

NATIONAL GEOSPATIAL-INTELLIGENCE AGENCY

Background suppression and feature based spectroscopy methods for subpixel material identification

Approved for Public Release 12-402

Robert S. Rand, National Geospatial-Intelligence Agency; John M.
Grossmann, The MITRE Corporation; Roger N. Clark and Eric Livo, U.S.
Geological Survey; Thomas Parr, BBN Technologies

8/3/2012

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the 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 comments regarding this burden estimate 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. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 03 AUG 2012	2. REPORT TYPE		3. DATES COVERED 00-00-2012 to 00-00-2012		
4. TITLE AND SUBTITLE Background suppression and feature based spectroscopy methods for subpixel material identification			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Geospatial-Intelligence Agency, Springfield, VA, 22150			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Feature-based imaging spectroscopy methods are effective for identifying materials that exhibit specific well-defined spectral absorption features. As long as a pixel contains a sufficient amount of material so that the absorption retains its predominant shape, a feature-based method can work well. However, there are occasions when a background material can mix with a material of interest, and significantly distort and maybe even remove the absorption. In such cases, the material identification capabilities of these methods are likely to be degraded. This effort proposes an approach to accommodate these conditions. The parameter values to determine fit of an absorption feature are selected to be more tolerant of distortions and the signal contributions of any detected sub-pixel backgrounds are removed by making use of a physically-constrained linear mixing model. This mixing model is used to remove any detected background spectra from the image spectra within the bounding locations of the spectral features. However, an expected consequence of loosening the parameter values and performing sub-pixel subtraction is an increase in false alarms. A statistically-based spectral matched filter is proposed as to reduce these false alarms. We test the individual and combined approaches for identifying full-pixel and sub-pixel Tyvek panels in an experiment using a HyMAP hyperspectral scene with ground truth collected over Waimanalo Bay, Oahu, Hawaii.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 15	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Background suppression and feature based spectroscopy methods for subpixel material identification

Robert S. Rand

National Geospatial-Intelligence Agency, Springfield, VA 22150

John M. Grossmann

The MITRE Corporation

Roger N. Clark and Eric Livo

U.S. Geological Survey, Denver, CO

Thomas Parr

BBN Technologies

ABSTRACT

Feature-based imaging spectroscopy methods are effective for identifying materials that exhibit specific well-defined spectral absorption features. As long as a pixel contains a sufficient amount of material so that the absorption retains its predominant shape, a feature-based method can work well. However, there are occasions when a background material can mix with a material of interest, and significantly distort and maybe even remove the absorption. In such cases, the material identification capabilities of these methods are likely to be degraded. This effort proposes an approach to accommodate these conditions. The parameter values to determine fit of an absorption feature are selected to be more tolerant of distortions and the signal contributions of any detected sub-pixel backgrounds are removed by making use of a physically-constrained linear mixing model. This mixing model is used to remove any detected background spectra from the image spectra within the bounding locations of the spectral features. However, an expected consequence of loosening the parameter values and performing sub-pixel subtraction is an increase in false alarms. A statistically-based spectral matched filter is proposed as to reduce these false alarms. We test the individual and combined approaches for identifying full-pixel and sub-pixel Tyvek panels in an experiment using a HyMAP hyperspectral scene with ground truth collected over Waimanalo Bay, Oahu, Hawaii.

Keywords: Hyperspectral, feature-based methods, imaging spectroscopy, sub-pixel, mixed-pixel analysis, spectral match filters, ACE, scene segmentation, background suppression.

1. INTRODUCTION

Feature-based (FB) imaging spectroscopy methods have been used successfully for both earth and planetary remote sensing to identify the mineralogy of terrestrial and planetary surfaces. The methods have been successful in detecting organics and also for mapping asbestos subsequent to the World Trade Center attack on September 11, 2001.¹ Recently the methods were used for a disaster relief effort to characterize and quantify oil that was spilled into the Gulf of Mexico during the Deepwater Horizon incident.² The primary software tool for feature-based analysis has been an Expert System called Tetracorder.³ During the Deepwater Horizon incident, Tetracorder was used to estimate oil thickness, the spill's extent, and the total amount of oil that was leaked from the incident site into the Gulf of Mexico as of May 17, 2010. A conservative lower-limit estimate of 66,000 barrels and an aggressive lower-limit estimate of 120,000 barrels at the surface were provided.²

Based on the Tetracorder methodology, an algorithm called the Spectral Feature Detector (SFD) has been developed. SFD is a feature-based imaging spectroscopy method that has been implemented in the C++ programming language and includes a graphical user interface (GUI) to simplify the selection of spectral features and interaction with a spectral library, as well as overall program execution. SFD was developed as a collaborative effort between the U.S. Geological Survey (USGS) and the National Geospatial-Intelligence Agency (NGA).

