Detection and Classification of Buried Dielectric Anomalies by Means of the Bispectrum Method and Neural Networks

Ajay N. Balan and Mahmood R. Azimi-Sadjadi, *Senior Member, IEEE*

Abstract—The development of a neural network-based system for detection and classification of buried landmines is the main focus of this paper. Shape-dependent features are extracted by means of the bispectrum method. These features are then applied to the neural network. A multilayer back-propagation-type neural network is trained and tested on the feature sets extracted from equally spaced radial slices of image windows. Simulation results obtained for two types of targets indicated good detection and classification rates.

I. INTRODUCTION

THE development of a target detection system capable of detecting various types of buried landmines, under different environmental conditions presents many technical problems. Some of the problems which inhibit detection are: targets have different sizes, compositions and dielectric properties; soil properties vary with location and moisture content; the depths at which targets are buried may vary, leading to variable sensor responses, and similarity of responses of targets and nontargets.

Various target detection schemes have been applied to this detection problem. In [1] a neural network-based approach was suggested which offers potentially powerful, robust and adaptive means of detecting and classifying targets in high cluttered backgrounds. The data representation scheme used was the method of principal components, or the Karhunen-Loeve (KL) transform. This method reduces and decorrelates the input data to the neural networks. The simulations performed using this method of data representation provided excellent results as far as detection is concerned. However, this method requires forming a generalized transformation matrix to obtain reduced images from weak and strong targets of different compositions. This may cause smearing effects in the classification performance. In addition, this method is energy dependent, and hence the detection results depend largely on the amplitude of the return.

In this paper, the bispectrum method [2]–[6] is used to extract shape-dependent features from target and nontarget images. These features are invariant to the effects of translation, amplification, and D.C. level variation, and further are not affected by additive Gaussian noise. An extension of the bispectrum algorithm to the 2-D case is suggested, in which

the features are extracted from the entire image window. These features are then applied to a multilayer back-propagation neural network to perform the detection/classification tasks. Simulation on targets with two different compositions is performed for single- and multiple-frequency data. The effects of increasing the number of slices on the performance of the detector/classifier are studied, and the results are then compared with those of the KL transform in [1].

II. BISPECTRUM IN SHAPE DISCRIMINATION

It is known that information regarding the signal or image shape lies primarily in the relative phases of the Fourier transform [7], and signals with similar magnitude, but different phases, will produce widely varying bispectra. This has been observed in the case of quadrature phase coupling in random signals with similar power spectra [5]. Hence, a bispectrum-based technique may be used to discriminate between signals with different phases. The next step would be to extract suitable features to classify different shapes based upon the similarity measures. The approach presented in [2] is based on extraction of features or parameters that are invariant under translation, amplification and D.C. level variations of the signal.

The bispectrum $B_x(\omega_1, \omega_2)$ of the deterministic discrete-time signal $x[n]$ can be calculated using

$$B_x(\omega_1, \omega_2) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} m_3^x(k, l) e^{-j(\omega_1 k + \omega_2 l)} \quad (1)$$

where $|\omega_1|$, $|\omega_2| \le \pi$ and $|\omega_1 + \omega_2| \le \pi$, and $m_3^x(k, l)$ is the third-order moment (or high-order correlations) [4] of $x[p]$ given by

$$m_3^x(k, l) = \sum_{p=-\infty}^{\infty} x[p]x[p + k]x[p + l]. \quad (2)$$

Alternatively, in the frequency domain we have

$$B_x(\omega_1, \omega_2) = X(-(\omega_1 + \omega_2))X(\omega_1)X(\omega_2) \quad (3)$$

where $X(\omega)$ is the Discrete Time Fourier Transform (DTFT) of $x[n]$. Fig. 1 shows the entire region of the bispectrum. Owing to the symmetry properties [5], [6], the computation of the bispectrum can be confined to the shaded region which is referred to as the "Primary Region" for which $0 \le \omega_1 \le \omega_2 \le \pi$. All other regions of the bispectrum contain
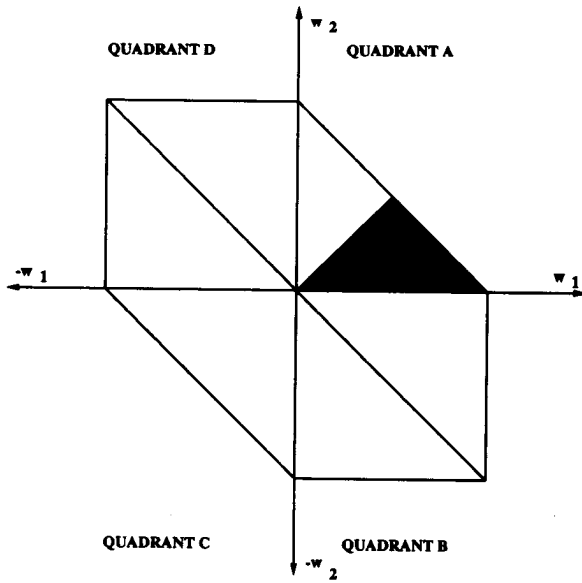
Fig. 1. Regions of the bispectrum.

