mixer used was manufactured by Mini-circuit and referenced ZAD 3-H. It features a high input signal level and low noise characteristics.

A sound projector designed for underwater sound generation provided the vibrating surface. It is an electromagnetic device actuating a flat circular piston of 100-millimeter diameter. A rubber membrane, for sealing purposes, covers the piston. This surface is effectively rigid for a good sound reflection. It is larger than the spot size achieved by the ultrasonic transducer, and vibrates uniformly.

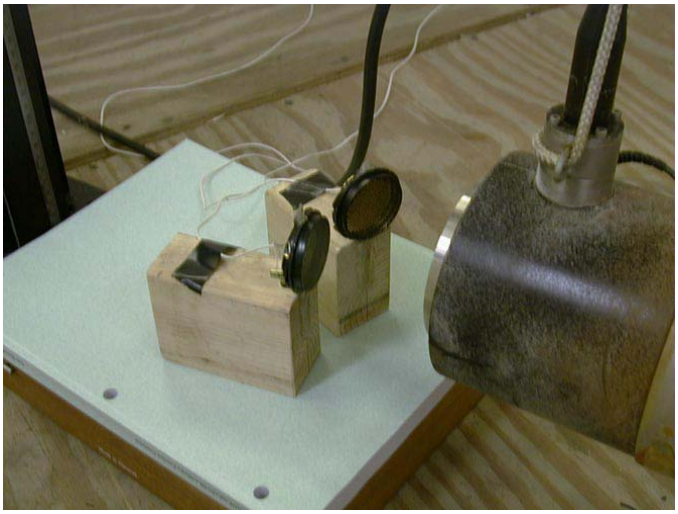


Figure 3.11 Picture of the transducers facing the sound projector

The vibration amplitude was controlled with an audio amplifier according to measurements from an accelerometer. The audio amplifier amplified the signal produced by a Wavetek signal generator and powered the sound projector. The intensity of the vibration could be set by either the signal generator and the

amplifier. The vibration amplitude was measured by an accelerometer attached to the center of the piston surface with bee wax. The accelerometer, of sensitivity 3.42mV/g, was used with a charge amplifier and a spectrum analyzer in order to determine the peak acceleration of the surface. From the voltage level U given in dBm by the spectrum analyzer, the peak voltage amplitude is:

$$V_{peak} = \sqrt{2\pi_{ref} R_{ref}}\, 10^{\frac{U}{20}}$$

Where $R_{ref} = 50\Omega$ is the reference impedance and $\pi_{ref} = 1\text{mW}$ is the reference power. The peak acceleration in ms^{-2} is then

$$\gamma = \frac{9.8 V_{peak}}{3.42 \times 10^{-3}}$$

As we consider one sinusoidal frequency, the peak amplitude of the vibration is:

$$z_{\max} = \frac{\gamma}{\omega^2}$$

Analysis of the spectrum of the modulated signal can provide a quality check for the vibration amplitude measurement. The side lobe induced by the vibration are down from the carrier amplitude by a factor of kx_0, where x_0 is the peak amplitude of the vibration. Therefore the vibration amplitude calculated from the modulated signal is:

$$x_0 = \frac{1}{k\cos(\alpha)} A \log \frac{V_{lobe} - V_{carrier}}{20}$$

During the tests the transducers were located at their natural focusing distance from the piston. They were aiming at the center of the piston from a distance of 60mm at a total angle of 54 degrees (α = 27degrees). A Krohn-Hite source repeater amplified the signal from the Analogic function generator to 150V peak. A Krohn-Hite active low pass filter filtered the output of the mixer. It performed a 3-poles Bessel filtering with a –3dB cutoff frequency at 3000Hz. The response of the system was studied for a fixed vibration amplitude between 200Hz and 1200Hz. In order to get the maximum output level, the mixer input levels were set at the maximum specified by the manufacturer (see below).

Maximum input on LO: 17dBm (2V)

Maximum input on REF: 7dBm (700mV)

For each frequency, the vibration was set to 40 nanometers according to the accelerometer measurement. The side lobe level was measured for the received signal. The mixer output and the noise floor level were measured after filtering. All the measurements were performed on the spectrum analyzer with a resolution bandwidth of 3Hz.

From the measurements, the sensitivity of the analog demodulation was computed. This sensitivity corresponds to the smallest modulation detectable, it is expressed by the ratio of the of the smallest side lobe detectable to the carrier level on the modulated signal (in dB). The value displayed in the table was computed from the side lobe level and the signal to noise ratio at the output of the mixer:

Demodulation sensitivity = tone level – side lobe level + output Signal/Noise

Freq	Accelerometer measurement (dBm)	Displacement From Acc (meter)	Side lobe/peak (dB)	Displacement From side lobe (meter)	Output level (dBm)	Noise Floor Level	S/N (dB)	sensitivity (dB)
200	-82.3	4.4E-08	-85	3.0E-08	-82.3	-95	12.7	97.7
400	-70.3	4.4E-08	-86.5	2.6E-08	-83.3	-98	14.7	101.2
600	-63.2	4.4E-08	-88.95	1.9E-08	-83.8	-101	17.2	106.15
800	-58.2	4.4E-08	-88.6	2.0E-08	-85.3	-101	15.7	104.3
1200	-51.2	4.4E-08	-88.35	2.1E-08	-87.8	-104	16.2	104.55

Table 3.1 results of the mixer test

Often, the two side lobes levels measured on the spectrum analyzer had different levels by one or two dBs. This phenomenon is not yet explicable. A possible reason is a the amplitude modulation of the signal because of the variation of the length of the acoustic path. Such modulation appears as side lobes in the spectrum of the modulated signal. Unlike for the case of a phase modulation, the side lobes have the same phase. Hence, the superposition of a phase and amplitude modulation can generate such uneven side lobes. The value in the table above is the average of the two side lobe levels. The displacement given by the accelerometer and the side lobe levels match pretty well. Differences could come from the inaccuracy of the accelerometer measurements. For a constant displacement of the piston, as the frequency is increased, the output level decreased a little. Part of this phenomenon is due to the filter, whose cutoff frequency is relatively close to the signals.

Freq	200	400	600	800	1200
Gain (dB)	0	0	-0.2	-0.6	-1.6

Table 3.2 Transfer function of the Krohn-Hite 3 poles Bessel filter for fc=3000Hz

The constant peak displacement was set according to the accelerometer signal. The side lobe levels seem to indicate that in the lower frequency range the actual vibration amplitude was higher. This seems to indicate that the relatively lower sensitivity of the accelerometer at low frequencies is the second factor responsible for an apparent drop of the output in the higher frequency range.
Due to the drop of the noise floor, the sensitivity increases with frequency. The average resolution for the mixer demodulation and over the frequency range 200-1200Hz is **103dB**. It corresponds to a displacement amplitude of about **9 nm** for a 50kHz carrier.

We can also observe a drop of the noise floor level when the frequency increases. This could be due to the spectral impurity of the 50kHz. The side lobe must be extracted from the skirt of the carrier. It was not possible to measure this, since the signal to noise ratios achieved are higher than the measurement capability of the spectrum analyzer, even for frequencies close to the carrier frequency.

Conclusion:

An average sensitivity of 103dB was obtained over the frequency range 200-1200Hz. This is close to the desired performance. A higher noise floor level was observed in the lower frequency range. A problem was encountered with the

mixer device; its low input impedance distorted the input signals. It was thus decided to implement multiplier chips that would give high input impedances, low output impedances and higher signal levels, rather than mixers. These devices have also better linearity. Crystal oscillators (which have a higher spectral purity) were bought in an attempt to reduce the noise floor level in the low frequency range.

Sensitivity tests of multiplier demodulation configurations

Tests were performed with the same experimental setup using multiplier circuits. A multiplier cannot be used on the radar system because the frequency is much too high. However, these active chips can at frequencies higher than 1Mhz, and so can perform operations around 50kHz very well. Due to their active nature, they can provide a high input impedance, higher signal levels, and a good linearity. Several multiplier circuits were tested. Three are presented in this report. The circuit TM3 uses a four-quadrant multiplier designed by Analog Devices. It accepts input levels up to 10Volts peak and an adjustable gain is available at the output. The four-quadrant multiplier AD534 and the chip AD630 manufactured by Analog devices were also tested.

The configuration of the test was identical to the mixer test. However the gain of the preamplifiers were increased so that the input levels are maximal.

Freq	Accelerometer measurement (dBm)	Displacement From Acc (meter)	Side lobe/peak (dB)	Displacement From side lobe (meter)	Output level (dBm)	Noise Floor Level	S/N (dB)	Sensitivity (dB)
200	-82.3	4.4E-08	-86.6	5.1E-08	-57.8	-82	24.2	110.8
400	-70.3	4.4E-08	-87.4	4.6E-08	-58.8	-88.6	29.8	117.2
600	-63.2	4.4E-08	-88.2	4.1E-08	-60	-91	31	119.2
800	-58.2	4.4E-08	-88.5	4.0E-08	-60.3	-92	31.7	120.2
1200	-51.2	4.4E-08	-89	3.8E-08	-62.8	-94	31.2	120.2

Table 3.3 Results for the TM3 configuration

Freq	Accelerometer measurement (dBm)	Displacement From Acc (meter)	Side lobe/peak (dB)	Displacement From side lobe (meter)	Output level (dBm)	Noise Floor Level	S/N (dB)	Sensitivity (dB)
200	-82.3	4.4E-08	-86.5	5.1E-08	-57.3	-83.1	25.8	112.3
400	-70.3	4.4E-08	-88.3	4.1E-08	-57.6	-92	34.4	122.7
600	-63.2	4.4E-08	-87.5	4.5E-08	-58.4	-94	35.6	123.1
800	-58.2	4.4E-08	-87.5	4.5E-08	-58.2	-95	36.8	124.3
1200	-51.2	4.4E-08	-88.1	4.2E-08	-60.4	-94	33.6	121.7

Table 3.4 Results for the AD534 configuration

Freq	Accelerometer measurement (dBm)	Displacement From Acc (meter)	Side lobe/peak (dB)	Displacement From side lobe (meter)	Output level (dBm)	Noise Floor Level	S/N (dB)	Sensitivity (dB)
200	-82.3	4.4E-08	-85	6.1E-08	-57.3	-86	28.7	113.7
400	-70.3	4.4E-08	-86.5	5.1E-08	-57.6	-90	32.4	118.9
600	-63.2	4.4E-08	-87.8	4.4E-08	-58.4	-91.4	33	120.8
800	-58.2	4.4E-08	-86.3	5.2E-08	-58.2	-91	32.8	119.1
1200	-51.2	4.4E-08	-88.0	4.3E-08	-60.4	-94	33.6	121.6

Table 3.5 Results for the AD630 configuration

For the three the multiplier configurations: The results show a great increase in the sensitivity relative to mixer configuration. The objective of 120dB resolution is reached for a large part of the frequency range. The same rise of the noise floor is observed in the low frequency range. The AD534 multiplier seems to give slightly higher output level and the same noise floor level as the TM3 circuit.

The configuration with one multiplication implemented for these tests is not very realistic for a use in the fields. The distance from the transducer to the soil surface would have to be adjusted so that it creates continual quadrature between the receive signal and the carrier. This is not practical, especially when using plane arrays of transducers over an irregular surface. Therefore a signal processing that can run with an arbitrary acoustic delay (2kd) has to be implemented. The schematic next page illustrates a possibility. This flow diagram follows closely the derivation exposed at the beginning of the chapter. The processing uses a quadrature circuit and two multiplier circuits in addition to the previous configuration. The analog part of this system is presented in the following section. The filtering and the digital operations are exposed in the next chapter.

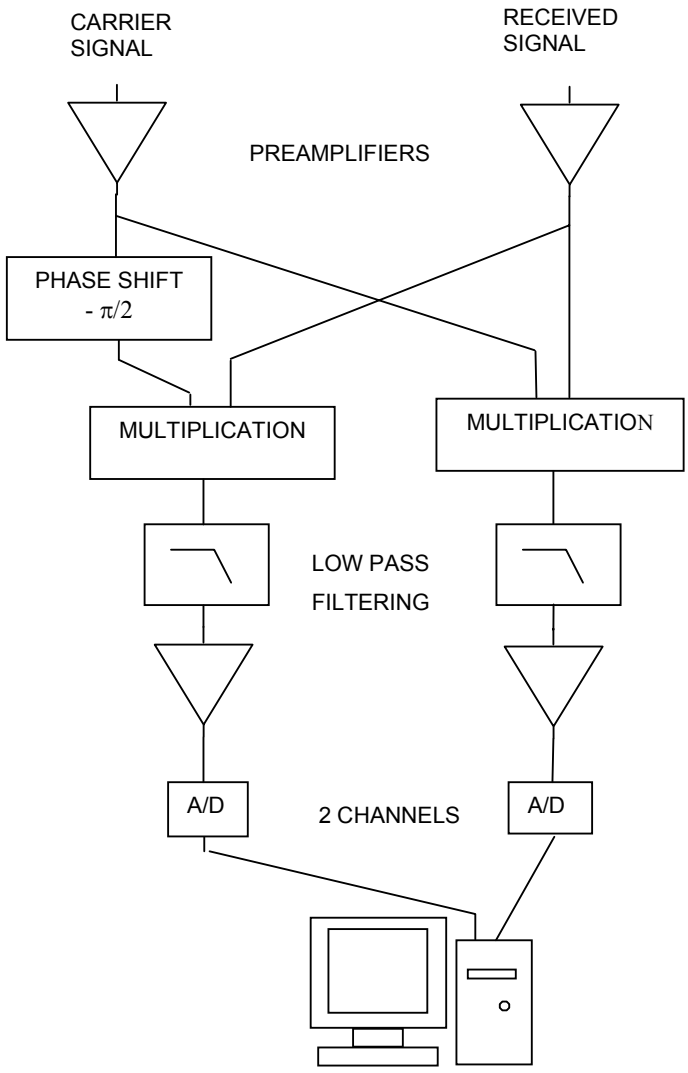
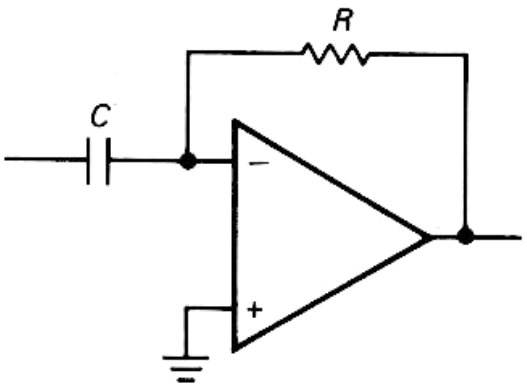


Figure 3.12 Schematic of the signal processing

The quadrature was obtained by realizing a differentiating circuit with an operational amplifier. This differentiation gives a phase shift of π/2 to the 50kHz carrier. The impedances of the feedback resistor and capacitor were chosen to get a gain of 1 at 50kHz. The circuit can process carrier level up to 10Volts peak. The analysis of the output signal on the spectrum analyzer showed a very pure signal. The signal to noise ratio was over 100dB.



R = 570 kΩ C = 5pF

Gain at 50kHz: $G_{50kHz} = 1/RC\omega$

=.9

Figure 3.13 Operational amplifier circuit for quadrature

Tests were performed with quadrature between the carrier and the received signal realized with this circuit. The results showed identical noise and output levels with the results obtained by achieving quadrature by adjusting the acoustic path length. This indicates that the performance of this circuit is suitable for the intended application.

Conclusion

The feasibility of digital phase demodulation was investigated. Such processing would require a very fast acquisition card featuring low clock jitter and at least 16bits digitization. Due to the cost of such a device, an analog solution was investigated. An analog phase demodulation was implemented and tested with several multiplication devices. The tests were performed with actual transducer signals reflected from a vibrating rigid surface. The sensitivity of the different configurations was computed over the frequency range [200,1200Hz] from the modulated signal and the output signal spectra. In every case, the sensitivity was found to be poorer by a few decibels in the lower part of the frequency range. The origin of this is a rise of the noise floor at low frequencies. The mixer configuration performed a resolution from 98dB at 200Hz to 104dB at 800Hz. The multiplier circuits gave better results thanks to their much higher input impedance and output levels. Sensitiviy from 110 to 120dB for the TM3 and sensitivity from 112 to 124dB was observed for the AD534 multipliers. Finally, a quadrature device having a dynamic range over 120dB was built with an operational amplifier.

