Unsupervised Target Detection in Hyperspectral Images Using Projection Pursuit

Shao-Shan Chiang, *Student Member, IEEE,* Chein-I Chang, *Senior Member, IEEE,* and Irving W. Ginsberg

***Abstract*—In this paper, we present a projection pursuit (PP) approach to target detection. Unlike most of developed target detection algorithms that require statistical models such as linear mixture, the proposed PP is to project a high dimensional data set into a low dimensional data space while retaining desired information of interest. It utilizes a projection index to explore projections of interestingness. For target detection applications in hyperspectral imagery, an interesting structure of an image scene is the one caused by man-made targets in a large unknown background. Such targets can be viewed as anomalies in an image scene due to the fact that their size is relatively small compared to their background surroundings. As a result, detecting small targets in an unknown image scene is reduced to finding the outliers of background distributions. It is known that "skewness," is defined by normalized third moment of the sample distribution, measures the asymmetry of the distribution and "kurtosis" is defined by normalized fourth moment of the sample distribution measures the flatness of the distribution. They both are susceptible to outliers. So, using skewness and kurtosis as a base to design a projection index may be effective for target detection. In order to find an optimal projection index, an evolutionary algorithm is also developed to avoid trapping local optima. The hyperspectral image experiments show that the proposed PP method provides an effective means for target detection.**

***Index Terms*—Evolutional algorithm, hyperspectral imagery, kurtosis, projection index, projection pursuit (PP), skewness, target detection.**

I. Introduction

DETECTING small targets in hyperspectral images can be very difficult, particularly in an unknown environment. This is primarily due to ground sampling distance generally larger than the size of targets of interest. In this case, target detection must be carried out at the subpixel scale. With no availability of prior knowledge, it is virtually impossible to validate designed algorithms. In this paper, we present a projection pursuit (PP) approach to unsupervised target detection. The PP was first developed by Friedman and Tukey [1], [2] to be used as a technique for exploratory analysis of multivariate data and has been studied extensively since then [3]–[5]. Unlike most of developed target detection algorithms that require statistical models such as linear mixture, PP is a linear mapping that searches for interesting low-dimensional projections from a high-dimensional data space via a projection index (PI). The PI is a measure used to explore projections of interestingness. In particular, it can be designed to characterize nonlinear structures in projected distributions. For example, if the desired direction of a PI is one pointing to data variance, PP is reduced to a well-known technique, principal components analysis (PCA). Using PP for hyperspectral image classification has been studied previously in [6], [7], where Jimenez and Landgrebe [6] designed a PI based on Bhattacharyya's distance to reduce the dimensionality of feature space and Ifarragaerri and Chang [7] used the information divergence (relative entropy) as a PI looking for interesting projections that deviates from Gaussian distributions. However, their techniques were developed primarily for image analysis, not specifically for target detection. In military applications, man-made target detection in hyperspectral imagery is important since their size is generally small and sometimes even smaller than the ground sampling distance (GSD). Such target detection must rely on subpixel spectral detection, not pixel-based spatial detection. From this point of view, an interesting structure of an image scene is the one caused by man-made targets in a large unknown background. So, the small targets can be viewed as anomalies in an image scene due to the fact that their size is relatively small in the spatial sense compared to the image background. As a result, detecting such small targets in an unknown image scene can be reduced to finding the outliers or deviations from the background distribution. It may also occur that a small region or set of background pixels can be also detected. Since there is no prior knowledge, the detected small targets may include man-made targets or natural objects such as tree, grass, rocks, and interferers. It is known that *skewness* defined by normalized third moment of the sample distribution measures the asymmetry of the distribution and *kurtosis* defined by normalized fourth moment of the sample distribution measures the flatness of the distribution. They both are susceptible to outliers. So, using skewness and kurtosis as a base to design a PI may be an effective means for target detection.

Once a PI is determined, finding optimal solutions for the desired PI is crucial. Unfortunately, there are generally no analytic solutions and they must be solved by numerical algorithms. Two major principles, "hill-climbing" and "random move" have widely used to design optimization algorithms. For a relatively smooth PI where the first derivatives exist, "hill-climbing"-based gradient descent methods are usually preferred. However, they may be trapped in local optimal solutions that are close to the initial starting points. In this case, a good guess of initial conditions is a key to success of these

Manuscript received September 20, 2000; revised December 22, 2000. This work was supported by Bechtel Nevada Corporation under Contract DE-AC08-96NV11718 through the Department of Energy, Washington, DC.

S.-S. Chiang and C.-I Chang are with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD 21250 USA (e-mail: cchang@umbc.edu).

I. W. Ginsberg is with the Remote Sensing Laboratory, U.S. Department of Energy, Las Vegas, NV 89191 USA.

Publisher Item Identifier S 0196-2892(01)05493-6.

methods. On the other hand, a random move-based method such as simulated annealing (SA) can escape from local optima to some degree, but its success is also determined by the initial starting points. In this paper, we consider a similar random move-based approach to SA, called evolutionary algorithm (EA) [8]–[10] which is more likely to lead solutions to global optima.

Prior to implementation of EA, it generally requires to preserve the translation invariance property of PI. This can be done by centering the original data matrix using a linear translation followed by an eigen-decomposition method such as singular value decomposition (SVD) to whiten the centered data matrix. The latter whitening process is referred to as sphering in the PP literature. There are four stage processes implemented in EA, which are population selection, crossover, mutation, and termination. Of particular interest are crossover and mutation processes. While the crossover process generates a descendant population by combining pairs of individuals randomly chosen from a parent population, the mutation process performs a local optimization that makes random changes in individual points to search for local optimum points. The EA proposed in this paper is little different from the EA referred in the literature in the sense that it includes a zero-detection thresholding to achieve target detection. Since we are interested in small man-made targets, the outliers resulting from skewness and kurtosis are attributed to these targets. In order to extract these targets, zero detection is proposed to threshold each projection image so that the different desired targets can be segmented by thresholding from a sequence of projection images. The algorithm implemented by EA in conjunction with PP and a zero-detection thresholding method is referred to as PPEA in this paper. In order to evaluate the performance of PPEA, two data sets of hyperspectral images, airborne visible/infrared imaging spectrometer (AVIRIS) and a standardized hyperspectral digital imagery collection experiment (HYDICE) data set are used. In particular, to make a quantitative study and analysis possible, a set of custom-designed criteria for tallying the number of detected target pixels in the HYDICE data introduced in [11] is also used for experiments.

This paper is organized as follows. Section II considers four projection indices to be used for projection pursuit. Section III presents a preprocessing to sphere the data matrix prior to optimization. Section IV describes the procedure of how to implement EA. Section V introduces a zero-detection thresholding method for target detection. Section VI conducts a series of experiments. Section VII concludes some comments.

II. Projection Pursuit

The term PP was first coined by Friedman and Tukey and was used as a technique for exploratory analysis of multivariate data. The idea is to project a high dimensional data set into a low dimensional data space while retaining the information of interest. It designs a PI to explore projections of interestingness.

Following the approach in [4], we assume that there are N data points $\{\mathbf{x}_n\}_{n=1}^N$, each with dimensionality K, $\mathbf{X}=[\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n]$ is a $K\times N$ data matrix, and $\mathbf{a}$ is an K-dimensional column vector, which serves as a desired projection. Then $\mathbf{a}^{\mathrm{T}}\mathbf{X}$ represents an N-dimensional row vector that is the orthogonal projections of all sample data points mapped onto the direction $\mathbf{a}$, where T is the matrix transpose. Now if we let $H(\cdot)$ be a function measuring the degree of the interestingness of the projection $\mathbf{a}^{\mathrm{T}}\mathbf{X}$ for a fixed data matrix $\mathbf{X}$, a projection index (PI) is a real-valued function of $\mathbf{a}$, $I(\mathbf{a}):R^K\rightarrow R$ defined by

$$I(\mathbf{a})=H(\mathbf{a}^{\mathrm{T}}\mathbf{X}). \tag{1}$$

The PI can be easily extended to multiple directions $\{\mathbf{a}_j\}_{j=1}^J$. In this case, $\mathbf{A}=[\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_J]$ is a $K\times J$ projection direction matrix, and the corresponding projection index is also a real-valued function $I(\mathbf{A}):R^{K\times J}\rightarrow R$ given by

