Detection of Mines and Minelike Targets Using Principal Component and Neural-Network Methods

Xi Miao, Mahmood R. Azimi-Sadjadi, *Senior Member, IEEE,* Bin Tian,
Abinash C. Dubey, *Associate Member, IEEE,* and Ned H. Witherspoon

***Abstract*— This paper introduces a new system for real-time detection and classification of arbitrarily scattered surface-laid mines from multispectral imagery data of a minefield. The system consists of six channels which use various neural-network structures for feature extraction, detection, and classification of targets in six different optical bands ranging from near UV to near IR. A single-layer autoassociative network trained using the recursive least square (RLS) learning rule was employed in each channel to perform feature extraction. Based upon the extracted features, two different neural-network architectures were used and their performance was compared against the standard maximum likelihood (ML) classification scheme. The outputs of the detector/classifier network in all the channels were fused together in a final decision-making system. Two different final decision making schemes using the majority voting and weighted combination based on consensual theory were considered. Simulations were performed on real data for six bands and on several images in order to account for the variations in size, shape, and contrast of the targets and also the signal-to-clutter ratio. The overall results showed the promise of the proposed system for detection and classification of mines and minelike tagets.**

***Index Terms*— Landmine detection, neural networks, pattern recognition, principal component analysis.**

I. Introduction

THE development of a mine detection system capable of detecting and classifying various types of mines under different environmental conditions presents many technical problems. Historically, a number of automatic target detection and recognition schemes have been applied to this problem. However, in practice, they have been only partially successful and have been shown [3], [4] to produce high false-alarm rates. Some contributing factors which inhibit the detection and classification are the diverse sizes and compositions of targets, variation of soil properties with location and moisture conditions, nonrepeatability of the target signature, competing clutter objects having similar responses as the actual targets, and partially obscured targets. In addition, the conventional target classification schemes lose their accuracy when the feature space is of high dimension and the classes cannot easily be separated. The later scenario typically occurs when multiple targets are present in the scene and their distribution functions or feature sets are overlapping with those of the nontarget anomalies. These, coupled with a number of other shortcomings such as sensor sensitivity to environmental conditions, temperature variations, and computational and speed limitations of the current digital processing hardware motivated this work to develop more reliable, robust, and accurate schemes for automatic target detection and classification from multispectral optical imagery.

Neural networks have been shown to offer potentially powerful, robust, and adaptive means of detecting and classifying targets under changing (or even new) signature and environmental conditions [4]. One of the most desirable properties of these networks is their ability to learn from examples and to generalize from the training set to similar data not contained in the training examples.

In the past few years several neural-network-based methods for the detection and classification of mines and minelike targets have been proposed. In [5]–[7], Azimi-Sadjadi *et al.* developed neural-network-based approaches for the detection and classification of buried landmines from microwave data. Several data representation schemes such as the principal component method [5], moment invariant [6] and bispectrum [7] were used to reduce the dimensionality of the data and extract the salient features of the targets and nontargets. The test results indicated the effectiveness of the neural-network-based detector/classifier systems. Telfer *et al.* [8], [9] used wavelet preprocessing followed by neural-network detection/classification and obtained good results on multispectral imagery. Daud [10] used feedforward neural networks for mine discrimination using multispectral imagery. A trainable classifier using probabilistic neural-network (PNN) approach was used in [11] for land mine detection using multispectral imagery. The approach by Holmes *et al.* [12] is based on exploiting the spatial and spectral signatures of the mines. The key features of this approach is the ability to adapt to unknown or changing background statistics and to operate with unknown spectral signatures. An approach based on center-surround filter was discussed by Schmalz *et al.* [13].

This paper investigates the development and use of various neural-network structures for real-time target feature extraction and detection/classification. The overall structure of our system consists of several feature extractor and detector/classifier subsystems in parallel which provide individual decisions corresponding to the events in each optical band. The outputs

Manuscript received June 22, 1996; revised April 8, 1997 and January 2, 1998.

X. Miao, M. R. Azimi-Sadjadi, and B. Tian are with the Department of Electrical Engineering, Colorado State University, Fort Collins, CO 80523 USA.

A. C. Dubey and N. H. Witherspoon are with the Naval Warfare System Center/Dahlgren Division, Coastal Systems Station, Panama City, FL 32407 USA.

Publisher Item Identifier S 1045-9227(98)02773-8.

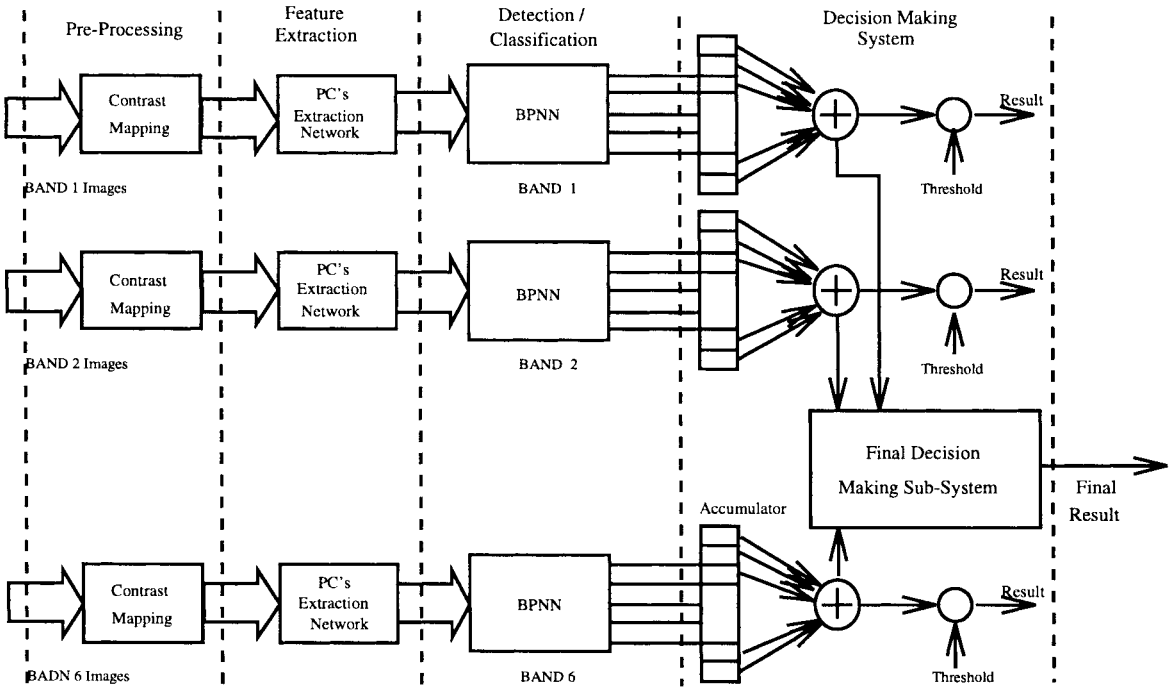


Fig. 1. Multichannel target detection/classification system.

of each detector/classifier network are then processed and fused together in a final decision making system. The final decision is made based upon the strength of the responses from different detectors/classifiers. This system and its constituent subsystems are introduced in the next section.