redundant information which can be obtained from the points in this region. Henceforth, any reference to calculation of the bispectrum refers to calculations in this triangular region only. Note that if ω is normalized frequency, then $|\omega| \leq 1$.

Now let parameters $I(a)$ be calculated for a set of predetermined values of 'a' as

$$I(a) = I_{\mathrm{re}}(a) + jI_{\mathrm{im}}(a) = \int_{\omega_1=0^+}^{1/(1+a)} B_x(\omega_1, a\omega_1)d\omega_1 \quad (4)$$

and $0 < a \leq 1$. Note that $I(a)$ is the integrated bispectrum along straight lines (each with a slope a) within the primary bispectral region, passing through the origin in bifrequency space. Next, parameters $P(a)$ are calculated using

$$P(a) = \arctan(I_{\mathrm{im}}(a)/I_{\mathrm{re}}(a)) \quad (5)$$

These $P(a)$ are the principal components of the phase of the integrated bispectrum $I(a)$ [2]. The set of these $P(a)$ is then used as the feature vector representing $x[n]$. It can be shown [2] that parameters $P(a)$ and $I(a)$ are translation and D.C. level invariant. Further $P(a)$ are amplification invariant [2].

The bispectrum performs well in noise as the bispectrum of Gaussian noise is zero. The data reduction capability of the algorithm, however, is variable, as the exact number of $P(a)$ needed for good discrimination can vary. In general, the number of $P(a)$ required is greater if different signals have similar shapes. This is due to the fact that, if the input pattern shapes are similar to each other, the spectral phases will be similar, which may lead to some similarities between different $P(a)$ feature vector sets. Typically, if a small number of $P(a)$ are used to represent a particular pattern, then the classifier has better generalization capabilities. On the other hand, better class separation is possible as the number of $P(a)$ used to represent a pattern increases. One way of overcoming the problem of class separation without increasing the dimension of the feature vector is to compute the variation of a particular $P(a)$, say $P(0.5)$ as the time scale (length) of the input signal

varies over a specified range of values [2]. While this approach may work for some cases such as automated robot vision [2], its applicability in a buried landmine detection scenario is questionable as a 2-D extension of this method is needed.

A. Slicing Algorithm

From polyspectral theory [3] it is evident that the bispectrum of a 2-D signal or image is four dimensional; hence its computation is cumbersome. In many 2-D applications such as image reconstruction [8] it is a common practice to compute a partial bispectrum, and work with this information. The approach taken in this section utilizes 1-D slices of the image window to compute the corresponding 2-D bispectral coefficients. The 1-D radial slices, originating at the center of each target, background or anomaly image window, are selected at regular equal angular spacing. It can be argued that a sufficient number of these slices provide an adequate representation of the image. A pattern set is then extracted from the Bispectrum of each 1-D slice, and these pattern sets are lumped together to obtain a composite feature vector.

The main reason for adopting the particular slicing strategy is that most target responses are symmetrical about the center of the window. Anomaly responses, on the other hand, are not necessarily symmetrical about the center of the image window. As a result, to obtain the most effective discrimination between targets and anomalies, this symmetry property of target responses must be exploited. If slices are obtained with a reference point other than the center, the shape of some target slices and anomaly slices may be similar. This can be proven to be detrimental to detection/classification, especially if a small number of slices are used.

In the following development the DFT [2] is used for practical implementation. The algorithm [9] consists of the following steps.

Step 1: Choose q equally spaced slices of the image $x[m, n]$, i.e. $x_k[p]$, $k \in [1, q]$, $q = 2, 4, 6 \ldots$ and $p \in [1, N - 1]$ where N is the total number of points on each slice.