$$I(\mathbf{A})=H(\mathbf{A}^{\mathrm{T}}\mathbf{X}). \tag{2}$$

The choice of the $H(\cdot)$ in (1) and (2) is application-dependent. Its purpose is to reveal interesting structures within data sets such as clustering. As mentioned in the introduction, in military applications, man-made target detection is of major interest. In this case, using skewness denoted by κ_3 and kurtosis denoted by κ_4 as a PI are appropriate criteria for target detection. Four projection indices defined as follows will be evaluated in this paper for target detection:

$$I_{\mathrm{skewness}}(\mathbf{a})=H_1(\mathbf{a}^{\mathrm{T}}\mathbf{X})=\kappa_3^2 \tag{3}$$

$$I_{\mathrm{kurtosis}}(\mathbf{a})=H_2(\mathbf{a}^{\mathrm{T}}\mathbf{X})=\kappa_4^2 \tag{4}$$

$$I_{\mathrm{mix}}(\mathbf{a})=H_3(\mathbf{a}^{\mathrm{T}}\mathbf{X})=\kappa_3^2+\frac{\kappa_4^2}{12} \tag{5}$$

where $I_{\mathrm{mix}}(\mathbf{a})$ is a linear mixture of κ_3 and κ_4

$$I_{\mathrm{product}}(\mathbf{a})=H_4(\mathbf{a}^{\mathrm{T}}\mathrm{X})=\kappa_3^2\cdot\kappa_4^2=(\kappa_3\kappa_4)^2 \tag{6}$$

where $I_{\mathrm{product}}(\mathbf{a})$ is a product of κ_3^2 and κ_4^2.

III. Sphering

Assume that $\{\mathbf{x}_n\}_{n=1}^N$ is a set of data samples, and $\mathbf{X}=[\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n]$ is the corresponding data matrix of which the i-th column vector is the i-th data sample $\mathbf{x}_i$. In order to satisfy the translation invariance property of PI, centering the original data matrix is required [4]. This can be done by the transformation $\mathbf{Y}=\mathbf{X}(\mathbf{I}-(1/N)\mathbf{1}\mathbf{1}^{\mathrm{T}})$, where $\mathbf{I}$ is an $N\times N$ identity matrix, and $\mathbf{1}=\underbrace{(1,1,\ldots,1)}_{N}^{\mathrm{T}}$ is an N-dimensional column vector with ones in all components. Assume that $\{\lambda_i\}_{i=1}^p$ are the eigenvalues of the sample covariance matrix Σ, and the transformed data matrix $\{\mathbf{v}_i\}_{i=1}^p$ are their corresponding eigenvectors. Since Σ is nonnegative definite, all eigenvalues are nonnegative and there exists a unitary matrix $\mathbf{Q}$ such that Σ can be decomposed into

$$\mathbf{Q}^{\mathrm{T}}\Sigma\mathbf{Q}=\Lambda \tag{7}$$

where $\mathbf{Q}=[\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_p]$ is a matrix made up of the eigenvectors $\{\mathbf{v}_i\}_{i=1}^p$, and $\Lambda=\mathrm{diag}\{\lambda_i\}_{i=1}^p$ is a diagonal matrix with $\{\lambda_i\}_{i=1}^p$ in the diagonal line.

If we let $\Lambda^{-1/2}=\mathrm{diag}\{\sqrt{\lambda_i}\}_{i=1}^p$, multiplying both sides of (7) by $\Lambda^{-1/2}$ results in

$$\Lambda^{-1/2}\mathbf{Q}^{\mathrm{T}}\Sigma\mathbf{Q}\Lambda^{-1/2}=\mathbf{I}. \tag{8}$$

From (7), we obtain the desired sphering matrix $\mathbf{A}$ given by

$$\mathbf{A} = \mathbf{Q}\Lambda^{-1/2} \tag{9}$$

so that $\mathbf{A}^{\mathrm{T}}\Sigma\mathbf{A} = \mathbf{I}$. The process of transforming the centered data matrix $\mathbf{Y}$ via the sphering matrix $\mathbf{A}$ outlined by (6)–(9) is called "sphering," which is also known as a whitening process to uncorrelate $\mathbf{Y}$.

IV. Evolutionary Algorithm (EA)

After the centered data matrix is sphered, an optimization algorithm will be used to find optimal solutions for a selected PI. For simplicity, the notation for sphered centered data matrix will be still represented by $\mathbf{X}$. The optimization algorithm proposed in this section is the EA developed in [8]–[10] and can be implemented by four stage processes, population selection, crossover, mutation, and termination, each of which is briefly described as follows. For details we refer to [8]–[10].

1) Initial population.

   Unlike gradient methods and simulated annealing, an EA requires an initial population to start with to perform a multidirectional search for optimal points simultaneously during each iteration. A population is made up of possible projection vectors, each of which can be viewed as individuals. To initialize EA, an initial population can be created randomly or with prior knowledge if there is any. A simple way to form an initial population is to include all basis vectors of the space to projected.

2) Population Selection Process.

   Once a population is formed, each individual in the population will be evaluated by a fitness function specified by the desired projection index $f(\mathbf{x})$. The individual that has better fitness will have more chances to generate new population in the next iteration. The procedure is described as follows.

   a) Calculate $f(\mathbf{v}_i)$ for each individual $\mathbf{v}_i$.
   b) Calculate the probability of selecting $\mathbf{v}_i$ by $p_i = f(\mathbf{v}_i)/\sum_j f(\mathbf{v}_j)$.
   c) Calculate the cumulative probability q_i for $f(\mathbf{v}_i)$ by $q_i = \sum_{j=1}^{i} p_j$.
   d) Generate a random number r in the range $[0, 1]$.
   e) Select the i-th individual $\mathbf{v}_i$ such that $q_{i-1} < r \leq q_i$.
   f) Repeat steps (d) and (e) until a desired population is formed.

   It should be noted that the above selection process is stochastic and some individuals may be selected more than once. This is in accordance with the natural principle of "survival of fitness," which says that the best individuals get more descendants while the average stays even and the worst dies off.

3) Crossover Process.

   The crossover process used in EA is called arithmetical crossover, which can be defined as follows. Assume there are two individuals, denoted by two vectors $\mathbf{v}$ and $\mathbf{w}$. Then the two vectors formed by $\mathbf{v}' = a\mathbf{v} + (1-a)\mathbf{w}$ and $\mathbf{w}' = (1-a)\mathbf{v} + a\mathbf{w}$ are called crossover offspring of $\mathbf{v}$ and $\mathbf{w}$ with $a \in [0, 1]$. Let p_c be the preselected crossover probability that will be used to determine the number of individuals needed to produce crossover offspring. This number is $N_c = \lfloor p_c \cdot N_k \rfloor$ with $p_c \cdot N_k - 1 < \lfloor p_c \cdot N_k \rfloor \leq p_c \cdot N_k$, where N_k is the size of population at the k-th iteration. The crossover is performed as follows.

   a) For each individual $\mathbf{v}$ in the current population, generate a random number r in the range $[0, 1]$. If $r < p_c$, the $\mathbf{v}$ will be selected as a candidate for crossover. The selection procedure will be continued to form a crossover vector group until the expectation of number of selected crossover vectors is equal to N_c.
   b) For each pair of two individuals $\mathbf{v}$ and $\mathbf{w}$ in the selected crossover group, form their two crossover offspring $\mathbf{v}' = a\mathbf{v} + (1-a)\mathbf{w}$ and $\mathbf{w}' = (1-a)\mathbf{v} + a\mathbf{w}$, where a is a random number generated from the uniform distribution in the range $[0, 1]$, since this requires N_c to be even. So if N_c is odd, we need select one more individual randomly to make it even.

   Through the crossover process, we get a more effective means to find more competitive projections other than the ones already in the population.

4) Mutation Process.

   Like the mutation defined in genetic algorithm, the mutation in EA is also performed on a bit-by-bit basis. For each individual $\mathbf{v} = [v_1\ v_2\ \ldots\ v_n]$, each component, i.e., v_i, is expressed in a binary expansion with l precisions by $v_i = v_{i1}v_{i2}\ldots v_{il}$. Then $\mathbf{v}$ can be represented by a binary string formed by concatenating the l-precision binary expansions of all components $\mathbf{v} = v_1v_2\ldots v_n = v_{11}v_{12}\ldots v_{1l}\ldots v_{n1}v_{n2}\ldots v_{nl}$, which is called a chromosome. Now the mutation process will be performed on the chromosome bit by bit. Let p_m a predetermined probability of mutation which determines the number of bits needed to be mutated, denoted N_m which is the expectation of the number of mutated bits is equal to $\lfloor p_m \cdot N_c \cdot n \cdot l \rfloor$.

   a) Encode each individual in the current population into a chromosome
   b) For each bit in the chromosome, generate a random number r from the uniform distribution in the range $[0, 1]$. If $r < p_m$, mutate the bit by flipping the bit between 0 and 1.
   c) Decode each chromosome into a new individual.

