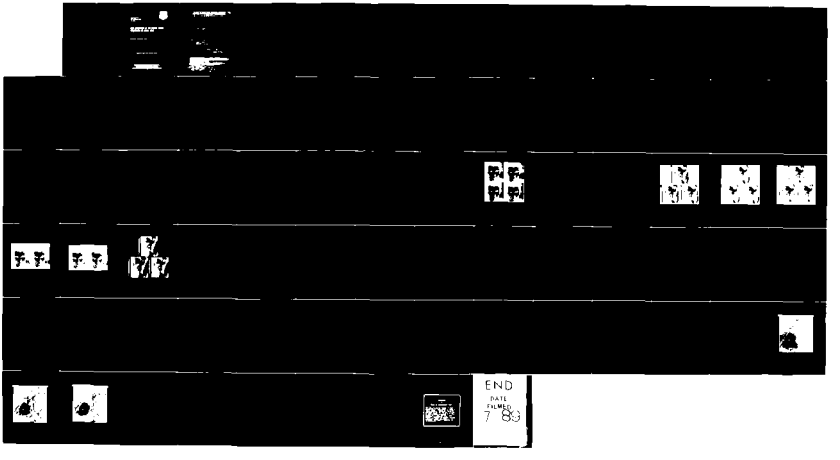
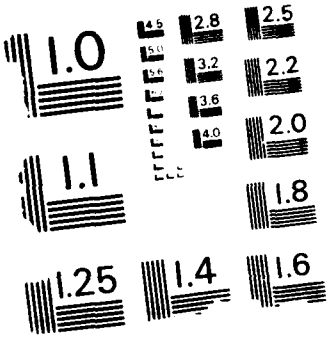
AD-A182 609 NEW DIRECTIONS IN THE DIGITAL SIGNAL PROCESSING OF IMAGE DATA(U) RENSSELAER POLYTECHNIC INST TROY NY H A PEARLMAN ET AL. MAY 87 RADC-TR-87-51 1/1

UNCLASSIFIED F30602-82-K-0151 F/G 8/2 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0

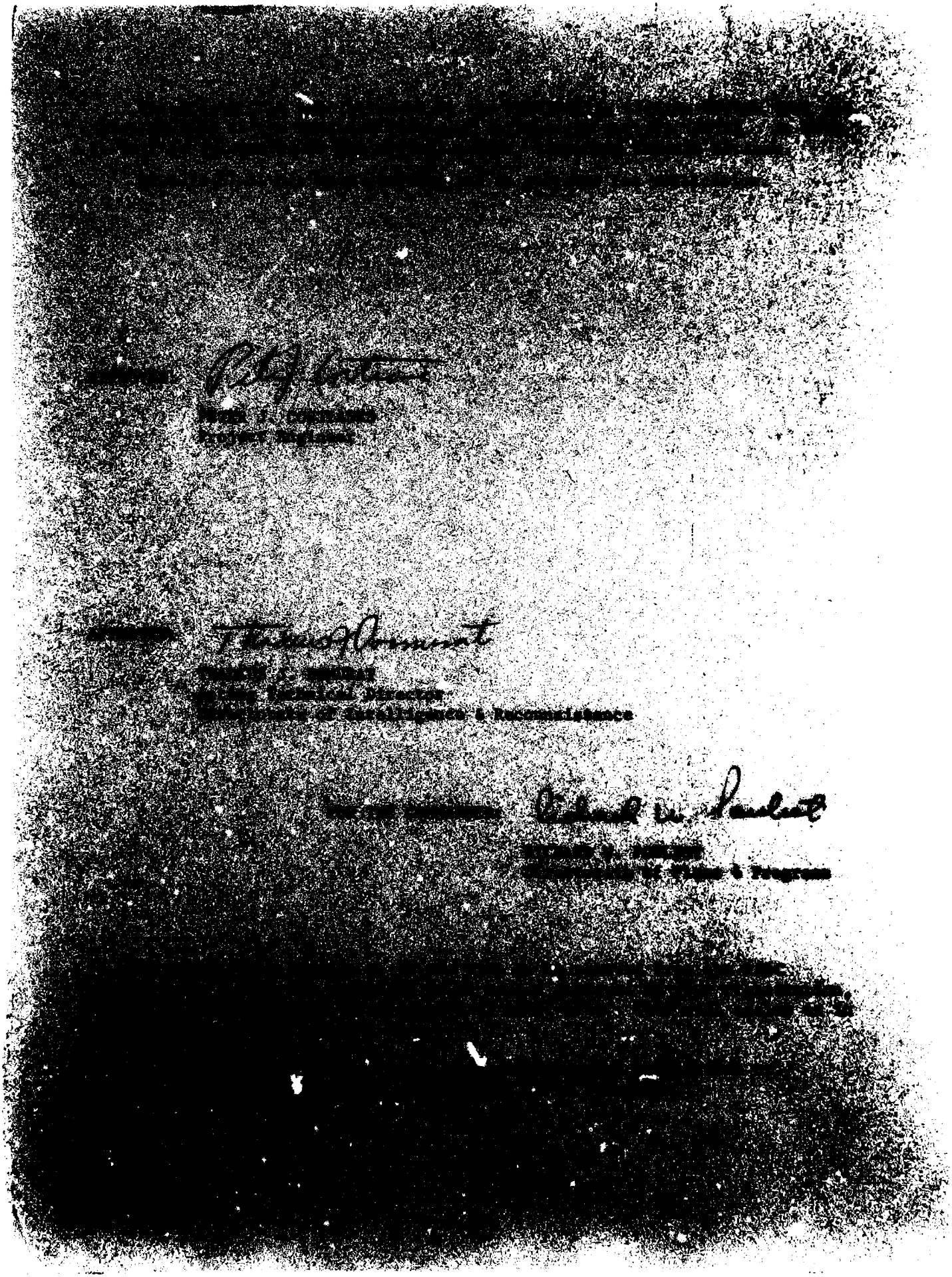
[redacted]

[redacted] IN DIGITAL SIGNAL PROCESSING OF IMAGE DATA

[redacted] Polytechnic Institute

[redacted]

[redacted]



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b RESTRICTIVE MARKINGS: N/A AD-A182 609
2a SECURITY CLASSIFICATION AUTHORITY: N/A	3 DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
2b DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A	
4 PERFORMING ORGANIZATION REPORT NUMBER(S): N/A	5 MONITORING ORGANIZATION REPORT NUMBER(S): RADC-TR-87-51
6a NAME OF PERFORMING ORGANIZATION: Rensselaer Polytechnic Institute; 6b OFFICE SYMBOL (If applicable):	7a NAME OF MONITORING ORGANIZATION: Rome Air Development Center (IRRE)
6c ADDRESS (City, State, and ZIP Code): Troy NY 12180-3590	7b ADDRESS (City, State, and ZIP Code): Griffiss AFB NY 13441-5700
8a NAME OF FUNDING/SPONSORING ORGANIZATION: Rome Air Development Center; 8b OFFICE SYMBOL (If applicable): IRRE	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F30602-82-K-0151
8c ADDRESS (City, State, and ZIP Code): Griffiss AFB NY 13441-5700	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO 62702F; PROJECT NO 4594; TASK NO 18; WORK UNIT ACCESSION NO 56

11 TITLE (Include Security Classification)

NEW DIRECTIONS IN THE DIGITAL SIGNAL PROCESSING OF IMAGE DATA

12 PERSONAL AUTHOR(S)

W.A. Pearlman, P. Burt, H. Freeman, J.W. Modestino, H. Stark

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM Sep 82 TO Oct 86	May 1987	80

16 SUPPLEMENTARY NOTATION

N/A

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Object detection and identification
12	01		restoration of photon noise limited imagery
15	04		image recovery from incomplete information (over)

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

A wide range of tasks in image processing have been treated under this contract. The categories of image processing covered are target detection, pattern recognition, image recovery from incomplete information, restoration of blurred images in additive and multiplicative noise, motion analysis with fast hierarchical algorithms, image texture modeling and discrimination, image segmentation and removal of cloud cover from aerial photographs. In all of these tasks, new, efficient, and fast algorithms have been developed, which comprise a significant contribution to the field of image processing. Within the report are descriptions of the tasks and the resulting processing algorithms and citations of publications produced under this contract.

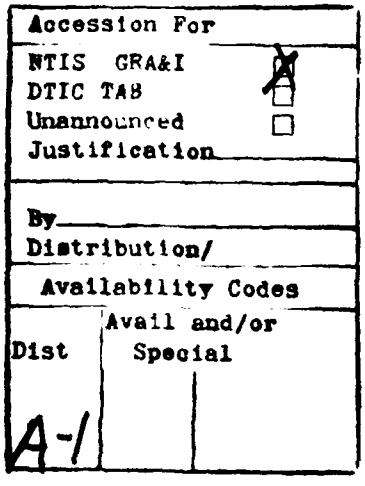
20 DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION	
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	UNCLASSIFIED	
22a NAME OF RESPONSIBLE INDIVIDUAL: Peter J. Costianes	22b TELEPHONE (Include Area Code): (315) 330-7780	22c OFFICE SYMBOL: RADC (IRRE)

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

Block 18. Subject terms (Cont'd)

motion analysis
stochastic image models
efficient algorithms for photo-interpretation



CONTENTS

0.0 Introduction to Final Report

The following is the final report on "New Directions in the Digital Signal Processing of Image Data (Image Modeling for Exploitation, Project No. 4594)" Contract Number F30602-82-K-0151. The report consists of two volumes. The first presents the work accomplished in descriptive, but summary form and lists the publications and presentations supported by the contract. This first volume fulfills the requirement of the final report on the contract. The second volume is optional and is a compilation of the published articles, theses or dissertations, and project reports produced from this contract. The first volume is divided into six major sections corresponding to the six major tasks comprising this report. Within each section are described the results of the relevant task along with its cited references and its publications.

The numbering of the compiled publications in the second volume is the same as that in the first volume for ease of reference. The accomplishments of the contract were voluminous and truly noteworthy, as evidenced by sheer magnitude of the entire report and the number of published articles and theses emanating from it.

1.0 TARGET DETECTION AND ROTATION-INVARIANT PATTERN RECOGNITION

1.1 TARGET DETECTION BY CORRELATION

In this task, we investigated the problem of scene matching with a reduced resolution reference scene. The completed work appears in the Master's degree project (thesis) report of H. Kidorf [1.3-1].

In this thesis we investigate the problem of scene matching by matched filtering when the target scene and the reference scene are of different resolutions. As is well known, the solution to the matched filter problem under additive white noise conditions is the correlation receiver. The effect on the performance of this target detector is desired under various imaging conditions: radial blurring, azimuthal blurring and with additive white noise. Additionally, we examine system performance when the reference scene resolution is reduced to that equivalent of the target.

This investigation is composed of the following steps:

1. Generation of a set of blurred test images
2. Correlation with a reference image (either high resolution or blurred)
3. Examination of peak correlation as a function of the degree of blur
4. Generation of noisy images and analysis as in Steps 2 and 3.

Blurred images are first produced in polar coordinates using a combination of the Circular Harmonic Transform (CHT) to implement azimuthal blurring, and convolution with a gaussian filter in implement radial blurring. The resultant image is then interpolated to values that comprise a cartesian grid.

After acquiring a set of images with varying radial and azimuthal blur, noise is added in various amounts. Analysis is then carried out to determine the effects of the varied parameters on the performance of image cross-correlation.