CHAPTER IV

NOISE AND FILTERING

A crucial part of this thesis work was to get a good sensitivity. Very small signals received by the transducer have to be enhanced by signal processing. Any noise in the system will tend to hide the small desired signal. The noise in the demodulated signal at the very end of the processing can have several origins. This chapter intends to make a list as complete as possible of the noise sources in the system and present the solutions implemented to reduce them. In particular, this chapter will present the filtering techniques chosen to remove the unwanted frequency components of the signals.

Noise measurement:

Throughout this study, noise was considered in the frequency domain. When trying to reduce or remove it, frequency is an important parameter. A spectrum analyzer is a very powerful tool for investigating noise. Its very high input impedance and great sensitivity allow tracking very small signal levels and noise floors. All the noise floor measurements stated in this report correspond to a resolution bandwidth of 3Hz. Though very powerful, this tool is limited by its dynamic range of 100dB. It can 'only' display signals that are 10^5 times smaller than the highest signal level input. The sensitivity wanted for the demodulation device is 20dB higher than this. As a result it is not possible to state that a

specific component of the system is good enough by direct measurement. The spectrum analyzer can only show whether or not this signal to noise ratio exceeds 100dBt. At the end of the process, after multiplication and filtering the sensitivity can be evaluated. However, it is difficult to determine the origin of the noise floor.

Acoustic noise:

When implementing an acoustic method for surface wave detection, one inevitably meets acoustic noise. The capacitance transducers function similarly to a speaker and a microphone. A background noise in the air would be detected by the receiver and thus becomes noise in the system. Spectral measurements were made on the receiver signal. When the transmitter is not running, the noise floor level around 50kHz is below –125dBm. During the operation of the system, the signal level is about 3dBm. Therefore the available dynamic range of the air media is at least 128dBm. Measurements of ultrasonic sound level in the lab showed that it was not significant. The projector's vibrating piston produces sound under 1200Hz that is detected by the receiver. This low frequency component of the received signal perturbs the multiplication operation. Mixers and multipliers do not work ideally and some of the input signal feeds through the chip. Therefore the low frequencies coming from the receiver have to be filtered. This problem is only related to the laboratory and the calibration test. The surface waves in the soil do not radiate sound in the air efficiently. They generate evanescent waves and the sound level decays exponentially as the observation point is moved away from the surface. However a high pass filtering device should still be installed to allow use near a low frequency sound source

(vehicle motor running…). It is easy to implement. The capacitance receiver is not very sensitive in this frequency range. Moreover this frequency range (100Hz-2kHz) is located at least 5 octaves away from the carrier frequency. Therefore a simple one or two pole passive high pass filter with Fc=30kHz can achieve a 30 or 60dB attenuation. On the system, the SRS pre-amp used to amplify the received signal performs this filtering.

Another noise problem related to the transducers was encountered. At some position close to the surface, the transducers showed interference. The signal to noise ratio on the receiver was deteriorated and showed interference maxima and minima at several frequencies very close to each other. Placing the transducers at their natural focusing distance from the vibrating surface solved this problem.

Electromagnetic Cross-talk

The tests where performed in the Space Science and Technology center of Georgia Tech. In the laboratory many power lines and forty-two 34W-fluorescent tubes provide the light. The power line generates an electromagnetic field at 60Hz and its harmonic frequencies (120Hz, 180Hz, 240Hz...). Similarly, the fluorescent light generates electromagnetic noise at frequencies such as 300Hz. In order to isolate the system from this noise, all the equipment used were shielded:

- All the wire connectors used were BNC shielded cables
- A brass back plate connected to the ground was added to the back of the transducers
- The multiplier circuit, the derivator circuit and the filter were placed in metal boxes.

The Power line can also talk through commercially available equipment. This problem was particularly met with the transformer supplying DC voltage to the Op-amp and the multiplier circuits. To solve this problem, the transformers were abandoned and batteries were placed inside the multiplier chassis.

The source:

The 'Analogic' digital function generator was used as a signal source. It can generate a pure tone, with a signal to noise ratio over a 100dB. This performance is hard to determine more accurately because most high performance spectrum analyzers can only discriminate signals with 100dB of amplitude difference. The modulation of the acoustic signal, generated by the vibration, appears in the spectrum as side lobes of the pure tone. These side lobes have frequency $\omega_c + \omega_L$ and $\omega_c - \omega_L$ and amplitude $k_H * u/\cos(\alpha)$, where u is the displacement amplitude of the ground surface. Hence, frequency ranges [ωc - 1200, ωc -100] and [ωc +100 ωc +1200] have to be clear of noise on the acoustic signal and therefore on the source signal. That is, the noise level has to be 120dB down from the 50kHz tone at these frequencies. In other works, the spectral purity of the signal generator is very important. Typically, this is a problem for frequencies close to the tone where the spectral density of the source phase noise is the highest. In particular, it could be the reason for the higher level of the noise floor for low frequencies on the multiplier output.

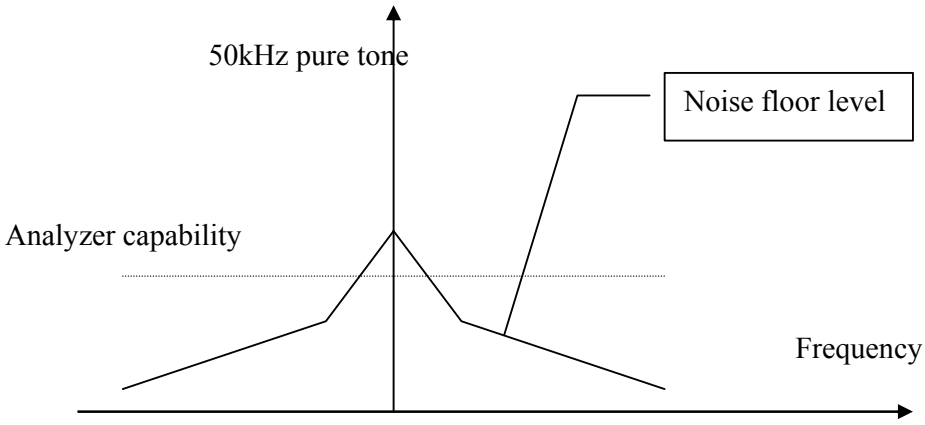


Figure 4.1 Spectral density of a pure tone

A test of the Analogic source with a spectrum analyzer seemed to show that the source signal to noise ratio is voltage level dependent. At some output levels, the source signal to noise ratio could be a little less than 100dB. Thus the signal to noise ratio of this device was no greater than 100dB

In order to increase this to 120dB, several techniques were studied. Voltage sources do not generate perfectly pure tones. Slight instantaneous phase or frequency shifts create a noise floor. In the same manner slight random variation of the amplitude will also create a noisy environment. These noise sources are respectively called phase and amplitude noise.

Phase noise cannot be reduced easily on an existing source. However, we can use a technique that annihilates its effect on a multiplier operation. The random phase shift induces a broadband noise on the output of the multiplier. This is because the phase shifts of the two inputs are not the same. The reason is that the carrier signal comes straight from the source whereas the modulated signal is delayed by the propagation of the acoustic wave from the transmitter to

the receiver. This can be corrected by implementation of a delay line on the carrier signal. The idea is to create a delay on the carrier signal that would equal the acoustic propagation time. Then, the instantaneous phase shifts originated by the source will match at the input of the multiplier. This delay can be performed electronically, or acoustically. An acoustic delay line composed of two additional capacitance transducers was implemented on the system. The transducers were placed facing each other and separated by a distance equaling the acoustic path of the two other transducers. Unfortunately, no positive change was observed on the output of the multiplier. This indicates that the highest noise source is the system is not phase noise from the source.

The broadband noise can also come from **amplitude noise**. An amplitude modulation of frequency ω_A can be modeled as a factor of $1+A\cos(\omega_A)t$ on the carrier signal. This can also be expressed as

$$(1+A\cos(\omega_A)t)*\cos(\omega_C t) = \cos(\omega_C t) + A/2[\cos(\omega_C+\omega_A)t + \cos(\omega_C-\omega_A)t]$$

Hence a modulation of the amplitude at frequency ω_A will appear in the spectrum as side lobes at frequency $\omega_C + \omega_A$ and $\omega_C - \omega_A$. When using such a signal as a carrier, the multiplier will demodulate these side lobes as if they were a modulation originated by the soil vibration. In practice, the amplitude noise on voltage sources is very small and broadband. Therefore it creates a broadband noise in the signal and in the output of the multiplier. Clipping the source signal with a hardware circuit can solve this problem. The sine-wave is transformed into a square-wave which amplitude will not vary. Then the signal can be filtered to get a sine-wave with no amplitude modulation.

Another alternative to these two techniques is to get a cleaner source. High noise performance can be obtained by using crystal oscillator technology. Fabricants usually guaranty a very low phase noise level. Vectron specifies phase noise level 130dB down to the carrier level, 100Hz away from the tone. This very pure tone is more than satisfactory for the seismic wave detection. A voltage-controlled oscillator was installed in the system. A 50kHz, TTL (square wave) output configuration was chosen in the attempt to limit amplitude noise. The square wave can be expressed as a Fourier series:

$$S = \sin(\omega_c t) + \sum_{n=1}^{n=\infty} \frac{1}{2n+1} \sin((2n+1)\omega_c t)$$

The harmonics have to be filtered in order to obtain a sinusoidal wave. A sharp low pass filter can perform this operation. A passive filter was chosen so that the very good noise performance is preserved. The three stages passive filter implemented is represented below.

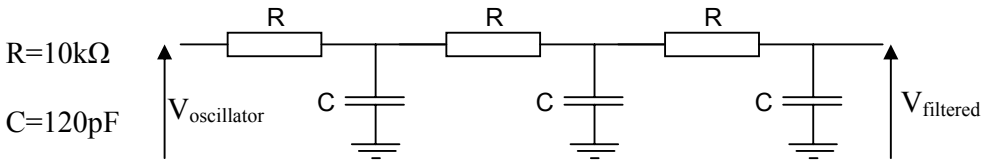


Figure 4.2 three stages passive low pass filter

At the first harmonic is frequency (150kHz) the filter provides an attenuation of 25dB. This is not enough to obtain a pure sinusoidal signal but it enables the transducers to run in good conditions. The remaining harmonics do not disturb the running of the multiplier operation: the first harmonic is at least 25dB lower

than the 50kHz component. As a result, the low frequencies resulting from the multiplication of the carrier first harmonic and the modulated harmonic have a level 50dB lower than the regular demodulated signal. The table below compares the noise floor level after multiplication and filtering for a multiplier configuration with and without crystal oscillator.

Frequency	200	400	600	800	1200
ANALOGIC Source	-82	-89	-91	-92	-94
Crystal oscillator	-86	-90	-91	-92	-94

Table 4.1 Noise floor levels in dBm after filtering of the multiplier output

The results are disappointing. The decrease of the noise floor level is very slight. At low frequencies, the noise floor still exhibits the same pattern with a higher level. Several components of the system can be responsible for this poor improvement. First a TTL output might not correspond to a very effective amplitude clipping. Some amplitude noise may still be present. Secondly, the filtered tone of the crystal oscillator is amplified to 10 volts for the processing and to 150V peak to power the transducers. The Ithaco preamplifier and the Kron-Hite power amplifier used for these purposes may add noise to the signal. Finally the transducers don't have a perfect linear response and may produce noise around the pure tone frequency.

Filtering of the output of the multiplier

Prior to acquisition on a computer, the output signal of the multiplier has to be filtered to remove the high frequency component. Such a filtering is necessary because aliasing would occur otherwise: the high frequency tone level is so high that the dynamic range of a common A/D card would not allow acquiring the low frequency component. Recall the expression of the output signal of the multiplier device:

$$0.5*A_1A_2(\cos(-k(d+\delta(t)) + \cos(2\omega_c t + k(d+\delta(t))$$

The component at $2\omega_c$ rad/s needs to be filtered. An attenuation of 120dB at 100kHz would decrease the level of this component to the level of a 1-nanometer vibration. The signal could then be acquired by a low sampling rate A/D card. Usually, filter manufacturers do not guaranty such high attenuations in the stop band, so a filter was designed and implemented in the laboratory. In order to get a flat amplitude response on the frequency range 100 to 800Hz, the cutoff frequency has to be relatively high. Fc=10 kHz was chosen as a reasonable value. Theoretically filters perform 6dB attenuation per octave and per pole. The cutoff frequency and the 100kHz component are separated by a decade. Then, a 6 poles Chebyshev (with 1dB ripple) filter theoretically gives a 120 dB attenuation at 100kHz [15]. Two filters built with three cascaded operational amplifiers circuits were implemented on the same board along with two AD534 multipliers.

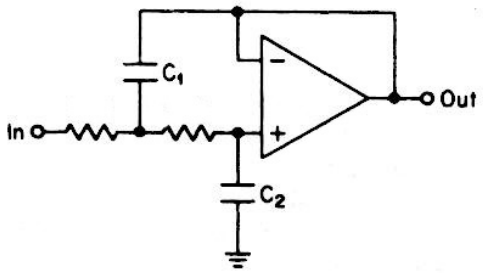


Figure 4.3 One stage of the Chebyshev active low pass filter

The resistors and capacitor were chosen to give a cutoff frequency of 10kHz.

	Resistor (kΩ)	Capacitor C1 (pF)	Capacitor C2 (pF)
Stage 1	330	50	47
Stage 2	220	50	25
Stage 3	39	1500	100

Table 4.2 values of the resistors and capacitors for the filter

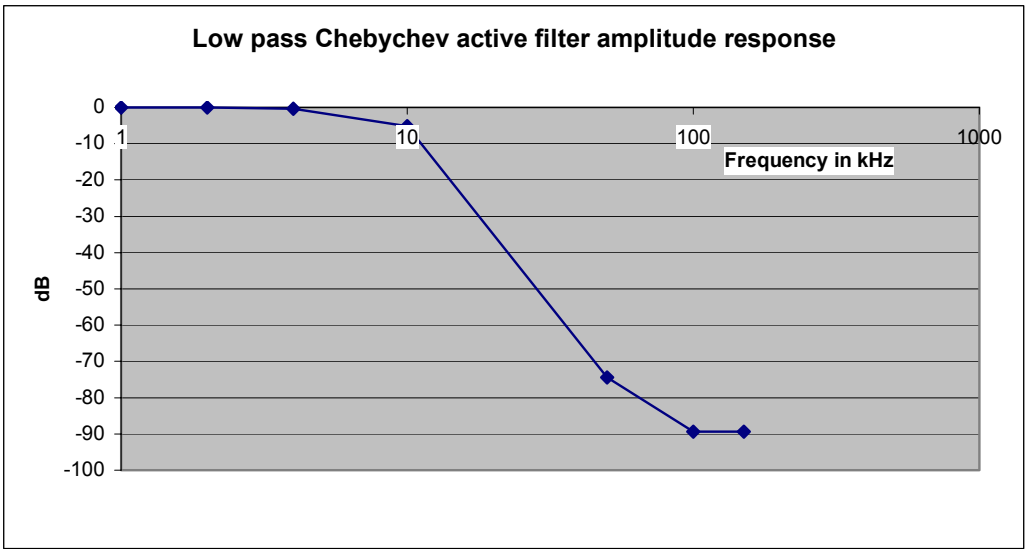


Figure 4.4 Amplitude response of the active filter implemented on board

Measurements indicate that this filter performs a maximum attenuation of 90 dB in the stop band. Spectrum measurement of the filtered signal while the system was running showed remaining 50kHz tones of a level of -60dBm. Having the multipliers and the filters mounted on the same board is probably responsible for this. A cross talk phenomenon can occur through the +/-15V power line. However such a low level does not have much consequence. The level of a nanometer vibration being -90dBm, the 50kHz tones are well in the vertical range of the acquisition card. The advantage of the Chebyshev design is that it provides a very sharp transition from the pass band to the stop band. On the other hand, it has poor phase delay characteristics. The phase delay is not uniform in the range 100Hz-800Hz and has to be corrected digitally. This can be performed after acquisition of the signal.

Digital signal processing

After filtering of the output of the multiplier, several operations remain to be computed to get the displacement of the ground. These operations can be performed digitally, after acquisition on a computer. In particular additional filtering of the high frequencies and the calculation presented in chapter III has to be done. In order to sample a signal component at 100kHz, Niquist rule requires that the signal should be acquired at a frequency over 200kHz. Also, the dynamic range of the acquisition must be adequate for a resolution of the system of 1nanometer.