   The goal of the mutation is to explore the neighborhood of each individual to find an optimal projection.

5) Termination Process.

   There are several ways to terminate EA. 1) The algorithm reaches an optimal solution. 2) No significant changes are made. 3) The number of iterations exceeds a predetermined limit.

V. Thresholding of Projection Images Using Zero-Detection

Like PCA, PPEA produces a sequence of projection images that show the information in the decreasing order of magnitudes of projection values produced by a PI. It should be noted that the magnitude referred to here is the absolute projection values

where the projection values of a PI can be positive or negative. These projection images can be used for target detection. Assume that $\mathbf{a}_1$ is the projection vector that maximizes a PI. Using $\mathbf{a}_1$ we can project the data matrix $\mathbf{X}$ into a projection space generated by $z_n^1 = \mathbf{a}_1^{\mathrm{T}}\mathbf{x}_n$, where the data matrix $\mathbf{X}$ is formed by all pixel vectors in the image, i.e.,

$$\mathbf{X} = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n].$$

For example, if we assume that the histogram of the projection values $\{z_k\}_{k=1}^N$ is described by a Gaussian distribution, the outliers are generally caused by small targets that create ripples on either side of tails of the Gaussian distribution. So, the first values in the histogram that occur at zero will be selected as the desired thresholds. Then those projected pixels with gray scales excessive the thresholds were extracted and considered to be target pixels. As a result, two images can be obtained by such zero-detection thresholding, one for target detection and the other for the use of next projection. The one used for target detection is a binary image obtained by setting those pixels exceed the thresholds to 1 and zero, otherwise. This binary thresholded image, denoted by $\mathbf{B}_1$ will be used to tally target pixels detected by this particular projection image. A second image is a gray scale image and obtained by assigning 0 to all the pixels in the binary image $\mathbf{B}_1$ while the gray scales of the pixels not in the $\mathbf{B}_1$ remaining unchanged. This resulting image is referred to as the first projection image $\mathbf{X}_1$ and will be used for the next projection. The reason for producing such projection images is to prevent the target pixels detected in $\mathbf{B}_1$ from being considered again in the new projection image $\mathbf{X}_1$. A second projection vector $\mathbf{a}_2$ is then found again by maximizing a PI based on the first projection image $\mathbf{X}_1$ and generated by a projection space formed by $\{z_k^1\}_{k=1}^N$ that are generated by $\mathbf{z}^2 = \mathbf{a}_2^{\mathrm{T}}\mathbf{X}_1$ via $\mathbf{a}_2$. Similarly, a binary image $\mathbf{B}_2$ threshold from $\mathbf{z}^2$ can be found by zero-detection in the same manner that $\mathbf{B}_1$ was generated. In analogy with finding $\mathbf{X}_1$, a second projection image $\mathbf{X}_2$ can be also formed by setting to zero the gray scales of pixels in the binary image $\mathbf{B}_2$ while the pixels not in the binary image $\mathbf{B}_2$ remain unchanged. The same procedure is continued to generate the third projection vector $\mathbf{a}_3$, which maximizes PI based on the second projection image $\mathbf{X}_2$ to form two projection images $\mathbf{z}^3$, $\mathbf{X}_3$ and a third binary thresholded image $\mathbf{B}_3$ for subsequent target detection, etc. until the PPEA converges. Accordingly, a set of the desired targets can be extracted from a sequence of $\{\mathbf{B}_j\}$ which are resulting from a sequence of projection images $\{\mathbf{z}^j\}$ by zero-detection thresholding.

As an illustrative example, Fig. 1(a) shows a Gaussian-like histogram generated by the first projection space formed by $\{z_k^1\}_{k=1}^N$ that were generated by $\mathbf{z}^1 = \mathbf{a}_1^{\mathrm{T}}\mathbf{x} = (z_1^l, z_2^l, \ldots, z_N^l)^{\mathrm{T}}$ via $z_n^l = \mathbf{a}_1^{\mathrm{T}}\mathbf{x}_n$ using the HYDICE data in Fig. 4(a). The histogram in Fig. 1(a) was plotted based on the number of projected pixels versus their projection values generated by $z_n^l = \mathbf{a}_1^{\mathrm{T}}\mathbf{x}_n$ for $1 \le n \le N$. Since the right tail of the histogram does not clearly show ripples, Fig. 1(b) enlarges its long and flat right tail. As we can see, there are many ripples and the first value that detects zero will be selected as the desired threshold value as pointed by an arrow. It should be noted that the histogram in Fig. 1(a) is a Gaussian-like distribution, but it is asymmetric. In this case, the projection index using the skewness was able to detect small targets in Figs. 7 and 8 that caused ripples.

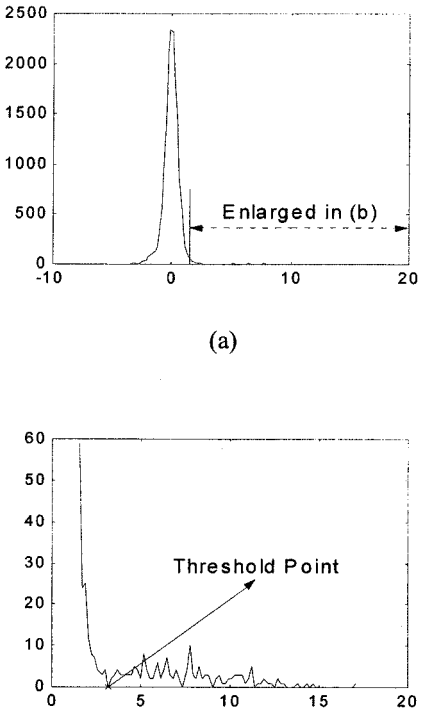


Fig. 1. (a) Gaussian-like histogram produced by $z_n = \mathbf{a}_1^{\mathrm{T}}\mathbf{x}_n$ using the HYDICE data in Fig. 4(a) and (b) enlarged right Gaussian tail of Fig. 1(a).

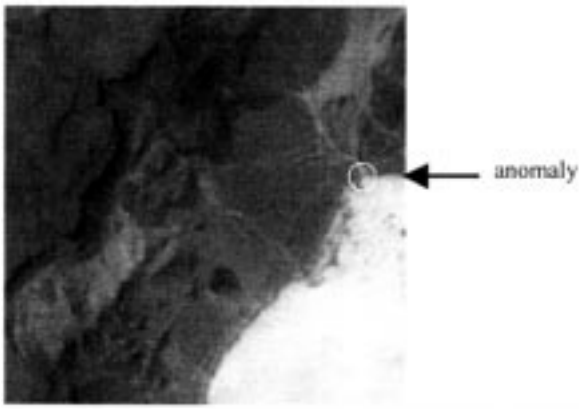


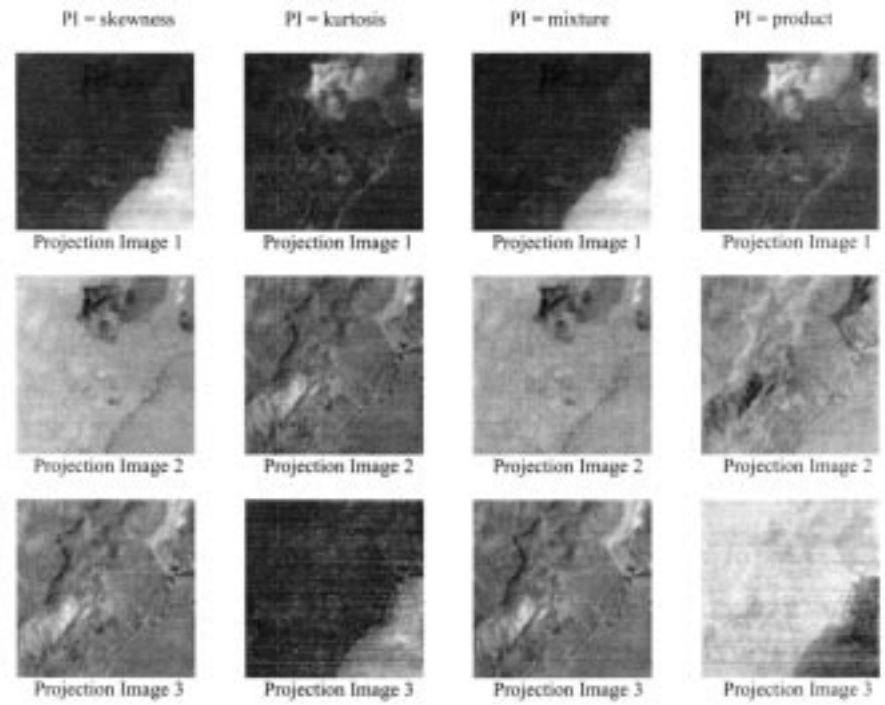
Fig. 2. AVIRIS scene.