In this thesis it was found that compared to the effects of azimuthal blur, radial blur has a lesser effect on the performance of scene matching by correlation. A comparison of the rates at which radial blur and azimuthal blur lessen the peak correlation of these blurred target images with the unblurred reference image shows that radial blur degrades this performance measure 58% slower than azimuthal blurring does. Additionally, the reduction of peak correlation as a function of radial blur was found to display a threshold-like behavior, i.e. correlation of radially blurred images with greater than 65% of their high frequency content removed results in the peak correlation being reduced by the effects of blurring at a rate more than four times faster than with those images with blur levels below this threshold.

Investigation into the effect on correlation performance of adding noise to the images was quantified and data presented so that trade-offs can be considered between the addition of noise and the acceptance of increased blur. With use of this data each of the three parameters are interchangeable (though not equivalent) in considering their effect on cross-correlation so, for instance, the reduction in correlation performance due to radial blur can alternatively be thought of in terms of the addition noise or in terms of equivalent azimuthal blur.

An additional area of investigation was the investigation of the effect of performing scene matching by correlation using a reduced resolution reference image rather than matching the blurred test image with a high

resolution one. It was discovered that when matching a slightly blurred image that the degradation suffered was only slight. But, as the degree of blur of the test image increased, the performance of correlation decreased significantly.

The work in this thesis also demonstrated the ability to simulate the real world images obtained with limited resolution sensors and sampled at less than optimal sampling rates. This tool proved useful in the investigation into the effects of lessened resolution on image correlation.

1.2 ROTATION-INVARIANT PATTERN RECOGNITION

The problem of recognizing a target pattern regardless of its orientation is the subject of this task.

In two recent interesting papers by Hsu and co-workers,[1,2] the method of circular harmonic function (CHF) expansion is suggested for achieving rotation-invariant pattern recognition by either optical or digital means. This promising approach indeed does detect targets regardless of orientation but seems to require - at least based on our experience - high SNR and/or high contrast imagery. Hsu's test statistic, i.e., the function of his observation, is a scalar, a situation essentially forced on him because of a property of his algorithm.

In the paper "Rotation-invariant Pattern Recognition Using a Vector Reference" by R. Wu and H. Stark [1.3-2] we present a modification of the CHF expansion algorithm that enables us to use vector statistics. In turn, using vector statistics we can detect and locate targets in poor SNRs and low-contrast imagery without sacrificing rotation invariance. Instead of using correlation with a single harmonic and establishing an optimum

expansion center, we use multiple harmonics combined with a less-than-optimum expansion center.

In another paper "Rotation-Invariant Pattern Recognition Using Optimum Feature Extraction" by R. Wu and H. Stark [1.3-3, 1.3-4] we consider the problem of detecting a target regardless of its orientation when it is known that the target must be from one of two classes. We assume significant random intraclass variability, a complication which requires techniques from statistical pattern recognition for amelioration. The Foley-Sammon transformation for selecting optimum features from random training samples is used to solve the problem.

By combining vector rotation-invariant signatures with sequential projection onto optimum subspaces for enhancing class separability we have been able to achieve rotation-invariant recognition with a high degree of accuracy. The use of vector signatures enables correct recognition in relatively poor signal-to-noise environments where other methods might fail.[3] Optimum feature extraction enables object-class recognition when strong statistical variations exist within the object classes. We have demonstrated this property even when one object is embedded in another (as P is embedded in B).

Some of these gains are brought about at the expense of greater computational complexity. The number of harmonics N needed to recognize an object depends on the complexity of the object and is determined from tests. Thus, to a first approximation, our algorithm requires N times as many computations as Hsu's. On the other hand, we do not need to search for a proper expansion center. The derivation of the optimum subspaces for feature selection is computationally intensive but is done only once, i.e., before the actual recognition problem begins. The projection of the

signatures $[Z_j]$ onto the optimum subspaces requires only a matrix multiplication. In the worst case, only three sequential tests need to be done for separating two classes. In situations where processing time must be kept low, a high-speed array processor of a microprocessor dedicated to the recognition problems can be used.

In the two earlier papers, [1.3-2, 1.3-3] we considered the problem of how a machine-vision system might recognize an object in a two-dimensional (2-D) scene (e.g., an image) regardless of the orientation of the object. We called this problem 2-D rotation-invariant pattern recognition. There have been numerous approaches to this problem including those of Hu,[4] Smith and Wright,[5] and Borg and Hall,[6] who used moment invariants. Fourier descriptors were used by Zahn and Roskies[7] and others.[8,9] Casasent and Psaltis[10] achieved position-, scale-, and rotation-invariant 2-D recognition by combining polar Fourier and Mellin transforms. Arsenault and co-workers[1,2] have approached their problem in a novel way, using the coefficients of a circular-harmonic function (CHF) expansion for rotation-invariant recognition. Indeed it is their approach that we have extended in our work. One of the purposes of this paper is to show that scale-invariant recognition can be easily and efficiently incorporated in the CHF algorithm to yield robust rotation/scale-invariant recognition.

A more complex problem is that of having a machine recognize a three-dimensional (3-D) object. In this area most of the proposed techniques seem to be extensions of results of 2-D investigations. For example, 3-D recognition using Fourier descriptors of the boundary curve was used by Richard and Hemoni.[11] Sadjadi and Hall[12] used 3-D moments that are

invariant under position, size, and orientation changes. 3-D aircraft identification has been investigated by using Fourier descriptors[13] and moment invariants.[14]

Numerous investigations in both 2-D and 3-D machine recognition are cited in the above references. A number of papers dealing with 2-D and 3-D recognition were read at a recent meeting on machine vision; brief abstracts of this meeting are available.[15] In general, however, research in 3-D recognition illuminates two major problems: the storage of vast amounts of reference data and the computational complexity involved in establishing a match between the test object and one of the references. Therefore extracting and matching features of images rather than matching the images themselves is the trend in 3-D recognition research.

Our major objective of this paper is to describe a 3-D machine-recognition procedure that uses the robust rotation- and scale-invariant recognition ability of the CHF/Mellin-transform approach, which will be described in the next section. Of course, to understand the 3-D problem we must necessarily be more restrictive in our assumptions regarding the degrees of freedom in the positions of the test object. In this regard, the most critical assumption is that there is only one degree of rotational uncertainty in the object's position. If we think of the object as being at the origin of a 3-D Cartesian coordinate system, then we allow a random angular displacement about the y axis but not about the x and z axes. Despite this restrictive assumption, we believe that our results can be useful in a number of robot-vision situations.

The central result of this research is a new data-mapping procedure from 3-D observation space to 2-D feature space based on pseudotomography.

The 2-D feature data are organized into a signature image from which scale- and rotation-invariant features are extracted. Scale invariance is obtained by using the Mellin transform; rotation invariance is obtained by using CHF coefficients. The particular signature image that we chose has the following properties: (1) it incorporates in one 2-D image information about all the views of the object, (2) it is easily processed to obtain invariance with respect to scale, (3) it is invariant with respect to brightness changes, (4) it is easily processed to obtain invariance with respect to rotation, (5) it is invariant with respect to lateral shifts of the rotation axis about which the 3-D object is viewed, (6) it is invariant with respect to vertical translation of the object, and (7) it gave robust performance. The work appears in the paper, "Three-Dimensional Object Recognition from Multiple Views" by F. Wu and H. Stark [1.3-5].

1. Y-N. Hsu, H. H. Arsenault, and G. April, Appl. Opt. 21, 4012 (1982).
2. Y-N. Hsu and H. H. Arsenault, Appl. Opt. 21,4016 (1982).
3. R. Wu and H. Stark, "Rotation-Invariant Pattern Recognition Using a Vector Reference," Appl. Oct. 23,838 (1984).
4. M. K. Hu. "Visual pattern recognition by moment invariants," IRE Trans. Inf. Theory IT-8, 179-187 (1962).
5. F. W. Smith and M. H. Wright, "Automatic ship photo interpretation by the method of moments," IEEE Trans. Comput. C-20, 1089-1094. (1971).
6. R. Y. Wong and E. L. Hall, "Scene matching with invariant moments, "Comput. Graphics Image Process. 8, 16-24 (1978).
7. C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," IEEE Trans. Comput. C-21, 269-281 (1972).
8. G. H. Granlund, "Fourier preprocessing for hand print character recognition," IEEE Trans. Comput. C-21, 195-201 (1972).

9. E. Persoon and K. S. Fu, "Shape discrimination using Fourier descriptors," IEEE Trans. Syst. Man Cyber. SMC-7, 170-179 (1977).

10. D. Casasent and D. Psaltis, "Position, rotation and scale invariant optical correlation", Appl. Opt. 15, 1795-1799 (1976).

11. C. W. Richard, Jr., and H. Hemani, "Identification of three-dimensional objects using Fourier descriptors of the boundary curve," IEEE Trans. Syst. Man Cybern. SMC-4, 371-378 (1984).

12. F. A. Sadjadi and E. L. Hall, "Three-dimensional moment in variants," IEEE Trans. Pattern Anal. Mach. Intell. PAMI-2, 127-136 (1980).

13. T. P. Wallace and P. A. Wintz "An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptors," Comput. Graphics Image Process. 13, 99-126 (1980).

14. S. A. Dudadi, K. J. Breeding, and R. B. McGhee, "Aircraft identification by moment invariant," IEEE Trans. Comput. C-26, 39-45 (1977).

15. Advance program of 1985 Winter Meeting on Machine Vision, Lake Tahoe, Nev. (Optical Society of America, Washington, D.C., 1985).

1.3 MAJOR PUBLICATIONS

1. H. D. Kidorf, The Effects of Reduced Resolution on Scene Matching by Correlation, Master of Engineering Project Report, Electrical, Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, Troy, NY 12180, May 1984.

2. R. Wu and H. Stark, "Rotation-invariant Pattern Recognition Using a Vector Reference", Applied Optics, Vol. 23, pp. 838-840, March 15, 1984.

3. R. Wu and H. Stark "Rotation-invariant pattern recognition using a vector reference," Pattern Recognition in Practice II, E. S. Gelsema and L. N. Kanal (Editors), Elsevier Science Publishers B. V. (North-Holland), 1986, pp. 401-410.

4. R. Wu and H. Stark, "Rotation-invariant pattern recognition using optimum feature extraction," Appl. Opt. 24, 179-184 (1985).

5. R. Wu and H. Stark, "Three-Dimensional Object Recognition from Multiple Views", J. Opt. Soc. Am. A., Vol. 3, pp. 1543-1557, Sept. 1986.

6. R. Wu, Rotation and Scale Invariant Recognition of Objects by Machines, Ph.D. dissertation, Electrical, Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, Troy, NY 12180, November 1985.

2.0 IMAGE RECOVERY FROM INCOMPLETE INFORMATION

2.1 DESCRIPTION OF COMPLETE WORK

Two major journal articles [2.2-1, 2.2-2] and the Ph.D. dissertation of A. Levi [2.2-3] describe the results of the investigation of image recovery from incomplete information. We shall present the abstracts and conclusions of these works. Further details, derivations and results are found in the body of these works.

The first paper is "Signal Restoration from Phase by Projections onto Convex Sets" (POCS) by A. Levi and h. Stark [2.2-1]. In this paper we have discussed the RFP (restoration from phase) problem using a new iterative algorithm known as POCS (projections onto convex sets). The method allows from any number of a prior known image constraints to be incorporated in the algorithm provided that these can be associated with convex sets. We have discussed methods of approximately optimizing the relaxation parameters and shown thereby that a significant improvement in performance can be obtained.

We have shown experimentally that the method of POCS failed to provide a unique restoration for signals that violated the uniqueness requirements discussed by Hayes[1] and herein. Finally, we have compared POCS with the well-known HLO RFP algorithm[2] and shown that the POCS algorithm performs essentially as well as the other while ensuring strong convergence in the finite-dimensional case. A demonstration in which POCS yielded convergence while the HLO method did not was furnished.