A VXI system performs acquisition of the signal, filtering and calculations. It consists of a PC board and a waveform analyzer TVS 625 card. The card acquires the two signals simultaneously at sampling frequency 250kHz. The signals are coded on 8bits and the card can record up to 15000 samples in a row on each channel. This allows a recording time of 6.10^{-2} second. During that time a Raleigh wave would travel a distance of 4.8 meters.

The waveform analyzer card is controlled through a Labview interface. The program developed performs an acquisition of the signals, filtering, and the calculations required to compute the soil surface displacement. The filters installed perform a 5 poles Bessel filtering with cutoff frequency of 5kHz. Correction of the phase shifts originated by the analog filtering was not implemented. The Labview code is attached.

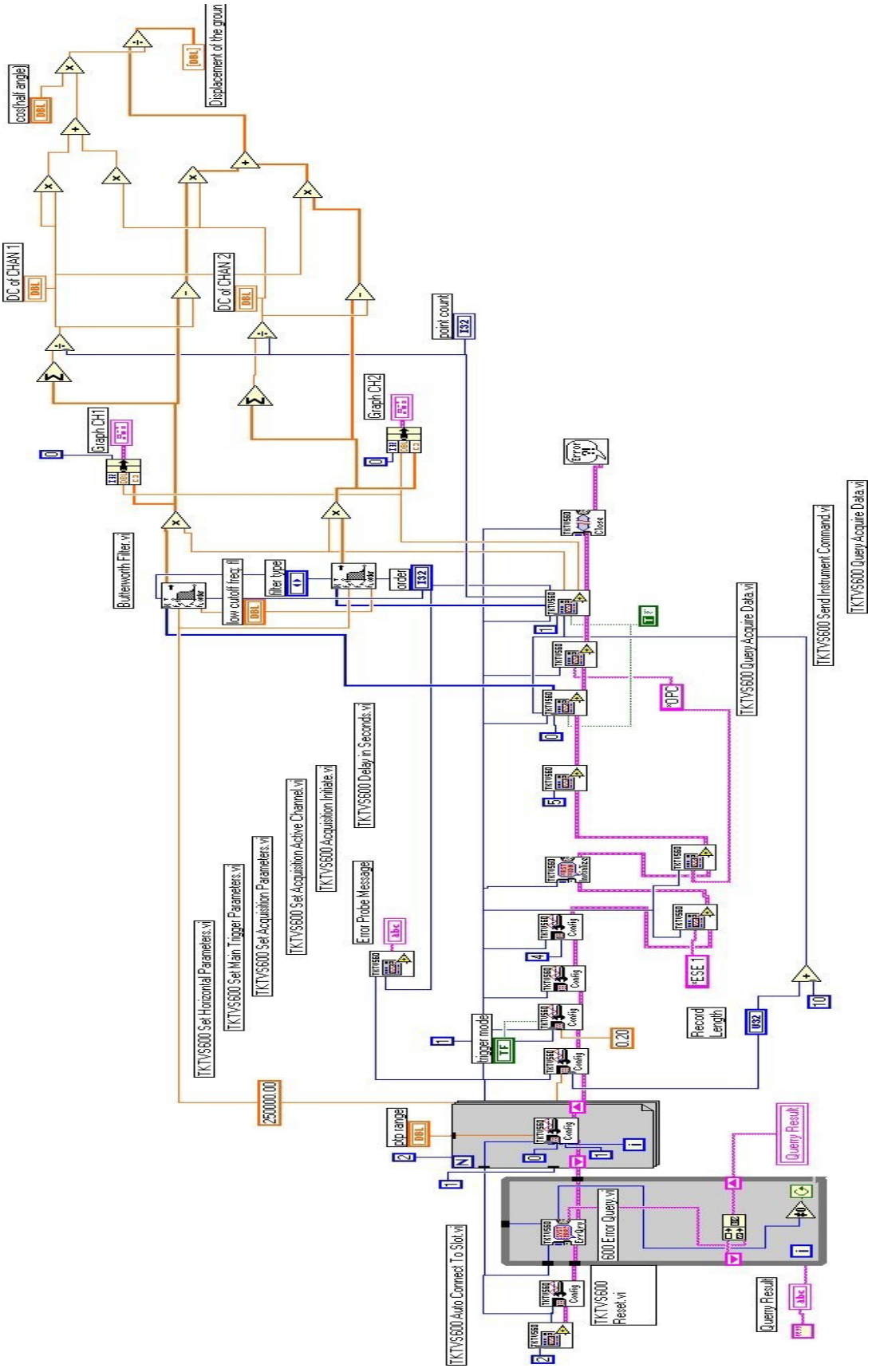


Figure 4.5 LabView diagram of the digital signal processing

Conclusion

Dynamic range of the different devices of the system and signal to noise of the signals could not be measured beyond 100dB. This limit, set by the measurement capabilities of the spectrum analyzer, makes it difficult to locate the element limiting the resolution of the whole system. Different sources of noise were studied. Acoustic noise level at the frequency of the transducer was found to be very low in the laboratory. These allow a dynamic range over 120dB when using the transducer with a driving voltage of 150V peak and a distance of 55mm. In order to limit electromagnetic cross talk, a shielded design was adopted for all the components of the system. The source noise can be amplitude noise or phase noise. The implementation of a delay line in the aim of improving the resolution did not give measurable result. A crystal oscillator was introduced. It gave only a slight reduction of the low frequency noise floor level. The remaining noise can either be coming from the transducer or amplitude noise of the crystal oscillator. For the filtering of the multiplier output, a sharp low pass filter is needed. An on-board active Chebyshev design filter was implemented. It performs a 90dB reduction at 100kHz. The remaining 100kHz tone was filtered after acquisition on a computer.

CHAPTER VI

CONCLUSIONS

A system using 50kHz electrostatic transducers and analog demodulation was implemented. It performs vibrations measurements adequate for use with surface waves in the ground used for mine detection. A displacement sensitivity on of 1-nanometer was obtained in the frequency range 400-1200Hz. In the low frequency range, the resolution deteriorates down to a few nanometers at 200Hz. The transducer system achieves a spot size of the order of a centimeter compatible with surface wave measurement in the frequency range presented.

The choice of electrostatic transducers allowed high acoustic pressures, over 120dB re 20μPa. Their very high receiving sensitivity permitted unamplified signal levels over 500mV. These conditions were favorable to high signal to noise ratio. The natural focusing of the transducers was used to get a small spot size and higher signal levels. Spherical focusing of the sound was considered between50kHz and 200kHz. A computer model based on the Raleigh integral was developed to calculate the sound field of spherically focused transducers. In particular, the results give focusing distances and pressure at the focused point for transducers running at different frequencies. At 50kHz it seems difficult to focus the sound more than a few centimeters away from the transducers. This will certainly be impractical when used in the fields. However, the calculations indicate that at 200kHz spherically focused transducers can focus at a distance of 13 centimeters with a pressure gain of four relative to the normalized pressure.

Such performances seem promising. Attenuation in the air should not be a problem for distances less than a meter and frequencies up to 200kHz. Capacitance transducers may not be ideal however due to the use of a high bias voltage of 150 to 200 Volts. Such transducers may malfunction in difficult environments. The bias voltage can easily be short-circuited accidentally. The transducers may also age rapidly in highly corrosive environments.

With the aim of determining the equipment needed to perform a digital demodulation of the signal, a study of data acquisition cards was done. A test on a recent card showed a noise floor level too high to permit phase demodulation with a sensitivity of 120dB. Computer models of the two noise sources for A/D cards were developed. The corresponding noise floor levels were evaluated with a resolution of one Hertz. Twelve and 16 bit coding allow resolutions of 115 and 140 dB respectively. The highest noise level on the card tests seemed to come from the clock jitter. If the clock jitter level (in seconds) is inversely proportional to the clock speed (or constant in degrees), a 1Gigahertz clock should be adequate. Finally, specifications for an acquisition card for demodulating of a 50kHz carrier signal with a sensitivity of 120dB are acquisition on 2 channels, 16bit coding and a 1GigaHz clock. Due to the cost of such a device, efforts were concentrated on an analog demodulation

Several analog demodulation configurations involving multiplication of the signal with the carrier tone and low pass filtering were investigated. The sensitivity of some multiplication devices was studied experimentally with the transducer system. The transducer were aimed at a vibrating piston, mixers and multipliers were used to demodulate the signal. Vibration from 200Hz to 1200 Hz

with amplitude of 40nanometer was used. A mixer setup was able to demodulate the signal with a sensitivity of 100dB. Multiplier circuits gave higher output levels with sensitivities up to 124dB. In all cases, a rise of the noise floor was observed in the lower frequency range. Finally a complete demodulation configuration was implemented with a quadrature circuit and on board filtering. This quadrature of the carrier signal was obtained with an operational amplifier differentiating circuit. It provides a gain of 1 at 50kHz and a phase shift of pi/2 with a dynamic range superior to 120dB.

Noise measurements were difficult to acomplish because of the high sensitivity desired for the demodulation. Acoustic noise was not a problem thanks to the very high source level of the transducer. Implementation of a delay line and a crystal oscillator showed that the rise of the noise floor in the low frequency range did not come from phase noise of the source. Transducer or source amplitude noise are believed to be responsible for this phenomenon. A good filtering of the multiplier output signal was obtained by implementation of on-board active filtering.

CHAPTER VII

RECOMMENDATIONS FOR FURTHER STUDY

Some more research could continue on reducing the noise and improving the sensitivity at low frequencies. The system will also have to be tested on transient waves in the sand. Then, some more work could be conducted for more practical outdoor use of the transducer system. In particular, this environment would require more robust transducers able to function further from the ground.

Although the filtering installed is satisfactory, it could be optimized. In particular, the active filtering of the multiplier output could be performed on a board distinct from that of the multiplier chips, using a different power supply. This configuration would reduce the cross talk of the 50kHz carrier tone and save some digital filtering. The rise of the noise level in the low frequency range could be further investigated. This rise reduces the resolution down to 112dB at 200Hz. It is believed that the demodulation is not responsible for this phenomenon. Therefore, Investigation should concentrate on the transducers and the source. Noise from the crystal oscillator can only be amplitude noise and clipping of its signal would reduce it. Using another type of transducer may provide a higher quality signal. However, the noise problem will be very minor if a higher frequency is chosen for the final system. To give an idea, a resolution of 114dB would give a minimum sensitivity of 1-nanometer for frequencies over 100kHz.

A more important step would be to test the system in the experimental sand tank. Some slight modification would have to be implemented for such experiments. Analog and digital filtering of the multiplier output signals induce a phase shift that is not uniform over frequency. This phase shifting has to be determined and a digital correction must be implemented so that the system can handle transient waves. This digital correction can be installed on the existing Labview program.

At the same time, some research could be conducted for a more practical "in situ" use of the system. The capacitance transducers are neither sealed, nor robust. They run with a High DC bias of 150-200V. These weaknesses could cause problems when using them in a humid or corrosive environment. Short circuits are likely to happen at such a high voltages, possibly damaging the system. Hence, long life of the batteries and the transducers is not guaranteed. More robust transducers, such as most piezoelectric transducers, are commercially available. They usually have a poorer receiving sensitivity, but better focusing may be able to compensate for the corresponding loss in signal level. Using the transducers in a typical outdoor environment would also require a localization of the transducers at a larger distance from the soil. In order to keep a small spot size, the sound beam will have to focus at this same distance. Some theoretical work on the focusing of the sound has been accomplished, but experimental tests remain to be conducted.

The most important issue for a practical use of this system could be the presence of vegetation. Most of the land mines remaining to be removed have been in the ground for a long time. Vegetation has had time to grow and hide the

soil surface. Herbs, plants and clutter could be at the same time, rigid enough for ultrasound reflection, and isolated from the ground and the surface wave. This issue would have to be investigated experimentally.

APPENDICES

```
% BIT NOISE SIMULATION:

echo off;clear all;
Fcarrier=49988;
fs=250E3; Ts=1/fs;T=1; N=fs*T;
Kc=2*50000*pi/340

t=(0:Ts:T-Ts);
% 50kHz Pure tone with 200Hz phase modulation
S=cos(2*pi*Fcarrier*(t)+2*Kc*1E-8*cos(2*pi*200*(t)));

%16 bits model
M=2^15; %+/-signal
S1=round(M*S)/M;

%12 bits model
M=2^11; %+/-signal
S2=round(M*S)/M;


%Fourier Transform of the three signals

freqs=(0:1:249999);
figure(1);
subplot(3,1,1);
plot(freqs,20*log10(abs(fft(S))));
axis([49000,51000,-260,120]);grid on
xlabel('frequency (Hz)'); ylabel('dB');
title('Fourier Transform of an Ideally Digitized
Signal')

subplot(3,1,2);
plot(freqs,20*log10(abs(fft(S1))));
title('Fourier Transform of Ideally Sampled Signal
with 16 Bits Coding')
axis([49000,51000,-60,120]);grid on
xlabel('frequency (Hz)'); ylabel('dB');

subplot(3,1,3);
plot(freqs,20*log10(abs(fft(S2))));
title('Fourier Transform of Ideally Sampled Signal
with 12 Bits Coding')
axis([49000,51000,-60,120]);grid on
xlabel('frequency (Hz)'); ylabel('dB');
```

```
% JITTER OF SAMPLING CLOCK SIMULATION:

echo off;clear all;
Fcarrier=50000;
fs=250E3; Ts=1/fs;T=1; N=fs*T;
Kc=2*50000*pi/340

t=(0:Ts:T-Ts);
S=cos(2*pi*Fcarrier*(t)+2*Kc*1E-8*cos(2*pi*200*(t)));

%clock jitter

%5E-8s jitter amplitude
err=.33333/20E6*randn(1,250000);
t2=t+err;
S1=cos(2*pi*Fcarrier*(t2)+2*Kc*1E-
8*cos(2*pi*200*(t2)));

%5E-9s Jitter amplitude
err=.33333/20E7*randn(1,250000);
t2=t+err;
S2=cos(2*pi*Fcarrier*(t2)+2*Kc*1E-
8*cos(2*pi*200*(t2)));

%Graphs of the Fourier Transform
freqs=(0:1:249999);
figure(1);
subplot(3,1,1);
plot(freqs,20*log10(abs(fft(S))));
axis([49000,51000,-260,120]);grid on
xlabel('frequency (Hz)'); ylabel('dB');
title('Fourier Transform of an Ideally Digitized
Signal')

subplot(3,1,2);
plot(freqs,20*log10(abs(fft(S1))));
title('Fourier Transform of Sampled Signal with 5E-8s
Jitter')
axis([49000,51000,-60,120]);grid on
xlabel('frequency (Hz)'); ylabel('dB')

subplot(3,1,3);
plot(freqs,20*log10(abs(fft(S2))));
title('Fourier Transform of Sampled Signal with 5E-9s
Jitter')
axis([49000,51000,-60,120]);grid on
xlabel('frequency (Hz)'); ylabel('dB')
```

```
%nearfield of planar rigid piston

clear all; echo off;close all;
figure;
w=2*pi*50000;c=340;
k=w/c; ro=1.2; U=9.3E-2;
a=19E-3;
N=20;
I=20;

p=zeros(I,110);
y=0;
dy=2*a/I;
dr1=a/N;


N2max=2*ceil(pi*(a+dr1)*N/a);
x2=zeros(N,N2max);
y1=x2;ds=x2;

r1=dr1;
for n=1:N,
   N2=2*ceil(pi*r1*N/a);
   dth=2*pi/N2; th=-pi/2;
   for n2=1:N2/2, %symetry
      x2(n,n2)=(r1*cos(th))^2;
      y1(n,n2)=r1*sin(th);
      ds(n,n2)=r1*dth*dr1;
      th=th+dth;
   end
   r1=r1+dr1;
end
```

```
z=1E-3; dz=.002;


for l=1:110,
   zscale(l)=z;
   y=0;
   for i=1:I,
      yscale(i)=y;
      r=sqrt(x2+(y-y1).^2+z^2);
      ptemp=exp(-j*k*r)./r.*ds;
      p(i,l)=sum(sum(ptemp));
      y=y+dy;
   end
   l
   z=z+dz;
end


p=2*j*p*ro*c*U*k/(2*pi);
p1=abs(p);
pdB=20*log10(p1/20E-6);
contour(zscale,yscale,pdB,30);
colorbar;
xlabel('Distance on Transducer Axis
(meter)','Fontsize',14)
ylabel('Radial Axis','Fontsize',14)
```

```
% On Axis pressure field generated by a rigid piston

clear all;close all;
w=2*pi*50000;c=330;
k=w/c; ro=1.2; U=9.3E-2;
a=19E-3;
%Distance on the transducer axis
z=(.0003:.0005:.12);
p2=-2*i.*sin(.5* (k*(z.^2+a^2).^.5-k*z));
p2=abs(p2);
plot(z,p2);

xlabel('Distance on Axis in Meter','FontSize',14);
ylabel('Normalized Pressure Amplitude','FontSize',14);
[y,x]=max(p2(30:150))
.0003+(29+x)*.0005
```

```
% nearfield of spherically focused transducer

clear all; echo off;close all;
figure;
w=2*pi*200000;c=340;
k=w/c; ro=1.2; U=9.3E-2;
%Radius of the transducer
a=19E-3;
%Geometrical focal distance
Rf=.05;

N=40;
I=35;

%Pressure on Plane limited by transducer axis
p=zeros(I,70);
y=0;z=5E-3;
%ordinate increment
dy=2*a/I;

dgam=a/(Rf*(N+1));
gam=dgam/2;

N2max=2*round(pi*Rf*sin(gam+N*dgam)/dgam);
x1=zeros(N,N2max);
y1=x1;z1=x1;ds=x1;

%discretisation of the transducer surface coordinate
(x1,y1,z1) area of surface element dS
for n=1:N,
   N2=2*ceil(pi*Rf*sin(gam)*N/a);
   dth=2*pi/N2; th=-pi/2+dth/2;
   dst=Rf^2*sin(gam)*pi*2/N2*dgam;
   %gap1(n)=Rf*sin(gam)*cos(th);
   for n2=1:N2/2, %symetry
      x1(n,n2)=Rf*sin(gam)*cos(th);
      y1(n,n2)=Rf*sin(gam)*sin(th);
      z1(n,n2)=-sqrt(Rf^2-x1(n,n2)^2-y1(n,n2).^2)+Rf;
      ds(n,n2)=dst;
      th=th+dth;
   end
   %gap2(n)=Rf*sin(gam)*cos(th);
   gam=gam+dgam;
end


% calculation of the pressure field with the Raleigh
integral
```

```
dz=.001;
z=0.001;
for l=1:70,
   zscale(l)=z;
   y=0;
   for i=1:I,
      yscale(i)=y;
      if z>(Rf-sqrt(Rf^2-y^2))
         r=sqrt(x1.^2+(y-y1).^2+(z-z1).^2);
         ptemp=exp(-j*k*r)./r.*ds;
         p(i,l)=sum(sum(ptemp));
      elseif z>.008
         r=sqrt(x1.^2+(y-y1).^2+(z-z1).^2);
         ptemp=exp(-j*k*r)./r.*ds;
         p(i,l)=sum(sum(ptemp));
      end

      y=y+dy;
   end
   z=z+dz;
end


p=2*j*p*ro*c*U*k/(2*pi);
p1=abs(p);
pdB=20*log10(p1/20E-6);
contour(zscale,yscale,pdB,30);
xlabel('distance on symmetry axis in
meters','FontSize',12);
ylabel('Radial Distance','FontSize',12);

colorbar;

figure;
pnorm=p1(1,:)/(ro*U*c);
plot(zscale,pnorm);
xlabel('Distance on Symmetry Axis in
Meter','FontSize',12);
ylabel('Radial Distance','FontSize',12);
```