SFD is similar though not identical to the imaging spectroscopic approach used in Tetracorder. The underlying logic in the SFD method attempts to capture the essence of Tetracorder's approach; however, it is embedded in a generalized development environment that enables the method to be combined and tested with other methods. As with Tetracorder, SFD was shown to be effective at characterizing oil thickness for the Deepwater Horizon incident.⁴

SFD and Tetracorder make use of the position and shapes of absorption bands in spectra to determine material composition and characteristics. Both methods use localized information to determine material composition with a small subset of the entire spectrum and also remove a local continuum prior to applying distance metrics for determining the similarity between image spectra and the reference spectra for materials of interest. The FB methods are intentionally meant to be sensitive to those regions of the spectrum relevant to a material's chemistry or help differentiate the material from another with which it shares some features, and to ignore the remaining regions that may be strongly influenced by other chemistry and do not help with differentiation.

Numerous variations of the basic linear model exist for modeling sites (pixels) in a scene as combinations of multiple endmembers.⁵⁻¹⁰ These approaches label each site in a scene as containing one, two, or perhaps many endmembers. Some approaches begin with what is sometimes referred to as a full model, where the full model contains all possible variables (endmembers), and subsequently eliminates variables that do not contribute to the statistical significance (e.g., using an F-statistic) of the model.¹¹ Alternatively, other approaches are synonymous with step-wise regression, where the process begins with a pair of variables (endmembers) and adds variables if they contribute significantly to the model.^{10,11} Although many studies exist for applying such models on spectra throughout the entire range of a hyperspectral sensor, few, if any studies exist for applying such models to specific spectral regions associated with specific absorption features of a material.

Feature-based imaging spectroscopy methods are effective for identifying materials whenever well-defined spectral absorption features exist. In sub-pixel cases, if a pixel contains a sufficient amount of material such that an absorption feature retains its predominant shape, existing feature-based methods can still work well. However, there are occasions when a background material can mix with a primary material, and significantly distort and maybe even remove the absorption in the observed mixture spectrum. In such cases, feature-based methods may not work well, if at all.

The current effort further investigates feature-based imaging spectroscopy methods and proposes an approach to improve the identification of a primary set of materials at sites in an image where sub-pixel phenomena are occurring. The parameter values in the FB method are selected to be more tolerant of distortions in absorption features and the signal contributions of any detected sub-pixel backgrounds are removed by making use of a physically-constrained linear mixing model (PCLMM) method⁷ that operates only on specific absorption features. The method estimates the contribution of the background spectra to the locally linear model and removes the appropriate fraction of the estimated background spectra from the image spectra within the bounding locations of the spectral absorption features. A feature-based method is then applied to the adjusted image spectra for identifying the sub-pixel materials of interest.

One effect of loosening the FB parameter values and trying to model the effect of estimated sub-pixel backgrounds is a broader variation of signals that become acceptable to the FB process as belonging to the primary set of materials of interest. We want this; however, an expected negative consequence is a likely increase in the false alarm rate. Statistically-based spectral matched filters are proposed to reduce these false alarms. We consider the Standard Matched Filter (SMF)¹³ and the Adaptive Coherence/Cosine Estimator (ACE)^{13,14} methods. In contrast to SFD, the proposed SMF and ACE use the information in a spectrum throughout the entire range (minus any bands excluded because of heavy atmospheric absorption or sensor artifacts).

This paper is organized as follows: Section 2 provides a description of the proposed algorithms. These include SFD, PCLMM, and two statistical matched filters. Section 3 provides a description of the experiment, consisting of a relatively large and diverse scene of HyMap hyperspectral data collected over Waimanalo Bay, Hawaii, and reference spectra collected on the ground as well as extracted from the scene. We test the capability of the proposed methodology to identify full pixel and sub-pixel Tyvek panels that were deployed during the collection. Section 4 provides a discussion of the results and Section 5 summarizes the conclusions of the study.

2. DESCRIPTION OF ALGORITHM

2.1 The Spectral Feature Detector (SFD) Approach

The chemistry of a material determines the underlying properties of its spectral signature, and often causes diagnostic features at very specific regions in the signature. Electronic processes result in absorptions primarily from 0.1 to 1.35 microns. Vibrational processes due to molecular stretching result in absorptions from about 0.9 to 1000 microns. Other factors, such as a material's grain size, areal (linear) mixing, and intimate (non-linear) mixing, are also very important in determining an observed spectrum. For example, suppose we consider the vibrational absorptions in hydrocarbons and the intimate mixing in oil/water emulsions. Figure 1 shows the spectra of three oil samples at differing thicknesses. Absorption features are highlighted in red and green on these spectra. The left edge of an absorption feature is labeled lc (left continuum edge) and the right edge is labeled rc .

Figure 1: Spectra of three oil samples at differing thicknesses are shown. Absorption features due to C-H with band centers (bc) at 1.2 microns and 1.73 microns are highlighted in red and green, respectively. The left edge of an absorption feature is labeled lc (left continuum edge) and the right edge is labeled rc (right continuum edge). The blue-colored diagonal line shows the inferred continuum spectra C_s of the 1 mm thickness oil sample. The edges and center for the feature near 1.2 microns are shown in the figure. Similar edges and center are defined near the 1.7 micron feature, but are not depicted.