The second paper is "Image Restoration by the Method of Generalized Projections with Application to Restoration from Magnitude" by A. Levi and H. Stark [2.2-2]. The method of projections onto convex sets can be used to

solve many problems in image restoration, e.g., restoration from phase, spectral extrapolation, and signal recovery in computer-aided tomography. However, image-restoration problems involving nonconvex constraints cannot be handled by the method of projection onto convex sets in a fashion that ensures convergence. The restoration-from-magnitude (RFM) problem is such a case. To handle the RFM as well as other nonconvex constraints, we describe an algorithm known as generalized projections and discuss its properties. When sets are nonconvex, it is possible for the algorithm to exhibit pathological behavior that is never manifest in convex projections. We introduce an error criterion called the summed-distance error (SDE) and show under what circumstances the SDE is a monotonically decreasing function of the number of iterations. Near-optimum performance of the algorithm is achieved by relaxation parameters. Comparisons with other RFM methods are furnished for synthetic imagery.

- We have found that the number of iterations required to restore an image may generally be reduced by 30% or more by using relaxation parameters compared with the method of pure projections.
- The variation of the SDE with relaxation parameter λ_2 shows the following behavior. For a small number of iterations, the minimum is relatively sharp, and relatively small deviations from the optimum value of λ_2 degrade the performance as measured by the SDE. As the number of iterations goes up, however, the minimum becomes shallower, and small deviations from the optimum value have relatively little effect. In nearly all cases, the range of the optimum value of λ_2 was in the range 1.5-3.0.

- o Unlike with the extrapolation problem, the use of the two-level constraints in the RFM problem gives much better results than those obtained without it.
- o An important observation is that the SDE $J(f_n)$ cannot be used either as a measure for the true error or to compare different algorithms. In fact, it can be shown that, for the algorithm, $J(f_n)$ is always smaller than the error $\| e_n \| = \| f - f_n \|$.
- o In many cases we observed tunnel behavior for large interation number n [i.e., the change in f_n from iteration to iteration and the corresponding change in $J(f_n)$ becomes negligible]. The poor restorations with relatively low $J(f_n)$ after 1000 iterations exhibited in some of the examples support the assumption that traps exist. These examples also demonstrate the importance of detecting traps and tunnels and of finding ways of getting out of them by changing the algorithm.

1. M. H. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform," IEEE Trans. Acoust. Speech Signal Process. ASSP-30, 140-154 (1982).
2. M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Signal reconstruction from phase or magnitude," IEEE Trans. Acoust. Speech Signal Process. ASSP-28, 672-680 (1980).

The abstract of A. Levi's Ph.D. dissertation [2.2-3] follows.

This dissertation deals with the problem of restoring images from incomplete information by the method of alternating projection onto convex sets (POCS) and its extension: the method of generalized projections (MGP). the investigation discussed herein are both theoretical and experimental.

We demonstrate the weak convergence property of the POCS algorithm when the relaxation parameters (RP) are varied from iteration to iteration. We then demonstrate how the RP's can be optimized to accelerate convergence. We illustrate the utility of these theoretical results by restoring a finite-size (space-limited) image from the phase of its Fourier transform. This problem is the well-known restoration from phase problem. We compare our result with the standard solution (the Hayes-Lim-Oppenheim, or HLO, method). The superior convergence properties of the POCS method over the HLO method is demonstrated by several examples in which the HLO method diverges while the POCS method does not.

The MGP and its application to problems which involve non-convex type constraints, is derived. The MGP is a two-step algorithm which possesses the property of **set-distance-reduction** for certain values of the RP's. We show that this property does not extend to algorithms with more than two steps (or two operators) in one iteration, but the restriction to two steps is not very restrictive in practice because several image constraints can be combined into one set. A performance measure called **summed distance error** (SDE) is defined and is shown to be useful for monitoring and controlling the algorithm. Pathological phenomena called **traps** and **tunnels** are defined and shown to occur in the MGP when non-convex sets are involved. They are responsible for the observed convergence to non-valid solutions or very slow convergence to valid ones.

The MGP algorithm is applied to the problem of restoring a signal when we are given only the magnitude of its Fourier transform (the restoration from magnitude (RFM) problem). We show that the MGP method incorporates other algorithms for the RFM problem such as the Gerchberg-Saxton (GS) method and Fienup's output-output algorithm. Methods of optimizing the RP's

on a per-step and per-cycle basic are described and the per-cycle method is shown to significantly improve the convergence rate. The method of MGP is shown to be useful for detecting traps and tunnels and for possibly getting out of them. Experimental demonstrations showing the effects of noise on the restoration are also furnished.

We close the dissertation by proposing several unresolved research problems that warrant further investigations.

2.2 MAJOR PUBLICATIONS

1. A. Levi and H. Stark, "Signal Restoration from Phase by Projections onto Convex Sets", J. Opt. Soc. Am., Vol. 73, pp. 810-822, June 1983.

2. A. Levi and H. Stark, "Image Restoration by the Method of Generalized Projections with Application to Restoration from Magnitude", J. Opt. Soc. Am. A., Vol. 1, pp. 932-943, Sept. 1984.

3. A. Levi, Image Restoration by the Method of Projections with Applications to the Phase and Magnitude Retrieval Problems, Ph.D. dissertation, Electrical, Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, Troy, NY 12180, December 1983.

3. POST DETECTION PROCESSING OF PHOTON-LIMITED, BLURRED IMAGES

3.0 INTRODUCTION

The focus of the investigation of means to restore photon-limited and blurred images was to develop new techniques where the image statistics were unknown and the processing could be realized in real-time with dedicated hardware. We describe here three new techniques which gave very good results with fast processing.

3.1 TWO STEPS ADAPTIVE, ROBUST ESTIMATOR

The first is a two-step procedure which restores low light-level images regarded by a linear space-invariant blur. The presentation of the results of the research occurred at the SPIE Annual International Symposium on Optics and Electro-Optics in San Diego, California in August 1984 with publication of a full article in the symposium's Proceedings [1]. As the blurred image is incident on the photo-detector array, the first step attempts to estimate the blurred image from the collection of photocounts from the array. The optimum linear minimum mean squared estimator (LMMSE) is obtained by solving an equation by Grandell [2] for estimating the intensity in a doubly stochastic Poisson field. The error of the optimum LMMSE estimator is orthogonal to the photocount process. Hence this error may be treated as additive and uncorrelated noise in the next step, which forms an estimate of the true object from the first steps estimate of the incident (blurred image) intensity. It is accomplished through an application of the constrained LMMSE (deconvolution) method set forth by Hunt [3]. An adaptive windowing technique generates sample statistics for the two-step estimator.

Hence the method requires no prior knowledge of image statistics to accomplish recovery of the object. The processing is very fast because the required sample statistics for the estimators and the window size logic are calculated recursively and the estimates are linear ones of point intensities. A full description of the estimators and their derivations is presented in the aforementioned paper [1]. Here we shall just describe some of the simulation results.

The system for generating a blurred, photon limited image and the two-step restoration procedure are illustrated in Figure 1. The linear blur matrix H acting upon the original image $\underline{f}$ has elements $h(i,j) = \frac{1}{JK}$, $i=0,1,\ldots,J-1$ and $j=0,1,\ldots,K-1$. The Poisson noise generator for a given parameter η produces the array of photocounts $\underline{g} = g(i,j)$ which are each scaled by η^{-1} for display of the photon-limited, blurred image $\underline{g_d}$. The received or input signal-to-noise ratio is defined to be

$$SNR_i = \frac{\langle (f-\langle f \rangle)^2 \rangle}{\langle (g_d-f)^2 \rangle}$$

where $\langle x \rangle$ denotes the spatial sample average over the array x. The restored or output signal-to-noise SNR_o is defined as above with the estimate f replacing g_d.

The two test images used in our simulations are the face images of a man and woman. For each of these images, Poisson noise degradation with different parameter η and degrees of lineal motion blur were simulated and then restored. Shown in Figure 2 is one typical example of a degraded image

with Poisson noise parameter η=0.125 and 4x4 (J=K=4) blur and its restoration. With blur, the restored image is not as sharp and appears slightly more noisy with an SNR improvement of 7.3 dB. The improvement with blur is higher than without blur, because the blurring procedure produces further degradation from the original image and the estimator partially deblurs it.

Comparison with other methods is complicated by the use of different test images and degrees of Poisson noise and by inconsistencies in definitions of signal-to-noise ratios. Another factor is the amount of processing or degree of computational complexity involved in different schemes. For example, Lo and Sawchuk used noisier test images and a different definition of signal-to-noise ratio [4]. Kasturi [5] compares several kinds of estimators on less noisy images than ours and his SNR calculations include the image means. The better estimators give good perceptual results, but use prior statistical knowledge and significant processing. The estimator here does compare favorably with these efforts, however, especially considering the robustness and adaptivity of the method and the relatively small processing requirements.

3.2 MINIMUM ERROR, MAXIMUM CORRELATION ESTIMATOR

One of the deficiencies of current image estimators is that the estimation error or noise, although uncorrelated with the data image has positive correlation with the true object. In principle, therefore, the estimation noise can be further processed to extract more information about the object. To remove this deficiency, a new estimator is proposed which attempts simultaneously to maximize the correlation between the estimate

(or, equivalently, minimize the correlation between the error and the object) and minimize the mean squared error. It turns out that this is an optimization problem with competing objectives. The solution is a compromise between the usual minimum mean squared error (Wiener) estimator and an inverse filter.

A simple example to illustrate the need of the double criterion is as follows. Consider a scalar case where the observed random variable Y equals the sum of the object variable (signal) X and independent noise N, i.e., $Y=X+N$ and the signal and noise variances, σ_x^2 and σ_n^2 are equal. Assume that the random variables have zero mean. Consider two possible estimates $\hat{X}$ of X, $\hat{X}=Y$ and $\hat{X}=0$. The mean squared error of the first estimate is σ_n^2 and that of the second is σ_x^2. Hence their estimation errors are equal. From a perceptual point of view, it is certainly better to see $\hat{X}=Y$ than $\hat{X}=0$, which is no information at all. One quantitive measure of the difference between the two estimates is the correlation of their error with the signal. For $\hat{X}=Y$ the correlation between the error and the signal X is zero, whereas, for $\hat{X}=0$, the correlation between the error and the signal is σ_x^2. Hence, for equal error, we choose the estimator which has the smaller correlation between its error and the signal.

The case of signal plus uncorrelated noise can be visualized vectorially as shown in Figure 3. When we select a linear estimator $\hat{X}=aY$, the variation of the constant a between 0 and 1 moves the estimator from the tail to the tip of the vector Y from a mean squared error of σ_x^2 to σ_n^2 with an interim minimum of $\varepsilon_{min}^2 = \sigma_x^2\sigma_n^2(\sigma_x^2+\sigma_n^2)^{-1}$ when the error vector is

perpendicular to Y. The correlation of the error with the signal moves from σ_x^2 at the tail to zero at the tip. Therefore, if one moves from the minimum error point toward the tip of Y, there is a trade-off of increased error for increased error-signal correlation. In Figure 4 the mean-squared error d(a) is plotted against the error-signal correlation c(a) as a varies from 0 to 1. The plot shows that a maximal trade-off point of error for correlation may be realized, depending on definitions of allowable range.