II. A Multichannel Target Detection/Classification System

The overall block diagram of our target detection/classification system is shown in Fig. 1. It consists of several preprocessing, feature extraction, detection/classification, and final decision making subsystems. As can be seen, six processing channels for six different optical bands are used and each channel consists of the same type of subsystems. Since images obtained from a scene at six different optical bands contain different intensity levels and the contrast of targets and nontargets is significantly different, fusing the information of each band is necessary for this problem. This is accomplished by combining the outputs of these channels in the final fusion subsystem. In what follows, the operations of the sensor, each subsystem and the relevant processes are discussed.

A. Multispectral Imaging Sensor

The data for this study was obtained from a multispectral video camera namely Xybion Model IMC-201 [14], [15]. The IMC-201 is intensified and gated for automatic exposure control. The intensifier, which allows for short exposure time does, however, limit the spatial resolution. The camera utilizes a spinning filter wheel located between camera lens and the imaging plane. The filter wheels are interchangeable and each contains six filters. The filter wheel rotation places a different filter in front of the camera imaging plane every 1/30th of a second which is also the frame rate of the camera. In this mode of operation, every video frame is a separate spectral band. The spectral range of the camera is from near UV (400 nm) to near IR (900 nm). The objective was to cover uniformly the entire 400–900 nm band during the field tests, hence leading to six spectral bands with an overlap of approximately 25 nm. Studies from these tests led to a bandwidth selection most suitable for mine detection in different environments. Real-time down-linking of the data to a base station was also an important factor in selecting as few bands as possible without sacrificing the performance of the system.

The camera functions are micro-processor controlled. The output from the camera is the standard RS-170 interlaced video, which is recorded on a Hi-8 video recorder. For standoff mine detection ground (SMDG) test [1], [2] the camera and targets are stationary and therefore the images from six spectral bands are coregistered. These images are digitized and then used for subsequent processing. The advantages of this camera are that it 1) is lightweight and compact; 2) consumes little power; 3) operates easily from a standoff distance; 4) collects target signatures distinct from the background; 5) possesses moderately high resolution; and 6) lends itself for applications in automatic target recognition (ATR). These advantages have resulted in substantial enhancement in the capability of the system for detection and localization of targets in background clutter.

B. Preprocessing Subsystem

The original images obtained from different scenes are found to have low contrasts. In addition to inconsistency in the intensity levels and the considerable variations in the

brightness and contrast within the images of different bands, targets may be obscured by the background clutter. At higher bands (4–6) the distributions of target and background do not overlap, while in lower bands (1–3) the distributions are significantly overlapping. This makes the detection and classification very difficult at these particular bands. On the other hand, the information in the lower bands can not be ignored, e.g., band 1 is found to be very effective in presence of shadows.

To overcome the variations of image intensity levels at different optical bands and improve the conspicuity of the targets in these images, an effective contrast enhancement procedure must be applied prior to feature extraction. Contrast mapping schemes [16] can be used to redistribute the gray levels so that the intensity of the targets would fall into a certain range and the background would be suppressed as much as possible.

C. Feature Extraction Subsystem Using Principal Component Method

Selection of an appropriate set of target features is one of the most important tasks for any target classification system. The primary goal of feature extraction is to obtain features which maximize the similarity of objects in the same class while at the same time maximizing the dissimilarity of objects in different classes. It also results in dimensionality reduction of data, computational efficiency and reduction of memory requirements of the classifier. This is particularly important when neural networks are used to perform the classification tasks as the dimensionality reduction of the input not only removes the redundancy of the data but also enables the use of a smaller size network structure which can be trained easier and has improved generalization capability.

The salient features of the data can be extracted through a mapping, such as *Fourier* transform, *discrete cosine* transform, Karhunen–Loève (KL) transform, or principal component (PC) method [16], from a higher dimensional input space to a lower dimensional representation space. The efficiency of a chosen mapping approach is judged based on the degree of data compaction subject to the constraint that the original data can be reconstructed with minimal distortion. Based on this criterion, the PC method is optimal since it packs most of the signal energy into the first few coefficients, and at the same time achieves complete decorrelation of the data [16]. This latter property is ideally suited for detection/classification purposes as these decorrelated features or components can be used to train the system effectively. Consequently, in our system we used PC method as the feature extraction method.

The conventional approach for PC extraction involves the computation of the input data covariance matrix and then the application of a diagonalization procedure to extract the eigenvalues and the corresponding eigenvectors [16]. This process can be computationally very inefficient especially when the dimension of the covariance matrix is large. In addition, all the eigenvalues and their corresponding eigenvectors have to be evaluated even though only the eigenvectors which correspond to the most significant eigenvalues are used in the data transformation process. These deficiencies make the conventional scheme inefficient for real-life applications.

In [17], a new PC extraction method based upon the recursive least squares (RLS) learning algorithm [18] was developed. The main advantage of this approach is that the step size for updating is not fixed and depends on the statistical characteristics of the input data. Thus, it does not have the accuracy versus convergence speed tradeoff problems [19], [20] and also provides a recursive way to determine the variance associated with each PC which is a deterministic factor in deciding the number of components needed for an accurate representation. This autoassociative network was used in this paper to implement the feature extraction subsystem. The trained network extracts the PC's of target and nontarget blocks of the image data in each channel. The training data set was chosen primarily from partial image data contained in some selected target blocks. Targets generally have variable sizes, shapes, and contrast. Thus, a block of target image can not present all the possibilities of target scenarios in different scenes and at different locations. In addition, for dim targets the contrast between targets and their surrounding background is pretty low, hence using this information will produce low detection rate and high false alarm rate. Subwindowing was used for both training and testing where the larger window localizes the target and its surrounding background, while a smaller subwindow is swept across the whole larger window by moving it one row or one column at a time. The operation of the subwindowing is shown in Fig. 2. This subwindowing approach provides a very effective tools in identifying blocks which contain very dim or partially obscured targets and at the same time eliminating the false positives at a great rate.