VI. Experiments

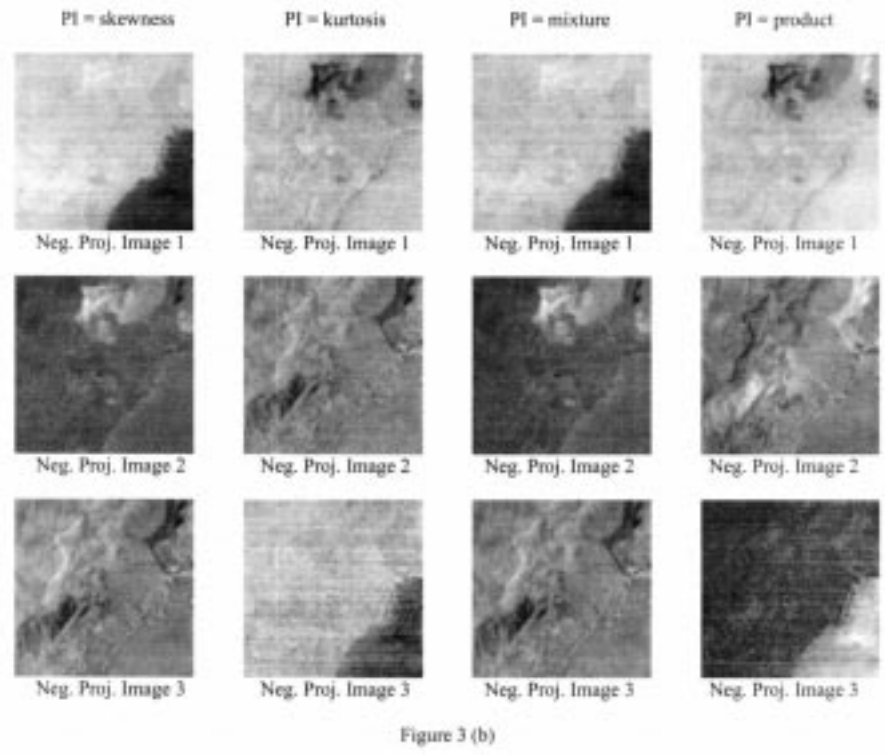
Two sets of hyperspectral data were used to evaluate the performance of PPEA in target detection.

A. AVIRIS Data Experiments

The 224-band AVIRIS data used in the experiments were the same data considered in [12]. It is a subscene of 200 × 200 pixels extracted from the upper left corner of the Lunar Crater Volcanic Field in Northern Nye County, NV, shown in Fig. 2, where five signatures of interest in these images were were demonstrated in [12], "red oxidized basaltic cinders," "rhyolite," "playa (dry lakebed)," "shade," and "vegetation." It



(a)



(b)

Fig. 3. (a) First three projection images of skewness, kurtosis, mixture and product. (b) First three negative projection images of skewness, kurtosis, mixture and product.

has been shown that in general there was no need to use all full bands for image classification [14], only 12 bands that were uniformly selected among 224 bands were used in the following experiments to ease computation of PPEA. Additionally, it was also shown in [13] that there was a single two-pixel anomaly located at the top edge of the lake marked by a dark circle in Fig. 2. This anomaly cannot be seen or detected visually, and its detection was also missed in [11], [12] because the method used, orthogonal subspace projection (OSP) was supervised. Fig. 3(a) shows the detection results of the first three projection images of all the four PIs proposed in this paper while Fig. 3(b) shows their negative counterparts of Fig. 3(b). The reason that

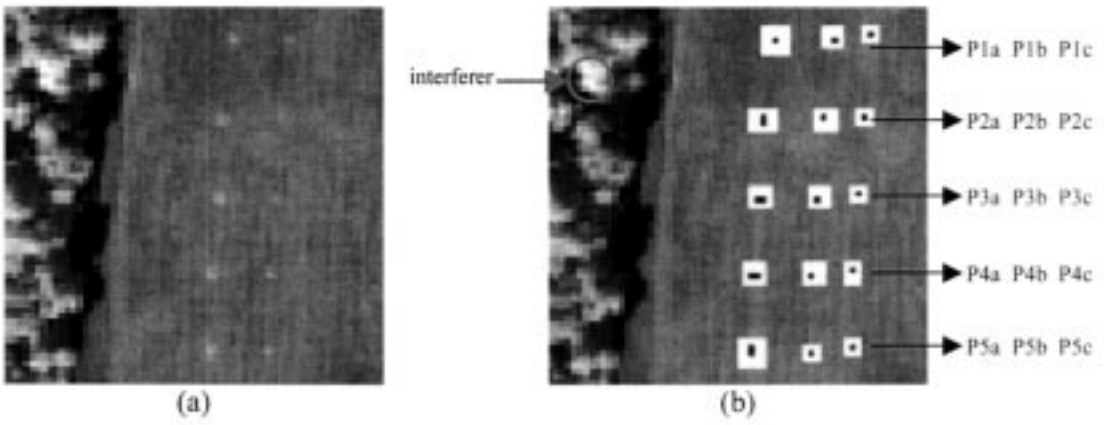


Fig. 4. (a) HYDICE image scene and (b) panel pixels in the scene in (a) masked by BLACK and WHITE.

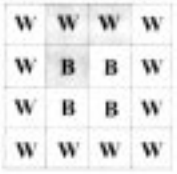


Fig. 5. Typical mask of B and W pixels.

we also included negative projections in Fig. 3(b) is because the projection values are not necessarily positive and they can be negative. The significance of interestingness of a projection is determined by its magnitude not signs. As shown, all the six signatures were detected in Fig. 3, specifically, the anomaly was detected in all the four PIs. Since kurtosis is more interesting than skewness in many other applications, we use kurtosis as an example for illustration. The cinders showed in the first projection positive image while the vegetation was detected in its negative counterpart. In the second projection, the vegetation and the rhyolite were extracted in the positive projection image, while the shade and anomaly were detected in its negative counterpart. In third projection, the playa was pulled out from the positive image. It turned out that both positive and negative projection images were very similar. We also noted that if a signature is interesting, it generally shows in more than one projection image such as the vegetation in Fig. 3.

B. Hydice Data Experiments

The HYDICE data used in the following experiments were directly extracted from the HYDICE image scene of size 64×64 shown in Fig. 4(a). There are 15 panels located on the field and arranged in a 5×3 matrix. The 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 for the experiments. The spatial resolution is 1.5 m and spectral resolution is 10 nm. The ground truth of the image scene is shown in Fig. 4(b) and provides the precise spatial locations of these 15 panels. Black pixels are panel center pixels and the pixels in the white masks are panel boundary pixels mixed with background pixels. Each element in this matrix is a square panel and denoted by p_{ij} with row indexed by $i = 1, 2, \ldots, 5$ and column indexed by $j = a, b, c$. For each row i, the three panels p_{i1}, p_{i2}, p_{i3} were made from the same material but have three different sizes. For each column j, the five panels p_1, p_2, p_3, p_4, p_5 have the same size but were made from five different materials. The sizes of the panels in the first, second and third columns are 3 m $\times$ 3 m, 2 m $\times$ 2 m, and 1 m $\times$ 1 m, respectively. The 1.5 m-spatial resolution of the image scene suggests that except for $p_{11}, p_{21}, p_{31}, p_{41}, p_{51}$,