The convexity of the c-d locus versus a was proved and the optimization problem of minimizing c(a) and d(a) simultaneously over a takes the form of

$$\min_a \left[w_1 c(a) + w_2 d(a)\right]$$

with weighting coefficients w_1 and w_2. The choices of w_1 and w_2 are subjective, but reasonable choices are the reciprocals of the ranges of the corresponding weighted variables. The ultimate objective function is

$$f(a) = \begin{cases} \sigma_n^2 c(a) + \sigma_x^2 d(a) & , \quad \sigma_x^2 \geq \sigma_n^2 \\ \sigma_x^2 c(a) + \sigma_n^2 d(a) & , \quad \sigma_x^2 < \sigma_n^2 \end{cases}$$

with the minimizing $a=a_o$,

$$a_o = \frac{1}{2} \frac{\sigma_x^2}{\max[s_x^2, \sigma_n^2]} + \frac{1}{2} \frac{\sigma_x^2}{\sigma_x^2 + \sigma_n^2}$$

resulting d and c of

$$d(a_o) = \begin{cases} \left(1+\dfrac{\sigma_n^2}{4\sigma_x^2}\right)\varepsilon_{min}^2 & \sigma_x^2 \geq \sigma_n^2 \\ \left(1+\dfrac{\sigma_x^2}{4\sigma_n^2}\right)\varepsilon_{min}^2 & \sigma_x^2 < \sigma_n^2 \end{cases}$$

$$c(a_o) = \begin{cases} \varepsilon_{min}^2/2 & \sigma_x^2 \geq \sigma_n^2 \\ \left[1-\dfrac{1}{2}\left(\dfrac{\sigma_x^2}{\sigma_n^2}\right)\right]\varepsilon_{min}^2 & \sigma_x^2 < \sigma_n^2 \end{cases}$$

with $\varepsilon_{min}^2 = \sigma_x^2\sigma_n^2\,(\sigma_x^2+\sigma_n^2)^{-1}$, the minimum mean squared error. When the signal-to-noise ratio σ_x^2/σ_n^2 is either very large or very small, the resulting mean squared error $d(a_o)$ is close to the minimum. However, when neither case applies, a small additional error is traded for a larger amount of error-signal correlation. This work and all subsequent derivations and results in image restoration with this double criterion appear in the Ph.D. dissertation of Woo-Jin Song [6] and [8]. In his dissertation Song also extended the previous scalar random variable case to stationary continuous and discrete image fields. The resulting digital linear restoration filter characteristic, which we call the minimum error, maximum correlation (MEMC) filter is

$$A(m,n) = \frac{1}{2} \frac{H^*(m,n)\, S_x(m,n)}{\max\left[\, |H(m,n)|^2 S_x(m,n),\ S_n(m,n)\right]} +$$

$$\frac{1}{2} \frac{H^*(m,n)\, S_x(m,n)}{|H(m,n)|^2 S_x(m,n) + S_n(m,n)} \qquad \begin{matrix} m=0,1,\ldots,M-1 \\ n=0,1,\ldots,N-1 \end{matrix}$$

where $H(m,n)$ is the sampled point spread function of the linear blur and $S_x(m,n)$ and $S_n(m,n)$ are the discrete power spectra of the object and noise, respectively.

Blurred images in additive Gaussian noise were simulated by the computer and restored through the above filter. The point spread functions of the blur were Gaussian with 8x8 support and lineal with 4x4 support. In Figures 5 and 6 are shown comparative results of estimation with the double-criterion (MEMC) and the Wiener Filters using lineal blur and signal-to-noise ratios of 13 and 10 dB. It is evident that the MEMC filter produces the sharper image at the expense of more "spotting". At the lower signal-to-noise ratio, the sharper image even with more noticeable "spotting" is preferred by most observers than the fuzziness in the Wiener filtered image. In Figure 7 are the same filter comparisons with no blur and 0 dB SNR. Here the relative sharpness of the double-criterion restored image is even more pronounced.

3.3 ADAPTIVE WINDOWING AND NONLINEAR FILTERING

Traditional and new estimators which derive their spectral estimates by averaging over regions of an image produce restorations of noisy images which are noisy in the flat areas due to insufficient smoothing and somewhat

fuzzy in the edge regions due to attempted noise smoothing. The problem arises partly because the regions over which the spectral parameters are estimated (the so-called analysis windows) contain both flat areas and edges and are appropriate for an average of both, but not separately for either. If the analysis window contains elements truly representative of the current picture element, then the estimate of the true object intensity at the point would be improved. A new scheme for adaptive windowing was tried in conjunction with subsequent non-linear filtering. We will describe the technique briefly and show some of the results, which are truly impressive.

Let us assume for the moment one-dimensional data where the degradation is additive, uncorrelated stationary noise of variance σ^2. Assume a window of length L_K has been set for the k^{th} data element. An activity index S_k is computed from the sample variance in that window. If $S_k < T_k$, a threshold inversely proportional to L_k, then the next window $L_{k+1} = L_k + 2$ unless it has reached a preset maximum. If $S_k \geq T_k$, then the window length is decremented by two. The sample mean m_k and sample variance V_k are calculated within the window straddling the k^{th} image element of intensity y_k to produce the signal estimate $\hat{x}_k$ by

$$\hat{x}_k = m_k + \frac{\max[0, V_k - \sigma^2]}{V_k} (y_k - m_k).$$

Although the estimate appears linear in its form, it is nonlinear because of the nonlinear dependence on V_k, which in turn is a nonlinear function of the

data. The mean, variance, window length, and error can be computed recursively, thereby reducing the computational requirements. The extension of the scheme to two dimensions is straightforward, except that the shape of the analysis window needs specification. For simplicity, square windows are chosen.

This adaptive, nonlinear estimator was applied successively to a noisy one-dimensional simulated waveform, a scan line of a noisy image, noisy images using one-dimensional windows, and images using square windows. In all cases, the edge or detail features are preserved and the flat contrast areas show almost imperceptible noise. In fact, the image restorations are remarkably sharp and clean. One example is presented here to give evidence to those assertions. Shown in Figure 8 is the original mans face image and the same image with additive, white Gaussian noise of variance $\sigma^2 = 489$. In Figure 9 appear restorations by the new adaptive window, nonlinear filter and the minimum mean-squared error or Wiener filter. Note the clarity and sharpness of the new filter's restoration compared to that of the Wiener filter. It is also surprising that the mean-squared error of the Wiener filter is smaller (82.22) than that of the new filter (87.76). Such a result is convincing evidence that mean squared error is not a good measure of visual error. Details and results appear in [9] and the Ph.D. dissertation of W. J. Song [6].

The adaptive window algorithm can also be utilized in other estimation paradigms . In this regard, we also incorporated it into the two step estimation procedure for blurred, photon-limited images described earlier. The estimate b_k of blurred intensity b_k at the k^{th} image point from the number of photodetector counts g_k is

$$b_k = \begin{cases} m_k/\lambda & , \text{ if } V_k = 0 \\ \dfrac{m_k^2 + \max\,[0,\, V_k - m_k]\, g_k}{\lambda \max\,[V_k,\, m_k]} & , \text{ otherwise} \end{cases}$$

where λ is the Poisson rate parameter, and m_k and V_k are the estimates of the mean and variance, respectively, of the image at point k taken in the analysis window. The mean-squared error, used as the noise in constrained deconvolution of the blur in the second step, is

$$d_k = \begin{cases} m_k/\lambda^2 & , \text{ if } V_k = 0 \\ \dfrac{m_k \max\,[0,\, V_k - m_k]}{\lambda^2 V_k} & , \text{ otherwise} \end{cases}$$

The derivations of these formulas and the subsequent results appear in Song's dissertation [6]. A sample of these results appear in Figure 10. In Fig. 10(a) is a test image degraded by 4x4 lineal motion blur and corrupted by Poisson noise with η=0.5. The result of the first-step estimation of the blurred image appears in Fig. 10(b). The blurred signal to noise ratio (SNR) increased by 6.6 dB from 10.9 to 17.5 dB. The second step constrained deconvolution of the blurred image plus the estimation noise appears in Fig. 10(c) for one value of a constraint parameter. There is considerable improvement in visual quality from both steps of the restoration procedure. It is evident that adaptive windowing produces superior reconstructions than the fixed window algorithm given in Section 3.1.

3.4 PRESENTATIONS AND PUBLICATIONS

3.4.1 CONFERENCE PRESENTATIONS

1. W.A. Pearlman and W-J. Song, "A Robust Method of Restoration of Photon-Limited, Blurred Images", SPIE Annual International Technical Symposium on Optics and Electro-Optics, Applications of Digital Image Processing VII, San Diego, California, August 20-24, 1984.

2. Woo-Jin Song and W.A. Pearlman, "Image Restoration with a Minimum Error, Minimum Correlation Estimator", Annual Meeting of the Optical Society of America, Washington, D.C., October 14-18, 1985.

3. Woo-Jin Song and W.A. Pearlman, "A Minimum-Error, Minimum Correlation Filter for Images", Applications of Digital Image Processing IX, SPIE's 30th Annual International Technical Symposium on Optical and Optoelectronic Applied Sciences and Engineering, San Diego, California, August 18-22, 1986.

4. Woo-Jin Song and W.A. Pearlman, "Restoration of Noisy Images with Adaptive Windowing and Nonlinear Filtering", Visual Communications and Image Processing, SPIE's Fiber LASE '86, Fiber Optics, Optoelectronics and Laser Applications in Science and Engineering, Cambridge, MA, September 15-16, 1986.

3.4.2 PUBLICATIONS

(a) Articles

1. W.A. Pearlman and W-J. Song, "A Robust Method for Restoration of Photon-Limited, Blurred Images", Proceedings of the SPIE, Vol. 504, Applications of Digital Image Processing VII, August 1984, pp. 270-276.

2. W-J. Song and W.A. Pearlman, "Image Restoration with a Minimum Error, Minimum Correlation Estimator", Abstract in J. Opt. Soc. Am. A, Optics and Image Sciences, 2 (13), December 1985, pp. P38-P39.

3. W-J. Song and W.A. Pearlman, "A Minimum-Error, Minimum-Correlation [illegible]ter for Images", Proceedings of the SPIE, Vol. 697, Applications of Digital Image Processing IX, August 1986.

4. W-J. Song and W.A. Pearlman, "Restoration of Noisy Images with Adaptive Windowing and Nonlinear Filtering", Proceedings of the SPIE, Vol. 707, Visual Communications and Image Processing, September 1986.

(b) Theses and Reports

5. Mark J. Morisi, Image Restoration in the Presence of Poisson Noise, Masters' Project Report, Electrical Computer and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, New York, May 1985.

6. Woo-Jin Song, New Robust Estimators for Image Restoration, Ph.D. dissertation, Electrical, Computer and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, New York, August 1986.

3.5 REFERENCES

[1] W.A. Pearlman and W-J. Song, "A Robust Method for Restoration of Photon-Limited, Blurred Images", Proceedings of the SPIE, Vol. 504, Applications of Digital Image Processing VII, August 1984, pp. 270-276.

[2] J. Grandell, "A Note of Linear Estimation of the Intensity in a Doubly Stochastic Random Field", J. Appl. Prob., Vol. 8, pp. 612-614, 1971.

[3] B.R. Hunt, "The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer", IEEE Trans. on Computers, Vol. C-22, pp. 805-812, 1973.

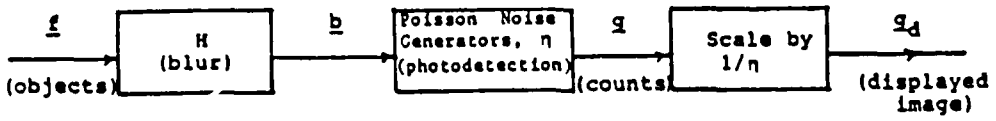
[4] C.M. Lo and A.A. Sawchuk, "Nonlinear Restoration of Filtered Images with Poisson Noise", Applications of Digital Image Processing III, SPIE Vol. 207, pp. 84-94, 1979.