D. Detector/Classifier Subsystem

Detection and classification tasks are performed using a multilayer backpropagation neural network (BPNN) [21]. As shown in Fig. 1, blocks of images in each channel were first applied to the corresponding PC extraction subsystem, the output of which is then fed to the detector/classifier network. The output of the network corresponding to each subwindow of a block is then stored in an accumulator and a decision is made after the presentation of all the subwindows of that block. Based upon the total number of "ones" in the accumulator, which corresponds to the total number of subwindows recognized as targets, a score will be determined for each block. A block is considered to be a target at a particular band, if its score is greater than a predetermined threshold.

E. Final Decision Making or Fusion Subsystem

After the images in each band are processed seperately by the associated neural-network detector/classifier, a final fusion making subsystem combines the results of each individual network. Two different schemes are considered. The first one is based upon the "majority voting" approach where the accumulated outputs in each particular channel corresponding to each subwindow are added up and compared with a threshold which is also determined empirically. If the sum is greater

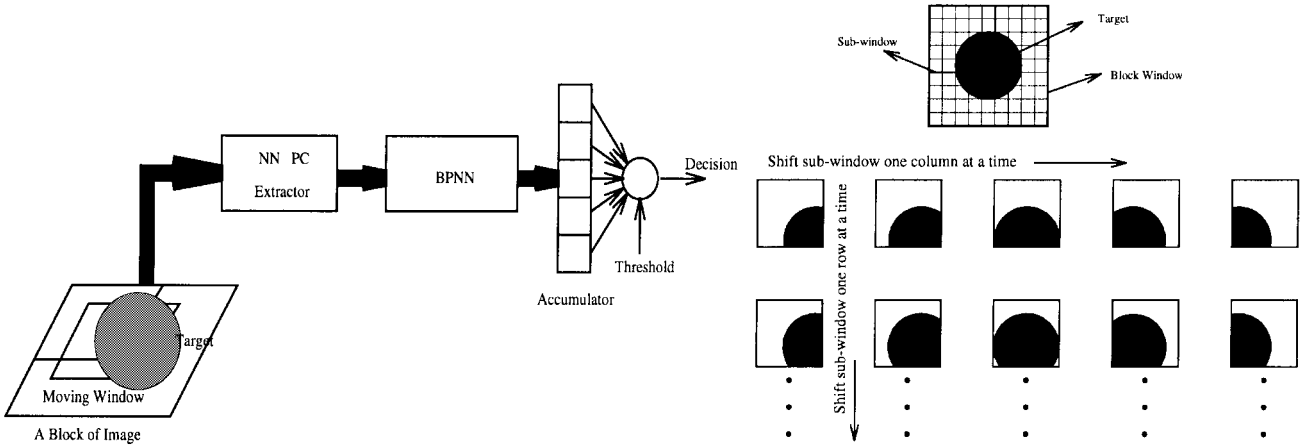


Fig. 2. Subwindowing operation.

than this threshold then the block is finally classified as a target. Although this approach is rather simple, it does not consider the relative importance of the individual data in each band. Improved results can be obtained by using methods based upon consensus theory which has been widely studied in the pattern recognition and neural-network fields [22]–[24]. The basic idea behind this theory is to combine the outputs of different neural networks with different weights that can reflect the relative importance of the classification result of that particular network. Several different implementation schemes have been studied and perhaps the most widely used one is the linear opinion pool [23].

Suppose for the jth training sample, the output of neural network for each band is $o_i^{(j)}, i = 1, \cdots, 6$, then the output of the final decision making system F_j for this sample will be

$$F_j = \sum_{i=1}^{6} \lambda_i o_i^{(j)} = \Lambda^t O^{(j)} \tag{1}$$

where $\Lambda = [\lambda_1, \lambda_2, \cdots, \lambda_6]^t$ is the weight vector and $O^{(j)} = [o_1^{(j)}, \cdots, o_6^{(j)}]^t$ is a vector containing all the six outputs from different bands for the jth training sample. If $\lambda_1 = \lambda_2 = \cdots = \lambda_6$, the final decision system will degenerate to the simple "majority voting" approach. One drawback of the linear opinion pool is that it is not Bayesian [23]. It was found [23], for some cases, that one neural network may dominate the final decision. Another kind of opinion combination scheme is the logarithmic opinion pool [23] for which the final output is given by

$$F_j = \prod_{i=1}^{6} [o_i^{(j)}]^{\lambda_i} \tag{2}$$

where generally we assume that $\sum_{i=1}^{n} \lambda_i = 1$. The logarithm opinion pool is Bayesian, but is more complicated than its linear counterpart. In this paper, the linear opinion pool was adopted for which the optimum weights is decided based on the least square (LS) criterion, i.e.,