A number of metrics can be used to characterize these features, including the band center (bc) location of its minimum, the band depth, and shape metrics of the feature. The SFD provides a choice between two shape metrics, a correlation coefficient and a spectral angle, which can be used to determine whether an observed spectrum $x(\lambda)$ is similar to a measured reference (material of interest) spectrum $h(\lambda)$. In this study, the correlation coefficient r is used.

$$r = \left[\frac{\left(\sum_{\lambda} h_c(\lambda) x_c(\lambda) - \left(\sum_{\lambda} h_c(\lambda) \right) \left(\sum_{\lambda} x_c(\lambda) \right) / n \right)^2}{\left(\sum_{\lambda} h_c(\lambda)^2 - \left(\sum_{\lambda} h_c(\lambda) \right)^2 / n \right) \left(\sum_{\lambda} x_c(\lambda)^2 - \left(\sum_{\lambda} x_c(\lambda) \right)^2 / n \right)} \right]^{\frac{1}{2}} \quad (1)$$

where the summations occur over the set of wavelengths between (and inclusive) of the left continuum wavelength λ_{lc} and the right continuum wavelength λ_{rc} ; more specifically $\lambda \in \{\lambda_{lc}, \dots, \lambda_{rc}\}$. The locations λ_{lc} and λ_{rc} are shown for the 1.2 micron feature in Figure 1. The variables $x_c(\lambda)$ and $h_c(\lambda)$ are the continuum-adjusted observed spectrum and the continuum-adjusted reference spectrum, respectively. The continuum adjusted values are computed as:

$$x_c(\lambda) = \frac{x(\lambda)}{C_x(\lambda)} \quad \text{and} \quad h_c(\lambda) = \frac{h(\lambda)}{C_h(\lambda)} \quad (2)$$

where $C_x(\lambda)$ is the estimated continuum for the observed spectrum, and $C_h(\lambda)$ is the estimated continuum for the reference spectrum. The $C_h(\lambda)$ for the 1.0 mm oil reference spectrum at the 1.2 micron feature is shown (the blue-colored line) in Figure 1. As indicated in the figure, the continuum is inferred as a straight line from the left continuum edge lc to the right continuum edge rc .

Another discriminating metric, called "shoulderness," developed by Clark *et al.*², is implemented in the SFD, for determining similarities between the image and reference spectra. This metric characterizes the relative positions of the

left and right edges of an absorption feature, prior to the removal of the continuum. The concept and metrics are shown in Figure 2.

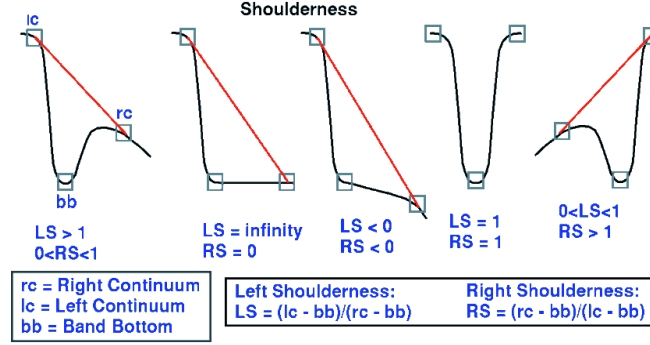


Figure 2: Illustration of the shoulderness metric. This metric characterizes relative positions of the shoulders of an absorption feature. [Courtesy of R.Clark, et al., USGS Open File Report 2010-1167²]

2.2 Localized Physically Constrained Linear Mixture Model (Localized PCLMM)

The Localized Physically Constrained Linear Mixture Model (Localized PCLMM) is based on the PCLMM that has been used in previous work.^{7,8,9,10} We remark that the notation in this section is adopted to be consistent with the later part this work.^{7,10} The PCLMM is a constrained step-wise regression implementation of the Spectral Mixture Analysis [SMA] technique for multiple endmembers, which can be posed in terms of the matrix equation

$$X_s = H\beta_s + \eta_s \quad (3)$$

where, at site s (pixel) in a scene, X_s is the observed reflected energy, β_s is the modeling parameter vector that (when certain assumptions are met) can be associated with the mixture proportions (abundances), and η_s is the random variable for model error. Both SMA and PCLMM can be treated as a special case of linear statistical modeling.¹¹ The columns of matrix H contain the spectra of the endmembers. We define

$$H = \{h^{(1)}(\lambda), \dots, h^{(N_{Ends})}(\lambda), \lambda \in \Lambda\} \quad (4)$$

as the set of material spectra, and $H_s \subset H$ as the set of material spectra at pixel site s . At a site s , $h^{(k)} = (h^{(k)}(\lambda_l), \lambda_l \in \Lambda)$ is the spectra for the k^{th} endmember, if $h^{(k)} \in H_s$, where Λ is the set of all wavelengths used in the model (e.g., the wavelengths corresponding to the bands of the sensor). In the Basic SMA where there is one partition, H_s equals H ; however, more generally there are multiple partitions, and H_s may be strictly contained in H . N_{Ends} is the total number of endmembers.