[5] R. Kasturi, Adaptive Image Restoration in Signal-Dependent Noise, Ph.D. Dissertation, Texas Tech University, 1982.

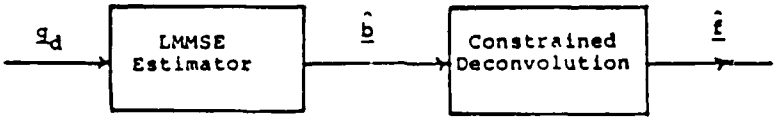
[6] W-J. Song, New Robust Estimators for Image Restoration, Ph.D. dissertation, ECSE Dept., Rensselaer Polytechnic Institute, Troy, New York, August 1986.

[7] W-J. Song and W.A. Pearlman, "A Minimum-Error, Minimum-Correlation Filter for Images", Proceedings of the SPIE, Vol. 697, Applications of Digital Image Processing IX, August 1986.

[8] W-J. Song and W.A. Pearlman, "Restoration of Noisy Images with Adaptive Windowing and Nonlinear Filtering", Proceedings of the SPIE, Vol. 707, Visual Communications and Image Processing, September 1986.



(a) Simulated received image



(b) Two-step estimator

Figure 1: **System for simulation and restoration of photon-limited, blurred mages.**

Figure 2: Degraded and restored Man images

(A) Noisy image, no blur η=.125, SNR_i=7.9dB

(B) Restoration of (A). SNR_r=14.5dB

(C) Noisy image, 4x4 blur η=.125, SNR_i=6.5dB

(D) Restoration of (C), SNR_r=13.8dB

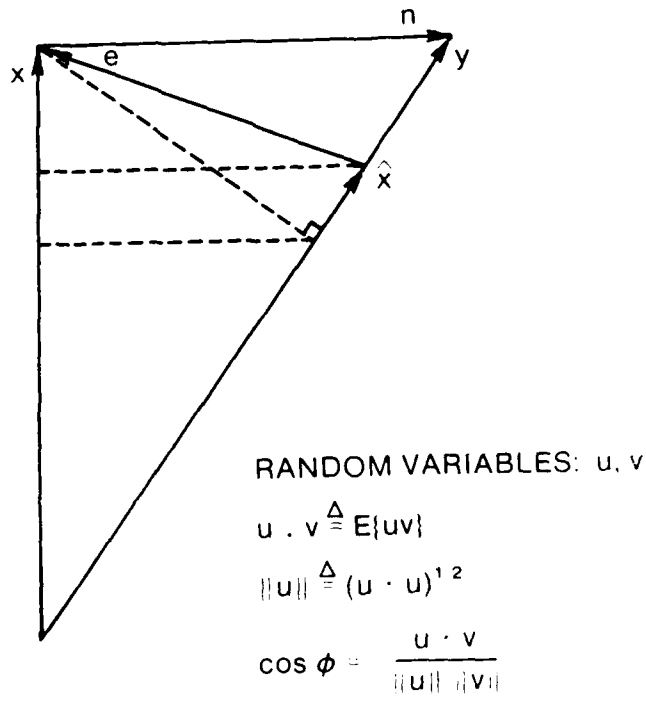


Figure 3: Vector representation of y=x+n and estimator $\hat{x}$=ay with x and n orthogonal

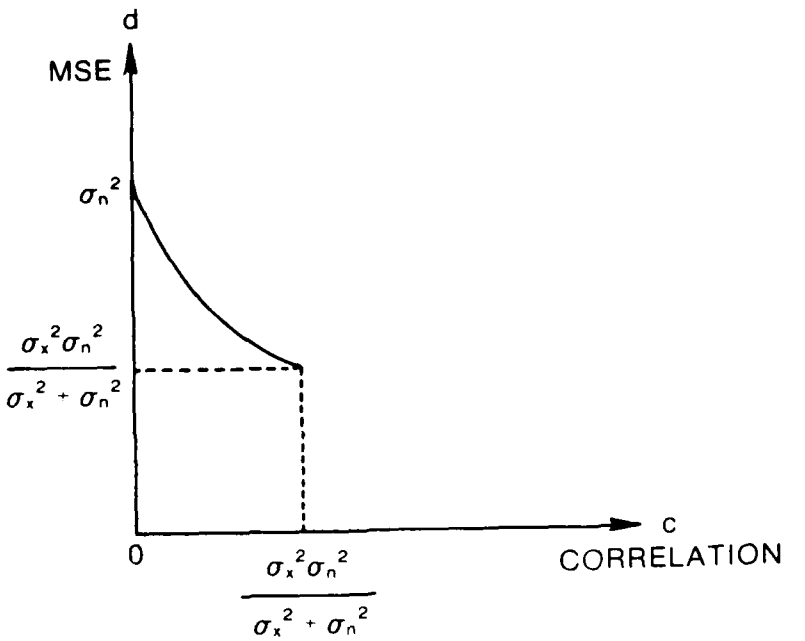


Figure 4: Locus of mean-squared error d versus error-signal correlation c as estimator parameter a in $\hat{x}=ay$ varies

(a)

(b) (c)

Figure 5: Restoration of image degraded by 4x4 lineal motion blur and additive white Gaussian noise of variance $\sigma_n^2=54$ (SNR=17dB): (a) blurred, noisy image; (b) image restored by Wiener filter; (c) image restored by MEMC filter.

(a)

(b) (c)

Figrue 6: Restoration of image degraded by 4x4 lineal motion blur and additive white Gaussian noise of variance σ_n^2=135 (SNR=13dB): (a) blurred, noisy image; (b) image restored by Wiener filter; (c) image restored by MEMC filter.

AWGN: SNR = 0 dB, no blur.

Figure 7: Simulated noisy image and restorations by Wiener and MEMC filters. Top is noisy image, bottom left is Wiener, and bottom right is MEMC restoration.

(a) (b)

Figure 8: Original and noisy mans face image.

(a) Original "mans face" image

(b) Image in additive noise of $\sigma^2 = 489$ (S/N=10dB).

(a) (b)

Figure 9: Comparison of restorations of image in Figure 8(b).
(a) Filtered "face" image by 2-D adaptive filter.
(b) "Face" image restored by Wiener Filter.

(a)

(b)

(c)

Fig. 10: (a) Photon-limited image with 4x4 motion blur, (λ=0.5)
(b) Estimated image from first step (before deblurring)
(c) Final restored image after second step.

4.0 FAST HIERARCHICAL ALGORITHMS FOR MOTION ANALYSIS

The results of this task are documented in detail in a Master's degree thesis [4.3-1], and a Masters degree project report [4.3 3]. The former presents a comparative study of correlation measures used in motion analysis and then proposes a multi-step approach. The latter report presents a method of motion representation based on image segmentation using pyramidal data structures and shows how the algorithm work for different image sequences. Excerpts from these reports follows.

4.1 CORRELATION AND THE MULTI-RESOLUTION, FLOW-THROUGH ALGORITHM

Correlation is commonly used in motion analysis to discover the correspondence among local image patterns in a sequence of image frames. However a variety of such measures can be employed and these may differ widely in performance and in computational cost. In this first half of this study [4.1-1] five basic correlation measures are compared. Performance is determined empirically as a function of image content (spectrum) and in the presence of various types of image degradation.

A multi-step approach to motion analysis is proposed in the second half of the study. Decomposition of the analysis into distinct steps and independent channels means that the computations may be performed in parallel pipelined hardware. Data is said to "flow-through" the system because analysis is uniform: a fixed sequence of operations is applied to all images and to all positions within each image, independent of image content.

4.1.1 PERFORMANCE OF LOCAL CORRELATION MEASURES

Correlation between local regions of two or more images may be used to detect pattern displacements, and hence object motion. However correlation can yield incorrect pattern matches. Of the five correlation measures examined, direct correlation consistently gave the highest error rates. Performance was particularly poor when the mean value of a local pattern was large compared to its variance, or when patterns to be matched differed in mean or contrast. Variance normalized correlation consistently gave the lowest error rates. No errors were obtained with this measure in the absence of relative image distortion. However, in the presence of noise its performance was only slightly better than other measures. Since the other measures require considerably less computation than variance normalized correlation these may be preferred under appropriate image conditions. Correlation following Laplacian filtering is particularly effective when the image contains high frequency components and any noise is restricted to low frequencies. Even the binary correlation performs well under these conditions. However both measures are very susceptible to high frequency noise. Therefore, the Laplacian correlation appears to be the most attractive of five measures studied. However this measure should not be used in the presence of high frequency noise.

4.1.2 PERFORMANCE OF MMF

We have described a system in which motion analysis is carried out in three distinct steps. First, each image in a motion sequence is passed through a set of filters to obtain a corresponding set of band-pass copies.

Second, a optimum nine "motion channels" are formed within each frequency band from original twenty-five channels, each "tuned" to motion in a particular direction and velocity. Finally, the outputs of these simple channels are combined in groups of three to form four compound motion channels, each tuned to a particular orientation, the outputs of the four orientation channels are compared and combined into a single motion estimate for each sample position within the frequency band. Each group of three channels has two outputs, a velocity estimate and a confidence measure. By comparing channel outputs, it is possible to determine whether motion is within textured image region, near a prominent image contour, or near an occluding edge. Channel outputs are combined in the manner appropriate for each of these cases in step 3.

4.1.3 CONCLUDING REMARKS

As demonstrated by the results, the combination of Laplacian pyramid structures together with the MMF strategy has proven to be a powerful technique for motion analysis. The power of the technique developed in the study lies in its simplicity, computational efficiency and its high accuracy.

4.2 MULTI-RESOLUTION SPLIT AND LINK SEGMENTATION ALGORITHM

The need for representing motion in digital images is motivated by applications in transmission of video and in artificial intelligence. In the first type of application, bandwidth compression in the image processing sense is sought; we attempt to transmit less data to communicate the

required information about the image. In the second type of application, we attempt to give the computer a primitive understanding of how objects move and an ability to track this motion without necessarily recognizing the objects. A method of motion representation based on image segmentation which lends itself to these applications is presented in this work [4.1-3] along with results from its software implementation.

The segmentation scheme presented here gives excellent results for a variety of images. The segments in the image are represented by links and segment values. The information provided by these links and segment values is contained in a pyramid structure which is suitable for analysis. Our motion algorithm uses the same representation for an image as the segmentation algorithm and is capable of updating the links and segment values given an estimate of displacement without the necessity of obtaining an entirely new segmentation. If we are willing to accept a processed image with quality not quite suitable for viewing but suitable for analytic purposes, then we have met our goals for representing motion.

4.3 PRESENTATION AND PUBLICATIONS

1. C. Yen, Local Correlation Measures Study and "Multi-Resolution, Flow-Through" Algorithm for Motion Analysis, M. Sc. Thesis, Electrical, Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, Troy, NY 12180, August 1983.
2. P. Burt, "Multi-Resolution Flow-Through Motion Analysis", Proceedings of Computer Vision and Pattern Recognition Conference, Washington, D.C., July 1983.

3. S. Schrier, A Multi-Resolution Split and Link Algorithm for Image Segmentation with Applications in Motion Representation, M.E. Project, Electrical, Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, Troy, NY 12180, May 1984.

