Automatic Target Recognition for Hyperspectral Imagery using High-Order Statistics

HSUAN REN, Member, IEEE
National Central University
Taiwan

QIAN DU, Senior Member, IEEE
Mississippi State University

JING WANG

CHEIN-I CHANG, Senior Member, IEEE
University of Maryland Baltimore County

JAMES O. JENSEN

JANET L. JENSEN
Edgewood Chemical and Biological Center
U.S. Army

Due to recent advances in hyperspectral imaging sensors many subtle unknown signal sources that cannot be resolved by multispectral sensors can be now uncovered for target detection, discrimination, and identification. Because the information about such sources is generally not available, automatic target recognition (ATR) presents a great challenge to hyperspectral image analysts. Many approaches developed for ATR are based on second-order statistics in the past years. This paper investigates ATR techniques using high order statistics. For ATR in hyperspectral imagery, most interesting targets usually occur with low probabilities and small population and they generally cannot be described by second-order statistics. Under such circumstances, using high-order statistics to perform target detection have been shown by experiments in this paper to be more effective than using second order statistics. In order to further address a challenging issue in determining the number of signal sources needed to be detected, a recently developed concept of virtual dimensionality (VD) is used to estimate this number. The experiments demonstrate that using high-order statistics-based techniques in conjunction with the VD to perform ATR are indeed very effective.

Manuscript received February 5, 2005; revised December 4, 2005; released for publication March 14, 2006.

IEEE Log No. T-AES/42/4/878281.

Refereeing of this contribution was handled by L. M. Kaplan.

This work was performed while H. Ren held a National Research Council Research Associateship Award at Edgewood Chemical and Biological Center, U.S. Army.

Authors' addresses: H. Ren, Center for Space and Remote Sensing Research, National Central University, 300 Jhong-Da Rd., Jhong-Li, Tao-Yuan, Taiwan 320; Q. Du, Dept. of Electrical and Computer Engineering, Mississippi State University, State College, MI 39762; C.-I Chang and J. Wang, Remote Sensing Signal and Image Processing Lab, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, E-mail: (cchang@umbc.edu); J. O. Jensen and J. L. Jensen, Edgewood Chemical and Biological Center, U.S. Army, APG, MD 21010.

I. INTRODUCTION

Recently, a new generation of remote sensing instruments with very high spectral resolution, called imaging spectrometers or hyperspectral imaging sensors have been developed to uncover subtle material substances that generally cannot be resolved by multispectral sensors. Therefore, hyperspectral sensors provide a new dimension in applications that cannot be addressed by classical spatial domain-based techniques. One of such applications is automatic target recognition (ATR) in finding targets that are relatively small and generally occur in low probabilities with no prior knowledge. More importantly, when these targets appear, their population is usually not too large. Of particular interest is that the size of a target may be smaller than the pixel size. In this case, the target is embedded in a single pixel and cannot be identified by its spatial presence. Unfortunately, many such targets exist, such as special species in agriculture and ecology, toxic/metal waste in environmental monitoring, rare minerals in geology, drug trafficking in law enforcement, and small combat vehicles in battlefield to name just a few.

Two types of targets are generally of interest for ATR and have been studied extensively in the literature. One is anomaly with signature spectrally distinct from its surroundings. Another is endmember that is defined as an idealized and pure signature for a class [1]. Depending upon which type of targets is of interest, different approaches have been investigated and developed. As for anomaly detection, a popular and well-known algorithm, generally referred to as RX algorithm was suggested by Reed and Yu [2] who formulated a binary composite hypothesis testing problem to derive a constant false alarm probability (CFAR) detector which turned out to be the Mahalanobis distance [3]. Others include Chang et al.'s RX algorithm-based anomaly detection and classification in [4], Ashton's adaptive Bayesian classifier in [5], Schweizer and Moura's Gauss-Markov random field (GMRF) in [6], independent component analysis (ICA)-based linear spectral random mixture analysis [7] and projection pursuit [8]. As for the endmember extraction, two most popular and widely used algorithms are pixel purity index (PPI) available in the Research Systems ENVI software [9] and N-finder (N-FINDR) developed by Winter et al. [10]. Additionally, many other approaches have been also developed, which include convex cone analysis (CCA) [11], unsupervised fully constrained least squares (UFCLS) [12], iterative error algorithm (IEA) [13], automated morphological endmember extraction (AMEE) [14], and projection pursuit [15]. A recently developed approach, called automatic target detection and classification algorithm (ATDCA)

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **OCT 2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE **Automatic Target Recognition for Hyperspectral Imagery using High-Order Statistics**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Maryland Baltimore County,Remote Sensing Signal and Image Processing Lab,1000 Hilltop Circle,Baltimore,MD,21250**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**14**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

[16–17], which made use of a sequence of orthogonal subspace projections (OSPs) [18] has been shown to be effective in automatic target extraction.

As mentioned above, since the targets of interest generally occupy a few pixels, they may not be able to constitute reliable second-order statistics. In this case, their presence may be more effectively captured through high-order statistics. In order to take advantage of high-order statistics, projection pursuit and ICA were introduced as alternatives for anomaly detection [7–8, 17] and endmember extraction [15, 19]. The idea was based on an assumption that if the image background can be characterized by second-order statistics, anomalies can be then viewed as outliers as opposed to the image background due to the fact that their sizes are relatively small and spectral features are very different compared with their surroundings. As a result, anomaly detection can be performed more effectively by searching for deviations from the background distribution. Similarly, the occurrence of pure signatures (endmembers) in a real image scene is usually very low and rare. Therefore, their existence can be more effectively characterized by high-order statistics. In doing so, the simplest higher order statistics are skewness and kurtosis which are the normalized third and fourth central moments and can be used to measure the asymmetry and flatness of distribution respectively. If the image background is assumed to be Gaussian distributed, the skewness and kurtosis can be used to measure the difference of a distribution from Gaussianity. Using this property as a criterion, skewness and kurtosis seem to be appropriate measures in detection of anomalies or small targets. This paper investigates ATR using high-order statistics and also explores approaches to finding optimal projection directions so that the projected data have the maximal high-order statistics. Since high-order statistics-based methods generally require calculation of a sequence of projections which can be very computationally expensive and cumbersome, this paper further develops efficient algorithms for finding these projections with greatly reduced computational complexity.

Another challenging issue for ATR is to determine the number of targets needed to be generated without prior knowledge. This is equivalent to determining how many projections are required to be generated for projection pursuit-based approaches [7–8], how many endmembers assumed to be present in the data for endmember extraction [9–15], or how many targets are extracted for target detection and classification [16–17]. A general approach is either to use trial and error or to compute the accumulated eigenvalues to account for a certain percentage of total energy. Unfortunately, neither method is effective as shown in [17] and [20]. In order to address this issue, a recently developed concept of virtual dimensionality (VD) is used to determine the number of projections required for ATR. The idea of the VD has its root in Neyman-Pearson detection theory [21] which uses the false alarm probability as a criterion to estimate the number of spectrally distinct signal sources present in the data. Using the VD as an estimate, the number of projections can be reasonably determined. Experiments conducted in the work presented here also support the use of the VD in determination of number of projections for ATR.

In order to demonstrate advantages of our proposed high-order statistics-based methods over second-order statistics-based methods, the RX algorithm [2] and the ATDCA [16–17] are selected for comparative analysis in two applications, endmember extraction and target detection and classification where two real hyperspectral image data sets are used for experiments. The reason for such selection is because these two techniques are second-order statistics-based methods and each of them uses a different type of criterion. For example, the RX algorithm is derived from a Gaussian kernel or Mahalanobis distance [2] and the ATDCA is developed on the signal-to-noise ratio (SNR)-based OSP [16–17].

The paper is organized as follows. Section II presents iterative methods to find skewness and kurtosis-based projections for ATR with extension to any high-order statistics. Section III conducts real hyperspectral image experiments to demonstrate the performance of the higher-order statistics-based ATR algorithms. Finally, Section IV draws some conclusions.