X^F is defined to be a wavelength-dependent random process governed by the spectral mixing relation at a site s :

$$X_s^F(\lambda) = \sum_{k: h^{(k)} \in H_s} \beta_s^{(k)} \cdot h^{(k)}(\lambda) + \eta_s(\lambda) \quad (5)$$

$\beta_s = (\beta_s^{(k)}, k: h^{(k)} \in H_s)$ is the vector of proportions at a site s on a label lattice. The function $\eta_s(\lambda)$ is associated with error due to variance in the endmember spectra. $\hat{\beta}_s^{(k)}$ are the estimates of $\beta_s^{(k)}$ obtained through some type of constrained least squares or other means. Because the hyperspectral cube G generates discrete samples of the spectrum at specific wavelength intervals, we can approximate the SMA process using standard Linear Modeling matrix notation:

$$X_s^F = H_s \beta_s + \eta_s, \quad s \in S_I \quad (6)$$

The physics of the spectral mixing phenomenon implies that the estimate of β_s should incorporate two constraints: a strict positivity constraint $\beta_s^{(k)} > 0$ for $k: h^{(k)} \in H_s$, and a sum to unity constraint $\sum_{k: h^{(k)} \in H_s} \beta_s^{(k)} = 1$.

Using (6), and the standard assumptions for linear models¹¹, X_s^F is normally distributed with a distribution $X_s^F \sim N\left(E(X_s^F), \sigma(X_s^F)\right)$, where $\hat{x}_s^F = E(X_s^F) = X_s^F - H_s \beta_s = g_s - H_s \beta_s$.

The term on the RHS of (3) has the distribution

$$\begin{aligned} \Pr(X_s^F | G) &= N(\hat{x}_s^F, \sigma(X_s^F)) \\ &= K e^{-u_1(g_s - H_s \beta_s)} \end{aligned} \quad (7)$$

Eq. (7) is maximized by the least squares solution.

The Basic SMA is the special case where only one global set, H , is used throughout the scene, i.e., all endmembers in the global set are included in each site's model, regardless of its location.

In the PCLMM, all combinations of endmember pairs are evaluated from a set of candidate endmembers. A pairwise endmember model is used:

$$x_i = \beta_1 \cdot h_{1i} + \beta_2 \cdot h_{2i} + \varepsilon_i \quad \text{for } i = 1, n \quad (8)$$

where, n is the number of bands used to model the mixing process; x_i is the observed intensity value of the mixture for the i th band, h_{1i} is the intensity value of the first endmember for the i th band; h_{2i} is the intensity value of the second endmember for the i th band and ε_i is the residual for the i th band. The two-endmember models are constrained in the following ways:

1. Only those endmember combinations that satisfy a minimum Degree of Compliance (DOC) are computed. The degree of compliance is defined as the number of spectral bands for which the observed spectrum lies between the two endmember spectra. Typically, the minimum DOC is very close (perhaps equal) to the number of spectral bands.

The DOC constraint tests the condition that

$$h_{1i} \leq x_i \leq h_{2i} \quad \text{or} \quad h_{2i} \leq x_i \leq h_{1i} \quad (9)$$

for all i . The measured DOC is the number of times this condition is satisfied.

2. For those endmember combinations that satisfy the DOC constraint, the LMM is computed. Tests are made for the positivity and sum-to-unity constraints: Specifically, $0 \leq \beta_1 \leq 1$; $0 \leq \beta_2 \leq 1$; and $\beta_1 + \beta_2 \sim 1$.

The proposed Localized PCLMM makes use of the same framework, except that equations (4) to (8) are modified by partitioning H_s and X_s^F into sets according to the wavelength regions Λ_i corresponding to absorption spectral features SF_i for $i = 1, N_{SF}$, where N_{SF} is the number of absorption features.

$$H_s = \{H_{si}, i = 1, N_{SF}\}, \quad \text{where } H_{si} = \{h_{si}^{(1)}(\lambda), \dots, h_{si}^{(1)}(\lambda), \lambda \in \Lambda_i\} \quad (10)$$

$$X_s^F = \{X_s^{Fi}, i = 1, N_{SF}\} \quad (11)$$

In the Localized PCLMM, H_{si} and $X_{si}^{F_i}$ are computed for each wavelength region Λ_i . These models are used to remove the estimated contribution of the background from the observed spectrum in the image to obtain a background-adjusted spectrum. Subsequently, the SFD method described in Section 2.1 is recomputed on the background-adjusted image spectrum in the continuum-removed space.

2.3 Spectral Matched Filters and the Adaptive Coherence/Cosine Estimator (ACE)

Spectral Matched Filters (SMFs) are statistically-based approaches that typically search for specific materials of interest against a predominant background. Some examples are Constrained Energy Minimization (CEM),¹² the standard Adaptive SMF,¹³ the Adaptive Coherence/Cosine Estimator (ACE),^{13,14} and the Kelley detector.¹³ SMFs are typically used for the detection of materials of interest that occupy a very small percentage of a scene. Relating to conventional signal processing terminology, these materials are called targets and the remainder of the image is called the background.