Step 2: For each of the slices, say the kth, the bispectrum $B_k(k_1, k_2)$ is calculated using

$$B_k(k_1, k_2) = X_k(k_1)X_k(k_2)X_k^*(k_1 + k_2) \quad (6)$$

where $X_k(l)$ is the DFT of $x_k[p]$, and the primary region for $B_k(k_1, k_2)$ satisfies $0 \leq k_2 \leq k_1 \leq k_1 + k_2 \leq ((N/2) - 1)$.

Step 3: Calculate $I_k(a)$ for each $B_k(k_1, k_2)$ using

$$I_k(a) = I_{k_r}(a) + jI_{k_i}(a) = \sum_{k_1=1}^{\mathrm{USUM}} B_k(k_1, ak_1) \quad (7)$$

where $0 < a \leq 1$, and $a = m\Delta a$, Δa being the uniform increment by which a increases. Also $m \in [1, A]$ where 'A' is the number of $I_k(a)$ values for the kth slice:

$$\mathrm{USUM} = \mathrm{floor}(((N/2) - 1)/(1 + a)) \quad \text{for } N \text{ even}$$
$$\mathrm{USUM} = \mathrm{floor}(((N - 1)/2 - 1)/(1 + a)) \quad \text{for } N \text{ odd}$$

where $\mathrm{floor}(x)$ is the integer obtained by truncating the fractional part of x, and USUM is chosen to ensure that

the summation includes bispectral coefficients within the primary region.

Step 4: For noninteger values of 'ak_1' the bispectrum is interpolated by

$$B(k_1, ak_1) = bB(k_1, \text{ceil}(ak_1)) + (1 - b)B(k_1, \text{floor}(ak_1)) \tag{8}$$

where $b = ak_1 - \text{floor}(ak_1)$, $\text{floor}(ak_1)$ is the integer obtained by truncating the fractional part of 'ak_1', and $\text{ceil}(ak_1) = \text{floor}(ak_1) + 1$.

Step 5: Calculate the principal components of the phase of $I_k(a)$, i.e. $P_k(a)$, using

$$P_k(a) = \arctan(I_{k_r}(a)/I_{k_i}(a)). \tag{9}$$

Step 6: Repeat for all the $x_k[p]$, $k \in [1, q]$, and obtain the feature vectors using all the $P_k(a)$ values.

These feature vectors are then used for training and testing of the multilayer neural network.

III. SIMULATION RESULTS

All the data used in this study was experimentally collected at the U.S. Army, Ft. Belvoir RD & E center [1]. The targets were rectangular blocks of size $12 \times 12 \times 3$ inches and were buried at various positions, angles and depths. They were laid out in rectangular lanes, and the sensor scanned across these lanes. A scan involved taking measurements at 27 positions, 1.5 inches apart for a total scan width of 39 inches. Once the sensor head had collected 27 horizontal data points, the cart was moved forward 1.5 inches and the process repeated. The data collected was represented as a rectangular matrix or array of data points. There were two types of buried targets, wood and nylon. For wood targets the collected image file had 27 columns and 504 rows of data, and the collected nylon image file had 27 columns and 487 rows.

Two types of simulations were performed. The first set of simulations incorporated data from sensor returns at 792 MHz which was the resonant frequency of the sensor. The second set of simulations involved data from sensor returns at five frequencies: 728, 760, 792, 824 and 856 MHz, i.e., the resonant and four side frequencies. All simulations were performed using three-layer back-propagation neural networks [10]. Feature vectors were obtained from windows of size 15×15 containing target and nontarget returns using the algorithm presented in Section II-A. For the purpose of practical implementation, the feature vector set was scaled to lie between 0 and 1. To avoid the effects of D.C. level fluctuations, $X_k(0)$ is set to zero for each slice. This is done to ensure that all the bispectral components $B(k_1, 0)$ are zero. If this was not done, values of $B_k(k_1, ak_1)$ where ak_1 was less than 1 would be influenced by $B(k_1, 0)$ due to the interpolation involved in (8). Thus, the algorithm would no longer be D.C. level invariant. This approach is similar to working with a mean subtracted signal, in which case $X_k(0)$ would be zero. Fig. 2 show the plots of feature vector sets using 4 slices for selected sets of nylon, wood and background images obtained at 792 MHz. As can be seen from these plots, while there are considerable similarities between the feature sets of nylon and wood targets,

TABLE I
PERFORMANCE RESULTS OF SINGLE-FREQUENCY NETWORKS FOR THE SLICING METHOD

Frequency	Network Architecture	Nylon			Wood		
		Detection %	Classification %	False Alarm %	Detection %	Classification %	False Alarm %
2 Slices							
792	20-15-10-3	100	66	25	100	75	17
4 Slices							
792	40-30-15-3	100	66	8	100	83	17
8 Slices							
792	40-30-15-3	100	80	8	100	83	8