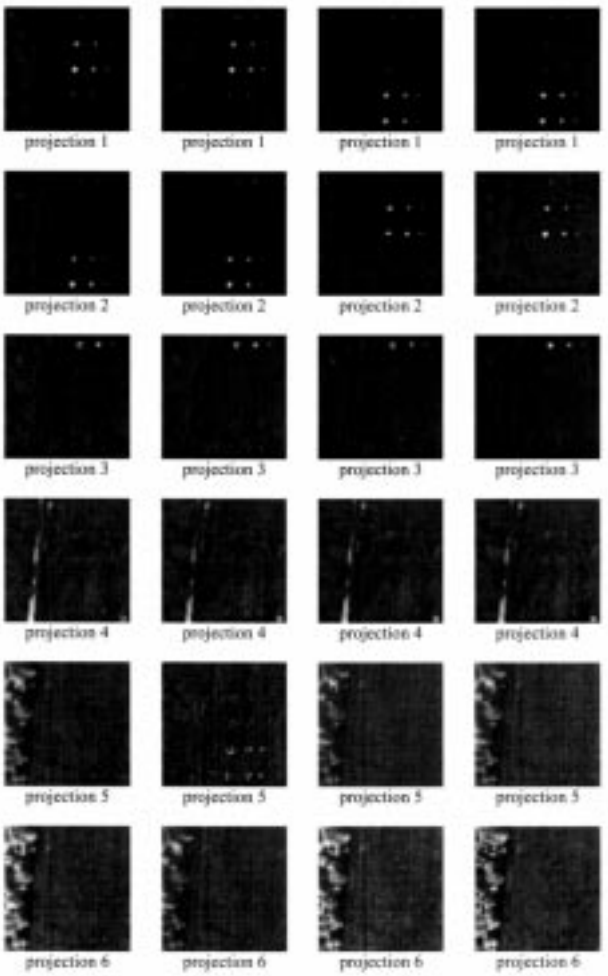


Fig. 6. Six gray scale projection images produced by four different PIs: (a) skewness, (b) kurtosis, (c) mixture, and (d) product.

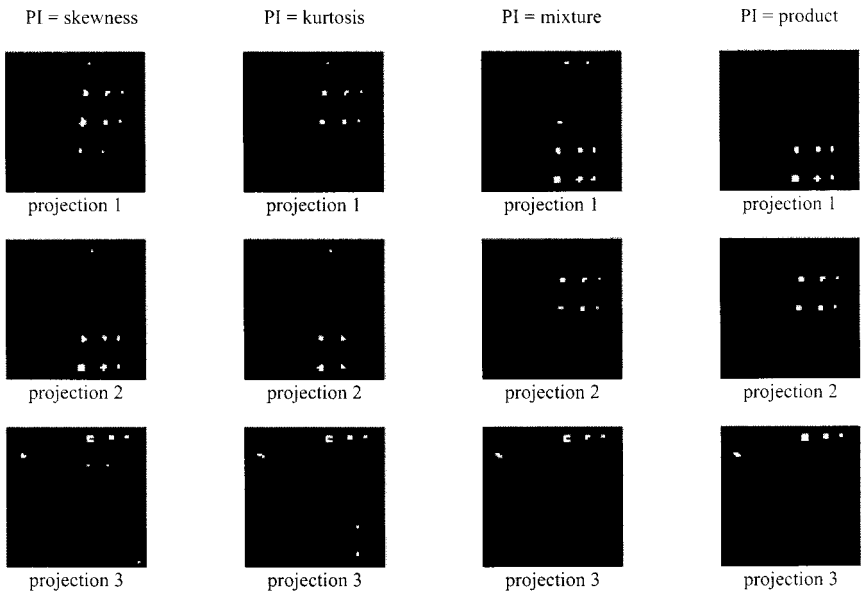


Fig. 7. Binary projection images resulting from thresholding the first three projection images in Fig. 6: (a) skewness, (b) kurtosis, (c) mixture, and (d) product.

which are two-pixel panels, all the remaining panels are only one-pixel in size. Apparently, without ground truth, it is difficult to locate these panels in the scene. Surprisingly, as will be shown in the experiments in Section VI, these panels can be accurately extracted by the proposed PPEA.

Like the AVIRIS data, we also used 12 bands selected uniformly for experiments to ease computation of PPEA. As will be shown in the following experiments, 12 bands are sufficient to retain all necessary information for PPEA to effectively detect target pixels. Since Fig. 4(b) provides a ground truth map, we can evaluate the performance of the PPEA by actually tallying number of panel pixels detected in the binary thresholded images B_j generated by the PPEA.

Following a similar study in [11], black and white pixels were used to tally results of target detection. The size of a mask used for a panel varies. A typical masked target of size 4×4 is shown in Fig. 5, with B pixels centered in the mask where a B and a W represent a B-masked pixel and a W-masked pixel, respectively. Two types of tally will be used for panel detection: 1) the number of panel pixels *hit* and 2) the number of panel pixels *detected*. Here we make a subtle difference between a panel *detected* and a panel *hit*. When a panel is detected, it means that at least one B target pixel is detected. When a panel is hit, it means that at least either one B or one W pixel is detected. As long as one of these B or W pixels is detected, we declare the panel is hit. So by way of this definition, a panel detected always implies a panel hit, but not vice versa.

In the following experiments, we will tally: 1) how many B panel pixels are detected; 2) how many W panel pixels are detected; 3) how many pixels are false alarmed for a panel, in which case neither a B-masked pixel nor a W-masked pixel is detected; and 4) how many B panel pixels are missed. In order to quantitatively study the detection performance of PPEA, the following definitions introduced in [11] were used.

N total number of sample pixel vectors;
p specific panel to be detected;
$N_{\mathrm{B+W}}(p)$ total number of B-masked and W-masked pixels in the panel p;
$N_{\mathrm{B}}(p)$ total number of B-masked pixels in the panel p;
$N_{\mathrm{W}}(p)$ total number of W-masked pixels in the panel p;
$N_{(\mathrm{B+W})\mathrm{D}}(p)$ total number of either B-masked or W-masked pixels detected in the panel p;
N_{BD} total number of B-masked pixels detected in the panel p;
$N_{\mathrm{WD}}(p)$ total number of W-masked p pixels detected in the panel p;
$N_{\mathrm{TPF}}(p)$ total number of false alarmed pixels, i.e., total number of pixels that are neither B-masked nor W-masked detected in the panel p;
$N_{\mathrm{TPM}}(p)$ total number of B-masked or W-masked pixels missed in the panel p.

With the help of the above notations, we can further define the B panel pixel detection rate $R_{\mathrm{BTD}}(p)$ for the panel p by

$$R_{\mathrm{BTD}}(p) = \frac{N_{\mathrm{BD}}(p)}{N_{\mathrm{B}}(p)} \tag{10}$$

and the W panel pixel detection rate $R_{\mathrm{WTD}}(p)$ for the panel p by

$$R_{\mathrm{WTD}}(p) = \frac{N_{\mathrm{WD}}(p)}{N_{\mathrm{W}}(p)}. \tag{11}$$

Since B pixels represent panel center pixels and W pixels may be panel pixels mixed by the background pixels, a good detection algorithm must have higher number of panel B pixels detected $N_{\mathrm{BD}}(p)$. On the other hand, detecting a W pixel does not necessarily mean a panel detected. Nevertheless, we can declare the panel to be hit. For this purpose, we define the panel hit rate $R_{\mathrm{TH}}(p)$ for panel p by

$$R_{\mathrm{TH}}(p) = \frac{N_{(\mathrm{B+W})\mathrm{D}}(p)}{N_{\mathrm{B+W}}(p)}. \tag{12}$$

So from (12), a higher panel hit rate $R_{\mathrm{TH}}(p)$ does not imply a higher panel detection rate $R_{\mathrm{BTD}}(p)$ or vice versa. This is because the number of W pixels are generally much greater than that of R pixels, thus it may eventually dominate the performance of $R_{\mathrm{TH}}(p)$.