II. ATR USING HIGH-ORDER STATISTICS

Assume that there are N data points $\{\mathbf{x}_i\}_{i=1}^N$, each of which has dimensionality L and $\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_N]$ is an $L \times N$ data matrix formed by $\{\mathbf{x}_i\}_{i=1}^N$. Let $\mathbf{w}$ be an L-dimensional column vector and assumed to be a desired projection vector. Then $\mathbf{z} = \mathbf{w}^T\mathbf{X} = (z_1, z_2, \ldots, z_N)^T$ is a $1 \times N$ row vector which represents the projection of data $\{\mathbf{x}_i\}_{i=1}^N$ mapped along the direction of $\mathbf{w}$, where ^{T} denotes the transpose of a vector or matrix. Now, assume that $F(\cdot)$ is a function to be explored and defined on the projection space $\mathbf{z} = \mathbf{w}^T\mathbf{X}$. The selection of the function F depends upon various applications. For example, in order to detect small targets in a large unknown background, skewness and kurtosis are generally used as criteria to measure asymmetry and flatness of a distribution, respectively. In this case, $F(.)$ can be defined by skewness (κ_3) with

$$F(z_i) = \kappa_3(z_i) = \frac{E[(z_i - \mu)^3]}{\sigma^3} = \frac{E[(\mathbf{w}^T\mathbf{x}_i - \mu)^3]}{\sigma^3} \quad \text{for each} \quad i = 1, 2, \ldots, N \qquad (1)$$

which is the normalized third central moment, kurtosis (κ_4) with

$$F(z_i) = \kappa_4(z_i) = \frac{E[(z_i - \mu)^4]}{\sigma^4} = \frac{E[(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i - \mu)^4]}{\sigma^4} \qquad \text{for each} \quad i = 1,2,\ldots,N \tag{2}$$

which is the normalized fourth central moment and any kth order statistic κ_5 with

$$F(z_i) = \kappa_k(z_i) = \frac{E[(z_i - \mu)^k]}{\sigma^k} = \frac{E[(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i - \mu)^k]}{\sigma^k} \qquad \text{for each} \quad i = 1,2,\ldots,N. \tag{3}$$

The μ and σ in (1)–(3) are the mean and standard deviation of random variable z_i, respectively. Since small targets can be characterized by those pixels that cause maximum magnitude of asymmetry and ripples of a distribution, finding a projection vector $\mathbf{w}$ that maximizes (1)–(3) is equivalent to finding a direction which these pixels are most likely aligned with. By projection of all data samples $\{\mathbf{x}_i\}_{i=1}^N$ on the projection vector $\mathbf{w}$, the desired small targets can be detected by those pixels that yield the largest projection along the direction of $\mathbf{w}$.

If we assume that most of image background can be described by second-order statistics and the statistical behaviors of targets of interest go beyond second-order statistics, a logical preprocessing for detecting such targets will remove the image background prior to target detection. In doing so, we first remove the sample mean and de-correlate the data matrix $\mathbf{X}$ by the sphering method described as follows.

A. Sphering

The idea of sphering is to centralize the mean of the data samples $\{\mathbf{x}_i\}_{i=1}^N$ at the origin while normalizing the data variances to one. In this case, two sets of data samples can be categorized. One set is made up of all data samples lying on the surface of a sphere centered at the origin with unit radius. The set of these data samples represents uninteresting data samples which may include most image background pixels. The second set of data samples contains all data samples which are not on the sphere, i.e., either inside or outside the sphere. Only these data samples are of major interest and can be further explored by orders of statistics higher than variance. So, working only on this set of data samples may exclude most of image background samples.

In order to perform sphering, we first remove the sample mean of data set by $\hat{\mathbf{X}} = \mathbf{X} - \boldsymbol{\mu}\cdot\mathbf{1}^{\mathrm{T}} = [\mathbf{x}_1 - \boldsymbol{\mu}, \mathbf{x}_2 - \boldsymbol{\mu}, \ldots, \mathbf{x}_N - \boldsymbol{\mu}]$, where $\boldsymbol{\mu} = (1/N)\sum_{i=1}^N \mathbf{x}_i$ is the sample mean vector and $\mathbf{1} = [\underbrace{11\cdots 1}_{N}]^{\mathrm{T}}$ is column vector with all ones in the components. Next step we de-correlate the zero-mean data sample matrix $\hat{\mathbf{X}}$.

Assume that $\{\lambda_l\}_{l=1}^L$ are the eigenvalues of the sample covariance matrix $\boldsymbol{\Sigma} = (1/N)\hat{\mathbf{X}}\hat{\mathbf{X}}^{\mathrm{T}}$ formed by $\hat{\mathbf{X}}$ and $\{\mathbf{v}_l\}_{l=1}^L$ are their corresponding eigenvectors. The covariance matrix can be decomposed into

$$\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{V} = \boldsymbol{\Lambda} \tag{4}$$

where $\mathbf{V} = [\mathbf{v}_1\mathbf{v}_2\cdots\mathbf{v}_L]$ is a matrix made up of the eigenvectors $\{\mathbf{v}_l\}_{l=1}^L$ and $\boldsymbol{\Lambda} = \mathrm{diag}\{\lambda_l\}_{l=1}^L$ is a diagonal matrix with L eigenvalues $\{\lambda_l\}_{l=1}^L$ as the diagonal elements. Let $\boldsymbol{\Lambda}^{-1/2} = \mathrm{diag}\left\{1/\sqrt{\lambda_l}\right\}_{l=1}^L$. Multiplying both sides of (4) by $\boldsymbol{\Lambda}^{-1/2}$ results in

$$\boldsymbol{\Lambda}^{-1/2}\mathbf{V}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{V}\boldsymbol{\Lambda}^{-1/2} = \mathbf{I}. \tag{5}$$

From (5), we obtain the desired sphering matrix $\mathbf{A}$, given by

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}^{-1/2} \tag{6}$$

so that $\mathbf{A}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{A} = \mathbf{I}$. The data set resulting from applying the sphering matrix $\mathbf{A}$ to the original data set, $\{\hat{\mathbf{x}}_i\}_{i=1}^N$ is denoted by $\{\mathbf{y}_i\}_{i=1}^N$ and the process of using (4)–(6) is called sphering which is also known as a whitening process of $\mathbf{X}$. In this case, the data matrix $\mathbf{Y}$ has zero mean and an identity matrix as its covariance matrix.

If we replace the original data samples $\{\mathbf{x}_i\}_{i=1}^N$ with sphered data samples $\{\mathbf{y}_i\}_{i=1}^N$ in (3) then $\mathbf{z} = \mathbf{w}^{\mathrm{T}}\mathbf{Y} = \mathbf{w}^{\mathrm{T}}[\mathbf{y}_1\mathbf{y}_2\cdots\mathbf{y}_N] = (\mathbf{w}^{\mathrm{T}}\mathbf{y}_1, \mathbf{w}^{\mathrm{T}}\mathbf{y}_2, \ldots, \mathbf{w}^{\mathrm{T}}\mathbf{y}_N)^{\mathrm{T}} = (z_1, z_2, \ldots, z_N)^{\mathrm{T}}$. Equation (3) can be reduced to

$$\kappa_k(z_i) = E[(\mathbf{w}^{\mathrm{T}}\mathbf{y}_i)^k] \qquad \text{for each} \quad i = 1,2,\ldots,N \tag{7}$$

with subscript k indicating k-order statistic for $k \geq 3$.

B. Algorithm for Finding Projection Images for ATR

After the data are sphered, the next task is to search a projection vector that is optimal in some sense. If the skewness is used as a criterion, the projection vector should be the one that points to the direction where the projected data has the most asymmetric histogram. If the kurtosis is used as a criterion, the projected data will yield the most heavy-tailed histogram. In this section, an iterative method is proposed to search for the optimal projection vector based on these criteria.

To find the projector which yields the maximum skewness, we impose a constrained problem as follows

$$\max_{\mathbf{w}}\left\{\frac{1}{N}\sum_{i=1}^N z_i^3\right\} = \max_{\mathbf{w}}\left\{\frac{1}{N}\sum_{i=1}^N \mathbf{w}^{\mathrm{T}}\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\mathbf{w}^{\mathrm{T}}\mathbf{y}_i\right\} \qquad \text{subject to} \quad \mathbf{w}^{\mathrm{T}}\mathbf{w} = 1 \tag{8}$$

where z_i is the projection resulting from the sphered data sample $\mathbf{y}_i$ via the projection vector $\mathbf{w}$. The constraint $\mathbf{w}^{\mathrm{T}}\mathbf{w} = 1$ is used for normalization such that the skewness of the resulting data after projection will not be affected by the magnitude of $\mathbf{w}$. Using the Lagrange multiplier method, an objective function is obtained by

$$J(\mathbf{w}) = E[\mathbf{w}^{\mathrm{T}}\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\mathbf{w}^{\mathrm{T}}\mathbf{y}_i] - \lambda(\mathbf{w}^{\mathrm{T}}\mathbf{w} - 1). \tag{9}$$

Differentiating (9) with respective $\mathbf{w}$ results in

$$\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} = 3E[\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\mathbf{y}_i^{\mathrm{T}}]\mathbf{w} - 2\lambda\mathbf{w} = 0. \tag{10}$$

Setting $\lambda' = (2/3)\lambda$ yields

$$(E[\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\mathbf{y}_i^{\mathrm{T}}] - \lambda'\mathbf{I})\mathbf{w} = 0. \tag{11}$$

Solving (11) is equivalent to finding the eigenvalue λ' of the matrix $E[\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\mathbf{y}_i^{\mathrm{T}}]$ and its corresponding eigenvector $\mathbf{w}^*$.

