Anomaly Detection and Classification for Hyperspectral Imagery

Chein-I Chang, *Senior Member, IEEE,* and Shao-Shan Chiang, *Member, IEEE*

***Abstract*—Anomaly detection becomes increasingly important in hyperspectral image analysis, since hyperspectral imagers can now uncover many material substances which were previously unresolved by multispectral sensors. Two types of anomaly detection are of interest and considered in this paper. One was previously developed by Reed and Yu to detect targets whose signatures are distinct from their surroundings. Another was designed to detect targets with low probabilities in an unknown image scene. Interestingly, they both operate the same form as does a matched filter. Moreover, they can be implemented in real-time processing, provided that the sample covariance matrix is replaced by the sample correlation matrix. One disadvantage of an anomaly detector is the lack of ability to discriminate the detected targets from another. In order to resolve this problem, the concept of target discrimination measures is introduced to cluster different types of anomalies into separate target classes. By using these class means as target information, the detected anomalies can be further classified. With inclusion of target discrimination in anomaly detection, anomaly classification can be implemented in a three-stage process, first by anomaly detection to find potential targets, followed by target discrimination to cluster the detected anomalies into separate target classes, and concluded by a classifier to achieve target classification. Experiments show that anomaly classification performs very differently from anomaly detection.**

***Index Terms*—Anomaly detection, Bhattacharyya distance, causal RXD (CRXD), classification, correlation matched-filter-based measure (RMFM), low-probability target detector (UTD), Mahalanobis distance, RX detector (RXD), target discrimination measure.**

I. Introduction

TARGET detection in remotely sensed images can be conducted spatially, spectrally, or both. The difficulty with using spatial image analysis for target detection in remotely sensed imagery arises from the fact that the ground sampling distance (GSD) is generally larger than the size of targets of interest. In this case, targets are embedded in a single pixel and cannot be detected spatially. Under such circumstances, target detection must be carried out at subpixel level. An anomaly detector enables one to detect targets whose signatures are spectrally distinct from their surroundings with no *a priori* knowledge. In general, such anomalous targets are relatively small compared to the image background and only occur in the image with low probabilities. Two approaches are of particular interest. One was developed by Reed and Yu [1]–[3] and is referred to as the RX detector (RXD), which has shown success in anomaly detection for multispectral and hyperspectral images [4], [5]. Another was proposed in [6], [7] and is referred to as low-probability detection (LPD), which was designed to detect targets with low probabilities in an image. Interestingly, both approaches result in the same form of a matched filter, but differ in two aspects. The RXD uses the pixel currently being processed as the matched signal, while the LPD makes use of the unity vector (i.e., it has all ones in its components) as its matched signal. Therefore, the matched signal used in the RXD varies pixel-by-pixel as opposed to the constant-matched signal used in the LPD. Since the RXD uses the sample covariance matrix to take into account the sample spectral correlation, it performs the same task as does the Mahalanobis distance [8]. On the other hand, the LPD makes use of the sample correlation matrix to account for spectral correlation among samples. As a result, it turns out to be a special case of the constrained energy minimization (CEM) filter developed in [6], [7], [9], [10] where the desired signal is designated as the unity vector.

In this paper, these two anomaly detectors are investigated and explored. In particular, several variants of the RXD and LPD are derived, and a real-time processing version of the RXD is also introduced where the sample correlation matrix instead of the sample covariance matrix is used. It is referred to as causal RXD (CRXD). The term "causal" borrows from signal processing terminology which means that the information used for data processing is up to the pixel being processed and updated solely based on the pixels that were already processed. "Real-time process" refers to a process that processes data samples when they come with no time delay. It can be implemented in two ways: line-by-line and pixel-by-pixel. As shown in [11], a line-by-line real-time process can be implemented by considering a line of pixels as an input array. It processes each line while the line is being scanned. A pixel-by-pixel real-time process considers each pixel as an input and processes a pixel while it being scanned. An experiment of a pixel-by-pixel CRXD will be demonstrated for illustration. Two advantages can be benefited from the CRXD. Since the computation of the inverse of a sample correlation matrix can be carried out in parallel via a QR-decomposition [11]–[13], the CRXD has an ability of processing data in a real-time fashion. In other words, unlike the RXD, which requires the knowledge of all the data samples to form the sample covariance matrix prior to processing, the CRXD processes and updates data either line-by-line or sample-by-sample. A second advantage results

Manuscript received November 22, 2000; revised September 24, 2001. This work was supported by the Bechtel Nevada Corporation under Contract DE-AC08-96NV11718 through the U.S. Department of Energy and Office of Naval Research under Contract N00014-01-1-0359.

C.-I Chang is 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.

S.-S. Chiang is with the Department of Electrical Engineering, Lunghwa University of Science and Technology, Taoyuan, Taiwan, R.O.C.

Publisher Item Identifier 10.1109/TGRS.2002.800280.

from the fact that the sample correlation matrix accounts for both the first-order and second-order statistics. For most remotely sensed images which are generally considered to be nonstationary, the CRXD can capture spectral variability more effectively than the RXD, which only takes care of second-order statistics.

Since the images produced by the RXD are generally grayscale, the target detection and discrimination are usually carried out by visual inspection. In order to avoid such human interpretation, an automatic thresholding method is designed to automatically select an appropriate threshold to segment targets from the image background. This method is adaptive and only depends on the grayscale image resulting from anomaly detection, but not on the original image.

It is often the case that an anomaly detector can detect different types of anomalies but cannot discriminate the detected targets from one another. So, in this paper, four target discrimination measures are further developed to resolve this dilemma. A target discrimination measure discriminates the detected anomalies and clusters them into separate target classes. Due to the fact that the number of the detected target pixels is generally small, the automatic thresholding methods commonly used in traditional spatial-based image processing may not be applicable. In this case, an automatic thresholding method is proposed and modified from the zero-detection thresholding technique proposed in [14] to cluster the detected targets into separate classes. The means of these target classes are then calculated and used as the target information for follow-up target classification. Because the resulting classification is supervised, the classification results may classify additional target pixels whose signatures are similar to the anomalies, but were not detected by the RXD.

In analogy with the CRXD, a real-time implementation of anomaly classification is also possible. However, in this case, a certain time delay is inevitable between anomaly detection and target discrimination, but is negligible if the image size is manageable, such as one used in the experiments conducted in this paper. In order to make distinction of this subtle difference, a process is referred to as an online process if it can be implemented in a timely manner with negligible time lag. So, with this definition, anomaly classification can be implemented as an online process by the CRXD in conjunction with a target discrimination measure and followed by the real-time linearly constrained minimum-variance (LCMV) classifier developed in [11], [15]. More specifically, anomaly classification can be accomplished by three-stage processes. The anomaly detection is initiated in the first-stage process, then followed by a clustering process using a target discrimination measure in the second stage and, finally, concluded by the LCMV classification in the third stage. All these three stages can be processed online with no appreciable time delay.

This paper is organized as follows. Section II develops techniques for anomaly detection which include the RXD, the LPD, and their variants. Section III demonstrates the real-time processing capability of the CRXD. Section IV designs a new computer-automated thresholding method for target detection. Section V describes an approach to anomaly classification where four target discrimination measures are introduced. Section VI presents a series of experiments to demonstrate the online process of anomaly classification. Section VII concludes with some remarks.

II. Anomaly Detection

In this section, we describe two types of anomaly detectors which are developed for different purposes. Nonetheless, both detectors perform the same functional form of a matched filter with different matched signals.

A. RXD

The RX detector is the RX algorithm referenced in [4], [5] which was developed by Reed and Yu [1]. Suppose that L is the number of spectral bands and $\mathbf{r}$ is an $L \times 1$-column pixel vector in a multispectral or hyperspectral image. Then the RXD implements a filter specified by

$$\delta_{\mathrm{RXD}}(\mathbf{r}) = (\mathbf{r}-\mu)^T \mathbf{K}_{L\times L}^{-1}(\mathbf{r}-\mu) \tag{1}$$

where μ is the global sample mean and $\mathbf{K}_{L\times L}$ is the sample covariance matrix of the image. The form of $\delta_{\mathrm{RXD}}(\mathbf{r})$ in (1) is actually the well-known Mahalanobis distance. In order to see how the RXD can detect anomalous targets, an exploration of how the RXD works is worthwhile.

It is known that principal components analysis (PCA) decorrelates the data matrix in such a manner that different amounts of the image information can be preserved in separate components images, each of which represents a different piece of uncorrelated information. So, PCA has been widely used to compress image information into a few major principal components specified by the eigenvectors of $\mathbf{K}_{L\times L}$ that correspond to large eigenvalues, but it is not designed to be used for detection or classification. However, if the image data contain interesting target pixels which occur with low probabilities in the data (i.e., the size of target samples is small), it is obvious that these targets will not be shown in major principal components, but rather in minor components specified by the eigenvectors of $\mathbf{K}_{L\times L}$ that are associated with small eigenvalues. This phenomenon was observed and demonstrated in [10]. It is interesting to note that the RXD can be considered to be an inverse operation of the PCA which searches for targets in minor components. This provides explanation of why the RXD works for anomaly detection.