5. TEXTURE MODELING AND DISCRIMINATION

5.0 INTRODUCTION

The focus of this task has been the development of stochastic models for texture and the related problems of developing and evaluating various texture classification and discrimination algorithms. Motivation for this work has been the following important image processing problem: segment an image field into disjoint regions which may possess the same average gray level but differ in the spatial distribution of gray levels. These two characteristics of an image are generally referred to as tone and texture, respectively. Our interest has been concerned with the latter. While somewhat elusive to define precisely, the most generally accepted definition of texture at present is that it consists of a basic local order or quasi-homogeneous pattern which is repeated in a "nearly" periodic manner over some image region which is large relative to the size of the local pattern.

Our work in this area can be conveniently classified into three broad categories:

1.) Develop useful stochastic texture models and characterize their properties.

2.) Develop model-based texture classification/discrimination algorithms on the basis of statistical decision theory concepts.

3.) Evaluate and characterize the performance of various texture classification/discrimination algorithms under identical conditions.

We will discuss each of these areas separately.

5.1 STOCHASTIC TEXTURE MODELS

We have considered several classes of stochastic texture models including:

1.) 2-D autoregressive moving average (ARMA) processes.

2.) 2-D random tessellation processes.

3.) The 2-D random grain model.

4.) 2-D Markov random fields (MRF's).

In what follows we will describe our work in each of these areas separately.

5.1.1 2-D ARMA MODELS

These models have been used extensively in one form or another in a variety of image processing applications although a comprehensive systematic study of this class has never been provided. Specifically, the 2-D ARMA process is generated recursively according to

$$S_{i,j} = \sum_{(k,\ell)\in \mathcal{D}'_a} a_{k,\ell} S_{i-k,j-\ell} + \sum_{(k,\ell)\in \mathcal{D}_b} b_{k,\ell} W_{i-k,j-\ell} \qquad i,j \geq 1, \tag{5.1}$$

where $\{W_{i,j}\}$ is a 2-D i.i.d. sequence and both[+] $\mathcal{D}_a$ and $\mathcal{D}_b$ are specified subsets of a non-symmetrical half plane $Z=\{k\geq 0, \ell \geq 0\} \cup \{k \geq 1, \ell \leq -1\}$. Assuming a row-by-row raster scanning pattern, the sequence $\{S_{i,j}\}$ is then generated as a causal filtering operation over past outputs and present and past inputs. A special case of (5.1) is the 2-D separable Gauss-Markov model with

$$S_{i,j} = \rho_1 S_{i-1,j} + \rho_2 S_{i,j-1} - \rho_1 \rho_2 S_{i-1,j-1} + W_{i,j}, \tag{5.2}$$

where $0 \leq |\rho_i| \leq 1$, i=1,2 represent horizontal and vertical pixel-to-pixel correlation coefficients, respectively, and $\{W_{i,j}\}$ is an i.i.d. zero-mean

---

+ The prime, as in $\mathcal{D}'_a$ in (2.1), is intended to denote exclusion of the point (0,0).

Gaussian sequence with common variance $\sigma_w^2 = \sigma_s^2(1-\rho_1^2)(1-\rho_2^2)$ defined in terms of the variance σ_s^2 of the assumed stationary sequence $\{S_{i,j}\}$. Similarly, it's possible to define a 2nd-order separable 2-D Gaussian process as

$$S_{i,j} = \sum_{k=0}^{2} \sum_{\ell=0}^{2} a_{k,\ell} S_{i-k,j-\ell} + W_{i,j}, \tag{5.3}$$

where again $\{W_{i,j}\}$ is an i.i.d. zero-mean Gaussian sequence. The corresponding coefficients can then be expressed in terms of pole locations in the vertical and horizontal spatial frequency planes as illustrated in Table 5.1. More specifically, r_i and θ_i, i=1,2, are the radial and angular positions, respectively, of the separable discrete system transfer functions, $H_i(z)$, generating $\{S_{i,j}\}$. This is illustrated in Fig. 5.1.

It can be shown that the input sequence $\{W_{i,j}\}$ in the 2nd-order case must possess variance

$$\sigma_w^2 = \sigma_s^2 \left(\frac{1-r_1^2}{1+r_1^2}\right)\left(\frac{1-r_2^2}{1+r_2^2}\right)(1-2r_1^2\cos 2\theta_1+r_1^4)(1-2r_2^2\cos 2\theta_2+r_2^4). \tag{5.4}$$

For example, if either $H_1(z)$ or $H_2(z)$ possesses a pole on the unit circle then $\sigma_w^2=0$. We assume, of course, that initial conditions are chosen to result in stationary conditions.

Due to the separability, the 2-D autocorrelation sequence is given by

$$R(m,n) = \sigma_w^2 R_1(m) R_2(n), \tag{5.5}$$

where

$$R_i(k) = \frac{1}{2\pi j}\oint_{C^1} H_i(z)H_i(z^{-1})z^{k-1}dz; \quad i=1,2. \tag{5.6}$$

In particular, it's easily seen that the correlation coefficients in the vertical and horizontal directions, respectively, are now given by

coefficient	value
$a_{0,1}$	$2r_2\cos\theta_2$
$a_{0,2}$	$-r_2^2$
$a_{1,0}$	$2r_1\cos\theta_1$
$a_{1,1}$	$-4r_1r_2\cos\theta_1\cos\theta_2$
$a_{1,2}$	$2r_1r_2^2\cos\theta_1$
$a_{2,0}$	$-r_1^2$
$a_{2,1}$	$2r_1^2r_2\cos\theta_2$
$a_{2,2}$	$-r_1^2r_2^2$

Table 5.1

Coefficient Values for 2-D Second-Order Auto-regressive Process.

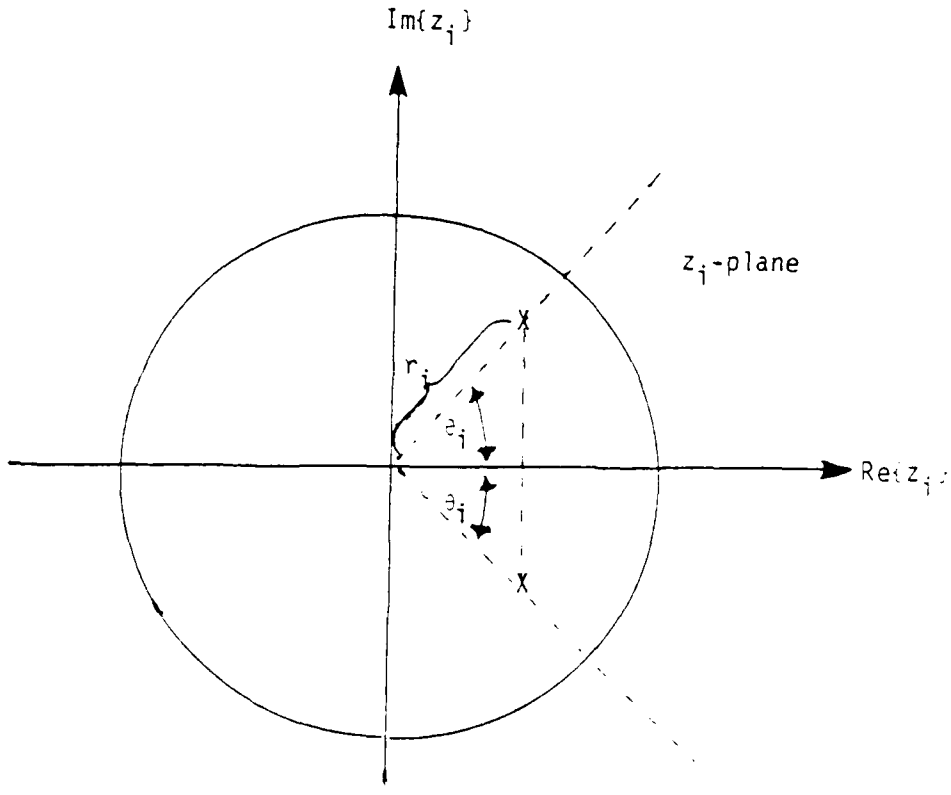


Figure 5.1

Pole Locations of Filter $H_i(z_i)$, i=1,2.

$$\rho_i(k) = \frac{r_i^{|k|}}{1+r_i^2}\left\{\frac{\sin[(|k|+1)\theta_i] - r_i^2\sin[(|k|-1)\theta_i]}{\sin\theta_i}\right\} \quad ; \; i=1,2. \tag{5.7}$$

Note, in particular, that $|\rho_i(k)| \leq 1$ for i=1,2. However, unlike the first-order case, where $\rho_i(k)=\rho_i^{|k|}$, the autocorrelation is no longer exponentially decreasing in k due to the sinusoidal terms within brackets. This property can be useful in modeling a wider range of textures than possible with the 1st-order model. Additional results on the 2nd-order model can be found in [5.1].

Our interest in ARMA models has been in developing their statistical properties, such as autocorrelation functions, joint p.d.f's, etc., and in developing methods for fitting these models to real-world texture samples. With regard to the latter activities we have been specifically concerned with developing maximum-likelihood (ML) estimates of the ARMA model parameters, particularly when observed in the presence of added noise.

5.1.2 RANDOM TESSELLATION PROCESS

This class of random fields has been studied extensively in the past (cf. [5.2]-[5.4]) and this aspect of the investigation has constituted an extension of past efforts. The construction procedure we have considered results in a tessellation of the plane into convex regions by an appropriately defined field of random lines which form the boundaries of these regions. The density of these random lines, or edges, is defined in terms of a rate parameter λ. Gray levels are then assigned within elementary regions to possess correlation coefficient ρ with gray levels in contiguous regions. For example, the gray levels can be generated by a 2-D ARMA model. Given a particular partitioning scheme then, the random fields are completely defined in terms of the two parameters λ and ρ. The parameter λ represents the edge business associated with an image while ρ is indicative, at least on an ensemble basis, of the edge contrast . For ρ large (in magnitude) and negative there is an abrupt almost black-to-white or

white-to-black transition across an edge boundary. If $\rho>0$, on the other hand, the transition across an edge boundary is much more gradual. It is relatively easy to define these parameters for wide classes of imagery data.

In previous work [5.2]-[5.4] we have been able to evaluate the second-order properties of this class of random field under the assumptions that the line processes generating the tessellation are themselves generated by either a Poisson or a periodic process. It can be shown that these two examples are special cases of renewal point processes with Gamma distributed interarrival distribution

$$f(x) = \frac{x^{\nu-1}}{\Gamma(\nu)\beta^{\nu}} \exp\{-x/\beta\};\ x \geq 0, \qquad (5.8)$$

with $\nu=1,2,\ldots$ and $\beta=1/\lambda\nu$ for fixed $\lambda>0$. For $\nu=1$ we have $f(x)=e^{-\lambda x}$, corresponding to the case of Poisson partitions, while for $\nu=\infty$, $f(x)=\delta(x-1/\lambda)$ corresponding to periodic partitions with fixed spacing $\nu=1/\lambda$. Clearly, the quantity ν plays the role of a homogeneity parameter. For small ν the mosaic appearance is that of a tiling of the plane by random rectangles. As ν increases the tiling becomes more and more regular or periodic.

In [5.5] we have been able to develop the second-order properties of the 2-D random tessellation process when the tessellation is generated by a pair of mutually independent renewal point processes with Gamma distributed interarrival distribution as given by (5.8). This work then represents a considerable extension of our abilities to characterize the second-order properties of the 2-D random tessellation process.