$$\Lambda_{\text{opt}} = \arg\min_{\Lambda} \|O\Lambda - D\|^2 \tag{3}$$

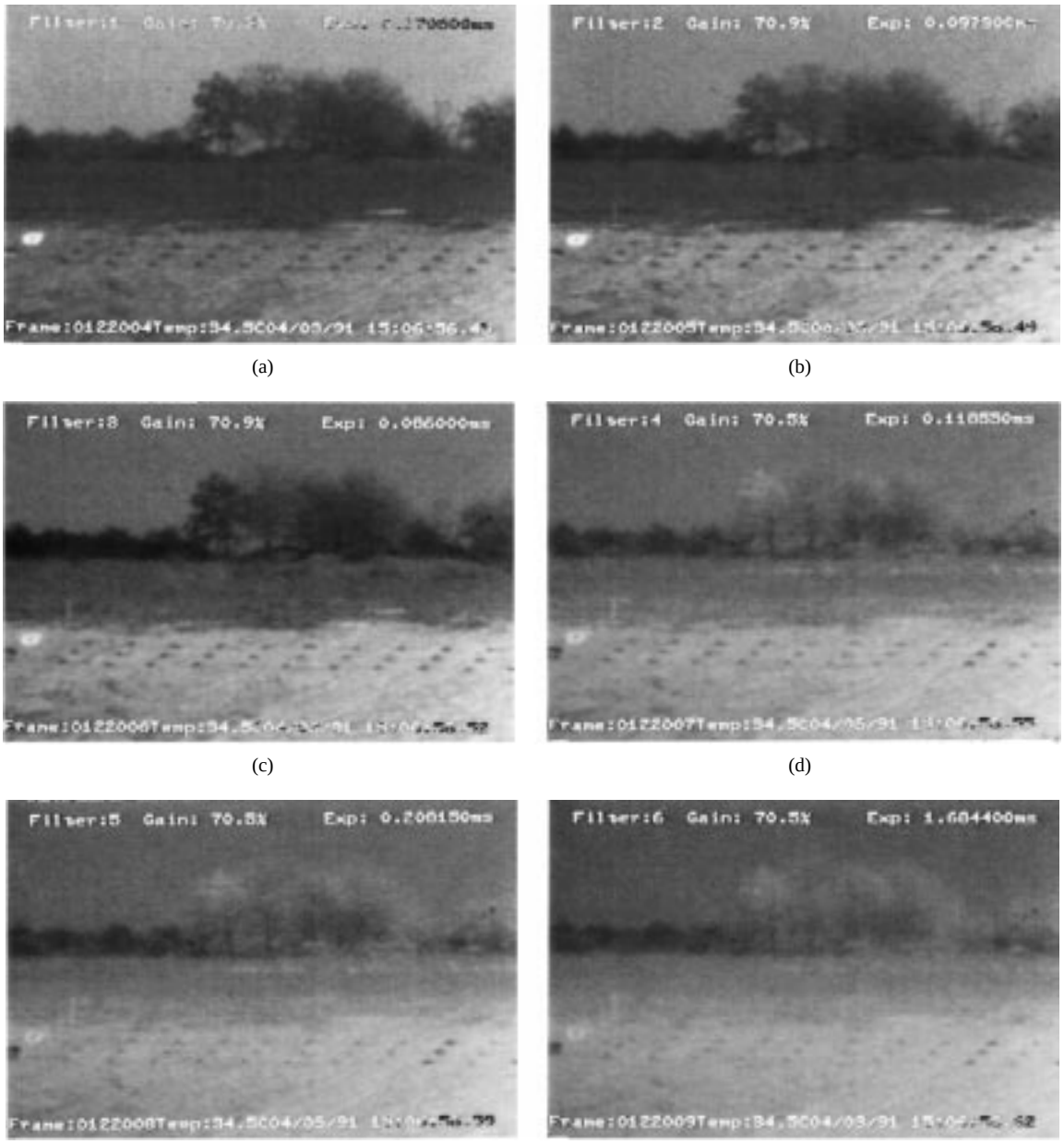
where $O = [O^{(1)}O^{(2)} \cdots O^{(N)}]^t$ is the output data matrix, $D = [d^{(1)}d^{(2)} \cdots d^{(N)}]^t$ is the desired output vector and N is the number of training samples. Note that for the jth training sample, $d^{(j)} = 1$ for target, while $d^{(i)} = 0$ for nontarget windows. The LS solution for this minimization problem yields [23] as

$$\Lambda_{\text{opt}} = [O^T O]^{-1} O^T D. \tag{4}$$

III. Implementation and Simulation Results

In this section, we will present the simulation results of the proposed mine target feature extraction and detection/classification methods on real data obtained from the Coastal Systems Station in Panama City, FL. This data consists of multispectral images collected from two different minefields at six different optical bands ranging from near UV (band 1) to near IR (band 6) regions. Almost half of the data was used for the purpose of training of the neural-network detection/classifier and the same number of cases or more were used to test the performance of the trained networks. The locations of targets verified by the ground truth maps were determined for proper training and testing of the neural networks. The training samples were chosen in a fashion so that a wide variety of targets and nontargets with different sizes, shapes, contrasts, textures, etc. were included. For testing, we have not only investigated how well the system can detect and classify these areas, but also processed the entire image block by block in order to determine how well the system was at finding the mines. In this case the entire image as opposed to localized areas was the input to the system. Fig. 3 shows the original images of a minefield for the six spectral bands.

First in order to improve the conspicuity of the targets in each band, a contrast mapping scheme was used as a preprocessor in each channel. To see the effects of this preprocessing, consider the original image in Fig. 3 corresponding to band 3. This image is contrast enhanced to give the image in Fig. 4. Comparing the two images clearly indicates that the contrast between the targets and background was significantly improved in the preprocessed image. As can be observed, especially in the encircled regions, one can



(a) (b) (c) (d) (e) (f)

Fig. 3. Multispectral images of a minefield for bands 1–6.

not tell the difference between the target and background from the original image, while in the processed image the targets appeared with better contrast in comparison with the surrounding background. These improvements facilitate the target feature extraction in the subsequent subsystem.

In this application, the size of the image blocks was fixed to 25×30. Several experiments with three different subwindow sizes namely 20×20, 15×15, and 10×10 pixels were conducted to investigate the effects of changing the subwindow size on the performance. The content of the subwindows for certain chosen target blocks were then used for training the PC extraction network. Clearly, the number and size of the input training data vector are variable depending on the size of subwindows. For example, for the subwindows size of 20×20 pixels, the data vector was arranged into a one-dimensional input vector of size 400 using row-ordering scheme. For each band, five target blocks were selected from each data set and eight different sets were used for training the feature extraction subsystem. Thus, the total number of target blocks used for training was 240. The mean of the training data vector was then subtracted from each individual data vector prior to PC extraction.