TABLE II
PERFORMANCE RESULTS OF FIVE-FREQUENCY NETWORK USING DATA FROM FOUR SLICES OF THE IMAGE WINDOW

Frequency (MHz) 4 Slices	Network Architecture	Nylon			Wood		
		Detection %	Classification %	False Alarm %	Detection %	Classification %	False Alarm %
728	40-30-15-3	87	69	8	92	55	8
760	40-30-15-3	93	86	33	92	73	8
792	40-30-15-3	87	92	8	92	73	33
824	40-30-15-3	87	92	8	75	78	8
856	40-30-15-3	93	93	0	92	82	8

the feature sets corresponding to background clutter do not generally follow a regular pattern. This property would allow the discrimination of weak targets from the competing clutter.

For the single-frequency case (792 MHz), a training data set consisting of 4 target and 4 background images was chosen from both nylon and wood databases. The training data set was chosen to provide a representation of strong, medium and weak targets. The network had three units in the output layer, and the desired outputs for nylon, wood and background images were $(1, 0, 0), (0, 1, 0)$ and $(0, 0, 1)$, respectively. The weights were initialized randomly in the range $(-0.1, 0.1)$. For all the simulations, the training took 25 000 epochs. The trained network was then used to test all the remaining target and background windows in the database. For the five-frequency case, similar training data sets as for the single-frequency case were chosen which consisted of 4 target and 4 background images from each frequency file. During the training phase, these vectors were applied sequentially to the neural network. The network parameters such as the initial range of random weights, step size and the desired outputs were similar to those of the single-frequency case. The five-frequency sequential network was trained with over 30 000 epochs.

For the single-frequency case simulations were performed with two, four and eight slices of the image window for each of the training sets. For the five-frequency case, however, simulations were performed with four slices of image windows using a training data set from each of the five frequencies. An array of ten $P(a)$ values was calculated from each of the slices. Thus, for four slices the input to the neural network was forty points long, and correspondingly for two slices the input was twenty points long. For the single-frequency simulation involving eight slices, five $P(a)$ values were calculated for each slice. This was done to keep the size of the network manageable. Tables I and II show the detection, classification, and false alarm rates obtained based upon the performance of the neural network detector/classifier on the testing data.

The method of slicing provided reasonable detection rates in most cases. It was noted that as the number of slices increased,
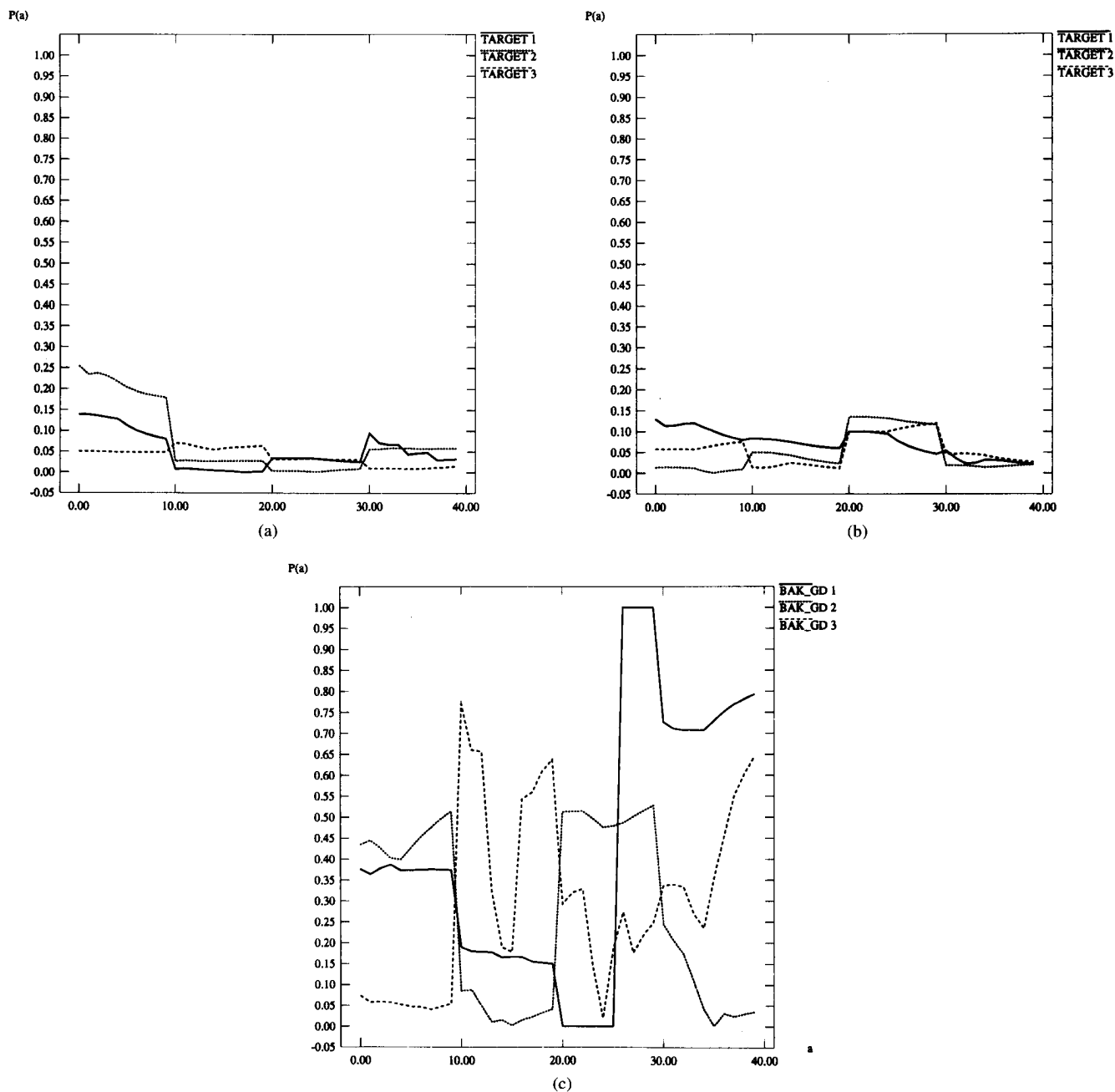
Fig. 2. A selected set of nylon target (a), wood target (b), and background (c) feature vectors for 4 slices of image window.