The ACE matched filter is based on a Generalized Likelihood Ratio Test (GLRT). ACE is derived with the assumption of an unstructured background model:

$$b \sim N(0, \Sigma_b) \quad (12)$$

where $N(0, \Sigma_b)$ denotes a multivariate normal distribution with a mean vector of zero and a covariance Σ_b . Typically Σ_b is estimated by the estimator $\hat{\Sigma}_b$, which is the covariance matrix of the background. Making the hypothesis test

$$H_0 : x = b \quad \text{Target is absent} \rightarrow \quad b \sim N(0, \Sigma_b) \quad (13a)$$

$$H_1 : x = ha + \sigma b \quad \text{Target is present} \rightarrow \quad b \sim N(ha, \sigma^2 \Sigma_b) \quad (13b)$$

the ACE detector can be obtained as

$$D_{ACE}(x) = \frac{x^T \Sigma_g^{-1} h (h^T \Sigma_g^{-1} h)^{-1} h^T \Sigma_g^{-1} x}{x^T \Sigma_g^{-1} x} \quad (14)$$

where, x is the data vector and h is the signal vector/matrix. We have made a small change in notation replacing $\hat{\Sigma}_b$ with Σ_g , where Σ_g represents the global background computed from sampling throughout the entire scene (the global background covariance matrix). In general, h can be a matrix containing multiple signals in a sub-space. However, in this investigation we do not consider signal subspaces, so h is a signal vector or equivalently a one column matrix.

Also considered in this study is the standard adaptive spectral matched filter (standard SMF), which has the simpler expression

$$D_{SMF} = \frac{h^T \Sigma_g^{-1} x}{h^T \Sigma_g^{-1} h} \quad (15)$$

where, x is the data vector, h is the signal vector, Σ_g is the estimate of the global background matrix. For further discussion on ACE or the standard SMF, see the publication by Manolakis.¹³

4. DESCRIPTION OF EXPERIMENT

4.1. Data

Numerous flight lines of HyMap imagery were collected over Oahu Hawaii from 24 January to 1 February 2009. The flight line selected for this experiment, Az148, contains a scene in the Waimanalo region of Oahu and was acquired on 1 February. The imagery consists of 125 bands, has a spectral range of 0.45-2.48 microns with a spectral resolution of about 0.01 microns, and a Ground Sampling Distance (GSD) of approximately 4 meters. For our experiment trials, a scene of size 500 pixels by 1250 lines was extracted from Az148. The selected image cube was calibrated from radiance to reflectance by the Navy Research Laboratory (NRL) using NRL's in-house atmospheric package called Tafkaa,¹⁵ which is a heavily modified version of ATREM.¹⁶ As part of the calibration, the imagery was scaled by a factor of 10,000 and converted to 16-bit integers for processing. A true-color RGB composite of this scene is shown in Figure 3.



Figure 3: An RGB composite of the HyMap scene used in the experiment. This subset contains 500 pixels by 1200 lines.
[Source of imagery: U.S. Naval Research Laboratory]

A total of 10 Tyvek fabric panels of varying size and surrounding backgrounds were placed at three different sites prior to the image collection. An aerial photograph of Site 1 is shown in Figure 4, displaying the arrangement for six of these panels (TY-4, TY-5, TY-6, TY-7, TY-8, and TY-9). Ground photos containing four of these panels are shown in Figure 5 (TY-4, TY-5, TY-6, and TY-7). Notice that TY-7 is difficult to see in the HyMap scene. It is expected to be difficult to detect because its sub-pixel size and being partially occluded by trees. TY-5 is expected to be difficult because of size and shading. Figure 6 shows a high resolution aerial photo of Sites 2 and 3. TY-1 and TY-10 are expected to be difficult because of size and shading. TY-2 is expected to be difficult because of its size. TY-3 is expected to be difficult because of its size and background.

Figure 4: Aerial photo of Site 1 within the HyMap scene showing the configuration of six Tyvek panels.
[Source of imagery: U.S. Naval Research Laboratory]



Figure 5: Ground photos of four Tyvek panels located at Site 1:
TY-6 (Top Left), TY-4 (Top Right), TY-5 (Lower Left), and TY-7 (Lower Right)
[Source of photos: USGS and NGA]

Figure 6. Aerial photos of Sites 2 and 3 are shown on the left and right, respectively. Site 2 contains TY-3, TY-2, TY-1. Site 3 contains TY-10. [Source of imagery: U.S. Naval Research Laboratory]

Spectral measurements of Tyvek were acquired on the ground from one of the panels by an ASD spectrometer, along with a number of other materials. One of the Tyvek spectra is shown in Figure 7. The red and green regions of this spectrum indicate two absorption features that are diagnostic of Tyvek. The absorption doublet near 2.3 microns is a combination C-H stretch and H-C-H bend, while the absorption doublet near 1.7 microns is the first overtone of the C-H stretch. These spectral features are typical of long chain C-C single bonded organic molecules, for example, like those found in aliphatic hydrocarbons.¹⁷ The absorptions near 1.4 and 1.9 microns are due to water in the Tyvek and are not uniquely diagnostic of Tyvek (also the Earth's atmosphere is too absorbing to accurately sense the position and shape of these absorptions in our data set). The absorption near 1.2 microns is the second overtone of the C-H stretch, but is influenced by the water absorption on the short-wavelength side. The sharp cut-off in the UV might also be used to detect this material, but the HyMap sensor does not include wavelengths that short.