When using kurtosis as the searching criterion, the constrained problem becomes

$$\max_{\mathbf{w}}\left\{\frac{1}{N}\sum_{i=1}^{N} z_i^4\right\} = \max_{\mathbf{w}}\left\{\frac{1}{N}\sum_{i=1}^{N}\mathbf{w}^{\mathrm{T}}\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\mathbf{w}^{\mathrm{T}}\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\right\} \quad \text{subject to} \quad \mathbf{w}^{\mathrm{T}}\mathbf{w} = 1. \tag{12}$$

In analogy with (11) we can also obtain

$$(E[\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\mathbf{w}^{\mathrm{T}}\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}] - \lambda'\mathbf{I})\mathbf{w} = 0 \tag{13}$$

which once again is to solve the eigenvalue λ' and its associated eigenvector $\mathbf{w}*$ of the matrix $E[\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\mathbf{w}\mathbf{w}^{\mathrm{T}}\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}]$. The obtained $\mathbf{w}^*$ is a desired the projector that yields the maximum kurtosis.

In order to extend the above treatment to any kth order central moment we solve the following eigen-problem

$$(E[\mathbf{y}_i(\mathbf{y}_i^{\mathrm{T}}\mathbf{w})^{k-2}\mathbf{y}_i^{\mathrm{T}}] - \lambda'\mathbf{I})\mathbf{w} = 0 \tag{14}$$

which is the eigenvector of $E[\mathbf{y}_i(\mathbf{y}_i^{\mathrm{T}}\mathbf{w})^{k-2}\mathbf{y}_i^{\mathrm{T}}]$. Using the property of eigen-decomposition, (14) can be reduced to

$$\mathbf{w}^{\mathrm{T}}(E[\mathbf{y}_i(\mathbf{y}_i^{\mathrm{T}}\mathbf{w})^{k-2}\mathbf{y}_i^{\mathrm{T}}])\mathbf{w} = \lambda' \tag{15}$$

because of $\mathbf{w}^{\mathrm{T}}\mathbf{w} = 1$. Equation (15) can be further simplified to

$$E[\mathbf{w}^{\mathrm{T}}\mathbf{y}_i(\mathbf{y}_i^{\mathrm{T}}\mathbf{w})^{k-2}\mathbf{y}_i^{\mathrm{T}}\mathbf{w}] = E[(\mathbf{y}_i^{\mathrm{T}}\mathbf{w})^k] = E[z_i^k] = \lambda' \tag{16}$$

which turns out to be the kth central moment of $\mathbf{z} = (\mathbf{w}^*)^{\mathrm{T}}\mathbf{Y}$.

Since a single projection vector $\mathbf{w}^*$ that solves (14) can only detect one type of anomaly. In order to detect more types of anomalies present in an image scene a sequence of projections must be performed. In doing so, when a projector vector $\mathbf{w}^*$ is found, the de-correlated data $\mathbf{Y}$ is then mapped into the linear subspace $\langle\mathbf{w}^*\rangle^{\perp}$ orthogonal to $\langle\mathbf{w}^*\rangle$ that is the space linear spanned by $\mathbf{w}^*$. The next projection vector $\mathbf{w}^*$ is then found by solving (14) in the space $\langle\mathbf{w}^*\rangle^{\perp}$. The same procedure is continued on until a stop criterion is satisfied such as the predetermined number of projections required to be generated. A detailed implementation of finding a sequence of projection vectors can be described as follows.

Projection Vector Generation Algorithm

1) Sphere the original data set $\mathbf{X}$. The resulting data set is denoted by $\mathbf{Y}$.

2) Find the first projection vector $\mathbf{w}_1^*$ by solving (13) to find the optimal projection vector based on maximizing the kth normalized central moment.

3) Using the found $\mathbf{w}_1^*$, generate the first projection image $\mathbf{Z}^1 = (\mathbf{w}_1^*)^{\mathrm{T}}\mathbf{Y} = \{\mathbf{z}_i^1 \mid \mathbf{z}_i^1 = (\mathbf{w}_1^*)^{\mathrm{T}}\mathbf{y}_i\}$ which can be used to detect the first type of anomaly.

4) Apply the OSP specified by $\mathbf{P}_{\mathbf{w}_1}^{\perp} = \mathbf{I} - \mathbf{w}_1(\mathbf{w}_1^{\mathrm{T}}\mathbf{w}_1)^{-1}\mathbf{w}_1^{\mathrm{T}}$ to the data set $\mathbf{Y}$ to produce the first OSP-projected data set denoted by $\mathbf{Y}^1$, $\mathbf{Y}^1 = \mathbf{P}_{\mathbf{w}_1}^{\perp}\mathbf{Y}$.

5) Use the data set $\mathbf{Y}^1$ and find the second projection vector $\mathbf{w}_2^*$ by solving (14).

6) Apply $\mathbf{P}_{\mathbf{w}_2}^{\perp} = \mathbf{I} - \mathbf{w}_2(\mathbf{w}_2^{\mathrm{T}}\mathbf{w}_2)^{-1}\mathbf{w}_2^{\mathrm{T}}$ to the data set $\mathbf{Y}^1$ to produce the second OSP-projected data set denoted by $\mathbf{Y}^2$, $\mathbf{Y}^2 = \mathbf{P}_{\mathbf{w}_2}^{\perp}\mathbf{Y}^1$ which can be used to produce the third projection vector $\mathbf{w}_3^*$ by solving (14). Or equivalently, define a matrix projection matrix $\mathbf{W}^2 = [\mathbf{w}_1\mathbf{w}_2]$ and apply $\mathbf{P}_{\mathbf{W}^2}^{\perp} = \mathbf{I} - \mathbf{W}^2((\mathbf{W}^2)^{\mathrm{T}}\mathbf{W}^2)^{-1}(\mathbf{W}^2)^{\mathrm{T}}$ to the original sphered data set $\mathbf{Y}$ to obtain $\mathbf{Y}^2 = \mathbf{P}_{\mathbf{W}^2}^{\perp}\mathbf{Y}$.

7) Repeat the procedure of steps 5 and 6 to produce $\mathbf{w}_3^*, \ldots, \mathbf{w}_k^*$ until a predetermined number of projection vectors is reached.

It should be noted that the implementation of step 2 in the above algorithm is not trivial. In order to solve (14) for the optimal projection vector $\mathbf{w}_1^*$ the following iterative procedure is proposed to execute the step 2.

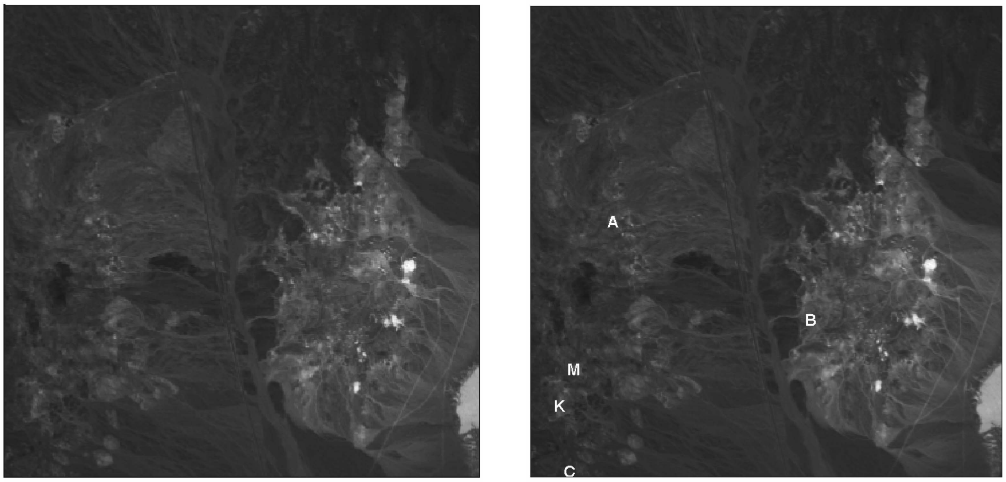
a) Initialize a random projector $\mathbf{w}_1^{(0)}$ and set $n = 0$

b) Calculate the matrix $E[\mathbf{y}_i(\mathbf{y}_i^{\mathrm{T}}\mathbf{w}_1^{(n)})^{k-2}\mathbf{y}_i^{\mathrm{T}}]$ and find an eigenvector $\mathbf{v}_1^{(n)}$ corresponding to the largest magnitude of eigenvalues of the matrix $E[\mathbf{y}_i(\mathbf{y}_i^{\mathrm{T}}\mathbf{w}_1^{(n)})^{k-2}\mathbf{y}_i^{\mathrm{T}}]$.

c) If the Euclidean distance $\|\mathbf{w}_1^{(n)} - \mathbf{w}_1^{(n+1)}\| > \varepsilon$ or $\|\mathbf{w}_1^{(n)} + \mathbf{w}_1^{(n+1)}\| > \varepsilon$, then let $\mathbf{w}_1^{(n+1)} = \mathbf{v}_1^{(n)}$ and $n \leftarrow n + 1$, go to step (b). Otherwise, $\mathbf{w}_1^{(n)}$ is the desired projector $\mathbf{w}_1^*$. Let $\mathbf{w}_1^* = \mathbf{w}_1^{(n)}$ and return to step 3 in the projection vector generation algorithm.