REFERENCES

[1] Kosta Tsipis,"Technological innovation in humanitarian demining" , Proceedings of the Human Factors and Ergonomics Society vol.1 1998, p750-753

[2] N. Goldfine, D. Clark " New quasistatic magnetic and electric field imaging arrays and algorithms for object detection identification and discrimination", Proceedings of the 1999 Detection and remediation Technologies for Mines and Minelike Targets IV.

[3] Smith, Preston S. Wilson, Fred W. Bacon, et al., "Measurement and Localisation of Interface Wave Reflections from a Buried Target", University of Texas at Austin, August 1997.

[4] Don, C.G. and D.E. Lawrence, "Detecting Buried Objects, Such as Land Mines, Using Acoustic Impulses" Echoes, Winter 1998.

[5] W. R. Scott, Jr., and J.S. Martin, "An acousto-electromagnetic sensor for locating land mines" in *Detection and Remediation Technologies for Mines and MinelikeTargets III*, Proc. SPIE 3392, pp 176-186, 1998.

[6] W. R. Scott, Jr., and J.S. Martin, "Experimental investigation of the acousto-electromagnetic sensor for locating land mines" in *Detection and Remediation Technologies for Mines and MinelikeTargets IV*, Proc. SPIE, 1999.

[7] W. R. Scott, C. Schroeder, J.S. Martin, G. Larson "Investigation of a technique that uses both elastic and electromagnetic waves to detect buried land mines" in proceedings of AP2000 -*Millennium conference on Antennas and propagation,* Davos, Switzerland, April 2000.

[8] W. R. Scott, J.S. Martin, G. Larson "Experimental Model for a Seismic Landmine Detection System" IEEE Transactions on Geoscience and Remote Sensing 2000

[9] C. T. Schroeder W. R. Scott, Jr., "Finite difference time domain model for elastic waves in the ground" Proceedings of SPIE Vol 3710, April 1999.

[10] P.H. Rogers, M. Cox, "Automated noninvasive Motion Measurement of Auditory Organs in Fish Using Ultrasound" from *Journal of Vibration, Acoustics, Stress, and Reliability in Design* January 1987.

[11] Kouwenhoven and Michael Hendrikus Laurentius, "High-performance frequency-demodulation systems" published by Delft University Press, 1998.

[12] Technical data “Ultrasonic Ranging System” Polaroid 1984

[13] C.Bilber, S. Ellin, E. Shenk, and J. Stempeck “The Polaroid Ultrasonic Ranging System” 1980

[14] Joe Adler “Jitter in Clock sources”,Vectron International
http://www.vectron.com/whatsnew/jitter.htm

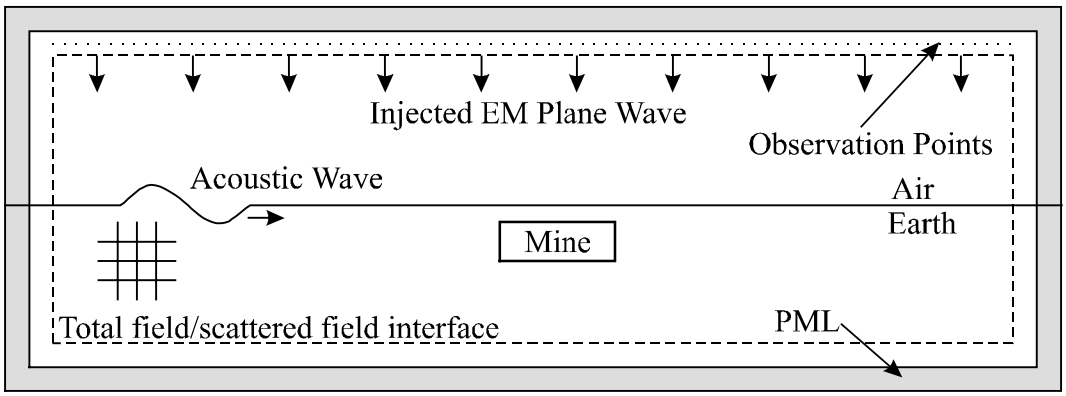
[15] Arthur B. Willams, Fred J. Taylor “Electronic Filter Design Handbook” third edition 1995.

Modeling Vibrating Boundaries with the Finite-Difference Time-Domain Method

Seung-Ho Lee* and Waymond R. Scott, Jr.
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

Over the past two decades, the finite-difference time-domain (FDTD) method has been widely used to model wave propagation, scattering, antennas, high-speed circuits, ground penetrating radar systems, etc. However, very little discussion has been given to modeling moving or vibrating boundaries. In this work, modeling of vibrating material boundaries is investigated using the FDTD method. This investigation is part of a project in which a land mine detection system that simultaneously uses both acoustic and electromagnetic waves is being studied. The system consists of an electromagnetic radar and an acoustic source. An acoustic wave is induced in the earth that travels through the earth and interacts with the mine; this causes the surface of the earth and the mine to vibrate. These vibrations are different near the mine. The radar is used to detect the vibrations and, thus, the mine. The vibrations are difficult to model with the FDTD method; because they are very small, and their frequency and propagation velocity are many orders of magnitude less than those for the electromagnetic waves.

A two-dimensional FDTD code has been written to study the feasibility of using the FDTD method for modeling the vibrating surfaces. A diagram of the model is shown below. An electromagnetic plane wave is injected and its reflection is recorded at the positions indicated. The acoustic wave travels across the surface and causes the surface to be displaced (vibrate). The reflected electromagnetic waves are recorded both when the acoustic wave is and is not present. The displacements are obtained by comparing the phase of these reflected electromagnetic waves. Displacements as small as 10^{-8} of a FDTD cell have been detected using this method, and displacements have been detected as a function of time and position for acoustic waves traveling across the boundary.



An Acousto-Electromagnetic Method for Detecting Buried Objects

Waymond R. Scott, Jr.*, Chistoph Schroeder*, and James S. Martin**
*School of Electrical and Computer Engineering
**School of Mechanical Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250
Tel: 404-894-3048, FAX: 404-894-4641, EMail: waymond.scott@ece.gatech.edu

A hybrid technique is presented that simultaneously uses both electromagnetic and acoustic waves in a synergistic manner to detect buried land mines. The configuration of the system currently being studied is shown in figure 1. The system consists of an electromagnetic radar and an acoustic source. A surface acoustic (Rayleigh) wave is induced in the earth by means of a stationary transducer on the surface of the earth. The transducer is placed near the region being scanned for mines and directs the surface wave into the region being scanned. The surface wave interacts with the mine and causes both the mine and the surface of the earth to be displaced. The displacement of the mine is different than the earth, because the acoustic properties of the mine are quite different than those of the earth. The displacement of the surface of the earth when a mine is present is different than when it is not present because of the waves scattered from the mine. The electromagnetic radar is used to detect these displacements and, thus, the mine.

Both experimental and numerical modeling is being performed to investigate this technique. The experimental model consists of an electrodynamic transducer to induce the surface wave, a tank filled with damp sand to simulate the earth, a simulated mine, and a radar mounted on an automated positioner to measure the vibrations. Experimental results obtained with this system are presented for both anti-personnel and anti-tank mines. The interaction of the acoustic waves with the mine can be seen in the results. The results look very promising; we have been able to detect both simulated anti-tank and anti-personnel mines.

A two-dimensional FDTD model for the acoustic waves has been developed. The results from the FDTD model are in fairly good agreement with those from the experimental model. The numerical model has been a useful aid in understanding the interaction of the acoustic waves with the mine. Results obtained with the FDTD model are presented that show the interactions of the acoustic waves with the mine.

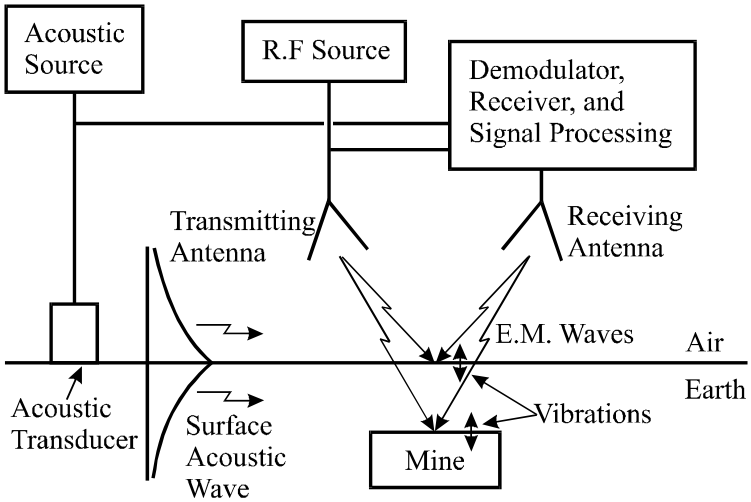


Figure 1. Acousto-electromagnetic mine detection system with the acoustic transducer placed on the surface of the earth.

Radar images of an ocean-like rough surface at high incidence angles

H. Kim* and J. T. Johnson

The ElectroScience Laboratory
The Ohio State University
1320 Kinnear Rd, Columbus, OH 43212
e-mail: hjk@esl.eng.ohio-state.edu

In recent years several analytical models for rough surface scattering have shown promising results for specified ranges of surface statistics. No approximate solutions, however, clearly explain all possible scattering mechanisms for a single surface realization since evaluation of analytical theories is typically based on results for averaged cross section over surface realizations. Due to lack of information on the physical behavior of scattering from rough surfaces, a more descriptive approach is necessary to understand the existing theories. Imaging techniques offer a unique tool for analysis and understanding of rough surface scattering phenomena. Recent development of efficient numerical methods for backscattering predictions enable radar image formation, and images formed from numerical scattering models can be used as a reference solution to evaluate the performance of the existing analytical models.

In this presentation, high resolution backscatter radar images of a 1-D ocean-like surface are formed through back-projection tomography using backscattered field calculations at high incidence angles. The surface profile is described by a Pierson-Moskowitz power-law spectrum and satisfies an impedance boundary condition (IBC). The approximate theories applied include the physical optics (PO) approximation, small perturbation method (SPM), and the local and non-local small slope approximations (SSA/NLSSA). The numerical methods applied are based on an iterative method of moments accelerated with a spectral acceleration technique.

Preliminary studies on a single surface with a Gaussian roughness spectrum have shown that multiple scattering effects can occur, and become more significant as the rms height increases. A simple ray tracing algorithm was used to specify the location of time-delayed spots due to double reflection between two points. Images also revealed that multiple scattering effects were captured successfully from the lowest order NLSSA results, especially in VV polarization, but not from the PO or the 1st order SSA which emphasize only single scattering and local interactions. At moderate incidence angles multiple scattering events were no longer observable. Unlike Gaussian surfaces, however, small multiple scattering effects are observed for Pierson-Moskowitz surfaces even at high incidence angles. These multiple scattering effects appear to be related to the Bragg scattering component in conjunction with the long wave portion of the surface spectrum. Detailed investigations of radar images will be described in terms of major and secondary scattering events, angular dependencies, and polarization effects to clarify these issues.

Anal. Chem. **2001,** *73,* 321−331

Comparison of the Performance of Different Discriminant Algorithms in Analyte Discrimination Tasks Using an Array of Carbon Black−Polymer Composite Vapor Detectors

Thomas P. Vaid, Michael C. Burl,[†] and Nathan S. Lewis*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

An array of 20 compositionally different carbon black−polymer composite chemiresistor vapor detectors was challenged under laboratory conditions to discriminate between a pair of extremely similar pure analytes (H_2O and D_2O), compositionally similar mixtures of pairs of compounds, and low concentrations of vapors of similar chemicals. Several discriminant algorithms were utilized, including *k* nearest neighbors (*k*NN, with $k = 1$), linear discriminant analysis (LDA, or Fisher's linear discriminant), quadratic discriminant analysis (QDA), regularized discriminant analysis (RDA, a hybrid of LDA and QDA), partial least squares, and soft independent modeling of class analogy (SIMCA). H_2O and D_2O were perfectly classified by most of the discriminants when a separate training and test set was used. As expected, discrimination performance decreased as the analyte concentration decreased, and performance decreased as the composition of the analyte mixtures became more similar. RDA was the overall best-performing discriminant, and LDA was the best-performing discriminant that did not require several cross-validations for optimization.

A. Background and Goals. Arrays of broadly responsive detectors, in conjunction with pattern recognition algorithms, have attracted significant recent attention for use in vapor detection.[1] Such detector arrays have been shown to allow identification, classification, and in some cases quantification of various organic vapors.[1] Unlike traditional "lock and key" chemical sensing, in the array approach an individual sensor need not be highly selective toward the analyte of interest. Instead, variations in the pattern of responses produced by the detector array are used to differentiate between various analytes.