More specifically, assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L$ are the eigenvalues of $\mathbf{K}_{L\times L}$ with zero mean, i.e., $\mu = \mathbf{0}$ so that $\mathbf{K}_{L\times L} = \mathbf{R}_{L\times L}$. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_L\}$ be a set of orthogonal unit eigenvectors (i.e., the length of $\mathbf{v}_l$ is one, $||\mathbf{v}_l|| = 1$) with $\mathbf{v}_l$ corresponding to λ_l. Using the spectral decomposition of a covariance matrix in [16], [17], we can form an eigenmatrix $\mathbf{A} = [\mathbf{v}_1\mathbf{v}_2\cdots\mathbf{v}_L]$ with the lth column specified by the lth eigenvector $\mathbf{v}_l$. The resulting matrix $\mathbf{A}$ is a unitary matrix and can be used to decorrelate $\mathbf{K}_{L\times L}$ into a diagonal matrix $\Lambda = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_{\mathrm{L}}\}$ such that $\mathbf{A}^T\mathbf{K}_{L\times L}\mathbf{A} = \Lambda$. If we let $\mathbf{y} = \mathbf{A}^T\mathbf{r}$, then

$$\begin{aligned}\mathbf{r}^T\mathbf{K}_{L\times L}\mathbf{r} &= (\mathbf{A}\mathbf{y})^T\mathbf{K}_{L\times L}(\mathbf{A}\mathbf{y}) = \mathbf{y}^T[\mathbf{A}^T\mathbf{K}_{L\times L}\mathbf{A}]\mathbf{y} \\ &= \mathbf{y}^T\Lambda\mathbf{y} = \sum_{l=1}^{L}\lambda_l y_l^2. \end{aligned} \tag{2}$$

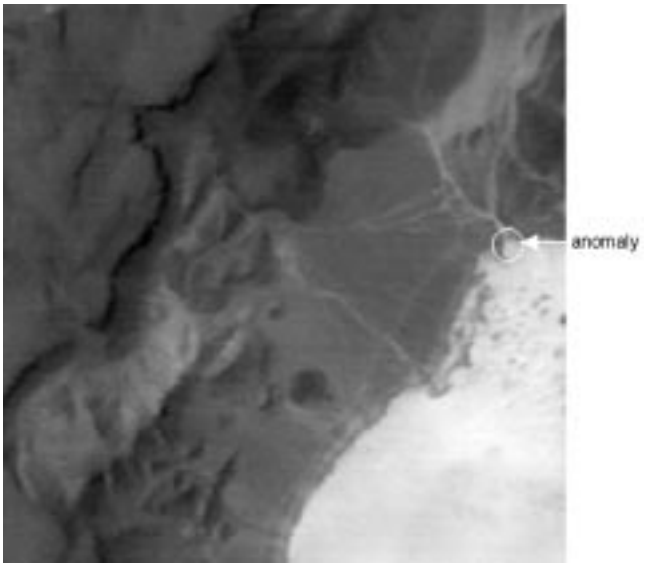

Fig. 1. An LCVF AVIRIS image scene.

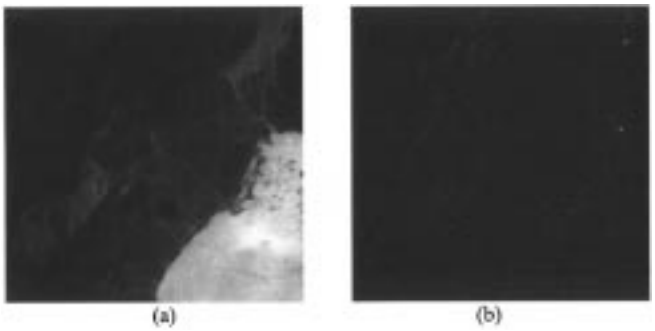

Fig. 2. (a) AVIRIS result produced by $\mathbf{r}^T\mathbf{K}_{L\times L}\mathbf{r}$ given by (2). (b) AVIRIS result produced by the RXD given by (1).

Using (2), expression (1) which specifies the RXD can be further expressed by

$$\mathbf{r}^T\mathbf{K}_{L\times L}^{-1}\mathbf{r}=\sum_{l=1}^{L}\lambda_l^{-1}y_l^2. \tag{3}$$

According to (2), the larger the eigenvalue is, the greater the value of $\mathbf{r}^T\mathbf{K}_{L\times L}\mathbf{r}$ is. So, (2) performs a similar task that PCA does. It can compress data by retaining only a few principal components specified by eigenvectors that corresponds to the first few largest eigenvalues. By contrast, using (3) the RXD detects anomalous targets with small energies that are represented by small eigenvalues. This is because the smaller the eigenvalue is, the greater the value of $\mathbf{r}^T\mathbf{K}_{L\times L}^{-1}\mathbf{r}$ is. In this case, it is crucial to determine p, the number of targets present in the image scene so that the eigenvalues beyond the first p eigenvalues will be considered to be noise energies. This issue is beyond the scope of this paper and was addressed in [18], [19].

In order to demonstrate these two concepts specified by (2) and (3), we consider an airborne visible/infrared imaging spectrometer (AVIRIS) image shown in Fig. 1 that was studied in [20]. It is a scene of size 200 $\times$ 200 and is a part of the Lunar Crater Volcanic Field (LCVF) in Northern Nye County, NV where five signatures of interest in these images were was demonstrated in [20]: "cinders," "rhyolite," "playa (dry lakebed)," "shade," and "vegetation." Additionally, it was also shown in [10] that there was a single two-pixel anomaly located at the top edge of the lake marked by a circle in Fig. 1.

Fig. 2(a) and (b) shows the results of the LCVF in Fig. 1 produced by (2) and the RXD respectively where Fig. 2(a) preserves the vast information of the image scene while the RXD detected a single two-pixel anomaly on the upper edge of the lake along with a small portion of vegetation on the upper right corner shown in Fig. 2(b).

It is worth noting from (1) that the RXD performs some kind of a matched filter specified by

$$\mathrm{M}_{\mathbf{d}}(\mathbf{r})=\kappa\cdot\mathbf{d}^T\mathbf{r} \tag{4}$$

where $\mathbf{d}$ is the matched signal and κ is a constant, but can be also a function of $\mathbf{r}$. The performance of (4) is completely determined by two parameters: the matched signal $\mathbf{d}$ and the scale constant κ that appears in front of the matched filter. By using (4), we can interpret the RXD as a matched filter operating on $\mathbf{r}-\mu$ with the matched signal $\mathbf{d}=(\mathbf{r}-\mu)^T\mathbf{K}_{L\times L}^{-1}$ by setting $\kappa=1$. By taking advantage of (4), two variants of the RXD—referred to as normalized RXD and denoted by $\delta_{\mathrm{NRXD}}(\mathbf{r})$, and modified RXD, denoted by $\delta_{\mathrm{MRXD}}(\mathbf{r})$—can be derived from (1) and (4) for anomaly detection as follows:

$$\begin{aligned}\delta_{\mathrm{NRXD}}(\mathbf{r})&=\frac{(\mathbf{r}-\mu)^T\mathbf{K}_{L\times L}^{-1}(\mathbf{r}-\mu)}{\|\mathbf{r}-\mu\|^2}\\&=\|\mathbf{r}-\mu\|^{-2}((\mathbf{r}-\mu))^T\mathbf{K}_{L\times L}^{-1}(\mathbf{r}-\mu)\\&=\left(\frac{\mathbf{r}-\mu}{\|\mathbf{r}-\mu\|}\right)^T\mathbf{K}_{L\times L}^{-1}\left(\frac{\mathbf{r}-\mu}{\|\mathbf{r}-\mu\|}\right)\\&=\overline{\mathbf{r}}^T\mathbf{K}_{L\times L}^{-1}\overline{\mathbf{r}}\end{aligned} \tag{5}$$

and

$$\begin{aligned}\delta_{\mathrm{MRXD}}(\mathbf{r})&=\left(\frac{\mathbf{r}-\mu}{\|\mathbf{r}-\mu\|}\right)^T\mathbf{K}_{L\times L}^{-1}(\mathbf{r}-\mu)\\&=(\overline{\mathbf{r}})^T\mathbf{K}_{L\times L}^{-1}(\mathbf{r}-\mu)\end{aligned} \tag{6}$$

where $\|\mathbf{r}-\mu\|=((\mathbf{r}-\mu)^T(\mathbf{r}-\mu))^{1/2}$ and $\overline{\mathbf{r}}=(\mathbf{r}-\mu)/\|\mathbf{r}-\mu\|$ is the normalized $\mathbf{r}-\mu$.

The $\delta_{\mathrm{NRXD}}(\mathbf{r})$ specified by (5) can be interpreted in two different ways. One is viewed as the normalized version of the RXD. Another interpretation of $\delta_{\mathrm{NRXD}}(\mathbf{r})$ is a matched filter with the matched signal $\mathbf{d}=(\mathbf{r}-\mu)^T\mathbf{K}_{L\times L}^{-1}$ as used in the RXD but a different scale constant $\kappa=\|\mathbf{r}-\mu\|^{-2}$ or a matched filter with the matched signal given by $\mathbf{d}=\|\mathbf{r}-\mu\|^{-2}(\mathbf{r}-\mu)\mathbf{K}_{L\times L}^{-1}$ with $\kappa=1$. The $\delta_{\mathrm{MRXD}}(\mathbf{r})$ specified by (6) can be interpreted as an RXD with the scale constant $\kappa=\|\mathbf{r}-\mu\|^{-1}$ or a matched filter with the matched signal $\mathbf{d}=(\overline{\mathbf{r}})^T\mathbf{K}_{L\times L}^{-1}$ and $\kappa=1$.

B. Uniform Target Detector

Another type of anomaly detector, referred to as the low-probability detector, was first derived in [6]. Its design was based on the sample correlation matrix $\mathbf{R}_{L\times L}$. If we replace $\mathbf{R}_{L\times L}$ in the

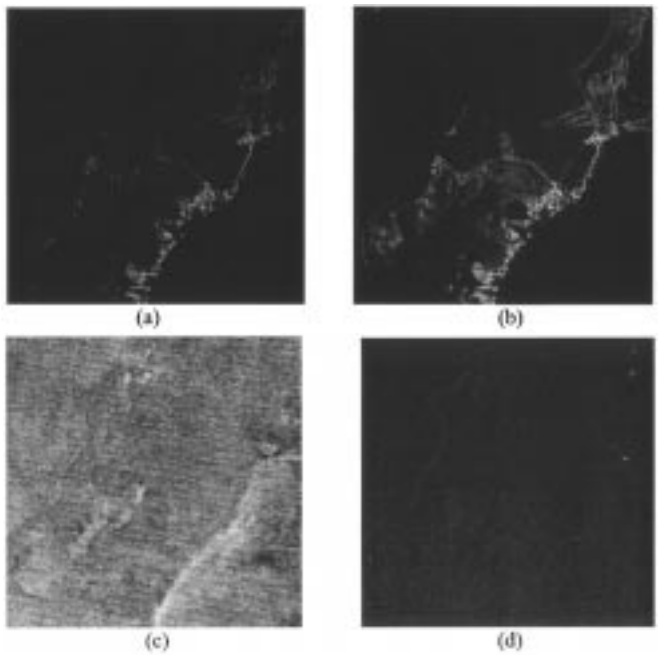

Fig. 3. (a) AVIRIS result produced by the NRXD. (b) AVIRIS result produced by the MRXD. (c) AVIRIS result produced by the UTD. (d) AVIRIS result produced by the RXD–UTD.