Image spectra were also extracted from the scene for purposes of providing endmember spectra for potential backgrounds in the PCLMM process. Mean spectra were computed from five regions of the scene representing gravel, sand, grass, trees, and stressed vegetation.



Figure 7: A reference spectrum of Tyvek collected on the ground by a field spectrometer is shown. The red portion of the spectrum indicates the 2.31 micron absorption region used by the SFD method, and the green portion of the spectrum indicates the 1.72 micron region used by the SFD. Note some water absorption is also occurring in the spectrum, indicating the Tyvek sample is wet. However, this shouldn't affect the detection accuracy of the SFD given that the Tyvek features of interest lie outside the water bands.

4.2. Methodology

The methods described in Sections 3.1 to 3.3 are investigated using the data described in Section 4.1. The ground-acquired spectral measurement of Tyvek shown in Figure 7 is used as the reference spectrum for the identification of Tyvek panels in the scene. The red and green regions of this spectrum indicate the two absorption features to be used by the SFD and PCLMM methods. In contrast to this, the SMF and ACE methods will use the entire spectrum, minus any bad regions defined in the bad-bands list of the HyMap header. These bad regions likely contain poor data that are due to heavy atmospheric absorption and/or sensor artifacts.

The results of six trials are reported. Trials 1 to 5 use a 250 by 400 subset of the scene shown in Figure 3, which includes all ten targets deployed for the experiment. Trial 6 uses the entire scene. Trial 1 tests the SFD method (by itself), where the parameter values for absorption fits are very stringent and set to accept very little difference in shape between reference and image spectra. The fit is measured by a correlation metric (r) defined by Equations (1) and (2). A fit of $r = 0.92$ is required in each absorption region, and the shoulderness of the reference and image spectra must match in each region. Another fitness threshold metric, fd (fit * band depth) is set to $fd \geq 0.1$, which is not overly stringent. For sake of consistency, this metric remains the same for all trials involving the SFD. Since very low sensor data values are often corrupted by noise, a noise threshold is used to avoid false detections. This threshold was set to a value of 400 which, based on the data range of 1:10,000, corresponds to a noise level of 4%. This value also remains the same throughout the trials. Trial 2 tests the SFD method by itself for parameter values that are more tolerant to differences in shape between reference and image continuum-removed spectra. The required fit is reduced to $r = 0.85$ and shoulderness is ignored. Trial 3 uses the background suppressed SFD, where estimated backgrounds are removed with the proposed localized PCLMM. The fit of the PCLMM is measured by the root mean square error (RMSE) between the estimated PCLMM model of the observed spectrum and the actual observed image spectrum. Trial 4 uses the ACE spectral matched filter by itself using 109 out of the 125 bands, where bands defined in a bad-bands list were eliminated. Trial 5 combines the SFD, PCLMM, and ACE processes, using the same parameters as Trial 3 for the SFD and PCLMM, and the same parameters as Trial 4 for the ACE method. Trial 6 applies the combined SFD, PCLMM, and ACE to the entire scene. Similar trials were made using the standard SMF as a replacement to the ACE part of the process. The results are not shown because they are almost identical to using ACE. The parameter values of the six trials are summarized in Table 1.

Table 1. The threshold parameters used in the six experiment trials are shown.

Trial	Fit	Fit*Depth	Noise Level	Shoulderness (Y/N)	PCLMM (max MSE)	ACE (low, high)
1	0.92	0.01	4%	Y		
2	0.85	0.01	4%	N		N
3	0.85	0.01	4%	N	40	N
4			4%			0.02, 0.4
5	0.85	0.01	4%	N	40	0.02, 0.4
6	0.85	0.01	4%	N	40	0.02, 0.4

5. RESULTS

The detection maps for Trials 1 to 3 are shown in Figures 8 and 9. Figure 8 also shows a true color RGB image of the subset of the scene used for Trials 1-5 (the entire scene is used in Trial 6). The detection map for Trial 1 (Figure 8, right-hand side) shows that three panels are identified (TY-6, TY-4, and TY-9) with no false alarms. These are the largest and easiest panels to detect. The remaining seven panels of smaller size are not detected. The detection map for Trial 2 (Figure 9, left-hand side) shows nine out of ten panels are identified by decreasing the required fit to a value of $r = 0.85$ and ignoring the shoulderness metric. The pink circle at the top of the detection map contains TY-10; the blue circle at the middle-right contains TY-9; the red circle in the middle contains TY-7; the yellow circle at the lower right contains TY-3; the purple circle at the lower middle contains two panels (TY-1 and TY-2); and the green circle contains three out of four deployed panels (TY-4, TY-6, and TY-8). One panel, TY-5 located within the green circle, remains undetected.

Unfortunately, relaxing the required fit and ignoring the shoulderness has the expected consequence of introducing a notable number of false alarms. Trial 3 introduces the PCLMM process into the analysis. Figure 9 (right-hand side) shows the PCLMM successfully identifies all ten targets, including TY-5 located within the green circle. Unfortunately, this success is also accompanied by a slightly higher number of false alarms as compared to Trial 2.