The ability of such detector arrays to discriminate between various analytes comprises one figure of merit for the sensing system as a whole. This figure of merit is analogous to the selectivity ratio of an individual, traditional chemical sensor for the target analyte relative to interferences, because when only one channel of data is available, the performance of a sensor system is identical to the performance of the sensor. However, one broadly responsive detector gives no information about an unknown analyte presented at an unknown concentration. In contrast, two differently responding, only partially correlated, detectors that each respond linearly to analyte concentration will yield a unique quantity, the ratio of their signals, for any given analyte. When an array of *n* detectors is exposed to an analyte, it generates *n* responses, which can be plotted as a single point in *n*-dimensional space. A set of exposures to a given analyte at a given concentration will yield a set of points in detector space, which are separated only by the variations in the detector responses. "Training" the array with many exposures to many known analytes will lead to several clusters, one for each analyte. Various discriminant algorithms can then be used to assign a single exposure of an unknown analyte to one of the clusters obtained from the training set, thus identifying the unknown with a specific probability of success.

Clearly, when an array approach to sensing is used, the system-level discrimination performance not only is a function of the detector performance but also is related to the performance of the accompanying data-processing algorithm. Expressions for the signal-to-noise ratio, sensitivity, and selectivity of a detector array system have been given by Lorber[2] and utilized by Kowalski and co-workers.[3,4] Previous studies have compared the performance of some of these algorithms on both real chemical data and simulated data.[5−7] The goal of this work is to evaluate the performance of various data-processing algorithms on a specific vapor detector array used in some several, relatively demanding discrimination tasks.

The detector arrays in the present study are formed from chemically sensitive resistors. Each detector material consists of regions of a conductor interspersed into regions of an insulating organic polymer. Typically the conductor is carbon black, and the dc electrical resistance of the composite is modulated by the swelling of the polymer that results from sorption of the analyte vapor.[8] Diversity in the response of the various detectors in the

[†] Jet Propulsion Laboratory, 4800 Oak Grove Blvd., Pasadena, CA 91125.

(1) Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E.; Vaid, T. P.; Walt, D. R. *Chem. Rev.* **2000**, *100*, 2595.
(2) Lorber, A. *Anal. Chem.* **1986**, *58*, 1167.
(3) Carey, W. P.; Kowalski, B. R. *Anal. Chem.* **1986**, *58*, 3077.
(4) Carey, W. P.; Beebe, K. R.; Kowalski, B. R. *Anal. Chem.* **1987**, *59*, 1529.
(5) Aeberhard, S.; Coomans, D.; de Vel, O. *J. Chemom.* **1993**, *7*, 99.
(6) Wu, W.; Mallet, Y.; Walczak, B.; Penninckx, W.; Massart, D. L.; Heuerding, S.; Erni, F. *Anal. Chim. Acta* **1996**, *329*, 257.
(7) Frank, I. E.; Friedman, J. H. *J. Chemom.* **1989**, *3*, 463.
(8) Lonergan, M. C.; Severin, E. J.; Doleman, B. J.; Beaber, S. A.; Grubbs, R. H.; Lewis, N. S. *Chem. Mater.* **1996**, *8*, 2298.

10.1021/ac000792f CCC: $20.00 © 2001 American Chemical Society
Published on Web 12/15/2000

array is achieved by using different insulating organic polymers to form the composite films of the chemiresistors.

Detector arrays formed from carbon black composite chemiresistors have been shown to exhibit excellent pairwise discrimination between even closely related sets of analytes when a statistically based, linear discriminant algorithm is used to analyze the responses of 10−20 chemically diverse conducting polymer composites.[9] To compare the relative performance of various discriminant algorithms in conjunction with these detector array data, the array must be presented with pairs of analytes that will not be perfectly classified by at least some of the discriminant methods. This was not the case with pairs of simple organic vapors, all of which were essentially perfectly separated from each other, including structural isomers such as *o*- and *m*-xylene.[10] As part of this work, we have challenged a carbon black−polymer composite detector array with a pair of compounds that are very chemically similar, H_2O and D_2O.

In addition, it is of interest to evaluate the array performance on analyte mixtures. The steady-state relative differential resistance responses of the carbon black composite detectors, which serve as the descriptors that form an *n*-dimensional odor space from an *n*-member detector array, are linear with analyte concentration, and the response of a binary mixture of analytes is the response of the pure analytes weighted by the mole fraction of analytes in the mixture.[11] For each exposure, the responses of the *d* detectors can be mapped to *d* orthogonal axes. In this space, the Euclidean distance between a binary vapor mixture that is 0.5 mol fraction of each constituent and a binary mixture that is a 0.6:0.4 distribution of these same analytes should be one-tenth of the Euclidean distance between the array responses of the individual pure analytes. Several different binary mixtures of 1-propanol and 2-propanol, and of *n*-hexane and *n*-heptane, were therefore utilized as part of the present work.

Another method to decrease the discriminating ability of a detector array is to decrease the signal-to-noise ratio of the individual detectors. Delivery of low concentrations of analytes will decrease the detector signal and therefore reduce the signal-to-noise ratio, broadening the clusters relative to their separation. A number of low-concentration (≤1.0% of the vapor pressure) exposures to 1-propanol, 2-propanol, *n*-hexane, and *n*-heptane were therefore studied, and the performance of different discriminant algorithms was also assessed for these specific sensing tasks.

B. Description of Selected Discriminant Algorithms. Discriminant algorithms generally fall into two categories: parametric methods, which assume that the data have a certain distribution (usually a normal Gaussian distribution), and nonparametric methods, which make no assumptions about the underlying structure of the data. The classical parametric methods include linear discriminant analysis[12,13] (LDA, also known as Fisher's linear discriminant) and quadratic discriminant analysis (QDA). A hybrid of LDA and QDA, termed regularized discriminant analysis (RDA), has been more recently introduced.[5,14] The classic nonparametric discriminant is *k* nearest neighbors (*k*NN),[13] which has been applied to chemical data as well as to other types of data.[15] Many other classifiers have been developed, including artificial neural networks (ANN),[16] partial least-squares methods (PLS),[17] and soft independent modeling of class analogy (SIMCA).[18,19] In this work, the performance of the *k*NN, LDA, QDA, RDA, PLS, and SIMCA discriminant algorithms was compared for various analyte discrimination tasks using data from the carbon black composite detector array. Brief explanations of the various discriminant algorithms are provided below.

1. *k*-Nearest Neighbor Discriminant. The *k*NN algorithm involves calculation of the distance between the response of a test analyte and the responses of all of the examples in the training set.[13] The most commonly used distance metric is the Euclidean distance, which in two and thre dimensions is the familiar spatial distance. For an arbitrary number of dimensions, the Euclidean distance is simply

$$\text{distance}_{ij} = [\sum_{n=1}^{d}(X_{in} - X_{jn})^2]^{1/2} \qquad (1)$$

where X_{in} and X_{jn} are the coordinates of the *i*th and *j*th point in the *n*th dimension, respectively, and *d* is the number of dimensions. The test sample is then assigned to the class having the largest number of nearest neighbors to the test data. For example, if $k = 3$, the classes of the three nearest neighbors are compared, and the unknown is assigned to the class with the majority of nearest neighbors. When choosing from more than two classes, any $k > 1$ allows the possibility of a tie. For this reason, and because it has been shown that $k = 1$ is the best method for a wide variety of distributions,[20] $k = 1$ has been used in our study. It has also been shown that any classification rule, including those with information about the statistical distribution of the data, can perform at best twice as well as *k*NN ($k = 1$) in the asymptotic case in which the training set includes a very large number of examples from each class.[20] The straightforward *k*NN classifier is therefore a good benchmark against which to measure other, more sophisticated, discriminants.

***2.* Linear Discriminant Analysis.** LDA is typically taken to mean Fisher's linear discriminant.[12] The orthogonal projection of points in a *d*-dimensional space onto a line reduces the classification problem from *d* dimensions to one dimension. When the data are projected onto one dimension, it is desirable to maximize the distance between the means of the two classes being separated, while minimizing their within-class variation. Such a ratio can be expressed as a resolution factor, RF (eq 2), where δ is the distance

$$\text{RF} = \frac{\delta}{\sqrt{\sigma_1^2 + \sigma_2^2}} \qquad (2)$$

between the two class means and σ_1 and σ_2 are the standard

(9) Doleman, B. J.; Lonergan, M. C.; Severin, E. J.; Vaid, T. P.; Lewis, N. S. *Anal. Chem.* **1998**, *70*, 4177.
(10) Vaid, T. P.; Lewis, N. S., unpublished results.
(11) Severin, E. J.; Doleman, B. J.; Lewis, N. S. *Anal. Chem.* **2000**, *72*, 658.
(12) Fisher, R. A. *Ann. Eugneics* **1936**, *7* (Part II), 179.
(13) Duda, R. O.; Hart, P. E. *Pattern Classification and Scene Analysis*; John Wiley & Sons: New York, 1973.
(14) Friedman, J. H. *J. Am. Stat. Assoc.* **1989**, *84*, 165.
(15) Kowalski, B. R.; Bender, C. F. *Anal. Chem.* **1972**, *44*, 1405.
(16) Burns, J. A.; Whitesides, G. M. *Chem. Rev.* **1993**, *93*, 2583.
(17) Geladi, P.; Kowalski, B. R. *Anal. Chim. Acta* **1986**, *185*, 1.
(18) Wold, S. *Pattern Recognit.* **1976**, *8*, 127.
(19) Wold, S.; Sjöström, M. In *Chemometrics: Theory and Application*; Kowalski, B. R., Ed.; ACS Symposium Series 52; American Chemical Society: Washington, DC, 1977; p 243.
(20) Cover, T. M.; Hart, P. E. *IEEE Trans. Inf. Theory* **1967**, *IT-13*, 21.

deviations of the two classes, respectively. Fisher's discriminant finds the vector **w** onto which the data are projected that maximizes the RF. The Fisher method does not prescribe how the resulting one-dimensional data should be separated into classes. In our work, we have used a simple threshold that is derived using the assumption that the projected (one-dimensional) distributions for each class are Gaussian.

3. Quadratic Discriminant Analysis. QDA assumes a multivariate normal distribution of the data for each class.[13] A data point $\boldsymbol{x}$ is placed in the class ω_k that minimizes the value of $D_k(\boldsymbol{x})$, as given by

$$D_k(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_k) + \ln|\boldsymbol{\Sigma}_k| - 2\ln[P(\omega_k)] \quad (3)$$

In this equation, $\boldsymbol{\mu}_k$ is the mean vector of class ω_k, $\boldsymbol{\Sigma}_k$ is the covariance matrix of class ω_k, and $P(\omega_k)$ is the a priori probability of membership in class ω_k. The value of $P(\omega_k)$ was taken to be equal to the quantity 1/(number of classes) for all of the classes. QDA effectively measures the distance from the unknown point to the mean of a class, while normalizing for the variance in the individual measurements (dimensions). The unknown is assigned to the class with the minimum "normalized" distance, $D_k(\boldsymbol{x})$. In practice, the class-conditional mean vectors and covariance matrices are not known in advance, so these parameters are typically estimated from training data using the conventional maximum likelihood (ML) estimators.[13]

4. Regularized Discriminant Analysis. RDA minimizes the same $D_k(\boldsymbol{x})$ as is done in QDA (eq 3), but the ML estimates of the class-conditional covariance matrices are replaced with regularized estimates, $\boldsymbol{\Sigma}_k(\lambda,\gamma)$.[14] The first regularizing parameter, λ, converts the class covariance matrix to a linear combination of the class covariance matrix and the pooled covariance matrix (i.e., that of all training samples) (eqs 4−6). The second regularizing

$$\boldsymbol{Q}_k = \frac{1}{N_k}\sum_{i=1}^{N_k} (\boldsymbol{x}_i^{(k)} - \boldsymbol{\mu}_k)(\boldsymbol{x}_i^{(k)} - \boldsymbol{\mu}_k)^T \quad (4)$$

$$\boldsymbol{Q}_p = \sum_{k=1}^{K} \frac{N_k}{N} \boldsymbol{Q}_k \quad (5)$$

$$\boldsymbol{\Sigma}_k(\lambda) = \frac{(1-\lambda)N_k\boldsymbol{Q}_k + \lambda N \boldsymbol{Q}_p}{(1-\lambda)N_k + \lambda N}, \quad 0 \le \lambda \le 1 \quad (6)$$

parameter, γ, shrinks the class covariance matrix toward a multiple of the identity matrix (eq 7). These regularizations correct for

$$\boldsymbol{\Sigma}_k(\lambda,\gamma) = (1-\lambda)\boldsymbol{\Sigma}_k(\lambda) + \frac{\gamma}{d}\mathrm{tr}[\boldsymbol{\Sigma}_k(\lambda)]\mathbf{I}, \quad 0 \le \gamma \le 1 \quad (7)$$

known discrepancies between the estimates of class distributions obtained from finite samples and the true population densities. The optimal values of λ and γ are determined by minimizing the misclassification in a leave-one-out cross-validation of all samples. The terms of eqs 4−7 are defined as follows: $\boldsymbol{Q}_k$ is the ML-estimated class-conditional covariance matrix of class ω_k, $\boldsymbol{Q}_p$ is the pooled covariance matrix, N_k is the number of objects in class ω_k, N is the total number of objects, K is the number of classes, $\boldsymbol{x}_i^{(k)}$ is the vector of the ith object in class ω_k, $\boldsymbol{\mu}_\kappa$ is the mean vector of class k, d is the number of variables (dimensions), $\mathrm{tr}[\boldsymbol{\Sigma}_k(\lambda)]$ is the trace of $\boldsymbol{\Sigma}_k(\lambda)$, and $\mathbf{I}$ is the identity matrix.

5. Partial Least Squares. A slightly different approach to classification is through the use of regression. Given a set of examples, we seek a weight vector **w** that will map each example to a desired target value. The target value is termed t_1 for class 1 and t_2 for class 2; t_1 is typically +1 and t_2 is typically −1. The parameter n_1 is defined as the number of examples in class 1, n_2 is defined as the number of examples in class 2, and n is defined as $n_1 + n_2$. If the examples from class 1 are arranged as rows in a matrix **X1** (each column is a detector) and the examples from class 2 are arranged as rows in a matrix **X2**, then **w** can be determined by solving the following multiple linear regression (MLR) problem:

$$\mathbf{t} = \mathbf{X}\mathbf{w} + \mathbf{e} \quad (8)$$

where **t** is a column vector containing n_1 rows of t_1 followed by n_2 rows of t_2. **X** is the vertical concatenation of the matrices **X1** and **X2**. The magnitude of the error vector, **e**, is minimized to solve the regression problem. Typically, the target vector and the measurements are mean-centered (and in some cases autoscaled as well). The minimum mean-squared error solution to the MLR problem is well known,[13] and is given by

$$\mathbf{w} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{t} \quad (9)$$

The effectiveness of **w** for classification can be determined by evaluating its predictive ability on new data (e.g., on a sequestered test set or on holdout examples in a leave-one-out cross-validation). If the target values are chosen as follows, $t_1 = n/n_1$ and $t_2 = -n/n_2$, then it can be shown that this approach reduces exactly to Fisher's linear discriminant.[13]

In some situations, such as when the measurements from different sensors are highly correlated or are noisy, obtaining a good weight vector through standard multiple linear regression is difficult due to the inverse appearing in eq 9. One method to resolve this problem is to perform a principal components analysis (PCA) on **X** to determine the directions that have the most variance. The data are projected onto this reduced dimensional subspace and directions with smaller variance are presumed to correspond to noise and discarded. The target values are then predicted from the projected subspace rather than from the original data. In the chemometrics literature, this approach is known as principal components regression (PCR).[21] The projected data are commonly referred to as the "score matrix".

PCR provides an alternative solution to the regression equation (eq 8) that may be better-behaved than the standard MLR solution. Partial least-squares regression is another method that provides an alternative solution to the regression equation.[17,21] The PLS method is similar to PCR, except that both the target vector and the measurements are used to determine a lower dimensional

(21) Livingstone, D. *Data Analysis for Chemists: Applications to QSAR and Chemical Product Design*; Oxford University Press: New York, 1995.

subspace from which the predictions will be made. Determination of the subspace is accomplished through an iterative procedure.[17]

6. SIMCA. The SIMCA algorithm, which was developed by Wold in the 1970s,[18,19] is based on representing each class with its own principal components model. If a class is viewed as a cloud of points in a d-dimensional space, PCA finds an orthonormal basis for the cloud. (Here we assume that the PCA is applied to mean-centered data.) The first principal component is the direction of maximum variance of the data. The second principal component is the direction of maximum variance in the subspace orthogonal to the first component, and so on. If the cloud is "thin" in some directions, the class can be accurately approximated as a linear combination of $k < d$ principal components.