The number of neurons p in the autoassociative network corresponds to the number of PC's that need to be extracted. This number can be determined based on the percentage of the energy of the image contained in the first p components which is measured by

$$\text{Energy\%} = \frac{1}{p}\frac{\sum_{i=1}^{p}\xi_i}{\sigma^2} \tag{5}$$

TABLE I
COVARIANCE MATRIX OF THE FIRST SIX COMPONENTS FOR THE RLS ALGORITHM

r	1	2	3	4	5	6
1	50493.0	11.73	1.73	-0.23	-1.55	-0.35
2	11.73	2706.7	-4.05	-3.88	-0.22	1.92
3	1.73	-4.05	1194.5	2.89	-0.41	1.11
4	-0.23	-3.88	2.89	578.9	13.79	-1.08
5	-1.55	0.22	-0.41	13.79	517.9	0.63
6	-0.35	1.92	1.11	-1.08	0.63	108.7

where ξ_i represents the variance of the ith component or the ith eigenvalue of the input covariance matrix and σ^2 is the variance of the image block. In our application, the first ten components were found to possess approximately 95% of the image energy. The other componenets were too small to have any effect on the performance. Thus, a single-layer autoassociative network with ten neurons was used. The weights of the network were randomly initialized between 0.5 and 1.0 and the initial value of $P(0)$ in the RLS algorithm was chosen to be 0.5. Upon completion of the training process, the weight vectors of the neurons converged to the relevant eigenvectors of the data covariance matrix and the outputs generate the associated components. For each component, convergence was achieved after only one epoch over the training data and the MSE dropped to almost zero. The distribution of the eigenvalues of the input covariance matrix obtained using the RLS algorithm, indicates that they decay very quickly for the lower order components. This can be demonstrated by studying the covariance matrix of the first six components, as shown in Table I. Additionally, from the values of the off-diagonal elements of this covariance matrix, it can be observed that the algorithm provided a good level of data decorrelation.

The same PC extraction network was used in each channel for the six optical bands. The network architecture, initial parameters, window size, and the training data sets were the same for all these channels. After the training was completed, the trained networks were then used to transform all the image blocks and extract their PC's. The extracted components were applied to the subsequent detection/classification subsystems. Several different BPNN structures were studied to perform detection/classification tasks. The optimal network structure was experimentally determined to be a three-layer BPNN with 25 neurons in the first hidden layer, ten neurons in the second hidden layer, and one output neuron. The neurons had sigmoidal type nonlinearity activation functions. The number of inputs was 11 which included ten PC's corresponding to a subwindow and the variance of the block. These 11 inputs were normalized between zero and one. All the connection weights were randomly initialized between 0.5 and 1.0. The initial learning step was 0.02 which decreased 90% in every five epochs and the momentum was 0.9. The desired output was set to one for targets and zero for nontargets.

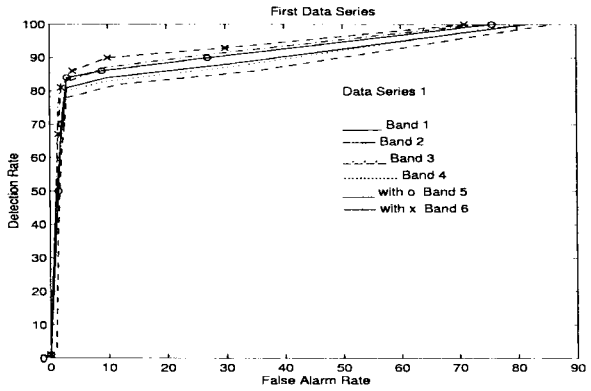
The training data set for the BPNN detector/classifier was chosen from two sets of eight different scenes; eight targets, and eight nontargets blocks were selected from each scene. These targets and nontargets had different sizes, shapes, contrasts, etc. The number of training data samples was dependent on the size of the subwindows. The training was completed after 3000 epochs and the final average mean square error (AMSE) was around 0.2. Once the training process was completed, the network was used to detect and classify all the testing image blocks.

Fig. 4. Image of band 3 after contrast enhancement.

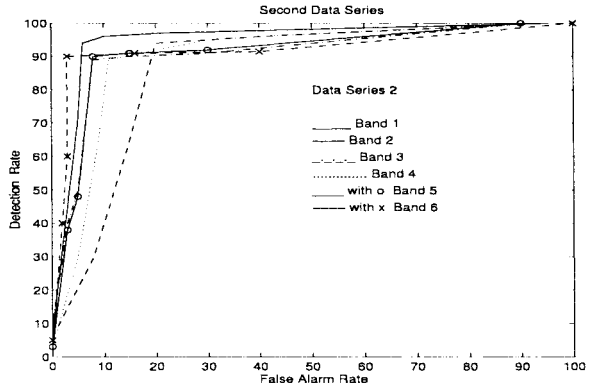
The performance of the network was tested based on all the blocks of the two sets of eight images which included both the training data and the other parts of the images which the network had not seen before. A block was considered to be a possible target at a particular band if its score in the accumulator was greater than a threshold which was determined empirically to be around 47% of the total number of subwindows.

In order to determine the performance of the BPNN detection/classification subsystem for different cases, the detection and false alarm rates were calculated. In addition, the receiver operating characteristic (ROC) curve of the results was also generated. The ROC curve [25] is the plot of the conditional probability of the detection versus the conditional probability of false-alarm responses. The results in [25] stated that the area under the ROC curve gives the probability of the correct detection.

Two series of images were used in this testing phase. The first series consisted of data, in which the size of the targets was small, e.g., typically of 7×7 pixels, and the contrast between the targets and background was low. However, background areas did not exhibit any rapid gray levels changes. In contrast to the first series, the target size was larger in the second series e.g., 12×12 pixels, whereas the changes in gray levels were sudden between the targets and their surrounding background. Fig. 5(a) and 5(b) show the ROC curves for these data sets for different thresholds ranging from 5% to 90%. The results indicated that the targets in bands 5 and 6 had very good detectability for the first series, while bands 1 and 6 were the best for the second series. The detection and false alarm rates were found to be much higher for the second series than for the first series. In the second series, background clutter in some areas had the same intensity variations as with the targets, hence causing higher false alarm rates in some bands. This was indicated in the results of the area under ROC curve which showed that the results of the first series were better than those of the second series for some individual optical bands.



(a)



(b)

Fig. 5. ROC curves for six optical bands and two sets of data.