Table 2 shows the numerical results for the panel TY-5 obtained during Trial 3. During this trial, the preliminary analysis of TY-5 (prior to invoking the PCLMM process) computed a fit of $r = 0.784$ for the first spectral feature (SF1) and $r = 0.969$ for the second spectral feature (SF2). Given that the fit for SF2 was acceptable no further analysis of SF2 is performed. However, the fit for SF1 was below the acceptable value of 0.85, and consequently the PCLMM process was invoked. The PCLMM estimated a background material consisting of gravel existed at this site with an abundance of 0.88, along with the target having an abundance of 0.05. The PCLMM achieved a fit of $RMSE = 18.9$ (on a scale of 0 to 10,000), which is an acceptable measure for the selected model.



Figure 8: A true color RGB of the subset of the scene containing Sites 1-3 (left) and the detection map of Trial 1 (right) are shown. Three panels are detected with no false alarms. [Source of imagery: U.S. Naval Research Laboratory]

Figure 10 shows the detection maps for Trials 4 and 5. The detection map for Trial 4 (Figure 10, left-hand side) shows ACE was successful in detecting all ten panels in the scene; however, just as the case with Trials 3 and 4, a notable number of false alarms are created when the algorithm parameters are set to accommodate increased distortions from a reference spectrum in the attempt to achieve high detection rates. A similar run was made using the standard SMF, which produced almost identical results. The resulting detection map is not shown because it is visually indistinguishable from Trial 4. The detection map for Trial 5 (Figure 10, right-hand side) shows the results of combining the SFD, PCLMM, and ACE processes. Identical results were found in a corresponding trial that used the standard SMF in place of ACE (results not shown). Excellent results were achieved: All ten targets are identified with no false alarms.

Figure 11 shows the detection map for Trial 6. In this trial, the entire scene is processed with the combined SFD, PCLMM, and ACE processes. As the case with Trial 5, all ten targets were identified. A few false alarms can be found

in the area of the scene corresponding to the sub-scene used in Trials 1-5. We might expect that the results in Trial 6 would be identical to those in Trial 5 for this area of the scene. However, this need not be the case. The ACE and standard SMF processes involve the estimation of a background covariance matrix. There is no reason to expect the covariance estimate computed from the sub-scene would be identical to the estimate computed from the full scene.

Table 2. Numerical results for the TY-05 panel computed during Trial 3 are shown.

	Fit	Depth	Fit*Depth	Accept
Pass 1: SF1 (ch 112-116)	0.784	0.069	0	No
Pass 1: SF2(ch 84-90)	0.969	0.031	0.03	Yes
Pass 2: PCLMM model for SF1 is accepted				
Pass 2: RMSE = 18.9				
Pass 2: Target abundance = 0.054				
Pass 2: Background abundance = 0.878				
Pass 2: SF1 (ch 112-116)	0.996	0.676	0.673	Yes
Pass 2: SF2(ch 84-90)	0.969	0.031	0.03	Yes

Figure 9: Detection maps of Trials 2 and 3 are shown on the left and right, respectively. The circles show the areas with targets.

We remark there are other regions outside of the study area where the detection map [in Figure 11] shows Tyvek material is present. These detections should not immediately be considered as false; particularly, in built-up areas where Tyvek or spectrally similar fabric materials may be common. For example, on the left-hand side of the map there are a number of bright colored points and patterns. In particular, there is a curious looking pattern within the red rectangle. On further investigation, this area of the scene contains the driving range of a golf course. The linear pattern coincides with a tree line. It turns out that adjacent to this tree line there is a fabric fence at the boundary of the driving range (See Figure 12). The fence on one side is more dominant than the other side in Figure 11, which is likely due to the viewing angle of the sensor at this location of the scene.

Figure 10: Detection maps of Trials 4 and 5 are shown on the left and right, respectively. The circles show the areas with targets.

Figure 11: Detection map of Trial 6 is shown. The circles show the areas containing targets. The rectangle shows a linear pattern that is associated with a sub-pixel fabric face adjacent to a tree line.



Figure 12: Fabric fences are shown at the boundaries of the Olamana Golf Course driving range.
(Web Photo: www.olamangolfinks.com)

6. CONCLUDING REMARKS

This study has proposed the combined use of the Spectral Feature Detector (SFD), the Physically Constrained Linear Mixing Model (PCLMM), and a spectral matched filter (ACE or standard SMF) to achieve improved detection and identification performance of materials that have diagnostic absorption features. The material of interest used to test the approach is Tyvek, which was selected because of two diagnostic absorption features.

SFD by itself is shown to be effective at identifying full-pixel panels of the test material with zero false alarms in the study area. However, in order to achieve such a low false alarm rate the parameters of the SFD needed to be set at values that excluded the detection of sub-pixel and occluded panels. If the parameters are loosened to allow the SFD to be more tolerant of distortions from a reference spectrum, most of the targets in the study area can be identified. However, the negative consequence of doing this is a notable number of false alarms.