TABLE I
VALUES OF FOUR PIS

	projection 1	projection 2	projection 3	projection 4	projection 5	projection 6
skewness	141.9172	198.5295	61.5250	22.1880	7.7661	6.1439
kurtosis	43665.6554	53827.0905	18125.6193	839.2410	1423.0965	194.3417
mixture	1141.7817	1137.9477	495.5014	25.8405	4.4355	1.7345
product	12880834	10769387	20991206	22674	1099	245

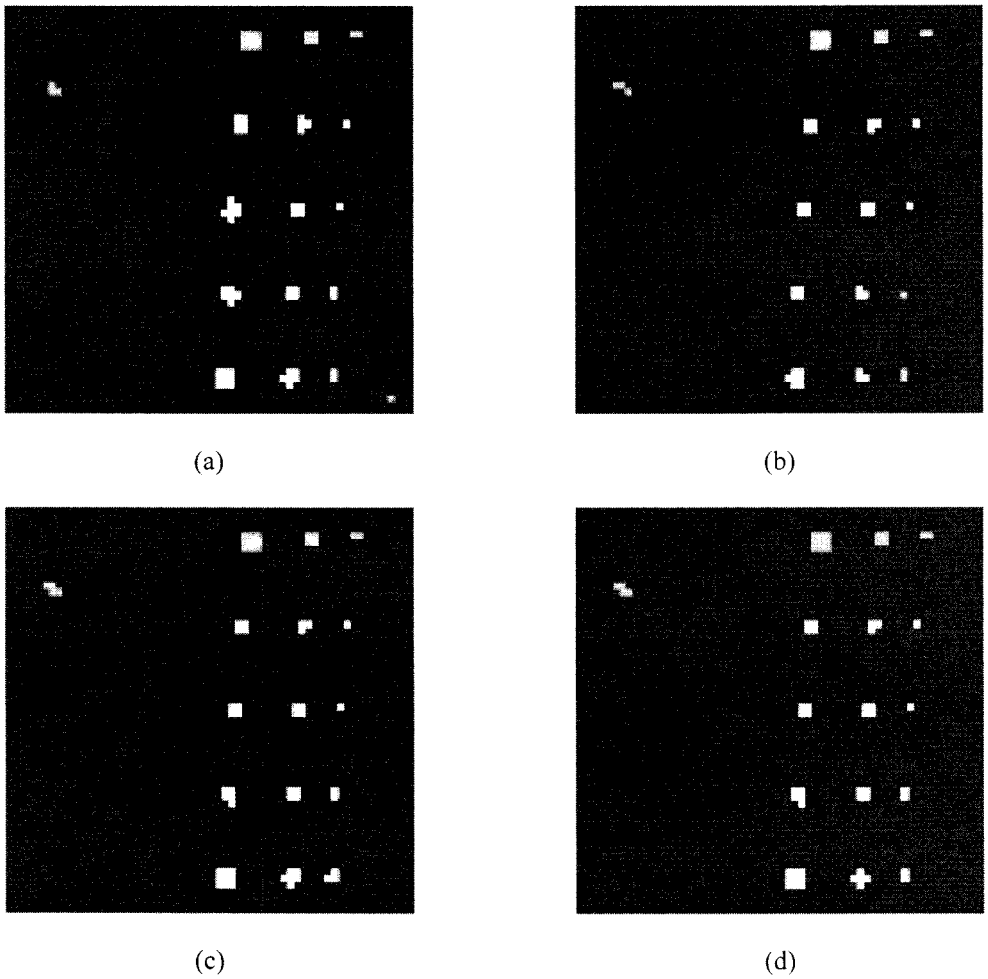


Fig. 8. Images resulting from combining the first three binary thresholded projection images in Fig. 7: (a) skewness, (b) kurtosis, (c) mixture, and (d) product.

In addition to (10)–(12), we are also interested in panel false alarm rate $R_{\mathrm{TPF}}(p)$ and panel miss rate $R_{\mathrm{TPM}}(p)$ defined as follows:

$$R_{\mathrm{TPF}}(p) = \frac{N_{\mathrm{TPF}}(p)}{N - N_{\mathrm{B+W}}(p)} \tag{13}$$

$$R_{\mathrm{TPM}}(p) = 1 - R_{\mathrm{TH}}(p) = \frac{N_{\mathrm{TPM}}(p)}{N_{\mathrm{B+W}}(p)} = \frac{N_{\mathrm{B+W}}(p) - N_{\mathrm{(B+W)D}}(p)}{N_{\mathrm{B+W}}(p)}. \tag{14}$$

If there are q panels, $\Gamma = \{p_i\}_{i=1}^{q}$ needed to be detected, the overall detection rate $R_{\mathrm{OD}}(\Gamma)$ for a class of panels, can be defined as

$$R_{\mathrm{OD}}(\Gamma) = \sum_{i=1}^{q} \mathrm{prob}(p_i) R_{\mathrm{BTD}}(p_i) \tag{15}$$

where $\mathrm{prob}(p_i) = N(p_i) / \sum_{k=i}^{q} N(p_k)$ for $1 \le i \le q$.

Fig. 6 shows the detection results based on the first six projections using the proposed four PIs where projection images were produced by skewness, kurtosis, mixture and product and are labeled by (a), (b), (c), and (d), respectively. As we can see from these images in Fig. 6, all the panels were extracted in the first three projection images. Fig. 7 shows the binary thresholded images of the first three projection images in Fig. 6 where all the four PIs detected the panels in the first row in their third projection images, the panels in the second and third rows in one projection image and the panels in the fourth and fifth row in another projection image. The reason that the panels in the second and third rows were detected by the same projection image is because their spectra are very similar. It is also true for the panels in the fourth and fifth rows. As shown in Fig. 7, all the 15 panels were successfully detected and extracted. These detection results are impressive by the fact that no panels of 1 m $\times$ 1 m in Fig. 4(a) can be detected by visual inspection. Table I tabulates

TABLE II
DETECTION RATES FOR THE PI USING SKEWNESS

T	N_{B+W}	N_B	N_W	$N_{(B+W)D}$	N_{BD}	N_{WD}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPM}
P1	50	3	47	15	3	12	35	1.0000	0.2553	0.3000	0.7000
P2	45	4	41	11	4	7	34	1.0000	0.1707	0.2444	0.7556
P3	41	4	37	12	4	8	29	1.0000	0.2162	0.2927	0.7073
P4	44	4	40	12	4	8	32	1.0000	0.2000	0.2727	0.7273
P5	43	4	39	17	4	13	26	1.0000	0.3333	0.3953	0.6047

TABLE III
DETECTION RATES FOR THE PI USING KURTOSIS

T	N_{B+W}	N_B	N_W	$N_{(B+W)D}$	N_{BD}	N_{WD}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPM}
P1	50	3	47	15	3	12	35	1.0000	0.2553	0.3000	0.7000
P2	45	4	41	8	4	4	37	1.0000	0.0976	0.1778	0.8222
P3	41	4	37	9	4	5	32	1.0000	0.1351	0.2195	0.7805
P4	44	4	40	8	3	5	36	0.7500	0.1250	0.1818	0.8182
P5	43	4	39	12	4	8	31	1.0000	0.2051	0.2791	0.7209

TABLE IV
DETECTION RATES FOR THE PI USING MIXTURE

T	N_{B+W}	N_B	N_W	$N_{(B+W)D}$	N_{BD}	N_{WD}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPM}
P1	50	3	47	15	3	12	35	1.0000	0.2553	0.3000	0.7000
P2	45	4	41	8	4	4	37	1.0000	0.0976	0.1778	0.8222
P3	41	4	37	9	4	5	32	1.0000	0.1351	0.2195	0.7805
P4	44	4	40	11	4	7	33	1.0000	0.1750	0.2500	0.7500
P5	43	4	39	18	4	14	25	1.0000	0.3590	0.4186	0.5814

TABLE V
DETECTION RATES FOR THE PI USING PRODUCT

T	N_{B+W}	N_B	N_W	$N_{(B+W)D}$	N_{BD}	N_{WD}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPM}
P1	50	3	47	15	3	12	35	1.0000	0.2553	0.3000	0.7000
P2	45	4	41	8	4	4	37	1.0000	0.0976	0.1778	0.8222
P3	41	4	37	9	4	5	32	1.0000	0.1351	0.2195	0.7805
P4	44	4	40	11	4	7	33	1.0000	0.1750	0.2500	0.7500
P5	43	4	39	16	4	12	27	1.0000	0.3077	0.3721	0.6279

TABLE VI
OVERALL DETECTION RATES FOR FOUR PIS

	$N = 4096$		$N_{B+W} = 223$		$N_B = 19$		$N_W = 204$			
PI	$N_{(B+W)D}$	N_{BD}	N_{WD}	N_{TPF}	N_{TPM}	R_{BTD}	R_{WTD}	R_{TH}	R_{TPF}	R_{TPM}
skewness	67	19	48	4	156	1.0000	0.2353	0.3004	0.0010	0.6996
kurtosis	52	18	34	3	171	0.9474	0.1667	0.2332	0.0008	0.7668
mixture	61	19	42	4	162	1.0000	0.2059	0.2735	0.0010	0.7256
product	59	19	40	4	164	1.0000	0.1961	0.2646	0.0010	0.7354

the values of PIs of the first six projection images for these four PIs. Table I also shows that the values of the first three values of all the four PIs are significantly greater than those of their remaining PIs. This observation is consistent with the detection results in Figs. 6 and 7 and provides a clue that only three projections may be sufficiently enough for target detection.