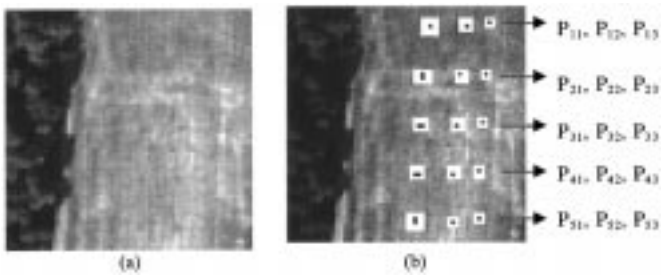

Fig. 4. (a) A 15-panel HYDICE image scene. (b) Ground truth map of Fig. 4(a).

LPD with the sample covariance matrix $\mathbf{K}_{L\times L}$, a uniform target detector (UTD) $\delta_{\mathrm{UTD}}(\mathbf{r})$ can be defined as

$$\delta_{\mathrm{UTD}}(\mathbf{r}) = (\mathbf{1}-\mu)^T\mathbf{K}_{L\times L}^{-1}(\mathbf{r}-\mu) \tag{7}$$

where

$$\mathbf{1} = \underbrace{(1,\,1,\,\ldots,\,1)}_{L}{}^T$$

is the unity vector with ones in all the components. Unlike the RXD specified by (1), which uses the image pixel vector $\mathbf{r}$ as the matched signal, the UTD uses the unity vector as its matched signal. The reason for choosing the unity vector is the following. Because there is no prior information available, the best scenario is not to introduce any information into the detector. In this case, an anomalous target is assumed to have radiance uniformly distributed over all the spectral bands. In this case, it is expected to extract background signatures which are uniformly distributed in the image. Fig. 3(c) shows the result of the LCVF in Fig. 1 produced by the UTD where it did not detect the anomaly shown in Fig. 2(b); instead, it detected most of the image background. However, if there is some partial knowledge available, the unity vector can be replaced by a certain specific vector. For example, if we are interested in short wavelengths, we can set zeroes for all visible bands while assigning ones to all the near-infrared bands.

In [4], Ashton and Schaum showed that background subtraction could enhance the RXD detection performance. This suggests that incorporating the UTD into the RXD may allow us to remove the background as was done in [4]. Such a detector, denoted by $\delta_{\mathrm{RXD-UTD}}(\mathbf{r})$, can be derived by subtracting the UTD from the RXD as follows:

$$\delta_{\mathrm{RXD-UTD}}(\mathbf{r}) = (\mathbf{r}-\mathbf{1})^T\mathbf{K}_{L\times L}^{-1}(\mathbf{r}-\mu). \tag{8}$$

Fig. 3(d) shows the result of the $\delta_{\mathrm{RXD-UTD}}(\mathbf{r})$ where no visible improvement was observed. This is because the total energy of the background in the LCVF scene is less than that of the anomaly. In this case, background subtraction using (8) did not improve much over the RXD.

As shown in Fig. 2(b), the RXD detected only one two-pixel anomaly in Fig. 2(b). One of main attributes may be due to the low 20 m spatial resolution. In order to further evaluate the performance of the RXD and the UTD, a high spatial resolution hyperspectral digital imagery collection experiment (HYDICE) scene considered in [21] was also used for experiments.

The HYDICE image shown in Fig. 4(a) has a size of 64 $\times$ 64 with 10 nm spectral resolution and 1.5 m spatial resolution. There are 15 panels located on the field, and they are arranged in a 5 $\times$ 3 matrix. The low signal/high noise bands (bands 1–3 and bands 202–210) and water vapor absorption band (bands 101–112 and bands 137–153) were removed. This results in a total of 169 bands. Fig. 4(b) shows the ground truth map of Fig. 4(a) and provides the precise spatial locations of these 15 panels. Black pixels are panel center pixels, considered to be pure pixels, and the pixels in the white masks are panel boundary pixels mixed with background pixels, considered to be mixed pixels. Each element in this matrix is a square panel and is denoted by p_{ij} with rows indexed by $i = 1, 2, \ldots, 5$ and columns indexed by $j = 1, 2, 3$. 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 $p_{1j}, p_{2j}, p_{3j}, p_{4j}, p_{5j}$ have the same size but were made from three materials and were painted by five different substances. 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 $p_{11}, p_{21}, p_{31}, p_{41}, p_{51}$ which are two-pixel panels, all the remaining panels are only one pixel in size.

Experiments similar to those conducted for the AVIRIS image were performed for the 15-panel HYDICE image scene in Fig. 4(a). Fig. 5(a)–(e) shows the results of RXD, NRXD, MRXD, UTD, and RXD–UTD respectively. In analogy with the results in Fig. 3, the RXD and the RXD–UTD performed nearly the same as they did in Figs. 2(b) and 3(d), while the NRXD, MRXD, and UTD extracted the image background, e.g., tree and grass.

The hyperspectral image experiments conducted in Figs. 3 and 5 may lead to a conclusion that the NRXD, MRXD, and

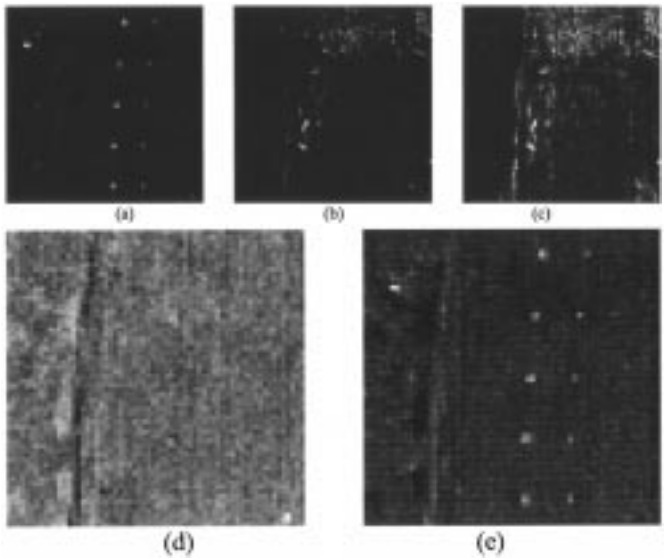

Fig. 5. (a) HYDICE result produced by the RXD. (b) HYDICE result produced by the NRXD. (c) HYDICE result produced by the MRXD. (d) HYDICE result produced by the UTD. (e) HYDICE result produced by the RXD–UTD.

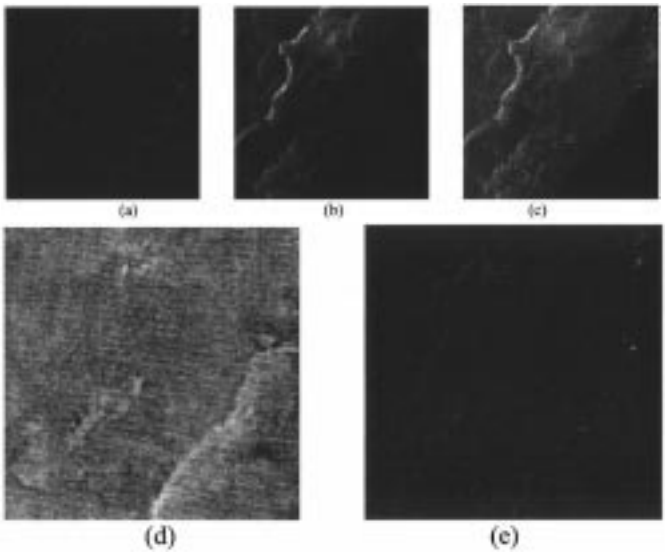

Fig. 6. (a) AVIRIS result produced by the RXD using $\mathbf{R}_{L\times L}^{-1}$, (b) AVIRIS result produced by the NRXD using $\mathbf{R}_{L\times L}^{-1}$, (c) AVIRIS result produced by the MRXD using $\mathbf{R}_{L\times L}^{-1}$, (d) AVIRIS result produced by the UTD using $\mathbf{R}_{L\times L}^{-1}$, (e) AVIRIS result produced by the RXD–UTD using $\mathbf{R}_{L\times L}^{-1}$.