the detection and false alarm rates were somewhat improved. This was expected, because an increase in the number of slices implies better signal representation. Simulations involving the single-frequency (792 MHz) data provided better detection than those involving five frequencies. This can be attributed to the fact that the process of target detection is dependent on the presence of each pattern class in the training samples, and representation of the same training set at different frequencies may not convey additional information pertaining to the shapes of the targets to the network. In addition, there are smearing effects caused by using data from 5 frequencies to train the neural network. In [1] a multiple neural network fusion system was suggested to improve the overall performance and avoid the smearing effects of the sequentially trained system. As

far as the classification rate is concerned, the five-frequency network performed better than the single one. This leads to the conclusion that classification capabilities depend on how well the network generalizes, which is a function of the variety of the training data set. The false alarm rate was at times higher than desired. However, the simulation demonstrated the fact that the network performed better as far as the false alarm rate was concerned when the number of slices increased. The number of slices, however, is restricted due to computational limitations. This seemed to be the major drawback of the proposed system. Another problem is the sensitivity of the bispectral estimator, which can cause abrupt fluctuations in the values of $I(a)$ and $P(a)$, at data points where there may have been sensor irregularities.

The effectiveness of the bispectrum method was compared with that of the KL method [1]. The results in [1] indicated that the KL method provided excellent detection and false alarm rates, for example, an average of 93% detection rate for zero false alarm rate. The detection results obtained by the bispectrum method are slightly better than those obtained by the KL method, but the false alarm rates are relatively higher. The KL method provides good false alarm rates because, in general, the energy present in a target window is higher than that of a background window. Thus, background images can easily be discriminated. By similar reasoning, the detection of very weak targets by means of the KL method is not efficient. In addition, the KL method requires forming a generalized transformation matrix to obtain the reduced images from weak and strong targets of different compositions. This obviously causes smearing effects in the detection and classification rates if a large number of targets with different shapes, sizes and compositions need to be detected and classified. In addition, this method is energy dependent, and hence the detection results depend largely on the amplitude of the return from a landmine. Thus, this method may not be able to distinguish between sensor returns from anomalies and targets with similar magnitudes and when the signal-to-noise ratio is very low. The bispectrum method, on the other hand, is a shape-dependent method which can offer discrimination between different targets on the basis of their shape and orientation, not their strength. Additionally, the bispectrum is invariant in the presence of Gaussian noise.