In the original SIMCA formulation,[18] the distance of a point from a class was determined by the out-of-space distance, i.e., by the Euclidean distance of the point from the subspace spanned by the k principal components used to model the class. The underlying assumption was that the variances in the directions orthogonal to the PCA subspace were all equal (e.g., due to white noise). By considering the out-of-space distance relative to the average out-of-space distance observed for the training set (the training examples do not all lie exactly on the PCA subspace), the SIMCA algorithm determined whether an unknown point was well-modeled by a particular class.

In more recent formulations,[19] the SIMCA distance includes an in-space distance, as well as an out-of-space distance. The in-space distance is a measure of how well the projection of the point into the principal components subspace agrees with the projections of the known class data. The maximum and minimum values of the projected training data along each dimension of the subspace define a bounding box. SIMCA uses a slightly larger box (one standard deviation wider along each principal component direction) to represent the in-space distribution. If the projected point falls within the SIMCA box, i.e., within the "normal bounds", the in-space distance is 0; otherwise, the in-space distance is given by the weighted Euclidean distance of the point from the SIMCA box, where the weights correspond to the inverse variance along each dimension. The in-space and out-of-space distances are then combined and the unknown test point is assigned to the nearest class.

With a different definition of the in-space distance that is not based on a bounding box, but is based instead on a Gaussian model of the in-space distribution, it is readily shown that SIMCA is similar to a form of regularized QDA known in the chemometrics literature as DASCO (discriminant analysis with shrunken covariances).[7] The maximum likelihood estimates of the class covariance matrices used in standard QDA are replaced by a principal components estimate in which variances along the directions of highest variance are retained, while variances along directions of lowest variance are replaced with a constant value (related to the average out-of-space distance of the training set, which is used as a normalizing factor in SIMCA). Frank and Friedman discussed the connection between LDA, QDA, RDA, SIMCA, and DASCO in more detail.[7]

EXPERIMENTAL SECTION

A. Materials. Poly(ethylene-*co*-vinyl acetate) (70% vinyl acetate), polycaprolactone, cellulose acetate, hydroxypropylcellulose, poly(4-vinylpyridine), poly(vinyl acetate), ethyl cellulose, poly(ethylene-*co*-acrylic acid) (86% ethylene), 1,2-polybutadiene, poly(methyloctadecylsiloxane), and poly(styrene-*co*-acrylonitrile) were purchased from Scientific Polymer Products. Poly(4-vinylphenol), poly(vinyl butyral), and poly(ethylene glycol) were purchased from Polysciences. Poly(ethylene oxide), poly(ethylene-*co*-vinyl acetate) (18% vinyl acetate), poly(styrene-*co*-maleic anhydride) (50:50), poly(vinylpyrrolidone), polystyrene, and poly(methyl methacrylate) were purchased from Aldrich. The carbon black was Black Pearls 2000 from Cabot Corp. Bis(2-ethylhexyl) phthalate was purchased from Aldrich. *n*-Hexane was 99+% from Aldrich, heptane was supplied by Mallinckrodt, and 1-propanol and 2-propanol were obtained from EM Science. The H_2O was filtered through a Barnstead 18 MΩ·cm resistivity filter. D_2O was 99.9 atom % deuterium, purchased from Aldrich and used as received.

Table 1. Polymers in the 20-Detector Array

detector	polymer[a]
1	poly(ethylene-*co*-vinyl acetate) (70% vinyl acetate)
2	poly(ethylene oxide)
3	poly(vinylpyrrolidone) *P*
4	1,2-polybutadiene
5	polycaprolactone
6	poly(4-vinylphenol) *P*
7	poly(vinyl acetate) *P*
8	cellulose acetate
9	poly(4-vinylpyridine) *P*
10	poly(methyl methacrylate) *P*
11	poly(styrene-*co*-maleic anhydride) *P*
12	poly(vinyl butyral) *P*
13	hydroxypropylcellulose
14	ethyl cellulose
15	poly(ethylene-*co*-acrylic acid) (86% ethylene)
16	poly(methyloctadecylsiloxane)
17	poly(ethylene glycol)
18	poly(ethylene-*co*-vinyl acetate) (18% vinyl acetate)
19	polystyrene *P*
20	poly(styrene-*co*-acrylonitrile) *P*

[a] *P* indicates plasticization with 8% by mass bis(2-ethylhexyl) phthalate

B. Detectors and Instrumentation. Polymers were generally dissolved in tetrahydrofuran, except for poly(4-vinylpyridine) and poly(vinylpyrrolidone), which were dissolved in ethanol, and poly(ethylene-*co*-vinyl acetate) (18% vinyl acetate) and 1,2-poly(butadiene), which were dissolved in toluene. Each polymer (160 mg) was dissolved in 20 mL of its respective solvent either at room temperature or by heating to 35−40 °C for several hours. Carbon black (40 mg) was then added and the suspension was then sonicated for at least 20 min.

Corning microscope slides were cut into 10 mm × 25 mm pieces to provide substrates for the detectors. A 7−8-mm gap across the middle of each slide was masked and 300 nm of chromium and 500 nm of gold were then evaporated onto the ends of the slides to form the electrical contacts. Detectors were formed by spin-coating polymer−carbon black suspensions onto the prepared substrates. The resulting films were then allowed to dry overnight.

C. Measurements. The instrumentation and apparatus for resistance measurements and for the delivery of vapors have been described previously.[9] The array of 20 polymers listed in Table 1 was used for the measurements. All exposures were performed for a duration of 300 s and were separated by periods of 600 s of

flowing laboratory air. The first several exposures in a long series tended to give responses that were different from those of the remainder of the exposures, so the initial 40 exposures were excluded from analysis for every data set evaluated in this work. The background air contained 1.10± 0.15 parts per thousand of water vapor, but no active auxiliary control over the humidity of the solvents or over the ambient temperature of the bubblers or the detectors (generally 21.5± 1.5 °C) was performed during data collection.

1. H_2O vs D_2O. Two bubblers were filled with D_2O (labeled 1 and 3) and two with H_2O (labeled 2 and 4). For all exposures, vapors were diluted to P/P^o=0.050, where P is the partial pressure of the analyte and P^o is the vapor pressure of the analyte at room temperature. Forty exposures alternating between H_2O and D_2O were performed, and then 200 additional exposures were performed, cycling 50 times sequentially through bubblers 1−4.

2. Pairwise Resolution of Similar Analytes at Low Fractions of their Vapor Pressure. A series of 120 exposures to 1-propanol and 2-propanol were performed, with exposures alternating sequentially between each member of the pair of analytes. All exposures were initially performed at a partial pressure, P, such that $P/P^o = 0.01$ for the analyte in a background of laboratory air. Similar data were collected at partial pressures of $P/P^o = 7.5 \times 10^{-3}$, 5.0×10^{-3}, and 2.5×10^{-3}, with 120 alternating exposures to each member of the solvent pair performed at each analyte concentration. An identical exposure sequence and protocol was performed for collection of the detector response data for *n*-hexane vs *n*-heptane. The first 40 exposures in each sequence were not included in the data analysis.

3. Mixtures of Analytes. Vapor was delivered from two bubblers, one containing 2-propanol and the other containing 1-propanol. The 40 initial exposures (which were not used in the data analysis) consisted of a combination of 2-propanol at $P/P^o = 2.5 \times 10^{-2}$ and 1-propanol at $P/P^o = 2.5 \times 10^{-2}$. For data collection, exposure 1 consisted of a combination of 2-propanol at $P/P^o = 2.5 \times 10^{-2}$ and 1-propanol at $P/P^o = 2.5 \times 10^{-2}$. Exposure 2 consisted of 2-propanol $P/P^o = 2.7 \times 10^{-2}$ and 1-propanol $P/P^o = 2.3 \times 10^{-2}$; exposure 3, 2-propanol $P/P^o = 2.1 \times 10^{-2}$ and 1-propanol $P/P^o = 2.9 \times 10^{-2}$; exposure 4, 2-propanol $P/P^o = 3.5 \times 10^{-2}$ and 1-propanol $P/P^o = 1.5 \times 10^{-2}$. The series of exposures 1−4 was repeated 100 times, for a total of 400 exposures. An analogous data set was collected for *n*-hexane and *n*-heptane.

D. Data Reduction. The average of resistance readings for the 60 s immediately prior to the beginning of the exposure was used as the baseline resistance, R_b, and the average of the resistance readings for the last 60 s of the exposure was taken as the steady-state response, R_{ss}. The quantity used in data analysis was the steady-state relative differential resistance change, $\Delta R/R_b$, where $\Delta R = R_{ss} - R_b$. Data were converted to $\Delta R/R_b$ form in Microsoft Excel, while all subsequent manipulations were performed using Matlab. Original Matlab code was written to analyze the data, but the SIMCA routine was based upon one by Donald B. Dahlberg, available on the Internet at ftp://ftp.cdrom.com/pub/MacSciTech/chem/chemometrics/Dahlberg SIMCA.text.

The $\Delta R/R_b$ data were evaluated in three different forms—unnormalized and normalized by two different methods. In the first normalization (n_a), for each exposure the signal ($X_i = \Delta R/R_b$) of the ith detector was divided by the sum of the X_i signals of all 20 detectors in the array (eq 10). In the second normalization

$$\boldsymbol{X}^{(n_a)} = \boldsymbol{X} / \sum_{n=1}^{d} X_n \tag{10}$$

$$\boldsymbol{X}^{(n_g)} = \boldsymbol{X} / [\sum_{n=1}^{d} (X_n)^2]^{1/2} \tag{11}$$

(n_g), signals were divided by the square root of the sum of the squares of the signals across the array (eq 11). In three dimensions, the first normalization method maps the data onto a plane, whereas the second normalization method maps the data onto the unit sphere. Because the responses of the carbon black composite detectors to various analytes have been observed to vary linearly with concentration of the analyte in the vapor phase,[9] either normalization results in a unique, concentration-insensitive signature for an analyte of interest. The two normalizations had a very similar effect on the classification accuracy of the discriminants studied herein; therefore, only the results from n_a are presented.

Except where otherwise specified, all the discriminants were evaluated using a leave-one-out cross-validation methodology. In this procedure, one exposure (data vector) is left out of the data set and the remaining exposures are used as a training set to create the classification boundary. The left-out exposure is then classified by this rule and the classification is checked against the analyte's true class. The procedure is repeated for each member of the data set, and the rate of correct classification is a useful measure of a particular discriminant's performance.

RESULTS

A. Discrimination Between H_2O and D_2O. Figure 1 presents the average responses and standard deviations of the detectors in response to 100 exposures of H_2O and 100 exposures of D_2O. Despite the similarities in response that were expected, and observed, for these two compounds, it was possible to discriminate robustly between the light and heavy water exposures based on the small differences in response patterns that were produced on the carbon black−polymer composite chemiresistor array.

Table 2 presents the resolution factors between D_2O and H_2O obtained from Fisher's linear discriminant when each bubbler is treated as a separate class. Bubblers containing H_2O were well-differentiated from bubblers containing D_2O, with resolution factors between 8.1 and 10.1.

Interestingly, the analyte exposures from bubbler 1 were resolved from analyte exposures from bubbler 3 by a factor of 2.1, even though both contained D_2O. Similarly, analytes from bubblers 2 and 4 were both nominally H_2O, yet were resolved by a factor of 1.8. Resolution factors obtained using the LDA algorithm will never be zero with a finite sample size. Additionally, small amounts of contamination in the bubblers and lines could possibly contribute to the differences in patterns from nominally identical analytes placed in different bubblers. As a test for differences between bubblers, the exposures were divided into four sets, two each of H_2O and D_2O, but with each set containing data from a combination of two bubblers. As shown in Table 3, resolution factors between H_2O and H_2O and between D_2O and D_2O were then only 0.8 and 0.9, clearly indicating that some of

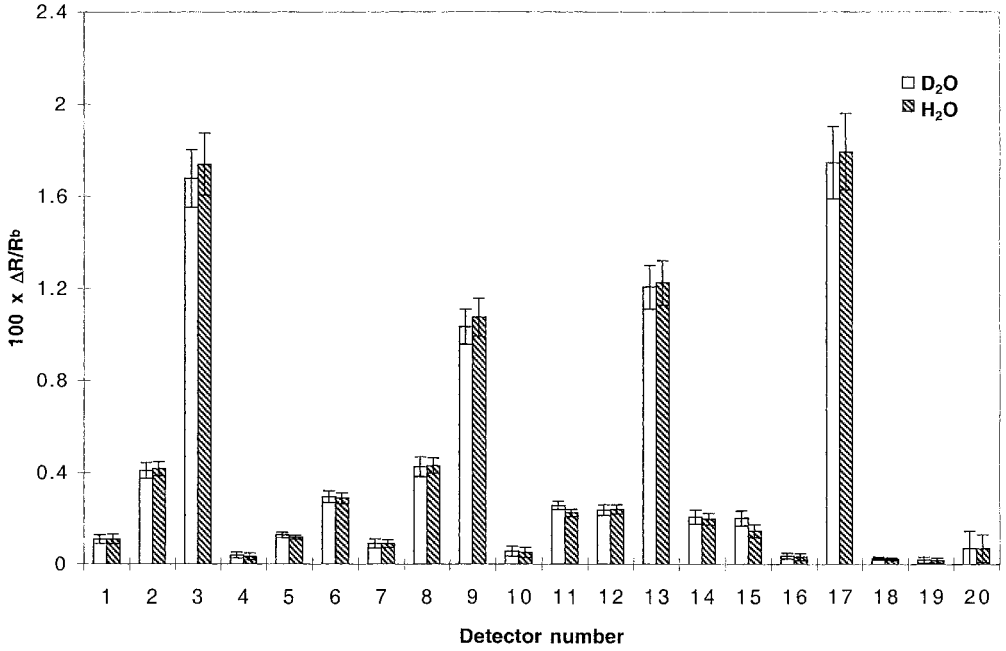


Figure 1. Steady-state relative differential resistance response, $\Delta R/R_b$, of carbon black−polymer composite vapor detectors to H_2O and D_2O (error bars are plus and minus one standard deviation). All exposures were at $P/P^o = 0.050$. Data represent means and standard deviations for 100 exposures to each analyte. The detector number indicates the polymer used to form the composite, with the detector numbering corresponding to the polymer composition listed in Table 1.

Table 2. Resolution Factors for H_2O versus D_2O Using LDA When Data from Each Bubbler Is Treated as a Separate Class

bubbler	1, D_2O	2, H_2O	3, D_2O	4, H_2O
1, D_2O	0.0			
2, H_2O	8.2	0.0		
3, D_2O	2.1	8.1	0.0	
4, H_2O	9.3	1.8	10.1	0.0

Table 3. Resolution Factors for H_2O versus D_2O Using LDA When Data Are Grouped into Four Classes, with the Two H_2O Classes Each a Random Combination of Half the H_2O Exposures and the Two D_2O Classes Each a Random Combination of Half the D_2O Exposures

analyte	D_2O	H_2O	D_2O	H_2O
D_2O	0.0			
H_2O	8.6	0.0		
D_2O	0.9	8.5	0.0	
H_2O	8.4	0.8	8.3	0.0

Table 4. Leave-One-Out Cross-Validation Error Rates for H_2O versus D_2O (Complete Data Set)

						SIMCA	
	*k*NN	LDA	QDA	RDA	PLS	12 PCs	best no. of PCs
unnormalized	0.125	0	0	0	0	0.015	0 (17)
n_a	0.37	0	0	0	0	0.005	0 (16)

the original discrmination was due to differences in what was delivered from the bubblers. The RF values for discrimination between these grouped exposures of H_2O and D_2O were still quite significant and fell in the range of RF = 8.3−8.6 (Table 3).