UTD can only perform background extraction. As a matter of fact, it is generally not the case. This is due to the fact that the information used in the detectors was $\mathbf{K}_{L\times L}$ which had removed the information of the first-order statistics. If a target's spectral properties can be characterized only by the first-order statistics, this target may not be detected by a detector solely designed based on the second-order statistics. This is particularly true for nonstationary images such as remotely sensed imagery. In order to account for both the first-order and second-order statistics, we replace $\mathbf{K}_{L\times L}$ with $\mathbf{R}_{L\times L}$ and $\mathbf{r}-\mu$ with $\mathbf{r}$ in (1), (5)–(8), and this yields the following sample correlation matrix $\mathbf{R}$-based detectors:

$$\tilde{\delta}_{\mathrm{RXD}}(\mathbf{r}) = \mathbf{r}^T\mathbf{R}_{L\times L}^{-1}\mathbf{r} \tag{9}$$

$$\tilde{\delta}_{\mathrm{NRXD}}(\mathbf{r}) = \left(\frac{\mathbf{r}}{\|\mathbf{r}\|}\right)^T \mathbf{R}_{L\times L}^{-1}\left(\frac{\mathbf{r}}{\|\mathbf{r}\|}\right) \quad \text{with } \|\mathbf{r}\| = \sqrt{\mathbf{r}^T\mathbf{r}} \tag{10}$$

$$\tilde{\delta}_{\mathrm{MRXD}}(\mathbf{r}) = \|\mathbf{r}\|^{-1}\mathbf{r}^T\mathbf{R}_{L\times L}^{-1} \text{ with } \|\mathbf{r}\| = \sqrt{\mathbf{r}^T\mathbf{r}} \tag{11}$$

$$\tilde{\delta}_{\mathrm{UTD}}(\mathbf{r}) = \mathbf{1}^T\mathbf{R}_{L\times L}^{-1}\mathbf{r} \tag{12}$$

$$\tilde{\delta}_{\mathrm{RXD-UTD}}(\mathbf{r}) = (\mathbf{r}-\mathbf{1})^T\mathbf{R}_{L\times L}^{-1}\mathbf{r}. \tag{13}$$

Figs. 6 and 7 show the detection results of the AVIRIS and HYDICE images, respectively. As we can see from these figures, the performance of the RXD and the RXD–UTD was nearly the same, with the UTD still detecting background and noise. To our surprise, the NRXD and MRXD performed quite differently. Unlike in Fig. 3, both NRXD and MRXD detected the shade in Fig. 6(b) and (c). In addition to detecting the shade, the MRXD also detected the anomaly in Fig. 6(c). Most interestingly is Fig. 7(b) and (c). Compared to Fig. 5(b) and (c), where the NRXD and MRXD extracted only image background, Fig. 7(b) and (c) showed that the NRXD and MRXD detected panels that were also detected by the RXD. In addition, they both also extracted some tree signatures and interferers. These two experiments simply demonstrated, that when images are not stationary, the sample correlation matrix $\mathbf{R}_{L\times L}$ is more appropriate and effective than the sample covariance matrix $\mathbf{K}_{L\times L}$ to capture image characteristics.

III. Real-Time Processing of RXD

By real-time processing referred in this paper we mean that the data can be processed in a timely manner with negligible time lag. It requires only one scan and can be used for online data processing. As a matter of fact, the real-time processing presented in this section for anomaly detection can be implemented in the same manner demonstrated in [11] that the computation of the sample correlation matrix $\mathbf{R}_{L\times L}^{-1}$ can be carried out via a QR-decomposition. In this case, the proposed anomaly detector (i.e., causal RXD described below) can be implemented in parallel as a parallel processor. It processes a pixel as it is received. The RXD defined by (1) cannot be implemented in such a way. This is because the computation of $\mathbf{K}_{L\times L}^{-1}$ requires to calculate the mean of the entire image. Comparing Fig. 2(b) against Fig. 6(a) and Fig. 5(a) against Fig. 7(a), we see that there is no visible and appreciable difference between using $\delta_{\mathrm{RXD}}(\mathbf{r})$ specified by (1) and using $\tilde{\delta}_{\mathrm{RXD}}(\mathbf{r})$ specified by (9). As demonstrated also in Fig. 6(b)–(c) and Fig. 7(b)–(c), it was suggested that $\mathbf{R}_{L\times L}$ might be better than $\mathbf{K}_{L\times L}$ in characterizing spectral properties of nonnatural targets. So, in order to implement the RXD in real time, it is advantageous to use the RXD with $\mathbf{K}_{L\times L}^{-1}$ replaced by $\mathbf{R}_{L\times L}^{-1}$ and $\mathbf{r}-\mu$ by $\mathbf{r}$. The resulting RXD is the $\tilde{\delta}_{\mathrm{RXD}}(\mathbf{r})$ specified by (9) and will be referred to as causal RXD. Since the

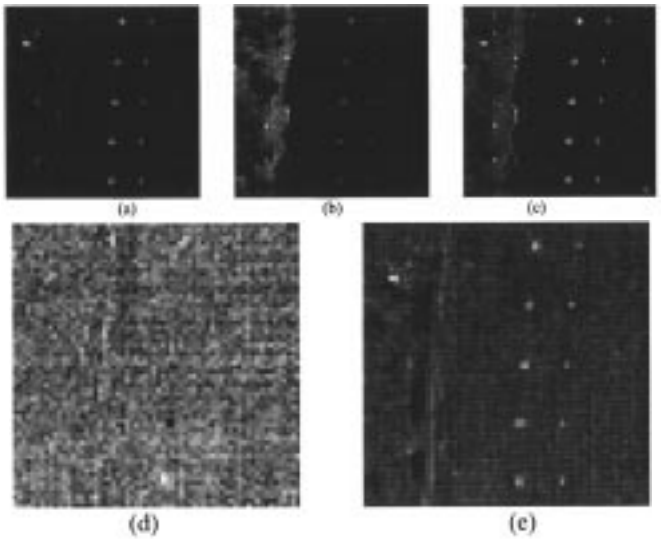

Fig. 7. (a) HYDICE result produced by the RXD using $\mathbf{R}_{L\times L}^{-1}$, (b) HYDICE result produced by the NRXD using $\mathbf{R}_{L\times L}^{-1}$, (c) HYDICE result produced by the MRXD using $\mathbf{R}_{L\times L}^{-1}$, (d) HYDICE result produced by the UTD using $\mathbf{R}_{L\times L}^{-1}$, (e) HYDICE result produced by the RXD–UTD using $\mathbf{R}_{L\times L}^{-1}$.

information used in the CRXD is updated in a causal manner, it may yield results slightly different from that produced by the RXD. As demonstrated in the following experiments, it is indeed the case. In order to demonstrate how the causal information is updated, Figs. 8 and 9 show the real-time implementation of the CRXD on the AVIRIS line-by-line as was done in [11] and HYDICE images pixel-by-pixel due to high spatial resolution, respectively. Fig. 8(a)-(c) show the results of the AVIRIS image. The vegetation was first detected in Fig. 8(a) and stayed unchanged in Fig. 8(b), then faded away after the anomaly was detected in Fig. 8(c) and finally vanished in Fig. 8(d) while the anomaly was enhanced. It is more interesting to note the results of the 15-panel HYDICE image shown in Fig. 9(a)–(g). If we compare the result of Fig. 9(h) against Fig. 7(a), the interferer shown in Fig. 7(a) faded away in Fig. 9(h). This experiment provides a good example to illustrate the difference between RXD and CRXD.

IV. Automatic Thresholding Method

In view of the fact that the images generated by RXD are generally grayscale, the detection is usually carried out by visual inspection. However, in order to avoid such human intervention and to make an objective assessment, we need to develop a computer-automated thresholding method that will automatically extract the anomalous target pixels and segment them from the background. In doing so, a threshold criterion is required to convert a grayscale image to a binary image where the detected targets can be extracted from the image background.

Recalling (1), RXD operates the form which allows one to detect anomalies in a large background by finding high peaks of gray levels in homogeneous regions. Therefore, the larger the gray values of the pixels, the more likely the pixels are anomalous pixels. This suggests that the gray-level values of anomalies should behave as outliers and fall in the right tail of the image distribution. For a given grayscale value α, we define a rejection

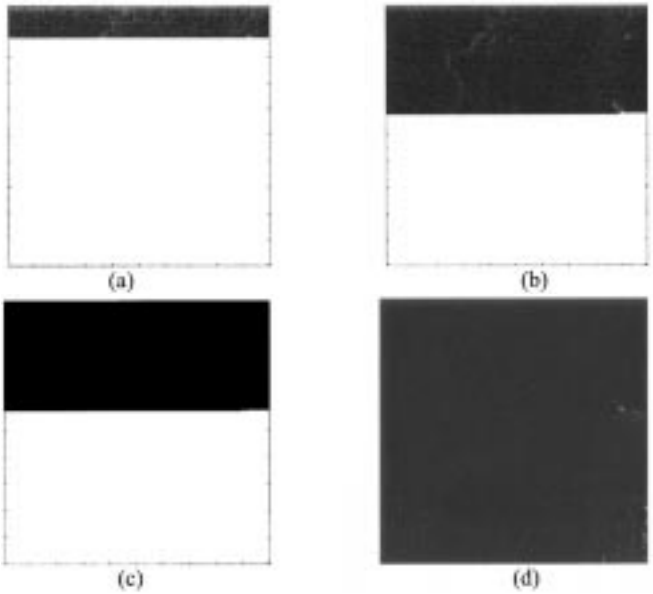

Fig. 8. (a) Detection of vegetation. (b) Before detection of anomaly. (c) Detection of anomaly. (d) Completion of the CRXD process.