In order to calculate various detection rates specified by (10)–(15), Fig. 8(a)–(d) was obtained by combining the first three binary thresholded projection images in Fig. 7 to tally detected panel pixels. Interestingly, three pixels in the tree line detected in Fig. 8(a) and (d) by skewness and kurtosis, and four pixels detected in Fig. 8(c) and (d) by mixture and product were not panel pixels but simply interferers. It should be noted that for each PI the sets of pixels detected by all

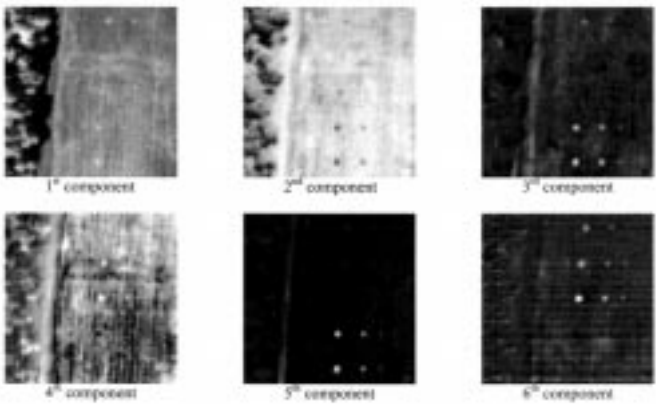


Fig. 9. Six component images resulting from principal components analysis.

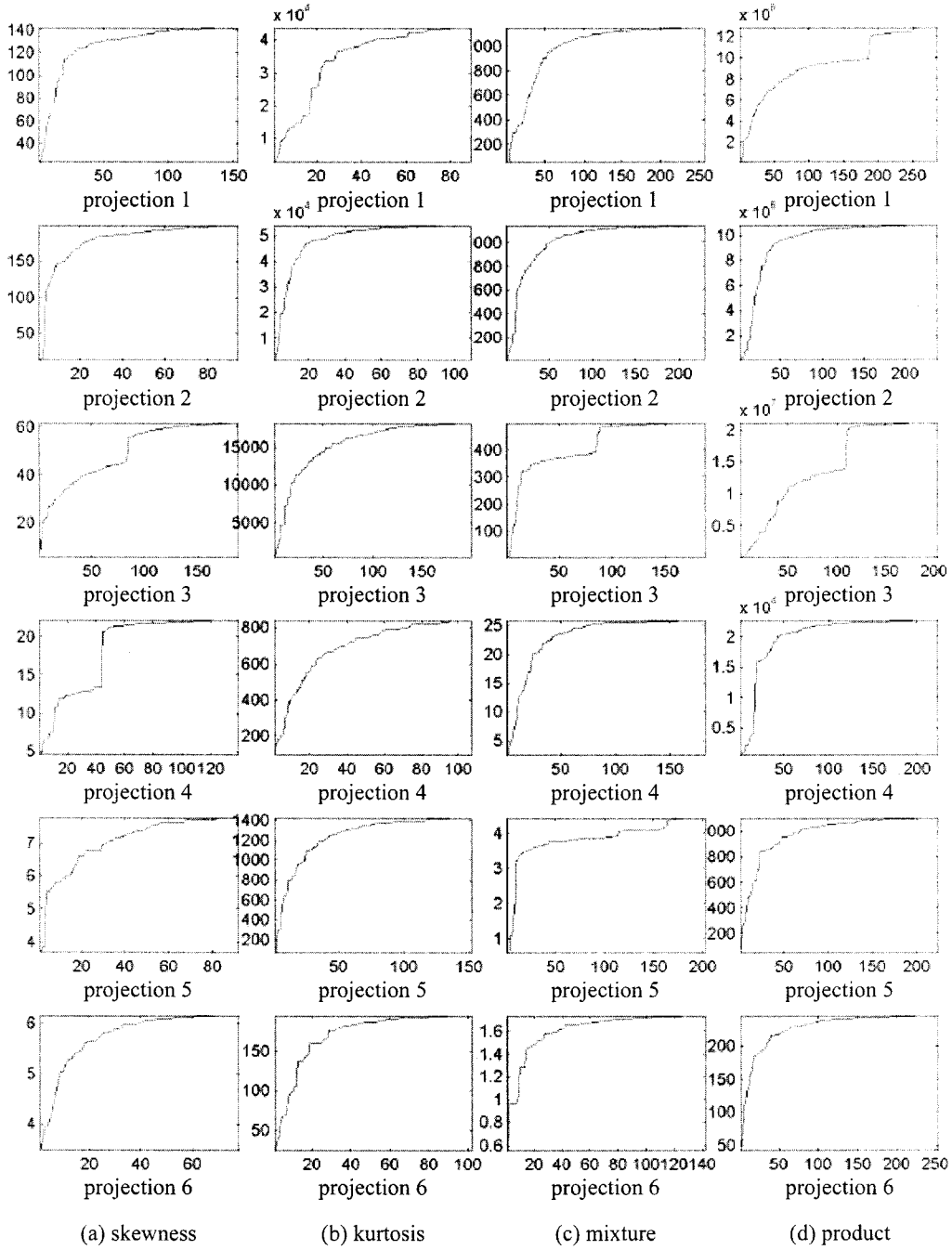


Fig. 10. (a) Plot of six learning curves of the PI using skewness corresponding to Fig. 6(a), (b) plot of six learning curves of the PI using kurtosis corresponding to Fig. 6(b), (c) plot of six learning curves of the PI using mixture corresponding to Fig. 6(c), and (d) plot of six learning curves of the PI using product corresponding to Fig. 6(d).

the three binary thresholded projection images were disjoint. So, the overall detected pixels by a particular PI is simply the sum of pixels detected by its first three binary thresholded projection images. All four PIs produced comparable results as shown in Tables II–VI. It is worth noting that no false alarm rate was calculated for Tables II–V and only Table VI has the overall panel false alarm rate $R_{\rm TPF}$. This is because we were interested in detecting all panel pixels in the scene rather than individual panel pixels. From Table VI, we can see that the overall B-target pixel detection rate $R_{\rm BTD}$ are 100% except of $I_{\rm kurtosis}(\mathbf{a}) = \kappa_4^2$, which produced 0.9474 and was also close to 1. On the other hand, all of the four PIs yielded low W-target pixel detection rates $R_{\rm WTD}$ and the target hit rates $R_{\rm TH}$. This shows that the proposed PPEA using four PIs are indeed very effective since a higher $R_{\rm BTD}$ reduces $R_{\rm WTD}$. In addition, and $R_{\rm TPF}$ is very low and less than or equal to 0.1%. By contrast, the panel miss rate $R_{\rm TPM}$ is very high and close to 70%-75%. This is due to the fact that $R_{\rm TPM}$ takes into account W panel pixels for target detection. In reality, W-pixels are generally not real panel pixels, and even if they are, they may be mixed by background pixels. A high $R_{\rm TPM}$ does not necessarily imply low $R_{\rm BTD}$. This

is particularly true when the number of W-pixels is relatively large compared to that of B-pixels.

In order to further demonstrate the performance of PPEA, PCA was also studied for comparison. It is known that PCA can be considered to be a special case of PP with data variance used as a projection index. The results of the first 6 component images are shown in Fig. 9. Comparing to Fig. 6, the four proposed PIs extracted all the 15 panels in their first three projection images while PCA picked up these panels in the fourth, fifth and sixth component images. This is because the panels are considered to be small targets in the scene and they do not generate significant information as opposed to that provided by the large grass field and tree line. As a consequence, the grass and tree line were extracted in the first three principal component images in Fig. 9. Interestingly, the interferers in the tree line were also shown in the first and third component images. This example shows that PCA can be used for preserving target information but not for the purpose of detecting targets. It should be noted that in order to account for nonstationarity of the data, the sample correlation matrix instead of sample covariance matrix was used for PCA.

To conclude our experiments, we further plotted in Fig. 10 the learning curves of the first six projections using PPEA for four PIs specified by (3)–(6) respectively with the x-axis being the number of iterations and the y-axis being the values of PIs. In all cases, no more 200 iterations are required in each optimal projection search. In particular, only approximate 150 iterations are needed to achieve the optimal search for the first three projections.