IV. CONCLUSION

The primary focus of this paper was on the application of the bispectrum method in conjunction with neural networks for the detection and classification of buried landmines. The advantage of the bispectrum method is that it provides translation, amplification and D.C. level invariant features. A data representation scheme was used in which features from equally spaced slices of target and background image windows were extracted. The extracted features of the slices were then used to represent the images for the training and testing of the multilayer back-propagation neural network detector/classifier. The feature extraction scheme used in this work utilized just the phase of the integrated bispectrum, i.e., the phases of $I(a)s$, $P(a)s$ were used. The magnitudes of the $I(a)s$ may contain additional information about the patterns. Additionally, the scheme used here could be extended to process complex-valued signals, which would lead to a bispectrum with two different nonredundant regions [5]. To improve the overall performance, the range of computations of $I(a)s$ and $P(a)s$ may be restricted to a sub-region of the primary Bispectral domain, in order to de-emphasize the spectral interactions between high-frequency noise (bispectral coefficients close to the outer perimeter of the shaded region in Fig. 1) and low-frequency signal components (bispectral components close to the origin in Fig. 1).

REFERENCES

[1] M. R. Azimi-Sadjadi, D. E. Poole, S. Seedvash, K. D. Sherbondy, and S. Stricker, "Detection and classification of dielectric anomalies using a separated aperture sensor and a neural network discriminator," *IEEE Trans. Instrum. Meas.*, vol. 41, no. 1, pp. 137–143, Feb. 1992.

[2] V. Chandran and S. Elgar, "Shape discrimination using invariants defined from higher order spectra," *Proc. of the 1991 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Toronto, May 1991, pp. 3105–3108.

[3] G. B. Giannakis and M. K. Tsatanis, "Signal detection and classification using matched filtering and higher order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, pp. 1284–1296, July 1990.

[4] C. L. Nikias and M. R. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proc. IEEE*, vol. 75, no. 7, pp. 869–891, July 1987.

[5] Y. C. Kim and E. J. Powers "Digital bispectral analysis and its applications to nonlinear wave interactions," *IEEE Trans. Plasma Sci.*, vol. PS-7, no. 2, pp. 120–131, June 1979.

[6] T. Matsuoka and T. J. Ulrych, "Phase estimation using the bispectrum," *Proc. IEEE*, vol. 72, pp. 1403–1411, Oct. 1984.

[7] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE*, May 1981, pp. 529–541.

[8] A. W. Lohmann, G. Weigelt, and Wirnitzer, "Speckle masking in astronomy: Triple correlation theory and applications," *Applied Optics*, vol. 22, no. 24. pp. 4028–4037, Dec. 1983.

[9] A. N. Balan and M. R. Azimi-Sadjadi, "Detection of buried anomalies by means of the bispectrum method," in *Proc. IEEE Instrum. Meas. Technol. Conf.*, 1992, pp. 80–83.

[10] J. A. Freeman and D. M. Skapura, *Neural Networks: Algorithms, applications and programming techniques.* Reading, MA: Addison Wesley, 1992.
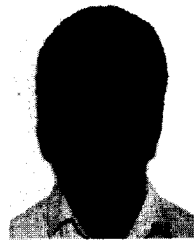
Ajay N. Balan was born in Bangalore, India, on October 23, 1966. He received the Bachelor of science degree in electrical engineering from Bangalore University in 1989. He received the M.S. degree in electrical engineering from Colorado State University in 1992. His research involved Image Processing and Pattern Recognition. His current interests include video compression and computer networks.

Mahmood R. Azimi-Sadjadi (S'81–M'81–SM'89) was born in Tehran, Iran, in 1952. He received the B.S. degree from University of Tehran, Iran in 1977, the M.Sc. and Ph.D. degrees from the Imperial College, University of London, England, in 1978 and 1982 respectively, all in electrical engineering.

He served as an Assistant Professor in the Department of Electrical and Computer Engineering, University of Michigan-Dearborn. Since July 1986 he has been with the Department of Electrical Engineering, Colorado State University, where he is now an Associate Professor. His areas of interest are digital signal/image processing, multidimensional system theory and analysis, adaptive filtering and system identification, target detection and tracking, and neural networks. Dr. Azimi-Sadjadi is a co-author of the book, *Digital Filtering in One and Two Dimensions*, Plenum Press, 1989.

Dr. Azimi-Sadjadi is the recipient of 1994 ASEE-Navy Summer Faculty Fellowship Award, 1990 Battelle Summer Faculty Fellowship Award, and the 1984 DOW Chemical Outstanding Young Faculty Award for the American Society for Engineering Education.