III. EXPERIMENTS

Two sets of real hyperspctral image data, airborne visible infrared imaging spectrometer (AVIRIS) Cuprite data and hyperspectral digital image collection experiment (HYDICE) data were used for experiments. The Cuprite data was used to demonstrate advantage of high-order statistics-based methods over second-order statistics-based methods in endmember extraction, while the HYDICE



(a) (b)

Fig. 1. (a) Spectral band number 50 (827 nm) of Cuprite AVIRIS image scene. (b) Spatial positions of five pure pixels corresponding to minerals: alunite (A), buddingtonite (B), calcite (C), kaolinite (K), muscovite (M).

data was used to show superior performance of high-order statistics-based methods to second-order statistics-based methods in target detection and classification.

A. Endmember Extraction

Endmember extraction is one of fundamental tasks in hyperspectral image analysis. It finds and identifies the purest signatures in image data for various applications such as image endmembers used for linear spectral unmixing, training samples for unsupervised image classification, etc. The proposed high-order statistics-based endmember extraction was evaluated for performance analysis where the widely used endmember extraction algorithm N-FINDR was used as benchmark comparison. The reason to choose N-FINDR over the PPI was due to the fact that the N-FINDR is available in the literature compared with the PPI which is only available in the ENVI software.

The first image data was collected over the Cuprite mining site, Nevada, in 1997, and is shown in Fig. 1(a). It is a 224 band AVIRIS image scene with a size of 350×350 pixels, is well understood mineralogically, and has reliable ground truth available at website [23] where the five pure pixels representing the five minerals, alunite (A), buddingtonite (B), calcite (C), kaolinite (K), and muscovite (M), referred to as endmembers are white-circled and labeled by A, B, C, K, and M in Fig. 1(b). This fact has made this scene a standard test site for endmember extraction.

It should be noted that in the Cuprite image data, bands 1–3, 105–115, and 150–170 have been removed prior to the analysis due to water absorption and low SNR in those bands. As a result, a total of 189 bands were used for experiments.

The VD estimated for this image scene with different values of false alarm probability P_{F} is given in Table I where the Harsanyi-Farrand-Chang (HFC) method in [17, 20] was used for estimation.

TABLE I
VD Estimated by HFC Method with Various False Alarm Probabilities

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	34	30	24	22	20

TABLE II
Four Endmembers Extracted by N-FINDR with MNF-DR

A	B	C	K	M	
A′	0.0235	0.1665	0.2143	0.1012	0.1542
C′	0.2235	0.1002	0.0511	0.2276	0.1222
K′	0.0812	0.1434	0.1771	0.0418	0.1010
M′	0.1675	0.0933	0.0971	0.1483	0.0381

In order to ensure that all mineral signatures of interest were included in high-order statistics-generated components, the false alram rate P_{F} chosen for the VD was set to 10^{-4} which resulted in VD = 22. However, it should be noted that this was an emiprical choice. Since each component image represents a specific class of targets it can be used to serve as abundance map for this particular class. Endmember signature is identified as the purest signature in the scene, in this sense, the brightest pixel in the image component (IC) can be extracted as an endmember. Using the spectral angle mapper (SAM) in [1] as a spectral similarity measure, the pixels whose signatures were identified to be closet to the ground truth mineral endmembers were extracted as candidate endmembers. Table II tabulated their SAM values between the four N-FINDR-extracted signatures, denoted by A′, C′, K′, M′ and the five ground truth endmembers, A, B, C, K, M, with no signature found to correspond to

TABLE IIIa
Five Signatures Extracted by Skewness

	A	B	C	K	M	Order of IC
A′	0.0172	0.1645	0.2115	0.0961	0.1476	15
B′	0.1591	0.0745	0.0793	0.1767	0.0969	13
C′	0.2192	0.1069	0.0362	0.2196	0.1174	11
K′	0.1029	0.1619	0.2025	0.0300	0.1183	2
M′	0.1650	0.0844	0.1162	0.1508	0.0808	17

TABLE IIIb
Five Signatures Extracted by Kurtosis

	A	B	C	K	M	Order of IC
A′	0.0172	0.1645	0.2115	0.0962	0.1476	10
B′	0.1591	0.0749	0.0793	0.1768	0.0969	11
C′	0.2192	0.1069	0.0362	0.2196	0.1174	18
K′	0.0888	0.1834	0.2283	0.0341	0.1453	7
M′	0.1458	0.0781	0.1141	0.1347	0.0706	21

TABLE IIIc
Five Signatures Extracted by 5th Moment

	A	B	C	K	M	Order of IC
A′	0	0.1576	0.2038	0.0961	0.1421	15
B′	0.1412	0.0589	0.0847	0.1521	0.0777	21
C′	0.2192	0.1069	0.0362	0.2196	0.1174	19
K′	0.0888	0.1835	0.2284	0.0341	0.1453	7
M′	0.1067	0.1293	0.1558	0.0688	0.0711	14

TABLE IIId
Five Signatures Extracted by FastICA

	A	B	C	K	M	Order of IC
A′	0	0.1576	0.2038	0.0961	0.1421	17
B′	0.1385	0.0623	0.079	0.1584	0.0826	13
C′	0.2192	0.1069	0.0362	0.2196	0.1174	21
K′	0.1029	0.1619	0.2026	0.03	0.1183	6
M′	0.1650	0.0844	0.1162	0.1508	0.0809	18

the mineral endmember C. It should be noted that when the N-FINDR was implemented, it required dimensionality reduction which was performed by the maximum noise fraction (MNF) [24] and the number of dimensions to be retained was set to p, in this case, $p = 22$.

In order to implement the proposed high-order statistics-based ATR for endmember extraction, the three criteria of skewness, kurtosis, 5th moment as well as the FastICA were evaluated for performance analysis.

There are three different ways to generate initial projection vectors to be used for the high-order statistics-based algorithms and FastICA:

1) random-generated initial projection vector: a unit vector with components randomly generated,
2) unity-based initial projection vector: a unit vector with all ones in its components,
3) eigenvector-based initial projection vectors: the eigenvectors corresponding to the p largest eigenvalues of the data sample covariance matrix.