Other methods have been studied as part of this effort, including the use of the Voronoi and Johnson-Mehl tessellations [5.6], [5.7], to provide the underlying partitioning of the plane. In the Voronoi tessellation the

centers of regions, called seeds, are distributed according to a 2-D spatial point process with rate parameter λ. The elementary geometric regions then consist of all points closer to a given seed than to any other seed. In the Johnson-Mehl tessellation, the seeds are distributed according to a space-time point process with spatial rate parameter λ and temporal rate parameter ν. The elementary regions then grow isotropically with spatial velocity . In this case some seeds may fail to generate regions if their spatial location is overcome by regions associated with previously distributed seeds. While these partitioning schemes represent interesting generalizations of the 2-D random tessellation process, we have met with little success in characterizing the resulting second-order properties.

5.1.3 2-D RANDOM GRAIN MODEL

This class of image models have alternatively been called random grain models, Boolean Models, bombing models, etc. Our approach has been to consider these models as the output of suitably defined spatial filters excited by a marked spatial point process [5.8] at its input. For example, suppose that the filter is space invariant and can be described in terms of its point spread function $h(\underline{x})$. Consider exciting this filter by a marked spatial point process described in terms of the triple $\{\underline{x}_i, a_i, \ _i\}$ where $\underline{x}_i$ are the event locations, and $\{a_i\}$ and $\{\ _i\}$ are <u>marks</u> representing amplitude and rotation to be assigned to the characteristic response of the filter to the spatial impulse at the corresponding locations $\{\underline{x}_i\}$. The field at the filter output is then described by

$$f(\underline{x}) = \sum_i a_i h_i(\underline{x}-\underline{x}_i), \qquad (5.9)$$

where $h_i(\underline{x})=h(\underline{\underline{A}}_i\underline{x})$ with $\underline{\underline{A}}_i$ the unitary matrix

$$\underline{\underline{A}}_i = \begin{bmatrix} \cos\ _i & \sin\ _i \\ -\sin\ _i & \cos\ _i \end{bmatrix} \qquad (5.10)$$

The net results are that $\{f(\underline{x}),\ \underline{x}\in R^2\}$ is given by the superposition of a number of randomly weighted and rotated versions of the original point spread function $h(\underline{x})$ replicated at random spatial positions. In this sense the resulting random field resembles the so-called low-density shot processes evolving according to a single temporal parameter. The description of the field in terms of filtering operations not only provides a phenomemological construction procedure for generating the field but allows the application of existing and well-developed techniques for analyzing the statistical properties of 1-D low-density shot processes.

This class of 2-D random fields can be extended to include nonisometric transformations of a basic point spread function $h(\underline{x})$. Here the field $\{f(\underline{x}),\underline{x}\in R^2\}$ is obtained as a randomly weighted superposition of versions of $h(\underline{x})$ which has undergone scaling in addition to the rigid body motions of translation and rotation. As an illustrative example, suppose the field is described according to (2.9) and the sequel were now

$$\underline{A}_i = b_i^{-1}\begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\zeta_i^{-1}\sin\theta_i & \zeta_i^{-1}\cos\theta_i \end{bmatrix} \quad ; \quad i=1,2,\ldots, \tag{5.11}$$

with b_i, ζ_i appropriately defined random variables subject only to the constraints $b_i \geq 0$ and $0 \leq \zeta_i \leq 1$. This transformation has an interpretation as a rotation by θ_i radians followed by separate scaling of the orthogonal cartesian coordinate axes. Furthermore, suppose that the basic point spread function $h(\underline{x})$ is described by

$$h(\underline{x}) = \begin{cases} 1 & ; \ \|\underline{x}\| \leq 1 \\ 0 & ; \ \|\underline{x}\| > 1 \end{cases} , \tag{5.12}$$

where $\|\underline{x}\|^2 = \langle \underline{x},\underline{x}\rangle = x_1^2 + x_2^2$ is the ordinary Euclidean norm. The corresponding response $h_i(\underline{x}) = h(\underline{A}_i\underline{x})$ is then elliptical in shape as illustrated in Fig. 5.2.

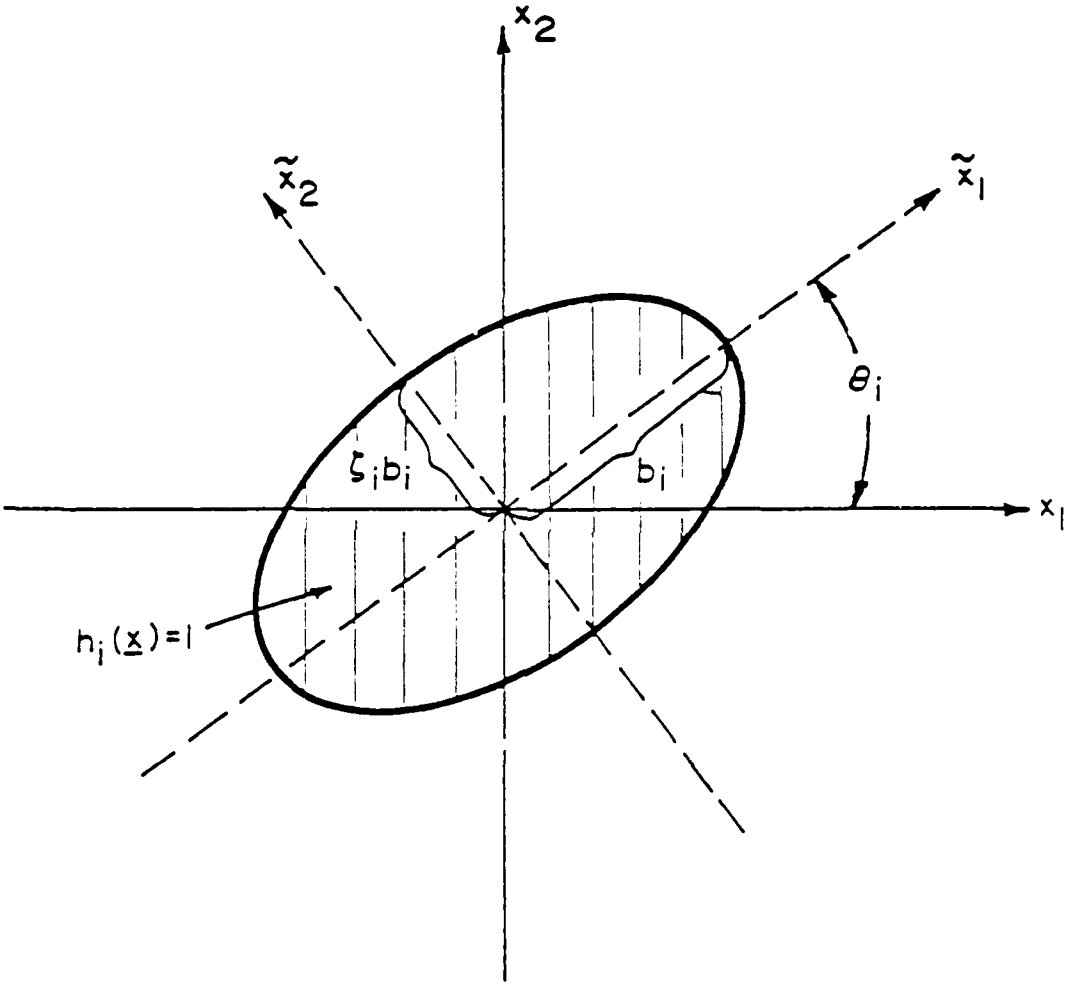


Figure 5.2

Characteristic Response of Random Grain

Stochastic Image Model

with major and minor axes b_i and $\zeta_i b_i$, respectively, each rotated by ϕ_i radians.

Other choices for the basic point spread function $h(\underline{x})$ are, of course, possible. For example, consider the sinusoidal response

$$h(\underline{x}) = \frac{\sin\{||\underline{x}||\}}{||\underline{x}||} \quad , \tag{5.13}$$

or even the exponential response

$$h(\underline{x}) = \exp\{-||\underline{x}||^2\} \ . \tag{5.14}$$

A reasonably complete characterization of the 2-D random grain model has been provided in [5.9]. This has included evaluation of second-order properties.

5.1.4 2-D MARKOV RANDOM FIELDS

Recent developments in the theory of 2-D MRF's provide a powerful alternative to existing stochastic fields as a texture model. The 2-D MRF attempts to extend the familiar 1-D Markov property to 2-D. More specifically, a 2-D MRF is defined by the property that the conditional p.d.f. of the gray level at any point in the image plane given values of all the other points depends only upon the gray levels at pixel positions in a specified neighborhood of the point in question.

Through its relationship with the so-called Gibbs random field (GRF), it can be shown that the joint p.d.f. of the MRF defined on a specified lattice assumes a precise functional form. Furthermore, this functional form can generally be defined in terms of a small number of physically meaningful parameters. These parameters can then be adjusted so that the resulting realizations of the MRF closely resemble real-world textures.

A detailed treatment of the construction and properties of MRF's appears in [5.10]. Here we demonstrated how the parameters of the MRF can be

matched to real-world textures by ML estimation techniques. In this work we also developed ML techniques for texture classification and segmentation and demonstrated the efficacy of these techniques through experiments on real-world texture samples.

5.2 OPTIMUM MODEL-BASED TEXTURE CLASSIFICATION/ DISCRIMINATION ALGORITHMS

Our efforts in this area have been concerned with developing ML-based texture classification/discrimination schemes based upon specific stochastic texture modeling assumptions. This work was begun in [5.11] where use was made of the 2-D random tessellation process described previously. The approach was to first reduce the data to its spatial gray-level co-occurrence matrix. Then ML classification/discrimination schemes were developed using as observations the elements of the spatial gray-level co-occurrence matrix. Under some ergodic assumptions on the underlying tessellation model, it was possible to explicitly compute the class-conditonal log-likelihood functionals. This approach has been shown to be quite effective when applied to real-world imagery.

In [5.12] this initial approach was extended to arbitrary stochastic texture models through an adaptive technique for estimating the unknown parameters appearing in the class-conditional log-likelihood functional. This approach was shown to provide excellent classification performance on real-world texture samples. The difficulty was that it required training data consisting of homogeneous observations from each of the possibly unknown texture classes. The approach then, while useful for texture classification, was not useful for texture segmentation since homogeneous training data is generally not available.

As part of the present effort we developed a technique for estimating the structural parameters of the underlying stochastic image model directly from the data itself. These estimates are then used in the technique of [5.11], with some additional simplifications, to provide a fully adaptive model-based approach to texture segmentation. The results are described in [5.13].

A second major contribution of this effort has been to develop ML-based texture classification/discrimination techniques using a MRF model of texture. This approach, as reported in [5.10], is shown to be quite effective in the classification/discrimination of real-world textures at a reasonably small computational expense. It's felt that this approach offers an attractive performance/complexity alternative to the schemes described in [5.11]-[5.13].

5.3 EVALUATION AND CHARACTERIZATION OF TEXTURE CLASSIFICATION/ DISCRIMINATION ALGORITHMS

In this work we have attempted to explore the performance of various existing texture discriminants operating on identical and carefully controlled stochastic texture models. More specifically, we have concentrated upon the texture features described by Haralick [5.14]. These features, which are extracted from the spatial gray-level co-occurrence matrix, were originally proposed on an ad hoc basis. Unfortunately, there is little information available on how to choose critical parameters, such as the separation distance, d. Furthermore, there has never been any comprehensive work performed to establish the conditions (e.g., the class of underlying stochastic texture models) under which these features provide effective discrimination and conditions under which they can be expected to fail. Finally, there has never been any work directed toward establishing the

relative performance of texture classification/discrimination procedures based on the Haralick features compared to optimum model-based algorithms, such as developed separately under this effort.

In [5.15] we describe the results of a comprehensive study of the relative performance of the Haralick features for a variety of stochastic texture modeling assumptions. This work has included ARMA models, as well as 2-D random tessellation processes. The results are quite useful in providing information on how to choose key parameters for the Haralick features and in identifying conditions under which the Haralick features either provide effective discrimination or fail to do so.