To further test that discrimination was occurring because of differences between H_2O and D_2O, and not because of various impurities in the bubblers or some other cause, the data were divided into two halves, one of which was used as a training set and the other of which was used a test set. The array was trained on the exposures from bubblers 1 (D_2O) and 2 (H_2O) and LDA was then used to classify the exposures from bubblers 3 (D_2O) and 4 (H_2O). All 100 of these exposures from bubbler 3 or 4 were correctly identified as either H_2O or D_2O using this procedure. Similarly, training on bubblers 3 and 4 and testing on 1 and 2 yielded 100 correct identifications. Training on 100 randomly selected exposures taken from all four bubblers and then testing on the other 100 exposures also produced perfect classification.

Table 4 presents the leave-one-out cross-validation error rates for all of the data obtained on this system. All the discriminants except for *k*NN and SIMCA (when a fixed number of principal components were used) were perfect in their classification. Normalization decreased the performance of the *k*NN algorithm, whereas it enhanced the performance of SIMCA. The degradation in performance of the *k*NN algorithm upon normalization of the response data occurred because the normalization produced less overall amplitude differences between the patterns, and the *k*NN algorithm utilized such differences in classifying the analytes.

B. Resolution of Analytes at Low Fractions of Their Vapor Pressure. 1. Form of the Data. Figure 2 shows the unnormalized response data for each detector in the array to hexane and to heptane, with each analyte at $P/P^o = 7.5 \times 10^{-3}$. Figure 3 displays similar data at an analyte partial pressure of $P/P^o = 2.5 \times 10^{-3}$. At a fixed fraction of the analyte's vapor pressure, the response patterns for hexane and heptane are quite similar, as would be expected from their similar chemical structure and properties. The mean magnitude of the response from detectors that showed significant signals when exposed to hexane (detectors 1, 2, 4, 5, 8, 12−19) decreased by a factor of 3.0 when the hexane partial pressure was decreased from $P/P^o = 7.5 \times 10^{-3}$ to $P/P^o =$

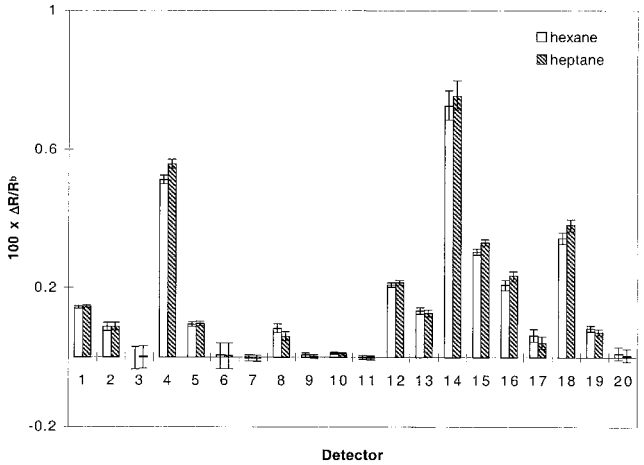


Figure 2. $\Delta R/R_b$ response of an array of carbon black−polymer composite vapor detectors to *n*-hexane or *n*-heptane at P/P^o = 0.0075. Means and standard deviations are for 100 exposures to each analyte, with exposures alternating sequentially between each member of the pair of analytes.

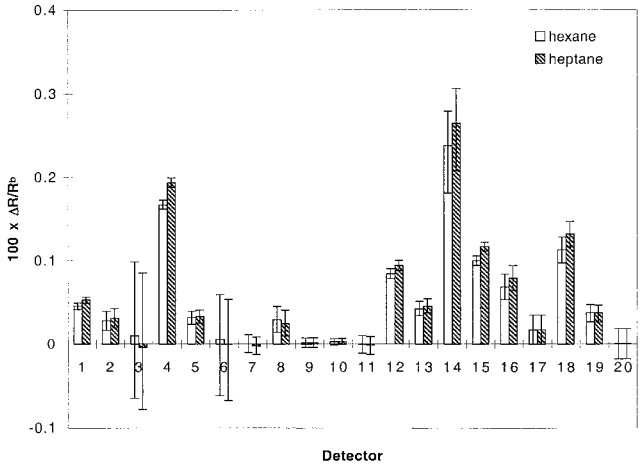


Figure 3. $\Delta R/R_b$ response of an array of carbon black−polymer composite vapor detectors to *n*-hexane or *n*-heptane at P/P^o = 0.0025. Means and standard deviations are for 100 exposures to each analyte, with exposures alternating sequentially between each member of the pair of analytes.

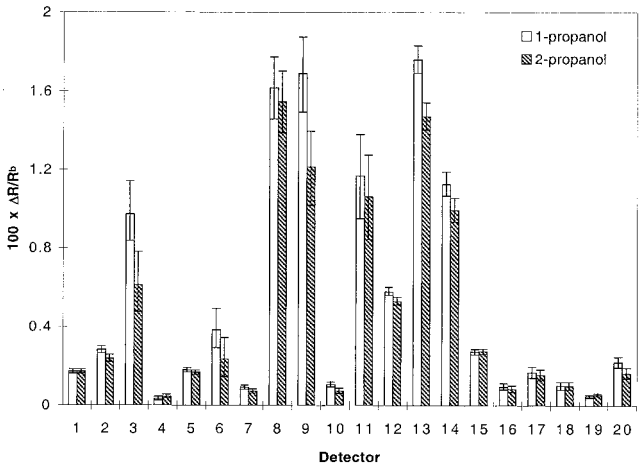


Figure 4. $\Delta R/R_b$ response of an array of carbon black−polymer composite vapor detectors to 1-propanol or 2-propanol at P/P^o = 0.010. Means and standard deviations are for 100 exposures to each analyte, with exposures alternating sequentially between each member of the pair of analytes.

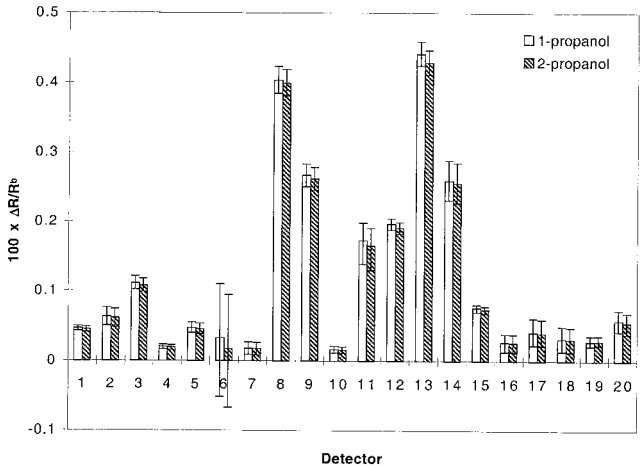


Figure 5. $\Delta R/R_b$ response of an array of carbon black−polymer composite vapor detectors to 1-propanol or 2-propanol at P/P^o = 0.0025. Means and standard deviations are for 100 exposures to each analyte, with exposures alternating sequentially between each member of the pair of analytes.

2.5×10^{-3}, whereas the same decrease in heptane partial pressure produced a mean signal decrease of 2.7 across the same set of detectors. These data are in accord with the linearity of response of carbon black composite detectors to analyte concentration that has been observed previously.[9]

In contrast, the absolute standard deviation of the responses across the set of 100 exposures was essentially constant as the analyte concentration was varied. For example, the ratio of the standard deviation of a detector's responses to hexane at P/P^o=7.5 $\times 10^{-3}$ to that at P/P^o=2.5 $\times 10^{-3}$ had an average of 1.15 across the set of detectors that responded well to hexane (1, 2, 4, 5, 8, 12−19), and this ratio had a value of 1.12 for heptane. Thus, the absolute signal strength decreased as the analyte partial pressure declined, but the absolute variance remained essentially constant, so the discrimination ability of the array is expected to become worse at lower analyte partial pressures.

A quite different situation was, however, observed for 1-propanol and 2-propanol. The absolute standard deviations decreased by an average of 3.91 for 1-propanol and by an average of 3.54 for 2-propanol when the partial pressure of these analytes was reduced from $P/P^o = 1.0 \times 10^{-2}$ to 2.5×10^{-3} (Figures 4 and 5). The main cause for the difference was not a change in random noise, but a steady drift in some of the detector responses over the course of this particular interval of data collection. The effect was more pronounced at $P/P^o = 1.0 \times 10^{-2}$ than at $P/P^o = 2.5 \times 10^{-3}$, accounting for the larger absolute standard deviation values observed at the higher analyte concentration. For illustration, Figure 6 shows the data for 100 responses of detector 8 to 1-propanol at $P/P^o = 1.0 \times 10^{-2}$ and $P/P^o = 2.5 \times 10^{-3}$, respectively. At the higher concentration, the signal drifted by 32% over 50 h, while at the lower concentration it drifted by only 10%. When a simple linear correction was applied to the data (Figure 6), the standard deviation of the higher concentration data decreased by a factor of 3.3, while that of the lower concentration data decreased by a factor of 1.3.

2. Performance of Various Discriminant Algorithms. Table 5 presents the leave-one-out cross-validation error rates for the different discriminant algorithms for the 1-propanol/2-propanol

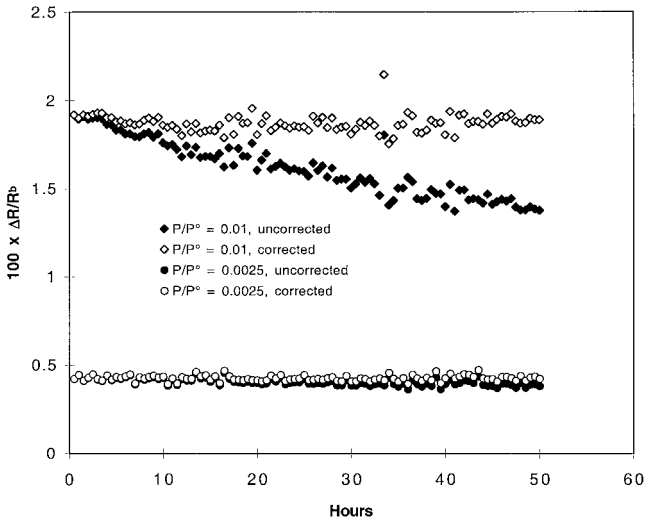


Figure 6. $\Delta R/R_b$ response of detector 8 to 1-propanol and 2-propanol at $P/P^\circ = 0.010$ (diamonds) and at $P/P^\circ = 0.0025$ (circles). Uncorrected, raw data are indicated by filled symbols, and data produced through the use of a linear correction to yield a regression line with slope of zero are indicated by unfilled symbols.

and hexane/heptane data sets. For both the 1-propanol/2-propanol and *n*-hexane/*n*-heptane classifications, the error rate increased for all discriminants at lower partial pressures of analyte. For the unnormalized data, LDA and RDA were the best discriminants (average error rates of 0.079 for hexane vs heptane) with RDA offering only a very slight improvement upon LDA. The PLS algorithm had an average error rate of 0.089, followed by QDA and optimized SIMCA at ~0.10. The *k*NN discriminant had an average error rate of 0.117, and the worst-performing discriminant was SIMCA, with an average error rate of 0.13.

The discriminants were more uniform in their leave-one-out cross-validated performance on normalized data. Once again, SIMCA and *k*NN were the worst classifiers. LDA and QDA were similar overall in their classification accuracy, but their classification performance differed somewhat in different tasks. Because RDA can vary between LDA and QDA, and necessarily chooses the best of these two limiting algorithms based upon cross-validation, RDA was the best discriminant for these normalized data.

C. Discrimination between Compositionally Similar Binary Analyte Mixtures. 1. Structure of Data. Figure 7 displays the average responses of each detector to the four different hexane/heptane binary mixtures. The detector responses exhibited a monotonic trend as the mole fraction of hexane was increased, as expected. In contrast, the response of some detectors was not monotonic for the 1-propanol/2-propanol vapor mixtures (Figure 8). Standard deviations of the detector responses for the 1-propanol/2-propanol vapor mixtures were also generally larger than those for the hexane/heptane mixtures. The larger standard deviations can be attributed to a steady change (usually a decrease) in the response of a detector observed over the course of that particular data collection interval, and the error introduced by the drift may account for the fact that the change across a series is not always monotonic, especially when comparing the very similar 50/50 and 54/46 binary mixtures of 2-propanol and 1-propanol.

2. Performance of Discriminant Algorithms. The leave-one-out cross-validation error rates for this data set are given in Table 6. For both the 1-propanol/2-propanol and *n*-hexane/*n*-heptane classifications, the error rate decreased for all discriminants as the separation in mole fraction between the analytes increased. Normalization did not have a large effect on discriminant performance. The LDA and RDA algorithms were the best-performing discriminants, with average error rates near 0.024. The RDA algorithm was nearly identical in performance to LDA and usually converged to the grid point $(\lambda,\gamma) = (1,0)$, equivalent to LDA. The PLS discriminant was almost as proficient as LDA and RDA, with average error rates of ~0.025. The other discriminants followed in the order, best to worst: QDA, optimized SIMCA, SIMCA, and kNN.

DISCUSSION

A. Discrimination between H_2O and D_2O. Although H_2O and D_2O have very similar physical properties, there are many quite measurable differences, including, for example, boiling point (100 vs 101.4 °C) and melting point (0 vs 3.8 °C).[22] Note that in Figure 1 the detectors with the largest responses (those that are most polar and hydrogen-bonding) tended to respond more strongly to H_2O than D_2O, while the converse is true of the relative responses of the less-polar polymers.

An examination of Figure 1 (and specifically the indicated standard deviations) reveals that most detectors would individually perform very poorly in distinguishing H_2O from D_2O. Detector 11 is the most discriminating individual detector, as reflected by the fact that the **w** vectors found between H_2O and D_2O always had their largest coefficients for 11. Even so, when 11 was removed from the data set, RFs of 8−10 were still obtained, and identification tests were perfect.

B. Performance of LDA and QDA. The H_2O and D_2O data do not provide an appropriate challenge for evaluating the performance of discriminant algorithms, because perfect classification was achieved for most of the algorithms investigated. Such comparisons could be made, however, for both of the experiments involving analytes at low fractions of their vapor pressure and for experiments involving compositionally similar binary analyte mixtures. In these tasks, LDA performed better than QDA. In RDA, where the floating parameter λ allows hybridization between LDA and RDA, a λ value near 1, corresponding to LDA, was generally found to be optimal. These results may at first seem surprising, because QDA is a more general classifier and because QDA reduces to LDA in the specific case when the class covariance matrices are equal. LDA simply uses the pooled covariance matrix, effectively assuming that all the class covariance matrices are equal.

If the true class covariance matrices are the same, then the two classifiers should perform identically in the asymptotic situation in which an infinite number of training examples are available and the class statistics are known exactly. However, in the present situation, the statistics must be estimated from a finite number of training examples. The QDA algorithm estimates a ($d \times d$) covariance matrix for each class, whereas LDA estimates a ($d \times d$) covariance matrix for the pooled data. The covariance estimates produced by QDA will be based on half as much data as in the LDA case and therefore are less likely to reflect the "true" covariance matrix. Also, as shown below, QDA emphasizes the differences in covariance structure between the two classes. From

(22) *CRC Handbook of Chemistry and Physics*, 67th ed.; Weast, R. C., Ed.; CRC Press: Boca Raton, FL, 1986.