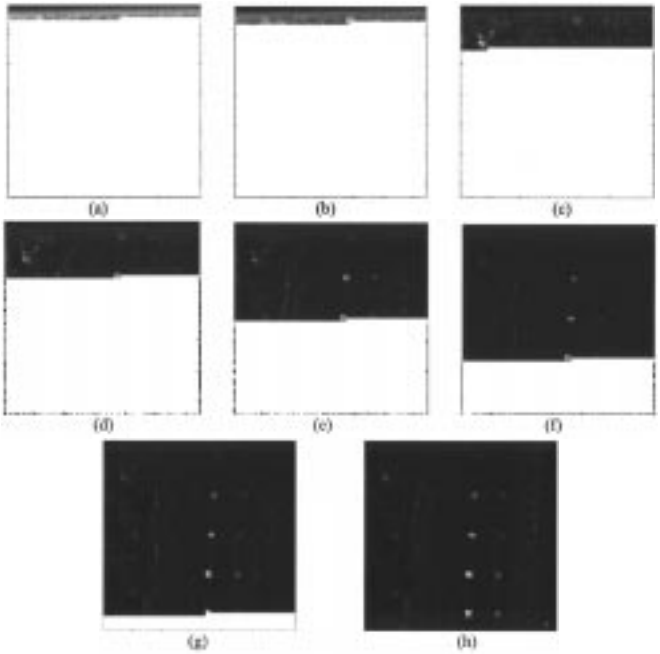

Fig. 9. (a) Nothing detected. (b) Detection of the panels in row 1. (c) Detection of the interferer. (d) Detection of the panels in row 2. (e) Detection of the panels in row 3. (f) Detection of the panels in row 4. (g) Detection of the panels in row 5. (h) Completion of the CRXD process.

region, denoted by $R(\alpha) = \{\mathbf{r}|\delta_{\mathrm{RXD}}(\mathbf{r}) < \alpha\}$, by the set made up of all the image pixels in the RXD-detected image whose gray-level values are less than α. We use the histogram of the RXD-detected image to define the rejection probability $P(\alpha)$ as

$$P(\alpha) = \Pr(R(\alpha)). \tag{14}$$

Then, a threshold α_0 for detecting the anomalies can be determined by setting a confidence coefficient γ such that $P(\alpha_0) = \gamma$. If $\delta_{\mathrm{RXD}}(\mathbf{r}) > \alpha_0$, then $\mathbf{r}$ will be detected as an anomaly. For example, Fig. 10(a) plots the gray-level values of

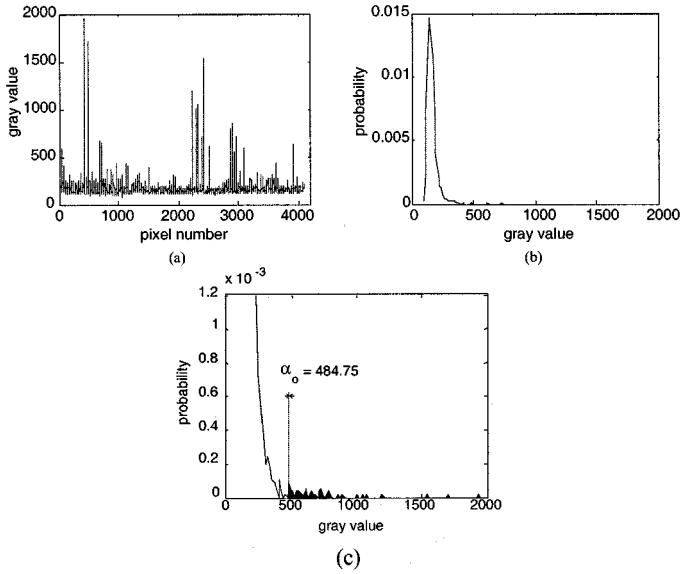

Fig. 10. (a) Plot of the gray values produced by RXD. (b) Histogram of (a). (c) An enlargement of right tail in (b).

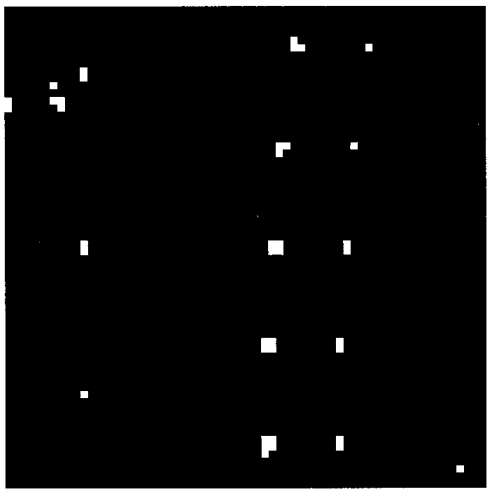

Fig. 11. A binary-thresholded image of Fig. 2(b) resulting from the automatic thresholding method with the confidence coefficient $\gamma = 0.99$.

the 15-panel HYDICE scene produced by the RXD in Fig. 5(a), and Fig. 10(b) shows its corresponding histogram. If the confidence coefficient was set by $\gamma = 0.99$, the corresponding threshold α_0 was found to be 484.75 as shown in Fig. 10(c). Using this α_0 to threshold the RXD-detected grayscale image in Fig. 5(a) resulted in a binary-thresholded image shown in Fig. 11 where there were 39 target pixels, $\{\mathbf{t}_j\}_{j=1}^{39}$, segmented by $\alpha_0 = 484.75$. It should be noted that the confidence coefficient of $\gamma = 0.99$ was selected empirically and can be adjusted. When the confidence coefficient γ is close to 1, only a few targets will be detected as anomalies. If the confidence coefficient γ is set too low ($\gamma \ll 1$), weak interferers and background signatures will be also extracted as anomalies. Fortunately, with the proposed thresholding technique, the sensitivity of selecting an appropriate threshold value can be reduced because it is based on the detected images rather than the original image.

V. Anomaly Classification

The goal of the anomaly detectors described in Section II was to find anomalies in an unknown image scene. However, there is a dilemma. RXD can detect anomalous targets but cannot differentiate them one another. The hyperspectral measures such as the spectral angle mapper (SAM) [22] and spectral information divergence [21] may be used for this purpose. However, these measures did not take advantage of the spectral correlation provided by the sample covariance matrix that is already available in the RXD. The Bhattacharyya distance defined in [23] by

$$\mathrm{B}_{ij} = \tfrac{1}{8}(\mathbf{m}_i - \mathbf{m}_j)^T((\Sigma_i + \Sigma_j)/2)^{-1}(\mathbf{m}_i - \mathbf{m}_j) + \frac{1}{2}\ln\left(\frac{|(\Sigma_i + \Sigma_j)/2|}{\sqrt{|\Sigma_i||\Sigma_j|}}\right) \quad (15)$$

seem a good candidate. It classifies two spectral signatures $\mathbf{m}_i$ and $\mathbf{m}_j$ using their respective class sample covariance matrices Σ_i and Σ_j to account for sample correlation. If we assume that $\Sigma_i = \Sigma_j = \Sigma$, then theBhattacharyya distance in (15) is reduced to the Mahalanobis distance

$$(\mathbf{m}_i - \mathbf{m}_j)^T\Sigma^{-1}(\mathbf{m}_i - \mathbf{m}_j). \quad (16)$$

As we recalled, RXD specified by (1) has the same identical detector structure $(\mathbf{x} - \mu)^T\mathbf{K}_{L\times L}^{-1}(\mathbf{x} - \mu)$ with $\mathbf{x} = \mathbf{m}_i$, $\mu = \mathbf{m}_j$, and $\mathbf{K}_{L\times L} = \Sigma$. This suggests that

$$(\mathbf{t}_i - \mathbf{t}_j)^T\mathbf{K}_{L\times L}^{-1}(\mathbf{t}_i - \mathbf{t}_j)$$

(which is referred to as the Mahalanobis distance in this paper) can be used to measure the discrepancy between two target pixels $\mathbf{t}_i$ and $\mathbf{t}_j$. So, the smaller the value is, the harder the discrimination between the two targets is.

The Mahalanobis distance specified by (16) is not the only distance measure that can be used for target discrimination. An alternative discrimination measure can be derived by the concept of a matched filter to calculate the match between two target pixels $\mathbf{t}_i$ and $\mathbf{t}_j$ after the global mean μ is removed. It is referred to as the covariance matched-filter measure (CMFM) and is given by

$$(\mathbf{t}_i - \mu)^T\mathbf{K}_{L\times L}^{-1}(\mathbf{t}_j - \mu). \quad (17)$$

Additionally, if the $\mathbf{K}_{L\times L}^{-1}$ in (16) and (17) is replaced by the inverse of the sample correlation matrix $\mathbf{R}_{L\times L}^{-1}$ two more measures can be modified from (16) and (17) and is referred to as the correlation Mahalanobis distance (RMD) given by

$$(\mathbf{t}_i - \mathbf{t}_j)^T\mathbf{R}_{L\times L}^{-1}(\mathbf{t}_i - \mathbf{t}_j) \quad (18)$$

and the correlation matched-filter measure (RMFM) given by

$$\mathbf{t}_i^T\mathbf{R}_{L\times L}^{-1}\mathbf{t}_j. \quad (19)$$

It is worth paying particular attention to (19). It is very close to the form of the orthogonal subspace projection (OSP) classifier (derived in [20]) and the form of the constrained energy minimization (CEM) detector (derived in [10]) where $\mathbf{t}_i$ is designated as the desired target signature and where $\mathbf{t}_j$ can be viewed as the image pixel vector $\mathbf{r}$ that was chosen from $\{\mathbf{t}_j\}_{j=1}^{39}$ in

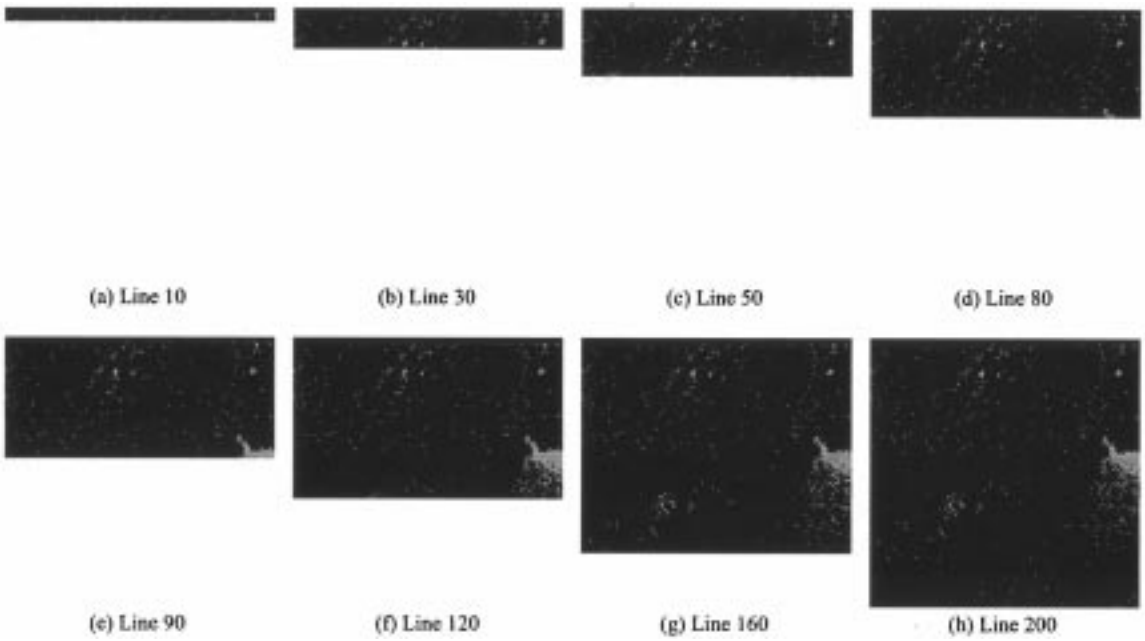

Fig. 12. Online target discrimination results of the AVIRIS scene resulting from the CRXD using RMFM.