5.4 PRESENTATIONS AND PUBLICATIONS

1. J. W. Modestino, 'Second-Order Two-Dimensional Autoregressive Model for Images', TM 84-3, Communications and Information Processing Systems Group, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NY, May 1984.

2. R. W. Fries, J. W. Modestino and A. L. Vickers, 'A Further Generalization of the Two-Dimensional Random Checkerboard Process', TR84-4, Communication and Information Processing Group, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NY, July, 1984.

3. F. Azadegan and J. W. Modestino, 'Construction and Properties of the Two-Dimensional Random Grain Model', Proc. 1986 Conf. on Information Sciences and Systems, Princeton, NJ, pp. 455-459, March 1986.

4. J. Zhang and J. W. Modestino, 'Markov Random Field with Applications to Texture Classification and Discrimination', Proc. 1986 Conf. on Information Sciences and Systems, Princeton, NJ, pp. 230-237, March 1986.

5. A. L. Vickers and J. W. Modestino, 'Further Results on Texture Discrimination Based Upon an Assumed Stochastic Texture Model', TR83-2, Communications and Information Processing Group, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NJ, Aug. 1983.

6. A. L. Vickers and J. W. Modestino, 'Evaluation of Selected Texture Discriminants Under Specific Stochastic Modeling Assumptions', TR84-5, Communication and Information Processing Group, ECSE Dept. RPI, Troy, NY, Aug. 1984.

7. J. Zhang, 'Markov Random Fields with Applications to Texture Classification and Discrimination', M. S. Thesis, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NY, Dec. 1985.

5.5 REFERENCES

5.1 J. W. Modestino, "Second-Order Two-Dimensional Autoregressive Model for Images", TM 84-3, Communications and Information Processing Systems Group, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NY, May 1984.

5.2 J. W. Modestino, R. W. Fries and A. L. Vickers, "Stochastic Image Models Generated by Random Tessellations of the Plane", Computer Graphics and Image Processing, vol. 12, pp. 74-98, 1980.

5.3 J. W. Modestino, R. W. Fries and D. G. Daut, "A Generalization of the Two-Dimensional Random Checkerboard Process", J. Opt. Soc. Am., vol. 69, pp. 897-906, June 1979

5.4 J. W. Modestino and R. W. Fries, "Stochastic Models for Images and Applications", in Proc. of NATO Advanced Study Institute on Pattern Recognition and Signal Processing, Ed. by C. H. Chen, Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands, 1979.

5.5 R. W. Fries, J. W. Modestino and A. L. Vickers, "A Further Generalization of the Two-Dimensional Random Checkerboard Process", TR84-4, Communication and Information Processing Group, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NY, July, 1984.

5.6 E. N. Gilbert, "Random Subdivisions of Space Into Crystals", Annals. Math. Stat., vol. 33, pp. 958-972, 1962.

5.7 B. J. Schachter, A. Rosenfeld, and L. S. Davis, "Random Mosaic Models for Textures", IEEE Trans. Syst., Man., and Cybern., vol. SMC-8, pp. 694-702, Sept. 1978.

5.8 D. L. Snyder, Random Point Processes, Wiley, New York, 1975.

5.9 F. Azadegan and J. W. Modestino, "Construction and Properties of the Two-Dimensional Random Grain Model", Proc. 1986 Conf. on Information Sciences and Systems, Princeton, NJ, pp. 455-459, March 1986.

5.10 J. Zhang and J. W. Modestino, "Markov Random Field with Applications to Texture Classification and Discrimination", Proc. 1986 Conf. on Information Sciences and Systems, Princeton, NJ, pp. 230-237, March 1986.

5.11 J. W. Modestino, R. W. Fries and A. L. Vickers, "Texture Discrimination Based Upon an Assumed Stochastic Texture Model", IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-3, pp. 557-580, Sept. 1981.

5.12 A. L. Vickers and J. W. Modestino, "A Maximum Likelihood Approach to Texture Classification", IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-4, pp. 61-68, Jan. 1982.

5.13 A. L. Vickers and J. W. Modestino, "Further Results on Texture Discrimination Based Upon an Assumed Stochastic Texture Model", TR83-2,

Communications and Information Processing Group, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NY, Aug. 1983.

5.14 R. M. Haralick, K. Shanmugan and J. Dinstein, "Textural Features for Image Classification", IEEE Trans. Syst., Man, Cybern., vol. SMC-3, pp. 610-621, Nov. 1973.

5.15 A. L. Vickers and J. W. Modestino, "Evaluation of Selected Texture Discriminants Under Specific Stochastic Modeling Assumptions", TR84-5, Communication and Information Processing Group, ECSE Dept. RPI, Troy, NY, Aug. 1984.

6. EFFICIENT ALGORITHMS FOR PHOTO-INTERPRETATION

The results of this task are reported in two documents, Removal of Cloud Shadows from Aerial Photographs by S-P. Shu [6.3-1] and ISES - An Image Segmentation Expert System for Aerial Photographs Based on the Use of Edge and Texture Features by S-P. She [6.3-2]. We shall present summaries of the work from the documents.

6.1 CLOUD COVER REMOVAL AND COMPENSATION

The objective of this project, the results of which are reported in [6.3-1], is to develop an algorithm compensating for extreme variation in scene illumination caused by cloud shadows in order to obtain a uniformly bright image display. A sample aerial photograph was provided for this purpose as well as two reference points, one underexposed due to cloud shadows and one overexposed.

An elegant algorithm has been developed successfully. The algorithm can determine cloud shadow edges under various criteria, divide the image into partial images along the cloud shadows edges, compensate for the brightness of each partial image, reconnect the compensated partial images into one high-quality image without visible connection edges and add heuristics to save computation time and improve the quality of the resultant image. The Cloud Shadow Modeling technique is used to model cloud shadows in aerial images by mathematical expressions and to add artificial cloud shadows into aerial images.

CLOUD SHADOW REMOVAL SCHEME

The Cloud Shadow Removal Scheme is illustrated by the following steps;

(1) Get Input Sample Image $I_i(i,j)$

An input sample image, such as Fig. 1, is obtained by digitizing the film of the sample aerial photograph using the Prime Computer System in the RPI Image Processing Lab.

(2) Determined Cloud Shadow Edges

A fuzzy-edge detection scheme, developed by the author, is used to determine the tight and loose bounds of cloud shadow edges in the image $I_i(i,j)$.

(3) Divide Sample Image Along Cloud Shadow Edges

The image $I_i(i,j)$ is divided into left and right part images, denoted by $I_L(i,j)$ and $I_R(i,j)$ respectively. The overlay image from $I_L(i,j)$and $I_R(i,j)$ is denoted $I_M(i,j)$.

(4) Compensate Image Brightness

The images $I_L(i,j)$, and $I_M(i,j)$ are compensated by algebraic gray-level transformation, histogram equalization, and mean & variance transformation techniques. The compensated images are denoted $I_{L1}(i,j)$ $I_{R1}(i,j)$ and $I_{M1}(i,j)$, respectively.

(5) Recombine Partial Image

The compensated images I_{L1}, I_{R1}, and I_{M1} are recombined into one image $I_g(i,j)$ as in Fig. 3 by the overlay image connection technique. This technique was developed by the author for recombining partial images into one image without visible connection edges.

The algorithm for the removal of cloud shadows from aerial photograph has been successfully implemented on the Prime Computer System in the RPI Image Processing Lab.

The provided sample image is shown in Fig. 1. By this algorithm, the resultant output image is obtained as shown in Figs. 2 and 3, in which the right part cloud shadow was removed and the entire image was enhanced to higher quality.

Fig. 1

Original Sample Image I_{1-1}

Resolution: 512 x 512 pixels

Fig. 4

Resultant Image I_{6-6}

Resolution: 512 x 512 pixels

Fig.

Resultant Image $I_{6,7}$

Resolution: 512 x 512 pixels

6.1 AN IMAGE SEGMENTATION EXPERT SYSTEM

Comparison:

	Advantages	Disadvantages
A	1. closed-loop region boundaries (no edge gaps) 2. smooth region inside a region	1. full of small holes 2. boundaries are ragged 3. spatially inaccurate
B	1. smooth region boundaries 2. spatially accurate 3. no small holes	1. edge gaps 2. micro edges

A: Based on Texture Classification

B: Based on Edge Detection

Fig. 4. Comparison of Edge Detection and Texture Classification for Image Segmentation

to try to develop an image-segmentation system that would have the same advantages but without the disadvantages. The ISES presented here comes close to achieving the goal, having achieved excellent experimental results.

In the ISES, the A* algorithm (AI minimum-cost-graph search) was applied for both heuristic edge extraction (HEE) as well as for region formation (RF). The A* algorithm is a powerful scheme in AI for searching desirable paths under a well-defined cost function. The cost function in RF was defined with both edge-features and texture-feature parameters, thus permitting the RF procedure to use edge and texture features simultaneously for optimal path search. This approach demonstrated the power of AI approach in simultaneously handling various image heuristics such as edge and texture features. Based on the simultaneous use of edge and texture features under a rule-based expert-system control, the RF presented the first solution for eliminating edge gaps and micro noisy edges. These edge gap and micro-edge problems had not been solved before by any kind of edge detection techniques.

The localized brightness compensation (LBC) in the ISES is not only a good preprocessing scheme for the ISES, but is also a useful general image enhancement scheme. This generalized LBC for image enhancement, is the first localized image-enhancement scheme. It is also the first image enhancement algorithm using rule-based expert-systems approach.

The HEE in the ISES is a new useful edge/boundary detector system for obtaining one-pixel-wide edges or boundaries from real images. Compared to texture classification, the HEE provided much more spatially accurate region boundaries. The one-pixel-wide edges/boundaries can be easily encoded into chain codes for various image processing applications. It is clear that a

one-pixel-wide edge/boundary is more useful than a non-one-pixel-wide edge boundary. The edge thinning operation in the HEE is a new efficient algorithm for obtaining one-pixel-wide edges or boundaries.

The knowledge-based controller in the ISES is a rule-based expert system model. It controls all the processing of the ISES, from early manipulating the image processing tools (LBC, HEE, and RF), through modifying the data set in the image fact base, to obtaining a high quality segmented output image. The inference engine in the ISES consists of the forward chaining, dynamic facts drive schemes. These rule-based expert system schemes are simulated human visual system used in segmenting images, therefore, the ISES can handle very complex and imperfect data set in the image fact base.

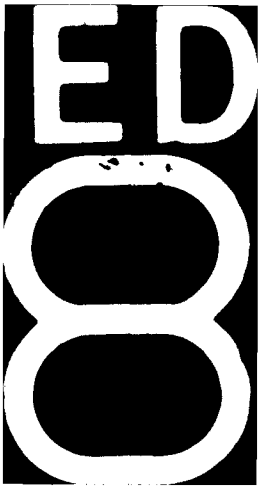
The LBC, HEE, and RF in the ISES were [illegible] texture features [illegible] expert system [illegible] image segmentation results. The HEE improved [illegible] the RF [illegible] edge gaps [illegible] rule-based expert system [illegible] improved the handling of complex [illegible] sets. They highly improved the ISES performance.

6.3 PUBLICATIONS

1. [illegible]

MISSION
of
Rome Air Development Center

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control, Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of C^3I systems. The areas of technical competence include communications, command and control, battle management, information processing, surveillance sensors, intelligence data collection and handling, solid state sciences, electromagnetics, and propagation, and electronic, maintainability, and compatibility.


ED
80