Adding the PCLMM improved the detection capability so that all the deployed test panels are identified. However, the process did not decrease the false alarm rate, but actually slightly increased it. Additional testing of this approach is needed because of the very limited number of cases (test panels) that actually needed to have a background removed for detection.

Using a spectral matched filter subsequent to the SFD and PCLMM process provides a very effective way of decreasing the number of false alarms generated when the SFD parameters were loosened and the PCLMM was used to increase detection capability. The choice of ACE or the standard SFM did not matter in this test.

7. ACKNOWLEDGEMENTS

Special thanks are given to Dr. Charles Bachmann and the U.S. Naval Research Laboratory for providing the HyMap imagery, as well as associated aerial photographs and atmospheric calibration.

8. REFERENCES

- [1] Swayze, G.A., **R.N. Clark**, S.J. Sutley, T.M. Hoefen, G.S. Plumlee, G.P. Meeker, I.K. Brownfield, K.E. Livo, and L.C. Morath, 2006, Chapter 3: Spectroscopic and x-ray diffraction analyses of asbestos in the World Trade Center dust: in *Urban Aerosols and Their Impacts: Lessons Learned from the World Trade Center Tragedy*, Jeff Gaffney and N. A. Marley (eds), *American Chemical Society*, Symposium Series 919, Oxford University Press, p. 40-65, (**Invited** book Chapter).

- [2] Clark R.N., Swayze G.A., Leifer I., Livo K.E., Kokaly R., Hoefen T., Lundeen S., Eastwood M., Green R.O., Pearson N., Sarture C., McCubbine I., Roberts D., Bradley E., Steele D., Ryan T., Dominguez R., "A Method for quantitative mapping of thick oil spills using imaging spectroscopy," *U.S. Geological Survey Open File Report*, 2010.
- [3] Clark R.N., Swayze G.A., Livo K.E., Kokaly R.F., Sutley S.J., Dalton J.B., McDougal R.R., Gent C.A., "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *J. Geophysical Research*, Vol. 108, No. E12, 5131, 2003.
- [4] Rand R.S., Clark R.N., and Livo K.E., "Feature-based and statistical methods for analyzing the Deepwater Horizon oil spill with AVIRIS imagery," *Proceedings of the SPIE Optics and Photonics*, San Diego, CA. August 2011.
- [5] Adams J., Smith M., and Johnson P., "Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 Site," *J. Geophysical Research*, Vol. 91, July 1986.
- [6] Boardman J., "Automating linear mixture analysis of imaging spectrometry data ," *Proceedings of the International Symposium on Spectral Sensing Research (ISSSR)*, San Diego, CA., July 1994.
- [7] Rand R.S., "A physically-constrained localized linear mixing model for TERCAT applications," *Proceedings of the SPIE Aerosense*, Orlando, FL., April 2003
- [8] Rand R.S., "Automated classification of built-up areas using neural networks and subpixel demixing methods on multispectral/hyperspectral data", *Proceedings of the 23rd Annual Conference of the Remote Sensing Society Remote Sensing Society (RSS97)*, Reading, United Kingdom, September 1997.
- [9] Rand R.S., "Exploitation of hyperspectral data using discriminants and constrained linear subpixel demixing to perform automated material identification." *Proceedings of the International Symposium on Spectral Sensing Research (ISSSR)*, Melbourne, Australia, November 1995.
- [10] Rand R.S. and Keenan D.M., "A Spectral Mixture Process Conditioned by Gibbs-based Partitioning," *IEEE Transactions on Geoscience and Remote Sensing: Special Issue on the Analysis of Hyperspectral Image Data*, Vol. 39, No.7, July 2001.
- [11] Montgomery D. and Peck E., *Introduction to Linear Regression Analysis*, 2nd Ed., Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, 1992.
- [12] Harsanyi J.C., "Detection and classification of subpixel spectral signatures in hyperspectral image sequences," Ph.D. dissertation, University of Maryland, 1993.
- [13] Manolakis D. and Shaw G., "Detection Algorithms for Hyperspectral Imaging Applications," *IEEE Signal Processing Magazine*, January 2002.
- [14] Rand R.S., "A neural network approach for improved detector performance of spectral matched filters in hyperspectral imagery," *Proceedings of the SPIE Optics and Photonics*, San Diego, CA. August 2009.
- [15] Gao B.-C.I., Montes M.J., Ahmad Z., and Davis C.O., "Atmospheric Correction Algorithm for Hyperspectral Remote Sensing of Ocean Color from Space" *Applied Optics*, Vol. 39, No. 6, 2000.
- [16] Gao B.-C.I. and Davis C.O., "Development of a line-by-line-based atmosphere removal algorithm for airborne and spaceborne imaging spectrometers" *Proceedings of SPIE Vol. 3118*, Vol. 3118, pp. 132-141, 1997.
- [17] Clark, R.N., Curchin, J.M., Hoefen, T.M., and Swayze, G.A., 2009, Reflectance spectroscopy of organic compounds I: Alkanes: *Journal of Geophysical Research*, 114, p. E03001, doi:10.1029/2008JE003150.