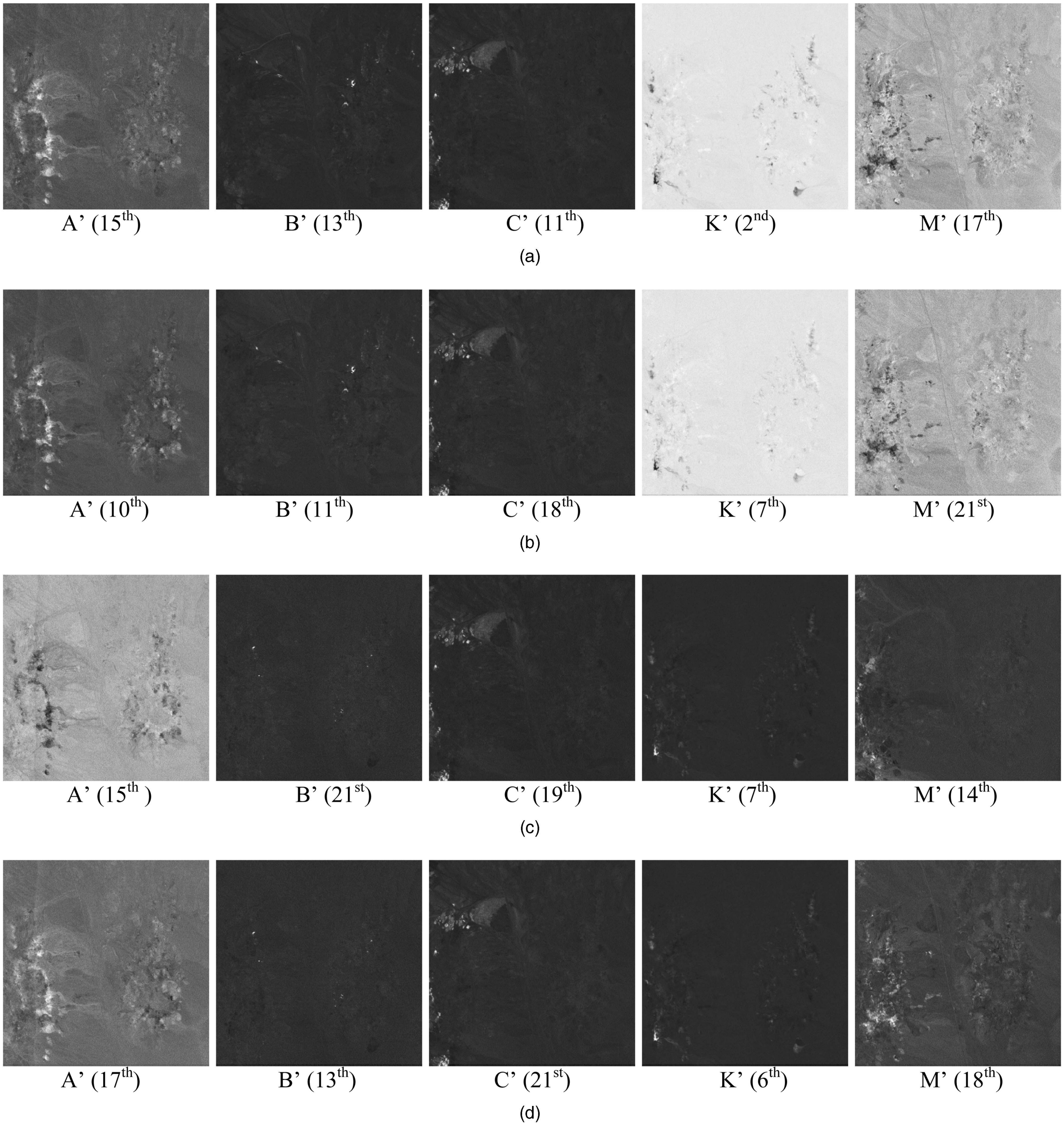
Nevertheless, according to our experiments, the results derived from these three different initial projection vectors were very similar even though their produced component images might appear in different orders. Therefore, only the SAM values between the ground truth endmembers and signatures extracted by algorithms using eigenvectors as initial projection vectors are tabulated in Tables IIIa–IIId where the ICs in the last columns indicate the orders of the ICs generated by algorithms. Figs. 2(a)–(d) also show the gray scale component images produced by the three algorithms and FastICA corresponding to the components identified in the last columns of Tables IIIa–IIId.

Four criteria of high-order statistics were used to produce image components for ATR, which are to maximize skewness, kurtosis, 5th moment, and statistical independency. The results are summarized in Table IV with the five minerals A, B, C, K, M listed in the 1st row. Since the ICs were generated sequentially, the orders of the ICs generated in different scenarios are usually different. In Table IV the numbers under each of the minerals indicate that the orders of the components that specified the particular minerals were generated. In addition, the computing time in seconds for each criterion is provided in the last column in Table IV.

According to Table IV, all the minerals were extracted within 30 components except two cases that were the skewness using the unity vector as the initial projection vector and the FastICA using a random vector as initial projection vector, each of which missed one type of mineral. As for CPU time, the FastICA required the least time. However, this may not be conclusive since our algorithms to run high-order statistics were not optimized in terms of source codes. Comparing the results in Tables III–IV produced by the high-order ATR to the results in Table II produced by the N-FINDR, it is very clear that the high-order statistics-based methods usually performed better than the N-FINDR in endmember extraction where the former extracted the pixels corresponding to all the five mineral endmembers compared with only four endmembers extracted by the N-FINDR.

B. Target Detection

The second mage data set is a HYDICE image scene shown in Fig. 3(a) that was used for experiments of target detection. It has a size of 64 × 64 pixel vectors with 15 panels in the scene and the ground truth map in Fig. 3(b) [17].



A’ (15th) B’ (13th) C’ (11th) K’ (2nd) M’ (17th)

(a)

A’ (10th) B’ (11th) C’ (18th) K’ (7th) M’ (21st)

(b)

A’ (15th) B’ (21st) C’ (19th) K’ (7th) M’ (14th)

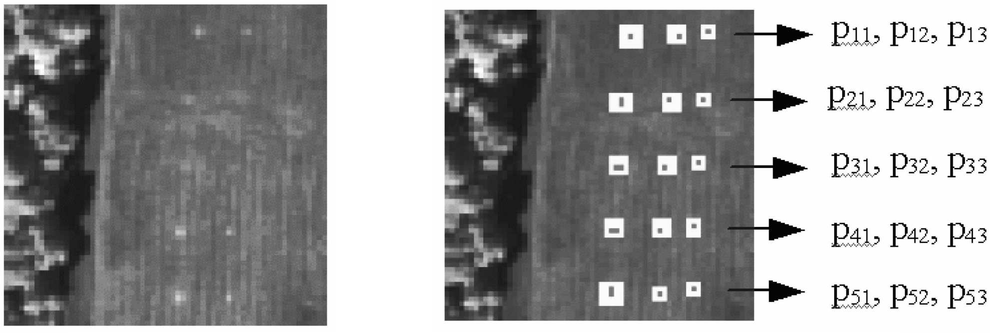
(c)

A’ (17th) B’ (13th) C’ (21st) K’ (6th) M’ (18th)

(d)

Fig. 2. (a) Image components produced by skewness. (b) Image components produced by kurtosis. (c) Image components produced by 5th moment. (d) Image components produced by FastICA.

It was acquired by 210 spectral bands with a spectral coverage from 0.4 μm to 2.5 μm. Low signal/high noise bands (bands 1–3 and bands 202–210) and water vapor absorption bands (bands 101–112 and bands 137–153) were removed. So, a total of 169 bands were used. The spatial resolution is 1.56 m and spectral resolution is 10 nm. Within the scene in Fig. 3(a) there is a large grass field background, and a forest on the left edge. Each element in this matrix is a square panel and denoted by p_{ij} with rows indexed by i and columns indexed by $j = 1,2,3$. For each row $i = 1,2,\ldots,5$, there are three panels p_{i1}, p_{i2}, p_{i3}, painted by the same material but with three different sizes. For each column $j = 1,2,3$, the 5 panels p_{1j}, p_{2j}, p_{3j}, p_{4j}, p_{5j} have the same size but with five different materials. So, panels in five different rows were painted by five different materials. The sizes of the panels in the first, second, and third columns are 3 m × 3 m, 2 m × 2 m and 1 m × 1 m respectively. Since the size of the panels in the third column is 1 m × 1 m, they cannot be seen visually from Fig. 3(a) due to the fact that its size is less



(a) (b)

Fig. 3. 15-panel HYDICE image. (a) 15-panel image scene. (b) Ground truth map of 15 panels.

TABLE IV
Order of Components that Specified Minerals were Generated

High-Order Statistics	Initial Projection Vectors	A	B	C	K	M	CPU Time (seconds)
skewness	random vectors	13	19	12	3	18	1548
	unity vector	17	Not found	11	3	21	1686
	eigenvectors	15	13	11	2	17	1516
Kurtosis	random vectors	19	15	20	4	17	4373
	unity vector	20	4	7	3	13	5736
	eigenvectors	10	11	18	7	21	3928
5th moment	random vectors	14	20	15	3	13	1853
	unity vector	15	17	21	3	14	1677
	eigenvectors	15	21	19	7	14	1981
FastICA	random vectors	14	13	12	7	Not found	293
	unity vector	21	14	20	4	13	340
	eigenvectors	17	13	21	6	18	306

TABLE V
VD Estimated by HFC Method with Various False Alarm Probabilities

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	14	11	9	9	7

than the 1.56 m pixel resolution. Fig. 3(b) shows the precise spatial locations of these 15 panels where red pixels (R pixels) are the panel center pixels and the pixels in yellow (Y pixels) are panel pixels mixed with the background. The 1.56 m spatial resolution of the image scene suggests that most of the 15 panels are one pixel in size except p_{21}, p_{31}, p_{41}, p_{51} which are two-pixel panels. This image scene provides an excellent example for ATR since it contains real subpixel targets and mixed pixels that cannot be simulated by synthetic images. This is because it is difficult to simulate a synthetic image with appropriate sample spectral correlation to reflect real data.

The VD estimated for this HYDICE image scene with different values of false alarm probability P_{F} is given in Table V.