Fig. 11 to match the desired target $\mathbf{t}_i$. So, the larger the matched value is, the more likely the two targets belong to the same class. It basically performs a similar task as a matched filter does. As a result, the measures specified by (17) and (19) are referred to as matched-filter-based measures.

The above three discrimination measures were proposed in [24]. They have been shown to perform very similarly in comparison with the Mahalanobis distance specified by (16). A detailed study of these measures along with the Mahalanobis distance proposed in this section can be found in [24], [25].

Using the RMFM as a target differentiation measure,we can implement an anomaly classifier as follows.

Anomaly Classification Algorithm

1) Apply an anomaly detector and the automatic thresholding method in Section IV to detect potential anomalous targets, denoted by $\{\mathbf{t}_j\}_{j=1}^{N}$, and segment these targets from the background.
2) Use the RMFM to group the targets detected in step 1) into separate target classes, denoted by $\{\omega_j\}_{j=1}^{p}$.
3) Find the mean of each target class, denoted by $\{\overline{\mathbf{t}}_j\}_{j=1}^{p}$.
4) Classify targets in the image using $\{\overline{\mathbf{t}}_j\}_{j=1}^{p}$ as the desired target information. The classifier used in this step can be selected at discretion, such as OSP in [20], LCMV in [11], [12]. It is worth noting that, in this step, the classification is done based on the target information provided in step 3). As a result, targets that match $\{\overline{\mathbf{t}}_j\}_{j=1}^{p}$ will be detected and also classified. These may also include those which were missed by the anomaly detector.

It should be noted that in order to implement step 2) it also requires a threshold value for the RMFM. Unfortunately, this knowledge is not available for anomaly classification. Since the number of detected targets $\{\mathbf{t}_j\}_{j=1}^{N}$, N, generated in step 1) is usually small, it is not sufficiently large to provide reliable statistics for (14). Consequently, the automatic thresholding method used for target detection may not be directly applicable to target discrimination. However, it can be modified as follows. Instead of plotting the gray-level values of detected target pixels as was done in Fig. 10(b), we plot the histogram of the RMFM values of $\{\mathbf{t}_j\}_{j=1}^{N}$ using (19) with each target designated as $\mathbf{t}_i$ for the desired target signature. Then, the first zero of the histogram is selected as the threshold value to determine if a target pixel and $\mathbf{t}_i$ should be clustered into the same class. As an example, if we want to cluster target $\mathbf{t}_1$, we let $\mathbf{t}_1 = \mathbf{t}_i$ and calculate the RMFM values of $\{\mathbf{t}_j\}_{j=1}^{N}$, $\mathbf{t}_1^T\mathbf{R}_{L\times L}^{-1}\mathbf{t}_j$ for all $1 \le j \le N$. Then we plot the histogram of the values of $\{\mathbf{t}_1^T\mathbf{R}_{L\times L}^{-1}\mathbf{t}_j\}_{j=1}^{N}$. Then, the first zero of the histogram is chosen as the threshold value. Those target pixels whose RMFM values are above or equal to the selected threshold value are clustered with $\mathbf{t}_1$ into the same class; those whose RMFM values are below the selected threshold value are clustered into a different cluster from that of $\mathbf{t}_1$. In case that the histogram does not have a zero, then all the target pixels $\{\mathbf{t}_j\}_{j=1}^{N}$ are clustered into one class. A similar approach was also investigated in [14] and has shown success in unsupervised subpixel target detection.

VI. Experiments

In order to demonstrate the anomaly classification, we will implement the CRXD–LCMV hybrid classifier using the RMFM specified by (19) as the target discrimination measure along with the automatic thresholding described in Section V. Since there was no appreciable difference between the Mahalanobis distance in (18) and RMFM as shown in [24], [25], we chose the RMFM over the Mahalanobis distance due to the fact that the RMFM can be easily implemented by a QR-decomposition. The whole processing consists of three stages: the implementation of CRXD in step 1), a target-clustering process using the RMFM, and a real-time processing of the LCMV classification developed in [11]. Like the RXD, which can be implemented in real time, the CRXD–LCMV anomaly

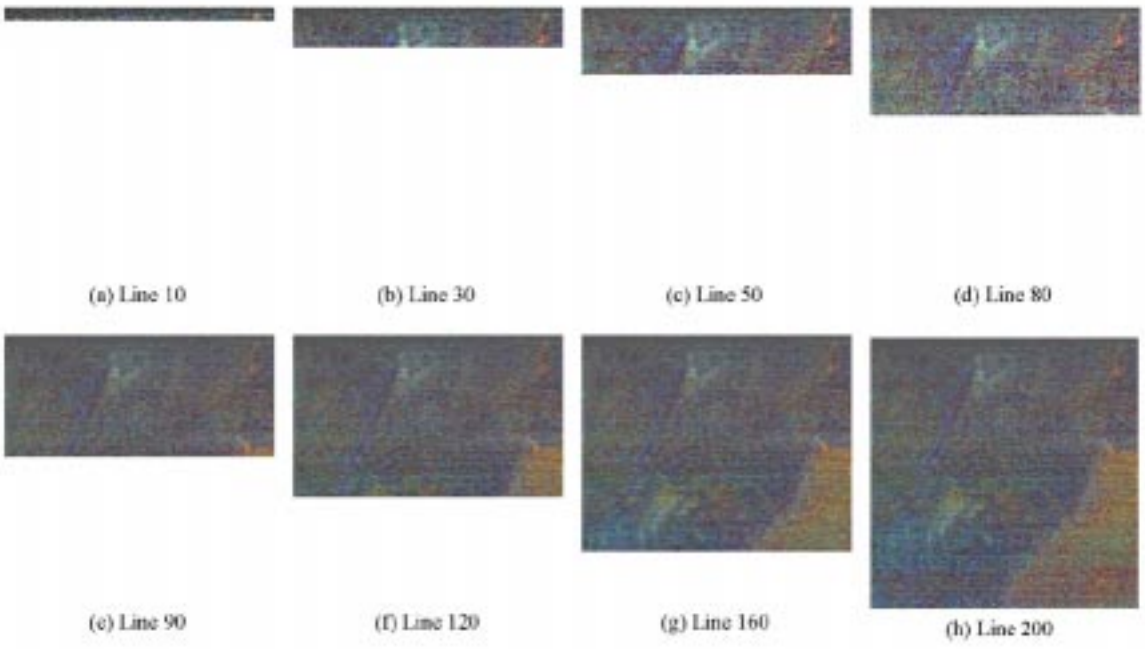

Fig. 13. Online target classification results of the AVIRIS scene resulting from the CRXD using RMFM in conjunction with the LCMV classifier.

detection and classification can be also implemented in online processing with negligible time lag. Since it requires a two-pass process, a time delay is inevitable. Nevertheless it can be minimized. The first process works in the same manner as does the CRXD does to detect anomalies in real time. The second pass of the discrimination and classification process takes place only a few-lines delay (or a set-of-pixels delay if it is implemented on a pixel-by-pixel basis) right after the CRXD took place. The algorithm takes advantage of a few-line time lag to generate sufficient target information for follow-up classification. Since the classifier used in the anomaly classification is the LCMV classifier that can be also implemented in real time as described in [11], the two-pass anomaly detection and classification can be essentially executed in a timely manner with a few-lines delay between the two passes. In order to distinguish from the real-time processing of the CRXD, such real-time processing with a few-lines time delay is referred to as an online process.

Fig. 12 shows the target discrimination result produced by the CRXD using the RMFM as the target discrimination measure where the line number underneath each image indicates that the result was obtained by using the causal information up to that particular line. As we can see, the results in Fig. 12 are completely different from that in Fig. 2(b). Other than the two-pixel anomaly detected in Fig. 2(b), many additional anomalies were picked up and discriminated by the RMFM. This is because the process of the target discrimination took place immediately after the detection process. As a consequence, a weak anomaly of one type would be dominated by a strong anomaly of another type. This explains why there was only one two-pixel anomaly detected by the RXD in Fig. 2(b), because its abundance dominated other anomalies which may belong to different types of targets, but had relatively low abundances. Fig. 13 further shows the target classification result produced by the complete process of anomaly classification which implements the CRXD using RMFM in conjunction with the LCMV classifier. This image actually classified the cinders, playa, rhyolite, shade, vegetation, and the two-pixel anomaly.