Table 5. Leave-One-Out Cross-Validation Error Rates[a] for 1-Propanol versus 2-Propanol and *n*-Hexane versus *n*-Heptane at Low Concentration

						SIMCA	
$100 \times P/P°$	*k*NN	LDA	QDA	RDA	PLS	12 PCs	best no. of PCs
			1-Propanol vs 2-Propanol				
0.01	0 (0.01)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0.025)	0 (0.005)
0.0075	0.01 (0.015)	0.005 (0)	0.005 (0.01)	0.005 (0)	0.015 (0)	0.02 (0.005)	0.01 (0)
0.005	0.41 (0.495)	0.26 (0.48)	0.335 (0.41)	0.255 (0.39)	0.3 (0.495)	0.47 (0.445)	0.36 (0.395)
0.0025	0.385 (0.465)	0.35 (0.495)	0.435 (0.47)	0.35 (0.415)	0.38 (0.55)	0.44 (0.515)	0.4 (0.42)
			n-Hexane vs *n*-Heptane				
0.01	0.03 (0.01)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.005 (0.005)	0.01 (0.005)	0 (0.005)
0.0075	0.035 (0.065)	0.005 (0.01)	0.01 (0.045)	0.005 (0.01)	0.005 (0.01)	0.065 (0.11)	0.03 (0.075)
0.005	0.02 (0.285)	0.005 (0.18)	0.005 (0.245)	0.005 (0.165)	0.005 (0.16)	0.01 (0.29)	0.005 (0.285)
0.0025	0.045 (0.41)	0.005 (0.35)	0.01 (0.32)	0.005 (0.29)	0.005 (0.305)	0.025 (0.36)	0.01 (0.345)
averages	0.134 (0.219)	0.079 (0.190)	0.101 (0.188)	0.079 (0.159)	0.089 (0.191)	0.130 (0.219)	0.102 (0.191)

[a] Error rate for unnormalized data; error rates for normalized data given in parentheses

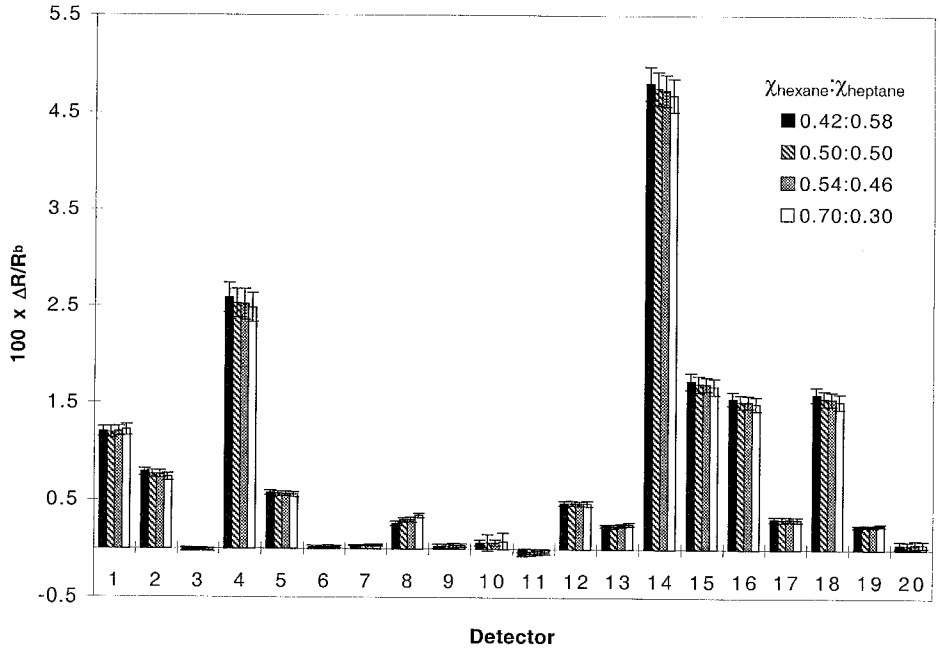


Figure 7. $\Delta R/R_b$ response of an array of carbon black−polymer composite vapor detectors to mixtures of hexane and heptane. The partial pressure of heptane is $[0.050-P/P°_{hexane}]P°_{heptane}$, where P is the partial pressure of hexane for a given exposure. The value of 20 $P/P°$ for each analyte in the mixture is indicated in the legend.

eq 3, we have

$$D_1(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_1^{-1} (\boldsymbol{x} - \mu_1) + \ln|\boldsymbol{\Sigma}_1| - 2\ln[P(\omega_1)] \quad (12)$$

$$D_2(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_2^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_2) + \ln|\boldsymbol{\Sigma}_2| - 2\ln[P(\omega_2)] \quad (13)$$

After some manipulation one obtains

$$D_2(\boldsymbol{x}) - D_1(\boldsymbol{x}) = \boldsymbol{x}^T(\boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_2^{-1})x + 2(\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1)^T \boldsymbol{x} + c \quad (14)$$

When $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are identical, the first term drops out and the LDA classifier is obtained. However, when $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are replaced with their estimated versions, which are not likely to be exactly equal, the first term remains, leading to suboptimal classification.

C. Performance of PLS and SIMCA. The performance of PLS tracked very closely with that of LDA. The PLS discriminant is fundamentally a form of multiple linear regression, and, as explained above, linear regression is equivalent to LDA. It is therefore not surprising that, through different algorithms for optimization, PLS and LDA give similar results. The LDA algorithm might be the preferred method because it is somewhat simpler to implement.

When compared to the other discriminants evaluated, SIMCA performed rather poorly on the discrimination tasks investigated in this work. When the model with the optimal number of principal components was chosen, 16 or 17 principal components were often found to give near-optimal (or optimal) classification accuracy. At these higher limits, SIMCA becomes somewhat similar to QDA, because it is using almost the full dimensionality of the data. Both SIMCA and QDA create a separate model for each class. In situations where the covariance matrices (size, shape, and orienta-

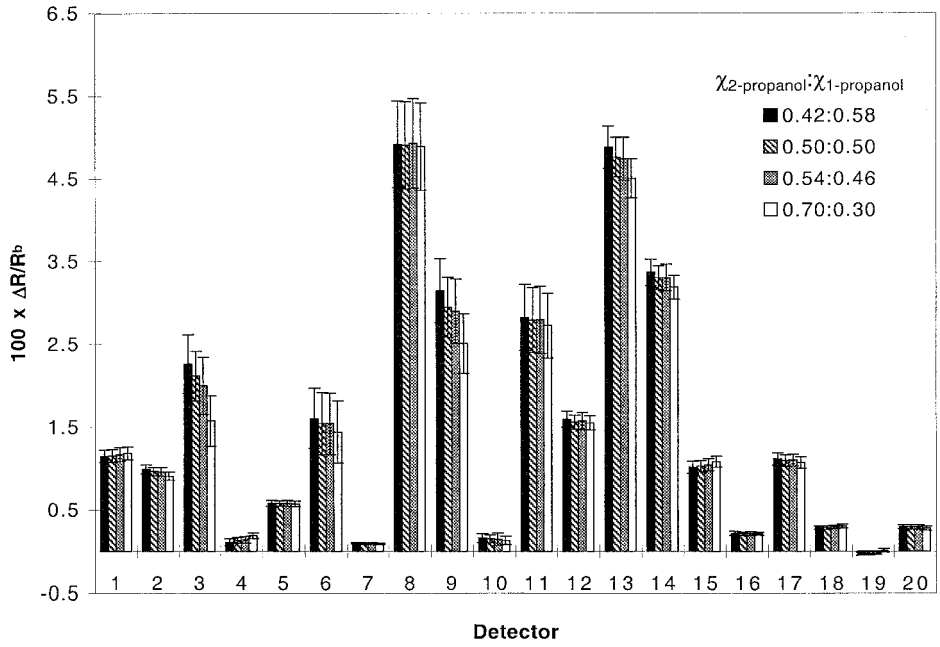


Figure 8. $\Delta R/R_b$ response of an array of carbon black−polymer composite vapor detectors to mixtures of 1-propanol and 2-propanol. The partial pressure of 2-propanol is $[0.050-P/P^\circ_{\text{1-propanol}}]P^\circ_{\text{2-propanol}}$, where P is the partial pressure of 1-propanol for a given exposure. The value of 20 P/P° for each analyte in the mixture is indicated in the legend.

Table 6. Leave-One-Out Cross-Validation Error Rates[a] for Compositionally Similar Analyte Mixtures of 1-Propanol/2-Propanol and *n*-Hexane/*n*-Heptane

						SIMCA	
Δ mixture[b]	*k*NN	LDA	QDA	RDA	PLS	12 PCs	best no. of PCs
			1-Propanol and 2-Propanol				
4	0.325 (0.165)	0.03 (0.025)	0.075 (0.08)	0.03 (0.025)	0.03 (0.025)	0.205 (0.145)	0.13 (0.12)
8	0.12 (0.105)	0.005 (0.01)	0.015 (0.02)	0.005 (0.01)	0.005 (0.01)	0.19 (0.21)	0.045 (0.065)
12	0.065 (0.045)	0 (0)	0 (0)	0 (0)	0 (0)	0.1 (0.05)	0.005 (0.005)
16	0.01 (0.01)	0 (0)	0 (0)	0 (0)	0 (0)	0.01 (0)	0 (0)
20	0.005 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.005 (0.005)	0 (0)
28	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
			n-Hexane and *n*-Heptane				
4	0.42 (0.45)	0.225 (0.21)	0.31 (0.285)	0.225 (0.21)	0.23 (0.21)	0.34 (0.365)	0.28 (0.305)
8	0.365 (0.295)	0.025 (0.04)	0.04 (0.06)	0.025 (0.01)	0.025 (0.025)	0.085 (0.13)	0.06 (0.105)
12	0.3 (0.27)	0.005 (0)	0.01 (0.05)	0.005 (0)	0.005 (0.005)	0.025 (0.045)	0.015 (0.025)
16	0.31 (0.265)	0.005 (0.005)	0.005 (0)	0.005 (0)	0.005 (0.005)	0.035 (0.025)	0.015 (0.025)
20	0.215 (0.135)	0 (0)	0 (0)	0 (0)	0 (0)	0.005 (0)	0 (0)
28	0.065 (0.01)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
averages:	0.183 (0.146)	0.025 (0.024)	0.038 (0.038)	0.025 (0.024)	0.025 (0.025)	0.083 (0.081)	0.046 (0.054)

[a] Error rate for unnormalized data; error rates for normalized data given in parentheses. [b] Δ mixture indicates the difference in mole fraction between the pairs of mixtures being discriminated, as follows: Δ = 4 refers to 50:50 (1-propanol/2-propanol) vs 46:54 (1-propanol/2-propanol); Δ = 8, 50:50 vs 58:42; Δ = 12, 46:54 vs 58:42; Δ = 16, 50:50 vs 30:70; Δ = 20, 50:50 vs 30:70; Δ = 28, 58:42 vs 30:70, and the total of P/P° was 0.050 for all mixtures. Analogous formulas apply to discrimination between mixtures of *n*-hexane and *n*-heptane, with the composition of all mixtures given in the Experimental Section.

tion of the data "cloud") of the two classes under study are very different, it is advantageous to have these separate models. However, as observed in the comparison of LDA with QDA, the data in our tasks generally consist of pairs of classes that have similar covariance matrices. There is therefore little advantage in forming two separate models.

When comparing SIMCA to the other discriminants, it is important to keep in mind the manner in which the models were formed. For LDA, QDA, and PLS, the model is created using the training data, and then unknown "test" data are classified according to the model. The situation is similar for SIMCA when 12 principal components was chosen as an approximately optimal number and used for both classes in all the tasks. In contrast, the optimized SIMCA model was customized for each classification by performing a leave-one-out cross-validation for models that used from 6 to 18 principal components. It is therefore most appropriate to compare the optimized SIMCA to RDA, which also built many models that were tested by cross-validation, and from which the best-performing model was chosen for each classification task.

Overall, QDA and RDA both outperformed SIMCA, whether it was optimized or not. Frank and Friedman discuss some shortcomings of SIMCA that may explain its relatively poor performance.[7]

D. Effects of Normalization. 1. Analytes at Low Fractions of Their Vapor Pressure. Because all pairs of vapors were delivered at the same fraction of their vapor pressure, to a first

approximation, the total response across the array should be similar for different analytes.[23] This is the case in our experiments, especially because the pairs of analytes investigated are so chemically similar. There are differences, however, with heptane giving a slightly larger total response than hexane, and 1-propanol producing a larger total response than 2-propanol. Normalization using eq 10 forces the total response across the array to be the same for every single exposure. If the response patterns of two analytes are similar but differ in magnitude, normalization will make their discrimination more difficult, and this was indeed found to be the case for both analyte pairs across all the discriminants (Table 5). However, normalization is necessary when one has no auxiliary information about the concentration of the analyte and is attempting to perform a classification/identification task for members of these analyte pairs.

2. Compositionally Similar Binary Analyte Mixtures. In contrast to the situation for pure analyte discrimination described above, for the binary mixture data, both normalization procedures led to an improvement in the performance of *k*NN, while the performance of the other discriminants was essentially unaffected by data normalization. This behavior occurs because each exposure is normalized individually, so the effects of variations in external parameters that influence all the detectors in the same way is eliminated though the normalization process. For example, if variations are present in the amount of analyte that is delivered to the array among nominally identical exposures, normalization will ideally correct all the response patterns to the same normalized pattern. Variance in detector response due to other external parameters (perhaps the temperature or the humidity of the background air) that affect the detector signals in the same direction, albeit by different relative magnitudes, will also be canceled to some extent by normalization. One large effect of this type is the drift of the detector signals over the course of the experiment. If the drift is in the same direction for all the detectors, it will be partially ameliorated by normalization. The standard deviations for individual detectors across a set of responses will decrease, but it is not clear how the classification accuracy of the discriminants will be affected.

The drift was much larger for the propanols than for the alkanes and decreased significantly for the propanols between $P/P^\circ = 0.01$ and 0.0025. The largest baseline resistance drifts of any of the sensors over the course of data collection was ~10%, and this appeared to have no correlation with the largest drifts in $\Delta R/R$. The largest downward drifts in $\Delta R/R_b$ (for propanols) were observed for hydrogen-bonding polymers, including poly(vinylpyrrolidone), poly(4-vinylphenol), cellulose acetate, poly(4-vinylpyridine), and poly(styrene-*co*-maleic anhydride), whereas the one polymer in which a significant upward drift in $\Delta R/R_b$ (for propanols) was observed was 1,2-polybutadiene.

E. Extension to Other Vapor Sensor Array Data Sets. Our experiments were carried out under controlled laboratory conditions using carbon black composite chemiresistors; thus, the conclusions regarding which discriminant performed optimally will not necessarily apply to other situations in which variations in detector responses can be produced by a variety of additional factors. For example, a hand-held detector array system that is utilized outdoors may encounter a variety of ambient temperatures, humidities, and background vapors. The resulting class covariance matrices may have a different form and relation to each other than those encountered in our experiments. We point out, however, that a 20-member array of polymer−carbon black detectors has little difficulty in distinguishing two analytes at significant fractions of their vapor pressure unless they are extremely similar (i.e., more similar than H_2O and D_2O). Therefore, the cases in which a choice of discriminant is important will occur only in classification of very similar vapors or at relatively low analyte concentrations. Training of such an array under the variety of conditions under which it will be expected to perform classifications of unknowns will presumably result in similar variances (and relationships between variances on different detectors, i.e., covariances) because the analytes themselves are so similar. The LDA algorithm, which assumes identical covariance matrices for both classes, will therefore likely perform well relative to the other discriminant algorithms evaluated in this work most situations in which the discrimination ability of such an array is challenged.

The conclusions described herein may well also apply to other polymer-based sensor arrays. Polymer-coated quartz resonators of either quartz crystal microbalance (QCM, also called thickness-shear mode resonators) or surface acoustic wave (SAW) devices[1] also utilize sorption of a vapor by the polymer film to detect an analyte. Because these methodologies also rely upon vapor sorption by a polymer film to produce a signal, the conclusions obtained above may apply to the data from such systems as well.

SUMMARY AND CONCLUSIONS

In summary, an array of 20 compositionally different carbon black−polymer composite chemiresistor vapor detectors was challenged under laboratory conditions to discriminate between a pair of extremely similar pure analytes (H_2O and D_2O), compositionally similar mixtures of pairs of compounds, and low concentrations of vapors of similar chemicals. H_2O and D_2O were perfectly separated from each other, and all 100 examples in a test set were correctly classified based on 100 examples in a training set. Discrimination performance decreased as the analyte concentration decreased, and for *n*-hexane and *n*-heptane, classification error rates on normalized data using a leave-one-out cross-validation method exceeded 18% when the analyte concentration was less than 0.005 P/P^o. Mixtures of chemically similar analytes were also robustly discriminated (error of 1% or less) when the analyte compositions differed by more than 0.006 P/P^o (and the total analyte concentration was 0.05 P/P^o), with classification error rates using the leave-one-out cross-validation method exceeding 20% only when the mole fractions of the hexane and heptane differed by less than 0.002 P/P^o in composition (and the total analyte concentration was 0.05 P/P^o). Excluding regularized discriminant analysis, which required the building and cross-validation of many models and which tended to become linear discriminant analysis under optimization, Fisher's classic linear discriminant was the best-performing method under the conditions evaluated in this work.

ACKNOWLEDGMENT

We acknowledge DARPA, the Army Research Office through a MURI grant, the Department of Energy, and NASA for support of this work.

Received for review July 10, 2000. Accepted October 5, 2000.

AC000792F

(23) Doleman, B. J.; Severin, E. J.; Lewis, N. S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5442.