A final comment is noteworthy. The proposed PPEA was first studied in [15] where AVIRIS data were used for analysis. Two major differences in the experiments were conducted in [15] and this paper. One is the spatial resolution of the data, 1.5 m for HYDICE data as opposed to 20 m for AVIRIS data. Since the fine spatial resolution generally uncovers small targets that caused the outliers of the background distribution, the PPEA approach is expected to perform better for HYDICE data than for AVIRIS data. In fact, this was the case supported by the above experiments. Another is that the spatial locations of the panels in HYDICE data in Fig. 4(b) can be specified by B and W pixels. In this case, we can actually tally how many of panel pixels detected to substantiate the PPEA approach.

VII. Conclusion

PP has received considerable interests in multivariate analysis because it can be used to explore interesting projections by which a high dimensional data set can be projected into a low dimensional space for various applications. This paper presents a new application of PP in unsupervised target detection for hyperspectral imagery. Since targets are small compared to their surrounding background, these targets can be viewed as pixels that cause outliers of the background distribution. Four projection indices are suggested based on the third and fourth moments to capture targets. In order to find the optimal projections, a revised PPEA is proposed where a zero-detection thresholding technique is introduced for the purpose of target detection. Targets, which cannot be detected by visual inspection, can be effectively extracted by PPEA using only a few projections. Although the experiments are conducted for PPEA for unsupervised target detection, it can be also used in conjunction with a database or spectral library for classification and identification. For this purpose, a spectral metric such as Euclidean distance [16], spectral angle [17] or spectral information divergence [18], [19] is required to measure the spectral similarity between the detected pixels and database to achieve classification. Finally, due to the nature of EA, the computational complexity of PPEA is generally very expensive. However, this can be compensated for high computer powers. As a concluding remark, it is worth noting that an independent component analysis-based linear mixture analysis recently developed in [20] took a similar approach, which also used the third and fourth moments as criteria for unsupervised subpixel detection. It can be considered as a special case of the projection pursuit presented in this paper where the background distribution is modeled by a Gaussian process and the abundance fractions of image endmembers were assumed to be random sources rather than nonrandom unknown quantities as are considered in most of linear unmixing methods such as orthogonal subspace projection in [12]. When an image scene is made up of several homogeneous background regions, each of which consists of a very large number of pixels, and in this case, we can assume that these background regions can be described by Gaussian distributions. Using this assumption, small targets can be detected effectively as demonstrated by experiments in [21].

Acknowledgment

The authors would like to thank Dr. J. C. Harsanyi for providing the AVIRIS data.

References

[1] J. H. Friedman and J. W. Tukey, "A projection pursuit algorithm for exploratory data analysis," *IEEE Trans. Comput.*, vol. C-23, pp. 881–889, Sept. 1974.

[2] J. H. Friedman, "Exploratory projection pursuit," *J. Amer. Statist. Assoc.*, vol. 82, pp. 249–266, 1987.

[3] P. J. Huber, "Projection pursuit," *Ann. Statist.*, vol. 13, no. 2, pp. 435–475, 1985.

[4] M. C. Jones and R. Sibson, "What is projection pursuit," *J. R. Statist. Soc. Assoc.*, pt. 1, vol. 150, pp. 1–36, 1987.

[5] M. E. Bullock, T. J. Patterson, and S. R. Fairchild, "Design of optimal transformation for multispectral change detection using projection pursuit," *Proc. SPIE*, vol. 2231, pp. 91–102, 1994.

[6] L. O. Jimenez and D. A. Landgrebe, "Hyperspectral data analysis and supervised feature reduction via projection pursuit," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2653–2667, Nov. 1999.

[7] A. Ifarragaerri and C.-I. Chang, "Multispectral and hyperspectral image analysis with projection pursuit," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 2529–2538, Nov. 2000.

[8] D. E. Golderg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

[9] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs 3rd. rev. and extended*. Berlin, Germany: Springer-Verlag, 1996.

[10] D. Dasgupta and Z. Michalewicz, *Evolutionary Algorithms in Engineering Applications*. Berlin, Germany: Springer-Verlag, 1997.

[11] C.-I. Chang and H. Ren, "An experiment-based quantitative and comparative analysis of hyperspectral target detection and image classification algorithms," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1044–1063, Mar. 2000.

[12] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 779–785, July 1994.

[13] C.-I. Chang and D. Heinz, "Subpixel spectral detection for remotely sensed images," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1144–1159, May 2000.
[14] C.-I. Chang, Q. Du, T. S. Sun, and M. L. G. Althouse, "A joint band prioritization and band decorrelation approach to band selection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2631–2641, Nov. 1999.
[15] S.-S. Chiang and C.-I. Chang, "Target subpixel detection for hyperspectral imagery using projection pursuit," in *EOS/SPIE Symp. Remote Sensing, Conf. Image and Signal Processing for Remote Sensing V*, Florence, Italy, Sept. 20–24, 1999.
[16] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[17] R. A. Schowengerdt, *Remote Sensing: Model and Methods for Image Processing*, 2nd ed. New York: Academic, 1997.
[18] C.-I Chang, "Spectral information divergence for hyperspectral image analysis," in *Int. Geoscience and Remote Sensing Symp.'99*, Hamburg, Germany, June 28, 1999.
[19] C.-I. Chang, "An information theoretic-based approach to spectral variability, similarity and discriminability for hyperspectral image analysis," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1927–1932, Aug. 2000.
[20] S.-S. Chiang, C.-I. Chang, and I. W. Ginsberg, "Unsupervised hyperspectral image analysis using independent components analysis," in *IEEE 2000 Int. Geoscience and Remote Sensing Symp.*, Honolulu, HI, July 24–28, 2000.
[21] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. New York: Kluwer, 2001.

Shao-Shan Chiang (S'98) received the B.S. degree from National Chengchi University, Taipei, Taiwan, R.O.C. in 1988, and the M.S. degree from National Taiwan University, Taipei, in 1990, both in mathematics. He is currently pursuing the Ph.D. degree with the Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County (UMBC), Baltimore, MD.

From 1992 to 1996, he was a Software Engineer in Taipei. From 1996 to 1998, he was a Research Assistant with the Anesthesiology Research Laboratory, UMBC. Since 1999, he has been a research assistant in the Remote Sensing Signal and Image Processing Laboratory at UMBC. His research interests include remote sensing, digital signal and image processing, and pattern recognition.

Mr. Chiang is a member of SPIE.

Chein-I Chang (S'81–M'87–SM'92) received the B.S. degree from Soochow University, Taipei, Taiwan, R.O.C., in 1973, the M.S. degree from the Institute of Mathematics, National Tsing Hua University, Hsinchu, Taiwan, in 1975, and the M.A. degree from the State University of New York, Stony Brook, in 1977, all in mathematics. He also received the M.S. and M.S.E.E. degrees from the University of Illinois, Urbana, in 1982, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 1987.

He was with the University of Maryland, Baltimore County (UMBC), Baltimore, as a Visiting Assistant Professor from January 1987 to August 1987, Assistant Professor from 1987 to 1993, as an Associate Professor in the Department of Computer Science from 1993 to 2001, and is now a Professor in the Department of Computer Science. He was a Visiting Research Specialist with the Institute of Information Engineering, National Cheng Kung University, Tainan, Taiwan, from 1994 to 1995. He has a patent on automatic pattern recognition and several pending patents on image processing techniques for hyperspectral imaging and detection of microcalcifications. He is currently on the editorial board of *Journal of High Speed Networks* and is the Guest Editor of a special issue of that journal on Telemedicine and Applications. His research interests include automatic target recognition, multispectral/hyperspectral image processing, medical imaging, information theory and coding, signal detection and estimation, and neural networks.

Dr. Chang is a member of SPIE, INNS, Phi Kappa Phi, and Eta Kappa Nu.

Irving W. Ginsberg received the B.A. and Ph.D. degrees in theoretical physics in 1960 from Wayne State University, Detroit, MI.

He is currently a Consultant to the U.S. government and industry on the application of spectral and infrared technologies and algorithms. He was the Chief Scientist of the Department of Energy's Remote Sensing Laboratory, Washington, DC, until 2001. Previously, he worked for the Willow Run Laboratories and the Environmental Research Institute of Michigan, University of Michigan. Ann Arbor. Much of his work has been on the phenomenology of reflectance and in the exploitation of multispectral and hyperspectral imagery.