The 15-panel HYDICE scene provides even more interesting results than the AVIRIS image after a target discrimination measure was incorporated into the classification process. According to Fig. 11, the 39 target pixels were detected: $\{\mathbf{t}_j\}_{j=1}^{39}$ which are labeled by the order that they were detected in a top-to-bottom and left-to-right manner. To further classify these 39 detected targets into separate target classes, we used the RMFM specified by (19) to plot their values as shown in Fig. 14(a)–(k) where $\mathbf{t}_i$ was designated as a seed pixel, with $\mathbf{t}_j$ running through all the 39 target pixels. For example, in Fig. 14(a), $\mathbf{t}_3$ was used as a seed pixel for $\mathbf{t}_i$, and $\mathbf{t}_j$ was chosen to be any target $\mathbf{t}$ from $\{\mathbf{t}_j\}_{j=1}^{39}$. As shown in the plot, the peak values of $\mathbf{t}_i^T\mathbf{R}_{L\times L}^{-1}\mathbf{t}_j$ produced by the RMFM were $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_4$. They were clustered together. So, $\mathbf{t}_3$ and $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_4$ were considered to be in the same class. It is clearly shown in Fig. 14(a)–(j) that the 39 detected targets were further clustered into ten target classes, denoted by $\{\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \mathbf{t}_4\}$, $\{\mathbf{t}_5, \mathbf{t}_6, \mathbf{t}_{17}, \mathbf{t}_{21}\}$, $\{\mathbf{t}_7, \mathbf{t}_9, \mathbf{t}_{10}, \mathbf{t}_{12}\}$, $\{\mathbf{t}_8, \mathbf{t}_{11}\}$, $\{\mathbf{t}_{13}, \mathbf{t}_{14}, \mathbf{t}_{15}, \mathbf{t}_{16}\}$, $\{\mathbf{t}_{18}, \mathbf{t}_{19}, \mathbf{t}_{20}, \mathbf{t}_{22}, \mathbf{t}_{23}, \mathbf{t}_{24}\}$, $\{\mathbf{t}_{25}, \mathbf{t}_{26}, \mathbf{t}_{27}, \mathbf{t}_{28}, \mathbf{t}_{29}, \mathbf{t}_{30}\}$, $\{\mathbf{t}_{31}\}$, $\{\mathbf{t}_{32}, \mathbf{t}_{33}, \mathbf{t}_{34}, \mathbf{t}_{35}, \mathbf{t}_{36}, \mathbf{t}_{37}, \mathbf{t}_{38}\}$, and $\{\mathbf{t}_{39}\}$ where 15 panels are classified in the images in Fig. 14(a) and Fig. (e)–(i). The results of Fig. 14 were further used for target discrimination. Fig. 15 shows the results of target discrimination produced by the CRXD using the RMFM. Finally, the means of these 10 target classes were calculated and used as desired target signatures for the LCMV classifier. The classification of these ten target classes is shown in Fig. 16 with different colors.

To verify our classification results, we need to find what these targets were. According to the ground truth map provided by

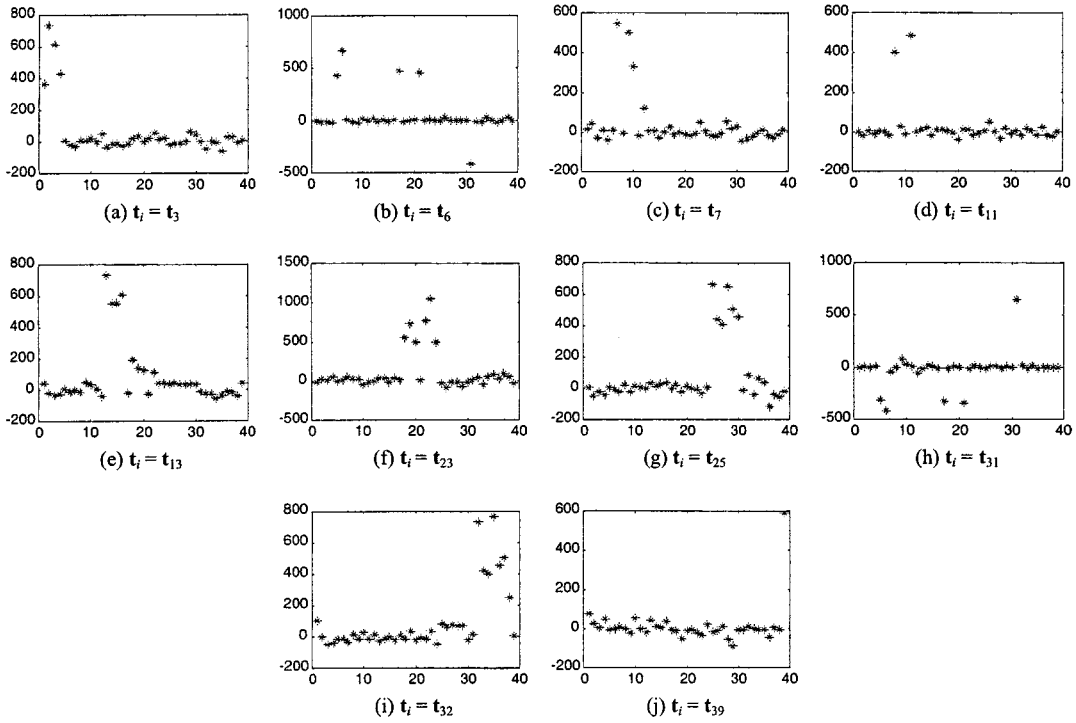

Fig. 14. Plot of the values of RMFM for 39 target pixels detected by the RXD in Fig. 2(b).

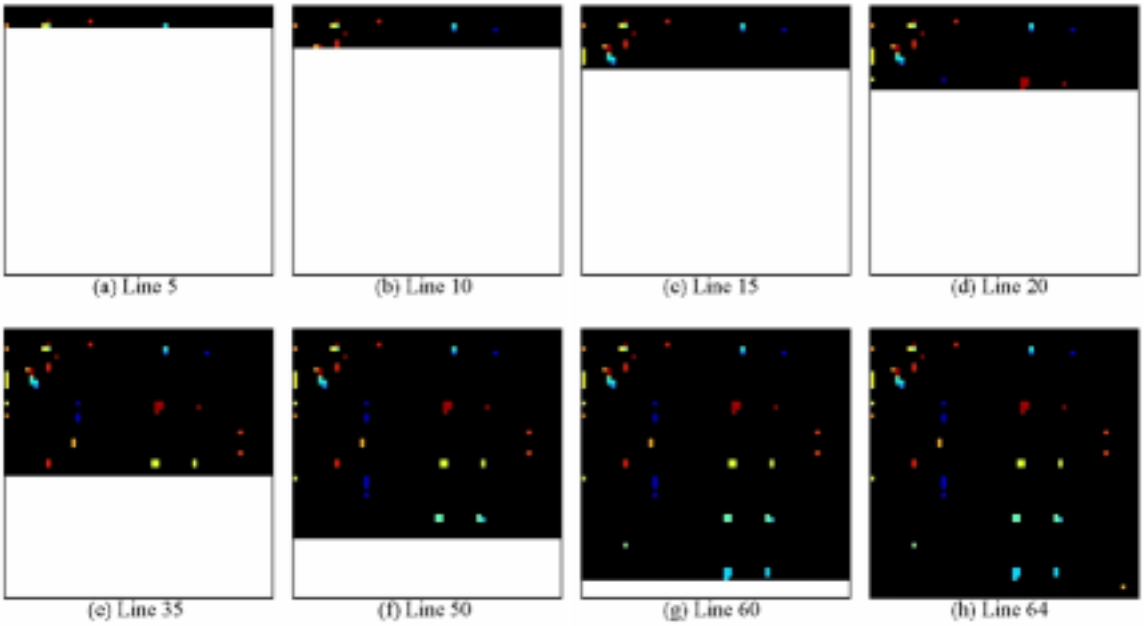

Fig. 15. Online target discrimination results of the 15-panel HYDICE scene resulting from the CRXD using RMFM.

Fig. 4(b), there are two B pixels in the panels of size 3 m $\times$ 3 m, $p_{21}, p_{31}, p_{41}, p_{51}$ except for p_{11} which has only one B pixel, and one B pixel in all panels of size 2 m $\times$ 2 m and 1 m $\times$ 1 m. Based on this information the 39 detected target pixels can be identified into ten classes as follows.

1) Panels in row 1: $\mathbf{t}_2 = B$ pixel of p_{11}; $\mathbf{t}_1$, $\mathbf{t}_3 = W$ pixels of p_{11}; $\mathbf{t}_4 = B$ pixel of p_{12}
2) $\{\mathbf{t}_5, \mathbf{t}_6, \mathbf{t}_{17}, \mathbf{t}_{21}\}$ = anomalies located in the forest
3) $\{\mathbf{t}_7, \mathbf{t}_9, \mathbf{t}_{10}, \mathbf{t}_{12}\}$ = anomalies located in the forest
4) $\{\mathbf{t}_8, \mathbf{t}_{11}\}$ = anomalies located in the forest
5) Panels in row 2: $\mathbf{t}_{13}$, $\mathbf{t}_{16} = B$ pixels of p_{21}; $\mathbf{t}_{14} = W$ pixel of p_{21}; $\mathbf{t}_{15} = B$ pixel of p_{22}
6) Panels in row 3: $\mathbf{t}_{22}, \mathbf{t}_{23} = B$ pixels of p_{31}; $\mathbf{t}_{18}, \mathbf{t}_{19} = W$ pixels of p_{31}; $\mathbf{t}_{24} = B$ pixel of p_{32}; $\mathbf{t}_{20} = W$ pixel of p_{32}
7) Panels in row 4: $\mathbf{t}_{28}, \mathbf{t}_{29} = B$ pixels of p_{41}; $\mathbf{t}_{25}, \mathbf{t}_{26} = W$ pixels of p_{41}; $\mathbf{t}_{30} = B$ pixel of p_{42}; $\mathbf{t}_{27} = W$ pixel of p_{42}