The VD was chosen to be 9 with the false alarm probability set to $P_{\mathrm{F}} = 10^{-4}$ and p was set to 18 to ensure that all the 15 panels of interest included components generated by high-order statistics. Additionally, the same three ways used to initial projection vectors for Table II were also used to initialize algorithms of four high-order statistics, skewness, kurtosis, 5th moment, and statistical independency.

The results are summarized in Table VI with panels in five rows listed in the 1st row where the numbers in Table VI indicate that the orders of the components that specified the panels of particular rows were generated. In analogy with Table IV, the least CPU time also came from the FastICA.

In addition to Table VI the ICs produced by the skewness, kurtosis, 5th moment, and statistical independency using eigenvectors as initial projection vectors are also provided in Figs. 4(a)–(d) to demonstrate the effectiveness of high-order statistics for ATR in target detection and classification where the number in parenthesis under each IC was the order in which that particular image component was generated. It should be noted that only the ICs corresponding to the five panels are listed.

These figures demonstrate that the ICs produced by high-order statitsics could be used for target detection and classification for ATR. On the other hand, since our proposed high-order statistics-based

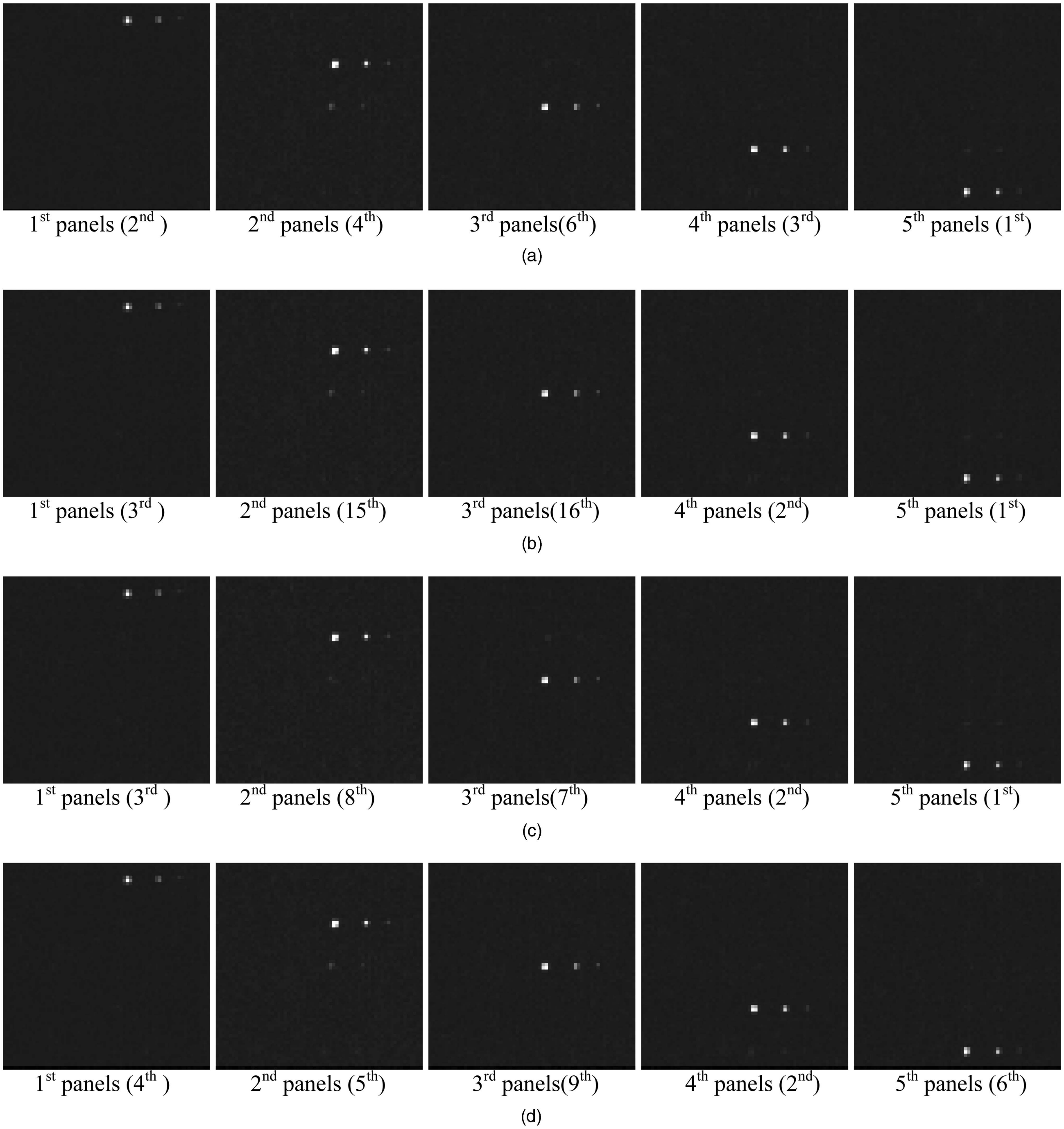
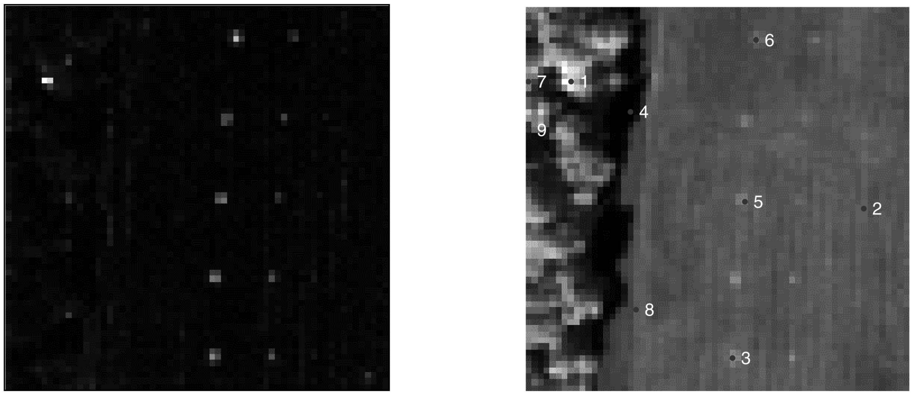


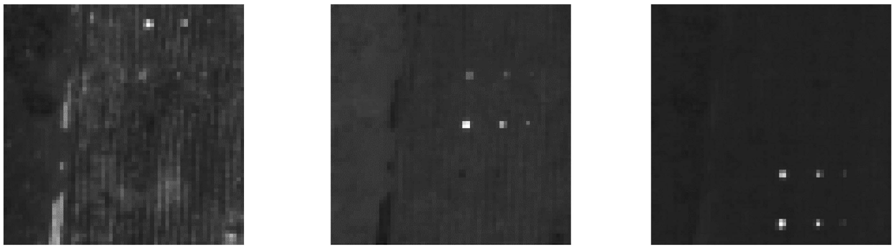
Fig. 4. (a) Image components produced by skewness that extracted panels in five rows for target detection. (b) Image components produced by kurtosis that extracted panels in five rows for target detection. (c) Image components produced by 5th moment that extracted panels in five rows for target detection. (d) Image components produced by FastICA that extracted panels in five rows for target detection.

methods are completely unsupervised and no prior knowldege is required, the second-order statistics-based RX algorithm for anomlay detection and the OSP-based ATDCA seem to serve as perfect candidates for comparison due to the fact that they both are completely unsupervised target detectors and require no prior knolwedge. Fig. 5(a)–(b) shows the results produced by the RX algorithm and the ATDCA, respectively, where the ATDCA-generated 9 target pixels $\{\mathbf{t}_i^{\mathrm{ATDCA}}\}_{i=1}^{9}$ were shown in Fig. 5(b) with numbers indicating the orders of target pixels being generated by the ATDCA. According to Fig. 5(b), only three ATDCA-generated target pixels, $\mathbf{t}_3^{\mathrm{ATDCA}}$, $\mathbf{t}_5^{\mathrm{ATDCA}}$, $\mathbf{t}_6^{\mathrm{ATDCA}}$ were identified to represent three R pixels in row 5, 3, and 1, respectively. The ATDCA then used these three panel pixels to perform panel classification. The results are shown in Fig. 6.



(a) (b)

Fig. 5. Detection results by (a) RX algorithm, and (b) ATDCA.



panels in row 1($\mathbf{t}_6^{\mathrm{ATDCA}}$) panels in row 3 ($\mathbf{t}_5^{\mathrm{ATDCA}}$) panels in row 5($\mathbf{t}_3^{\mathrm{ATDCA}}$)

Fig. 6. Classification results by ATDCA using these panel pixels.