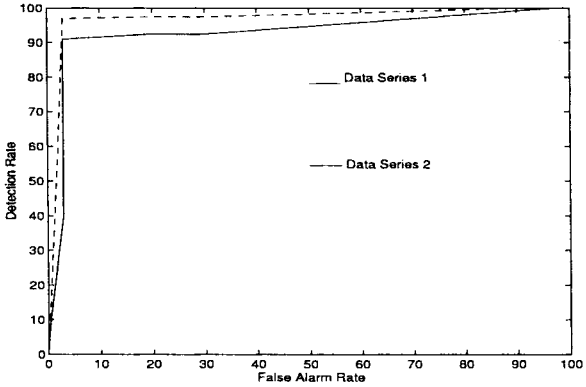
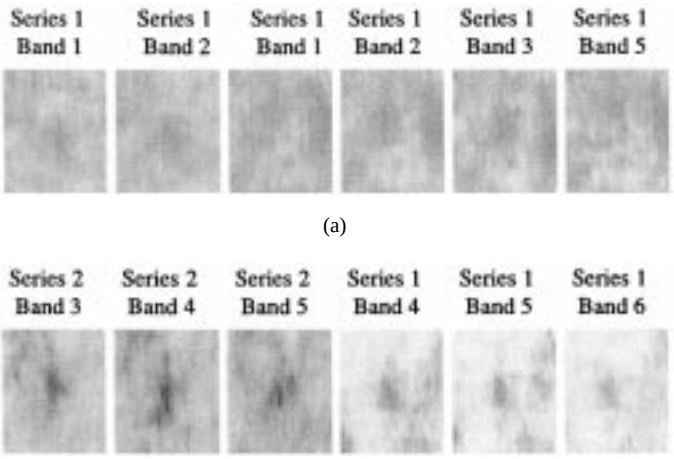


Fig. 6. The ROC curve of combining six optical bands.



(a)

(b)

Fig. 7. Some typical misclassified target and nontarget blocks.

In order to improve the performance of the detector, the results of the individual bands were fused together in a final decision making system. Two different final decision making systems were tested on this problem. The first approach combines the contents of the accumulators in each particular band corresponding to each subwindow. If the sum of these was greater than a threshold which was also determined empirically to be around 55% of the total number of subwindows of six optical bands, then the block was finally classified as target. Fig. 6 shows the ROC curves of the combined detection results from the six bands and for both data series. The area under ROC curve was measured to be 0.9722 for the second series and 0.9469 for the first series which showed considerable improvements over the individual results. The false alarm rate was 2% for the second series and 3% for the first series for threshold level of 47% and 45%, respectively. These results showed substantial decrease in the false alarm rate, especially for the second series, when the channel outputs are fused. One of the second series images, was tested to investigate the performance of the fusion system for each individual target and nontarget. Although in certain bands some targets were detected as nontargets, and some nontargets were detected as targets, the output of the final decision making system correctly classified them. This is due to the fact that even though these targets and nontargets were not correctly classified in some individual bands, they were correctly picked up in the other bands. Thus, the combined results of all the six optical bands takes care of the misclassificatios in certain individual bands. The same results could be obtained for other images. As a result, the final decision making subsystem plays a very important role in detecting dim and/or partially obscured targets. Fig. 7 shows some typical target and nontarget blocks that were misclassfied. Closer investigation revealed that the target misdetections mostly occurred for data set 1, in which the targets were smaller in dimension, and for spectral bands 1–3. The latter can easily be explained taking into account the fact that this particular data set was collected during the morning (8:58–10:32 am) on a sunny day in April 1991. Consequently, depending on the location and exposure to sunlight some of the targets may not show up in spectral bands 1–3 hence leading to misdetections.

Besides the majority voting, the linear opinion pool described in Section II-E was also tested as the final decision subsystem. The optimal weight vector was calculated based upon the training samples of all the six bands using (4). Interestingly enough, it was found that the weight associated with band 6, i.e., near IR, was the largest among all the bands due to the importance of this band in sunny conditions. Fig. 8 gives the ROC curves for both the optimal weighted combination or linear pool and the simple voting scheme on

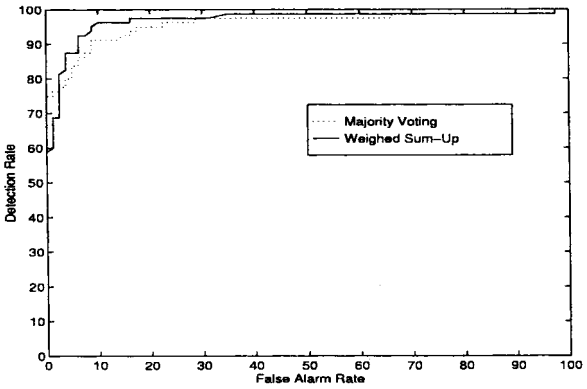
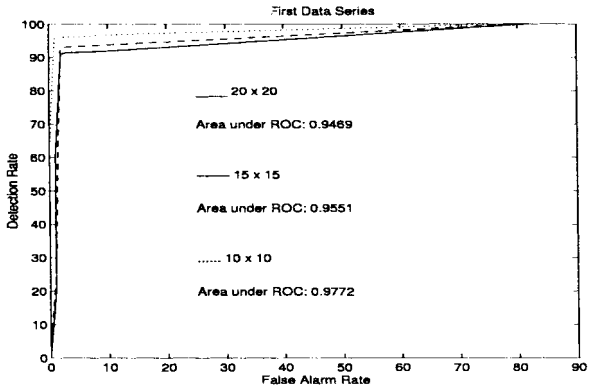


Fig. 8. ROC curves for the two different decision systems.

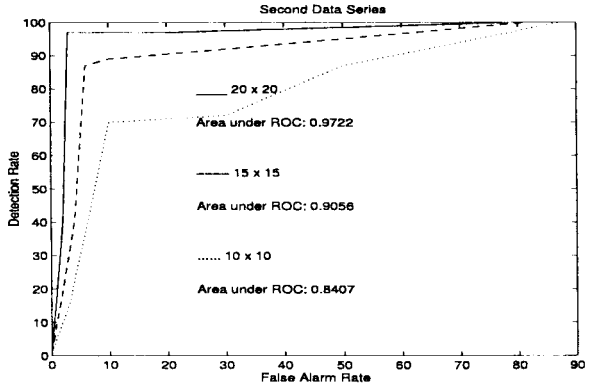
the combined data sets 1 and 2. As can be observed from these ROC curves, linear opinion pool approach outperforms the voting scheme in terms of both detection and false alarm rates. This is mainly due to the fact that unlike the voting scheme which uses the uniform weights for all the six bands, this approach adaptively places more emphasis on those significant spectral bands depending on the conditions during the data collection.