TABLE VI
Order of Components that Specified Minerals were Generated

High-Order Statistics	Initial Projection Vectors	Panels in Row 1	Panels in Row 2	Panels in Row 3	Panels in Row 4	Panels in Row 5	CPU Time (seconds)
skewness	random vectors	1	10	2	7	3	67.59
	unity vector	3	13	1	7	4	76.59
	eigenvectors	2	4	6	3	1	55.89
Kurtosis	random vectors	3	7	2	8	4	110.03
	unity vector	5	8	4	6	2	135.47
	eigenvectors	3	15	16	2	1	155.73
5th moment	random vectors	8	2	5	4	3	103.53
	unity vector	9	not found	3	4	2	103.35
	eigenvector	3	8	7	2	1	95.17
FastICA	random vectors	1	6	9	3	2	6.96
	unity	9	7	5	4	1	9.07
	eigenvectors	4	5	9	2	6	6.22

Since the panels in rows 2 and 3 were made by the same materials with different paints, the ATDCA which used the $\mathbf{t}_5^{\mathrm{ATDCA}}$ for classification also classified panels in row 2 as shown in Fig. 6. Similarly, it was true for panels in rows 4 and 5. If we compare Fig. 6 with Fig. 5(a), the ATDCA performed significantly better than the RX algorithm. However, if we further compare Fig. 6 with Fig. 4(a)–(d), it is obvious that the ATDCA could not compete against the high-order statistics-based methods because the latter could detect panels in the five rows correctly in five individual and separate components.

In order to make further comparison, the receiver operating characteristics (ROCs) analysis [21] was used to evaluate the detection performance of the four high-order statistics based methods, skewness, kurtosis, FastICA against the RX algorithm and ATDCA based on the results in Figs. 4(a)–(d), 5(a), and 6. Fig. 7(a)–(d) plot the ROC curves of each of high-order statistics based methods relative that

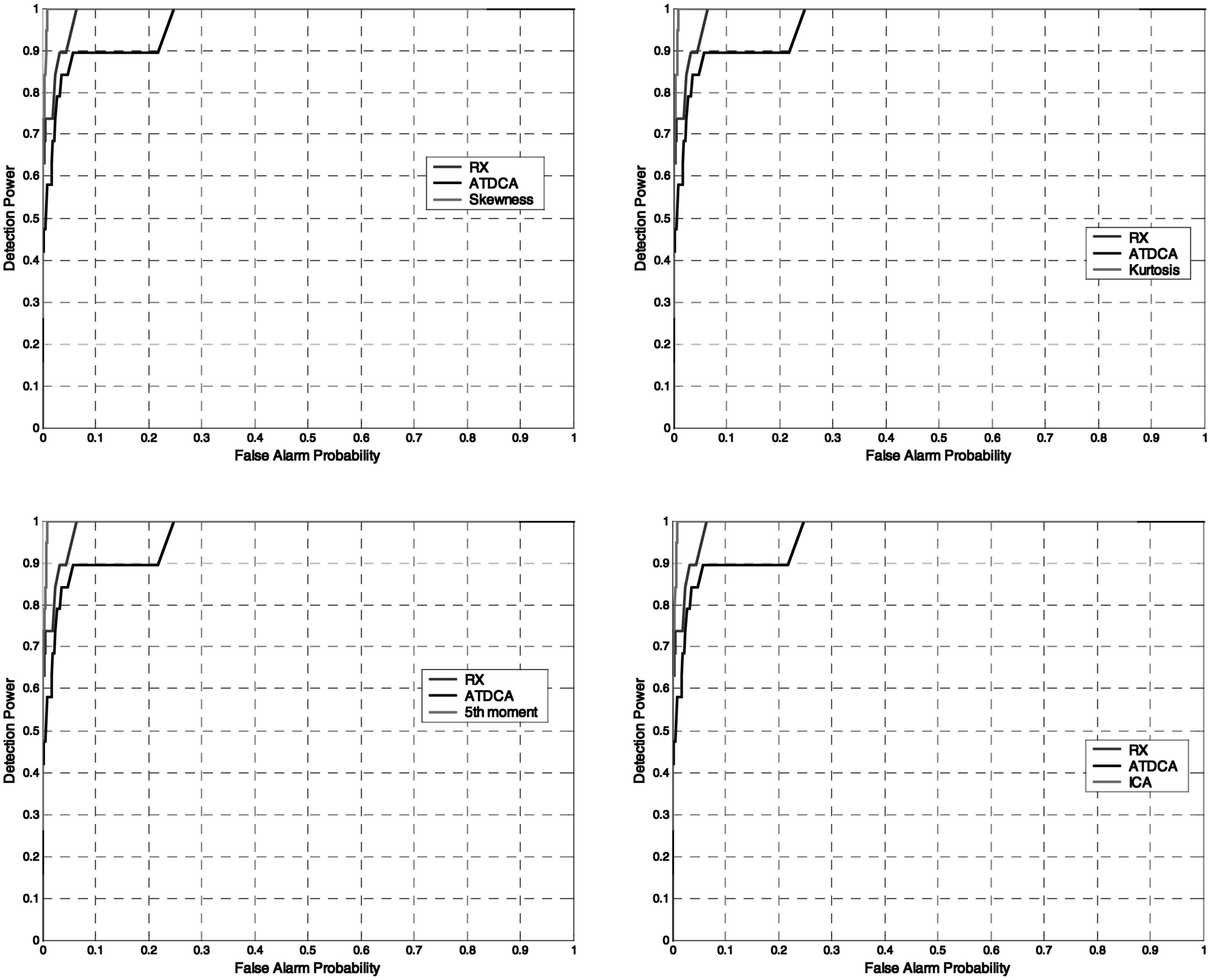


Fig. 7. ROC curves of high-order ATR, ICA and RX, ATDCA.

produced by the RX algorithm and ATDCA where all the four high-order statistics-based methods performed significantly better than the RX algorithm and ATDCA.

Four comments on Fig. 7 are noteworthy.

1) Since the detection images produced in Figs. 4(a)–(d), 5(a), and 6 were gray scale, a threshold must be applied to segment panels from the image background for panel detection. In this case, the ROC curves in Fig. 7 were produced by varying a threshold from the lowest gray scale to highest gray scale for panel detection.

2) Second, the same threshold was used and applied to the image in Fig. 5(a) and Fig. 6 and all the ICs in Fig. 4(a)–(d).

3) Third, due to close performance among all the four high-order statistics-based methods, it is very difficult to discriminate their ROC curves one from another. So, the ROC curves were plotted by each of the four high-order statistics based method against the RX algorithm and ATDCA for clarity.

4) Finally, it should be noted that high-order statistics-based methods take advantage of their generated components to perform target detection and classification, while the RX algorithm cannot discriminate the targets it detects. As a result, the RX algorithm can be only used for target detection. On the other hand, the ATDCA can be used for target classification as it did in Fig. 6. But its second-order statistics has limited its ability in detecting only three panel pixels, not five panel pixels for classification.

IV. CONCLUSION

This paper investigates high-order statistics-based methods for ATR in hyperspectral imagery where third-order normalized central moment (skewness), fourth-order normalized central moment (kurtosis), fifth normalized central moment, and infinite normalized central moment (statistical independency) are studied for comparative analysis. Algorithms for implementing criteria of high-order statistics are also developed and considered to be new except the FastICA developed in [22] which is used to implement the criterion of statistical independency. In order to demonstrate the utility of our proposed high-order statistics-based methods using skewness, kurtosis, 5th moment, and statistical independency as criteria, two applications in endmember extraction and target detection and classification are explored by experiments. The experimental results show that high-order statistics-based methods have clear advantages over the second-order statistics-based methods such as PCA.

REFERENCES

[1] Schwengerdt, R. A.
Remote Sensing: Models and Methods for Image Processing (2nd ed.).
New York: Academic Press, 1997, 447.

[2] Reed, I. S., and Yu, X.
Adaptive multuple-band CFAR detection of an optical pattern with unknown spectral distribution.
IEEE Transactions on Acoustic, Speech and Signal Processing, **38**, 10 (1990), 1760–1770.

[3] Fukunanga, K.
Statistical Pattern Recognition (2nd ed.).
New York: Academic Press, 1991.

[4] Chang, C.-I, and Chiang, S-S.
Anomaly detection and classification for hyperspectral imagery.
IEEE Transactions on Geoscience and Remote Sensing, **40**, 6 (2002), 1314–1325.

[5] Ashton, E.
Detection of subpixel anaomalies in multispectral infrared imagery using an adaptive Bayesian classifier.
IEEE Transactions on Geoscience and Remote Sensing, **36**, 2 (1998), 506–517.

[6] Schweizer, S. M., and Moura, J. M. F.
Efficient detection in hyperspectral imagery.
IEEE Transactions on Image Processing, **10**, 4 (2001), 584–597.

[7] Chang, C.-I, Chiang, S-S., Smith, J. A., and Ginsberg, I. W.
Linear spectral random mixture analysis for hyperspectral imagery.
IEEE Transactions on Geoscience and Remote Sensing, **40**, 2 (2002), 375–392.

[8] Chiang, S-S., Chang, C.-I, and Ginsberg, I. M.
Unsupervised target detection in hyperspectral images using projection pursuit.
IEEE Transactions on Geoscience and Remote Sensing, **39**, 7 (2001), 1380–1391.

[9] Boardman, J. W., Kruse, F. A., and Green, R. O.
Mapping target signatures via partial unmixing of AVIRIS data.
In *Summaries of JPL Airborne Earth Science Workshop*, Pasadena, CA, 1995.

[10] Winter, M. E.
N-finder: An algorithm for fast autonomous spectral endmember determination in hyperspectral data.
In *Proceedings of SPIE*, Image Spectrometry V, vol. 3753, 266–277.

[11] Ifarraguerri, A., and Chang, C.-I
Hyperspectral image segmentation with convex cones.
IEEE Transactions on Geoscience and Remote Sensing, **37**, 2 (1999), 756–770.

[12] Heinz, D., and Chang, C.-I
Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery.
IEEE Transactions on Geoscience and Remote Sensing, **39**, 3 (2001), 529–545.

[13] Neville, R. A., Staenz, K., Szeredi, T., Lefebvre, J., and Hauff, P.
Automatic endmember extraction from hyperspectral data for mineral exploration.
In *4th International Airborne Remote Sensing Conference and Exhibition/21st Canadian Symposium on Remote Sensing*, Ottawa, Ontario, Canada, 1999.

[14] Plaza, A., Martinez, P., Perez, R., and Plaza, T.
Spatial/spectral endmember extraction by multidimensional morphological operations.
IEEE Transactions on Geoscience and Remote Sensing, **40**, 9 (2002), 2025–2041.

[15] Ifarraguerri, A., and Chang, C.-I
Unsupervised hyperspectral image analysis with projection pursuit.
IEEE Transactions on Geoscience and Remote Sensing, **38**, 6 (2000), 2529–2538.

[16] Ren, H., and Chang, C.-I
Automatic spectral target recognition in hyperspectral imagery.
IEEE Transactions on Aerospace and Electronic Systems, **39**, 4 (2003), 1232–1249.

[17] Chang, C.-I
Hyperspectral Imaging, Spectral Techniques for Spectral Detection and Classification.
New York: Kluwer Acdemic/Plenum Publishers, 2003.

[18] Harsanyi, J. C., and Chang, C.-I
Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach.
IEEE Transactions on Geoscience and Remote Sensing, **32**, 4 (1994), 779–785.

[19] Wang, J., and Chang, C.-I
Independent component analysis approach to endmember extraction.
To be submitted.

[20] Chang, C.-I, and Du, Q.
Estimation of number of spectrally distinct signal sources in hyperspectral imagery.
IEEE Transactions on Geoscience and Remote Sensing, **42**, 3 (2004), 608–619.

[21] Poor, H. V.
An Introduction to Detection and Estimation Theory.
New York: Springer-Verlag, 1991.

[22] Hyvarinen, A., Karhunen, J., and Oja, E.
Independent Component Analysis.
New York: Wiley, 2001.

[23] http://speclab.cr.usgs.gov/cuprite.html

[24] Green, A. A., Berman, M., Switzer, P., and Craig, M. D.
A transformation for ordering multispectral data in terms of image quality with implications for noise removal.
IEEE Transactions on Geosciences and Remote Sensing, **42** (1998), 65–74.

Hsuan Ren (S'98—M'00) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1994 and the M.S. and Ph.D. degrees from University of Maryland Baltimore County, Baltimore, MD, in 1998 and 2000, respectively, all in electrical engineering.

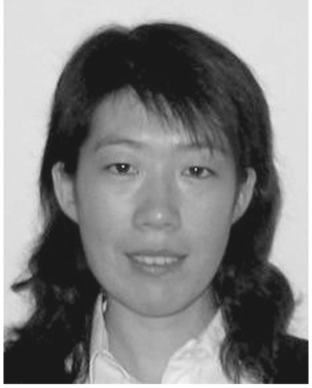
He received a National Research Council (NRC) Associateship award and was supported by the U.S. Army Edgewood Chemical Biological Center from 2000 to 2003. He is currently an assistant professor at the National Central University, Jhong-Li, Taiwan, R.O.C. His research interests include data compression, signal and image processing, and pattern recognition.

Dr. Ren has two patents pending on hyperspectral target detection and image classification. He is a member of SPIE and the honor society of Phi Kappa Phi.

Qian Du (S'98—M'00—SM'06) received her Ph.D. degree in electrical engineering from University of Maryland Baltimore County, Baltimore, MD, in 2000.

She was an assistant professor in the Department of Electrical Engineering and Computer Science at Texas A&M University–Kingsville from 2000–2004. Since Fall 2004, she has been with the Department of Electrical and Computer Engineering at Mississippi State University as an assistant professor. She has been working on remote sensing image analysis for many years with expertise in hyperspectral imaging.

She is a member of SPIE, ASPRS, and ASEE.

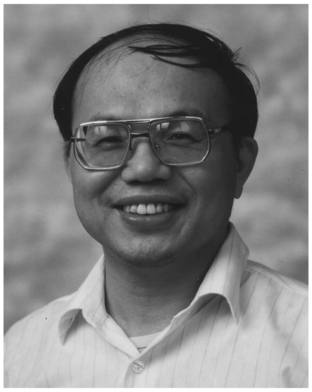


Jing Wang received the B.S. degree in electrical engineering and the M.S. degree in computer engineering from the Beijing University of Post and Telecommunications in 1998 and 2001. She also received the M.S degree in electrical engineering from the University of Maryland Baltimore County (UMBC), Baltimore, MD, in 2005, where she is currently a Ph.D. candidate.

She is currently a research assistant in the Remote Sensing Signal and Image Processing Laboratory, UMBC. Her research interests include signal and image processing, pattern recognition, and data compression.

Chein-I Chang (S'81—M'87—SM'92) received his B.S. degree from Soochow University, Taipei, Taiwan, M.S. degree from the Institute of Mathematics at National Tsing Hua University, Hsinchu, Taiwan and M.A. degree from the State University of New York at Stony Brook, all in mathematics. He also received his M.S., M.S.E.E. degrees from the University of Illinois at Urbana–Champaign and Ph.D. degree in electrical engineering from the University of Maryland, College Park.

Dr. Chang has been with the University of Maryland, Baltimore County (UMBC) since 1987 and is currently professor in the Department of Computer Science and Electrical Engineering. He was a visiting research specialist in the Institute of Information Engineering at the National Cheng Kung University, Tainan, Taiwan, from 1994–1995. He received an NRC (National Research Council) senior research associateship award from 2002–2003 sponsored by the U.S. Army Soldier and Biological Chemical Command, Edgewood Chemical and Biological Center, Aberdeen Proving Ground, Maryland. Additionally, Dr. Chang was a distinguished lecturer chair at the National Chung Hsing University sponsored by the Department of Education and National Science Council in Taiwan from 2005–2006 and is currently holding a chair professorship with National Chung Hsing University. He has three patents and several pending on hyperspectral image processing. He is on the editorial board of the *Journal of High Speed Networks* and was the guest editor of a special issue of the same journal on telemedicine and applications and will co-guest edit two special issues on Broadband Multimedia Sensor Networks in Healthcare Applications, *Journal of High Speed Networks* and High Performance Computing of Hyperspectral Imaging, *International Journal of High Performance Computing Applications*, December 2007. Dr. Chang has authored a book, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification* published by Kluwer Academic Publishers in 2003 and edited two books, *Recent Advances in Hyperspectral Signal and Image Processing*, Trivandrum, Kerala: Research Signpost, Trasworld Research Network, India, 2006 and *Hyperspectral Data Exploitation: Theory and Applications*, John Wiley & Sons, 2007. He is currently working on his, second book, *Hyperspectral Imaging: Signal Processing Algorithm Design and Analysis*, John Wiley & Sons, 2007 and is co-editing with A. Plaza on a book on *High Performance Computing in Remote Sensing*, CRC Press, 2007.



Dr. Chang is an associate editor in the area of hyperspectral signal processing for *IEEE Transactions on Geoscience and Remote Sensing* and a Fellow of SPIE and a member of Phi Kappa Phi and Eta Kappa Nu.

His research interests include multispectral/hyperspectral image processing, automatic target recognition, medical imaging, information theory and coding, signal detection and estimation and neural networks.