Since the target sizes are variable even in the same image, the suitable choice of the subwindow size can greatly improve the detection and false alarm rates. Choosing the right subwindow size turned out to be one of the most important factors in our target detection system. Three different subwindow sizes i.e., 20×20, 15×15, and 10×10 pixels were tried and the ROC curves were plotted in Fig. 9. The results revealed that for the first series images, the detection and false alarm rates were not very sensitive to the choice of subwindow size while these rates significantly changed for the second series. The best performance for the second series was achieved for 20×20 subwindow size since the targets were larger in size in this series. As the subwindow size decreased, the detection performance deteriorated. Generally speaking, if the size of the targets is larger than subwindow size, poor results are obtained as the subwindows will not contain the surrounding background. As a result, the extracted features would look like those of the plain background after the mean extraction process. This result can also be verified by evaluating the area under the ROC curve which shows that for the smallest subwindow size the area was smaller than that of the larger subwindow size. Thus, the 20×20 was found to be the best choice of subwindow size for this application.

Finally, the performance of the 11-25-10-1 BPNN was compared with those of a two-layer BPNN with architecture 11-15-1 as well as the standard ML classification scheme. The training of the two-layer network was completed after 3000 epochs. All the network parameters were chosen to be the same as those of the three layer. In order to overcome the issue of local minima, four networks with different randomly chosen initial weights were examined. The best neural network was chosen for which the ROC curve after the final decision making process is given in Fig. 10. Under the assumption that the input features have multidimensional Gaussian distribution, the ML classifer can be easily implemented. In the training



(a)



(b)

Fig. 9. ROC curves for different subwindow sizes.

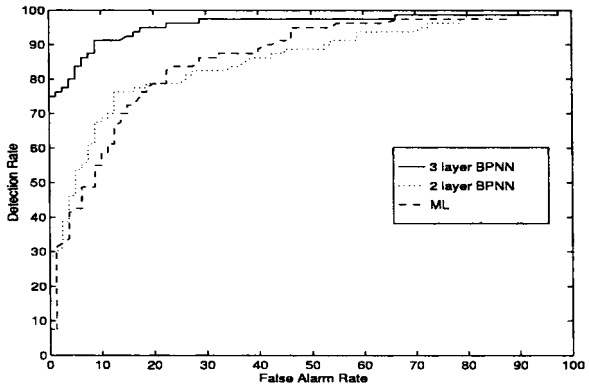


Fig. 10. ROC curves for the three classifiers.

phase, the mean and covariance of the target and nontarget were calculated for each spectral band from the training data set. In the testing phase, the likelihood function for each test sample is calculated and compared with a threshold. The ROC curves of the two BPNN structures and that of the ML scheme are shown in Fig. 10. As can be observed, the result of ML is inferior to that of the three-layer BPNN. One reason for this poor performance of ML scheme is attributed to the fact that the normal distribution assumption for the features is not an accurate assumption for this problem. Using more complicated Gaussian mixture models may indeed yield better performance. Additionally, the two-layer BPNN did not produce comparable results either.

IV. Conclusion

This paper presents a neural-network-based scheme for performing surface-laid mine target detection/classification. The proposed strategy involved the use of various neural-networks schemes for performing feature extraction and detection/classification tasks. An RLS-based learning rule for target feature extraction was used which provided a fast and accurate means of PC extraction. It was shown that the use of a neural network in each channel provided a useful tool for target detection/classification without requiring any prior statistical information. Two different final decision making systems were also considered to combine the detection and classification results of the neural networks in all the six spectral bands. Simulation results on six different optical bands were provided which indicated the effectiveness of the proposed schemes. The combined results of all the channels using the optimal weighted combination method showed excellent overall performance.

Acknowledgment

The authors would like to thank U.S. Marine Corps, Systems Command, and Naval Surface Warfare Center, Dahlgren Division, Coastal Systems Station, Panama City, FL, for providing the data needed for this study.

References

[1] N. H. Witherspoon and J. Holloway, Jr., "Video based multispectral detection of land mines, a technology application for use in law enforcement," in *SPIE Proc.*, Surveillance Technologies II, 1992, vol. 1693, pp. 185–194.

[2] N. H. Witherspoon *et al.*, "Coastal battlefield reconnaissance and analysis (COBRA) program for minefield detection," in *SPIE Proc.*, Detection Technologies for Mines and Minelike Targets, 1995, vol. 2496, pp. 500–508.

[3] B. Bhanu, "Automatic target recognition: State of the art survey," *IEEE Trans. Aerospace Electron. Syst.*, vol. AES-22, pp. 354–379, July 1986.

[4] M. W. Roth, "Survey of neural-network technology for automatic target recognition," *IEEE Trans. Neural Networks*, vol. 1, pp. 28–43, Mar. 1990.

[5] M. R. Azimi-Sadjadi *et al.*, "Detection and classification of dielectric anomalies using a separated aperture sensor and a neural network discriminator," *IEEE Trans. Instrum. Meas.*, vol. 41, pp. 137–143, Feb. 1992.

[6] M. R. Azimi-Sadjadi and S. A. Stricker, "Detection and classification of dielectric anomalies using neural networks—Further results," *IEEE Trans. Instrum. Meas.*, vol. 43, pp. 34–39, Feb. 1994.

[7] A. N. Balan and M. R. Azimi-Sadjadi, "Detection and classification of buried dielectric anomalies by means of the bispectrum method and neural networks," *IEEE Trans. Instrum. Meas.*, vol. 44, pp. 998–1002, Dec. 1995.

[8] B. Telfer *et al.*, "Detecting blobs in multispectral electrooptical imagery using wavelet techniques," in *SPIE Proc.*, Visual Information Processing II, 1993, vol. 1961, pp. 377–385.

[9] B. Telfer *et al.*, "Adaptive wavelet classification of acoustic backscatter and imagery," *Opt. Eng.*, vol. 33, no. 7, pp. 2192–2203, July 1994.

[10] T. Daud *et al.*, "Mine discrimination using multispectral imagery with feedward neural networks," in *SPIE Proc.*, Detection Technologies for Mines and Minelike Targets, 1995, vol. 2496, pp. 614–621.

[11] G. A. Clark *et al.*, "Multispectral image fusion for detection land mines," in *SPIE Proc.*, Detection Technologies for Mines and Minelike Targets, 1995, vol. 2496, pp. 850–864.

[12] Q. A. Holmes *et al.*, "Adaptive multispectral CFAR detection of land mines," in *SPIE Proc.*, Detection Technologies for Mines and Minelike Targets, 1995, vol. 2496, pp. 421–432.

[13] M. S. Schmalz *et al.*, "Center-surround filter for the detection of small targets in cluttered multispectral imagery I&II," in *SPIE Proc.*, Detection Technologies for Mines and Minelike Targets, 1995, vol. 2496, pp. 637–648 and 756–766.

[14] D. Flynn *et al.*, "Intensified camera modeling and validation," in *5th Annu. Ground Target Modeling and Validation Conf.*, Houghton, MI, Aug. 23–25, 1994.

[15] H. R. Suiter, "Methods used in evaluating multispectral camera resolution for land-mine detection," in *SPIE Proc.*, Detection Technologies for Mines and Minelike Targets, 1995, vol. 2496, pp. 249–258.

[16] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[17] S. Bannour and M. R. Azimi-Sadjadi, "Principal component extraction using recursive least squares learning method," *IEEE Trans. Neural Networks*, vol. 6, pp. 457–469, Mar. 1995.

[18] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[19] E. Oja, "A simplified neuron as a principal component analyzer," *J. Math. Biol.*, vol. 15, pp. 267–273, 1982.

[20] T. D. Sanger, "Optimal unsupervised learning in a single-layer linear feedforward neural network," *Neural Networks*, vol. 2, pp. 459–473, 1989.

[21] J. A. Freeman and D. M. Skapura, *Neural Networks: Algorithms, Applications and Programming Techniques*. Reading, MA: Addison-Wesley, 1991.

[22] J. A. Benediktsson and P. H. Swain, "Consensus theoretic classification methods," *IEEE Trans. Syst. Man, Cybern.*, vol. 22, pp. 688–70, July/Aug. 1992.

[23] J. A. Benediktsson *et al.*, "Parallel concensual neural networks," *IEEE Trans. Neural Networks*, vol. 8, pp. 54–63, Jan. 1997.

[24] R. A. Jacobs, "Methods for combining experts' probability assessments," *Neural Computa.*, vol. 3, pp. 867–888, 1995.

[25] M. L. Meistrell, "Evaluation of neural-network performance by receiver operating characteristic (ROC) analysis: Examples from biotechnology domain," *Computer Methods and Programs in Biomedicine*, vol. 32, pp. 73–80, 1990.

Xi Miao received the B.S. degree from Huazhong University of Sciences and Technology, Wuhan, China, in 1983, and the M.S. degree from Colorado State University, Fort Collins, CO, in 1995, both in electrical engineering.

He was a Control System Engineer in the Guangzhou Institute of Energy Conversion, Chinese Academy of Sciences, China, from 1983 until 1992. He is now with Cirrus Logic Inc. as a Software Engineer in the System Engineering Department.

Mahmood R. Azimi-Sadjadi (S'81–M'81–SM'89) received the B.S. degree from University of Tehran in 1977 and the M.Sc. and Ph.D. degrees from the Imperial College, University of London, U.K., in 1978 and 1982, respectively, all in electrical engineering.

He served as an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Michigan, Dearborn. Since July 1986 he has been with the Department of Electrical Engineering, Colorado State University, Fort Collins, where he is now a Professor. He is also the director of the Multi-Sensory Computing Laboratory (MUSCL) at Colorado State University. His areas of interest include digital signal/image processing, target detection and tracking, multidimensional system theory and analysis, adaptive filtering, system identification, and neural networks. He has more than 100 journals and refereed conference publications. He is a coauthor of the book, *Digital Filtering in One and Two Dimensions* (New York: Plenum, 1989).

Dr. Azimi-Sadjadi is the recipient of 1993 ASEE-Navy Senior Faculty Fellowship Award, 1991 CSU Dean's Council Award, 1990 Battelle Summer Faculty Fellowship Award, and the 1984 DOW chemical Outstanding Young Faculty Award of the American Society for Engineering Education. He is an Associate Editor of the IEEE Transactions on Signal Processing.

Bin Tian received the B.S. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1991 and 1993, respectively.

Since 1995, he has been with the Image/Signal Processing Laboratory, Department of Electrical Engineering, Colorado State University, Fort Collins, where he is currently a Ph.D. candidate. His main research interests are focused on signal/image processing and the theory and application of neural networks.

Abinash C. Dubey (S'87–A'87) received the B.S. degree in physics and mathematics from Bihar University in 1958 and the M.S. degree in geophysics from Western Ontario University, London, Ont., Canada, in 1966.

He has been working as a Research Physicist at the Naval Surface Warfare Center (NSWC), Coastal Systems Station, since 1987. His research areas have been in shallow underwater imaging using active laser source, multispectral imaging for surface and buried targets using passive electrooptic source, and underwater target detection using volume search sonar. His main research interests are in target detection and classification, data compression, and visualization. Prior to joining NSWC he worked in the oil industry in 2-D and 3-D signal processing, target detection, and reservoir characterization.

Dr. Dubey is a member of SPIE, the American Institute of Physics, and the Society of Exploration Geophysicists.

Ned H. Witherspoon received the B.S.degree in mathematics and the M.S. degree in physics from Florida State University, Tallahassee, FL.

In 1969 he joined the faculty at Gulf Coast Community College, Panama City, FL as an Assistant Professor of Physics and left in 1982 as a Full Professor of Physics. From 1982 to 1997 he was a Research Physicist at the Naval Surface Warfare Centre Dahlgren Division-Coastal Systems Station (CSS), Panama City. Currently, he is the Manager of the U.S. Marine Corps Project Office at CSS for mine detection and the Project Engineer for the Coastal Battlefield Reconnaissance and Analysis (COBRA) program. He is also an Adjunct Faculty Member at Embry-Riddle Aeronautical University. He has spent his career in developing electrooptic systems for the military remote sensing applications.

Mr. Witherspoon was Outstanding Professor of the Year at Gulf Coast Community College. He received the Meritorious Civilian Service Award the Marine Corps, and the Commanding Officer/Technical Director's Award at CSS.