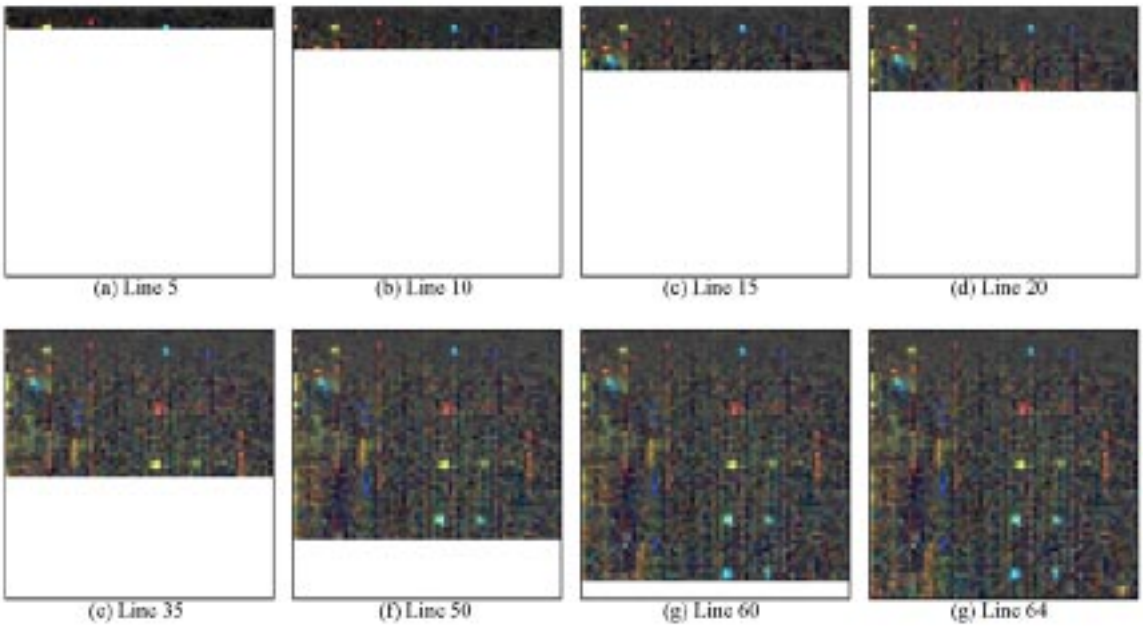

Fig. 16. Online target classification results of the 15-panel HYDICE scene resulting from the CRXD using RMFM in conjunction with the LCMV classifier.

8) $\mathbf{t}_{31}$ = anomaly in the forest
9) Panels in row 5: $\mathbf{t}_{32}, \mathbf{t}_{35}$; = B pixels of p_{51}; $\mathbf{t}_{27}$, $\mathbf{t}_{33}, \mathbf{t}_{36}, \mathbf{t}_{38} = W$ pixels of p_{51}; $\mathbf{t}_{37} = B$ pixel of p_{52}; $\mathbf{t}_{34} = W$ pixel of p_{52}
10) $\mathbf{t}_{39}$ = anomaly in grass pixel.

These ten classes coincide with those classified by Fig. 14. It should be also noted that the RXD failed to detect all the five 1 m $\times$ 1 m panels p_{i3}, $1 \leq i \leq 5$, due to small amounts of abundance contained in these five-panel pixels and their size smaller than the 1.5 m spatial resolution. This was also due to the fact that the RXD does not require any prior knowledge. As noted above, these missing panels can be extracted if we used the Mahalanobis distance target discrimination measure to generate the target information for target classification. As shown in Fig. 14(a), (f)–(h), and (j), it was indeed the case. This advantage results from the use of the target discrimination measure.

As a final comment, in order for the four target discrimination measures specified by (17)–(19) to work effectively, the number of samples that form the sample covariance matrix must be sufficiently large to avoid the ill-rank singularity problem resulting from calculating the inverse of the sample covariance matrix. In this case, the number of data samples must be greater than or equal to the number of spectral bands. Since the size of the remotely sensed imagery is generally larger than the total number of spectral bands used for acquisition, this requirement is generally satisfied. For the real-time implementation of the CRXD described in Sections III and VI which uses the sample correlation matrix, the real-time processing does not take place until the number of samples is greater than the number spectral bands. In other words, there is no real-time processing in the first few lines before we collect enough data samples to form the nonsingular sample correlation matrix.

VII. Conclusions

Anomaly detection has been widely studied in literature. Unfortunately, very little work on anomaly classification has been reported. In this paper, both anomaly detection and classification are investigated where four contributions are made. First, two types of anomaly detectors—RXD and LPD, along with their variants—were developed. In particular, a causal RXD was presented for real-time processing. Since an anomaly detector does not necessarily classify or discriminate the anomalies it detects, a second contribution is to design Mahalanobis distance-based and matched-filter-based target discrimination measures for target discrimination. This is then followed by a third contribution: that an automatic thresholding method was proposed for target detection after target discrimination. Finally, it is concluded by a fourth contribution where an LCMV classifier was employed in conjunction with the thresholding method so as to achieve anomaly classification. In summary, anomaly classification is made up of these four contributions. Its entire process can be implemented by four stages: anomaly detection, target discrimination, thresholding, and classification, each of which can take place nearly simultaneously with negligible time delay. This was demonstrated by experiments.

Acknowledgment

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

References

[1] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1760–1770, Oct. 1990.
[2] X. Yu, I. S. Reed, and A. D. Stocker, "Comparative performance analysis of adaptive multispectral detectors," *IEEE Trans. Signal Processing*, vol. 41, pp. 2639–2656, Aug. 1993.

[3] X. Yu, L. E. Hoff, I. S. Reed, A. M. Chen, and L. B. Stotts, "Automatic target detection and recognition in multispectral imagery: A unified ML detection and estimation approach," *IEEE Trans. Image Processing*, vol. 6, pp. 143–156, Jan. 1997.
[4] E. A. Ashton and A. Schaum, "Algorithms for the detection of sub-pixel targets in multispectral imagery," *Photogram. Eng. Remote Sens.*, pp. 723–731, July 1998.
[5] C. M. Stellman, G. G. Hazel, F. Bucholtz, J. V. Michalowicz, A. Stocker, and W. Scaaf, "Real-time hyperspectral detection and cuing," *Opt. Eng.*, vol. 39, pp. 1928–1935, 2000.
[6] J. C. Harsanyi, "Detection and classification of subpixel spectral signatures in hyperspectral image sequences," Ph.D. dissertation, Dept. Elect. Eng., Univ. Maryland-Baltimore County, Baltimore, MD, 1993.
[7] J. C. Harsanyi, W. Farrand, and C.-I. Chang, "Detection of subpixel spectral signatures in hyperspectral image sequences," in *Proc. Amer. Soc. Photogram. Remote Sens.*, Reno, NV, 1994, pp. 236–247.
[8] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[9] W. Farrand and J. C. Harsanyi, "Mapping the distribution of mine tailing in the coeur d'Alene river valley, Idaho, through the use of constrained energy minimization technique," *Remote Sens. Environ.*, vol. 59, pp. 64–76, 1997.
[10] C.-I Chang and D. Heinz, "Constrained subpixel target detection for remotely sensed imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1144–1159, May 2000.
[11] C.-I Chang, H. Ren, and S. S. Chiang, "Real-time processing algorithms for target detection and classification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 760–768, Apr. 2001.
[12] G. H. Golub and G. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: John Hopkins Univ. Press, 1989.
[13] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.
[14] S. S. Chiang, C.-I Chang, and I. W. Ginsberg, "Unsupervised subpixel target detection for hyperspectral images using projection pursuit," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 1380–1391, July 2001.
[15] C.-I. Chang and H. Ren, "Linearly constrained minimum variance beamforming for target detection and classification in hyperspectral imagery," in *Proc. IEEE IGARSS*, Hamburg, Germany, June 28–July 2, 1999, pp. 1241–1243.
[16] H. V. Poor, *An Introduction to Detection and Estimation Theory*, 2nd ed. New York: Springer-Verlag, 1994, pp. 58–59.
[17] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.
[18] J. Harsanyi, W. Farrand, and C.-I Chang, "Determining the number and identity of spectral endmembers: An integrated approach using Neyman–Pearson eigen-thresholding and iterative constrained RMS error minimization," in *Proc. Ninth Thematic Conf. Geologic Remote Sens.*, Feb. 1993.
[19] C.-I Chang and Q. Du, "A noise subspace projection approach to determination of intrinsic dimensionality for hyperspectral imagery," in *EOS/SPIE Symp. Remote Sens., Conf. Image Signal Process. Remote Sens. V*, vol. SPIE 3871, Florence, Italy, Sept. 20–24, 1999, pp. 34–44.
[20] 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.
[21] 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.
[22] R. A. Schwengerdt, *Renote Sensing: Models and Methods for Image Processing*, 2nd ed. Orlando, FL: Academic, 1997.
[23] K. Fugunaga, *Statistical Pattern Recognition*, 2nd ed. Orlando, FL: Academic, 1990.
[24] S.-S. Chiang and C.-I Chang, "Discrimination measures for target classification," in *Proc. IEEE IGARSS*, Sydney, Australia, July 24–28, 2001.
[25] S.-S. Chiang, "Automatic target detection and classification in hyperspectral imagery," Ph.D. dissertation, Dept. Comput. Sci. Elect. Eng., Univ. Maryland-Baltimore County, Baltimore, MD, May 2001, .

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 is currently a Professor in the Department of Computer Science and Electrical Engineering at the University of Maryland-Baltimore County (UMBC) since 2001. He has been with UMBC as a Visiting Assistant Professor from January 1987 to August 1987, Assistant Professor from 1987 to 1993, and as an Associate Professor in the Department of Computer Science and Electrical Engineering from 1993 to 2001. He was a Visiting Research Specialist in the Institute of Information Engineering at the National Cheng Kung University, Tainan, Taiwan, from 1994 to1995. He has patents on automatic pattern recognition and detection of microcalcifications and several pending patents on image processing techniques for hyperspectral imaging. He is currently on an National Research Council (NRC) Senior Research Associateship sponsored by the U.S. Army Edgewood Chemical Biological Center (ECBC). His research interests include automatic target recognition, multispectral/hyperspectral image processing, medical imaging, information theory and coding, signal detection and estimation, neural networks.

Dr. Chang is a Fellow of SPIE and a member of INNS, Phi Kappa Phi, and Eta Kappa Nu. He is currently Associate Editor of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and also on the editorial board of *Journal of High Speed Networks*, of which he was the Guest Editor of a special issue on Telemedicine and Applications.

Shao-Shan Chiang (S'98–M'01) received the B.S. degree in mathematical science from National Chengchi University, Taipei, Taiwan, R.O.C., in 1988, M.S. degree in mathematics from National Taiwan University, Taipei, in 1990, and M.S. and Ph.D. degrees, both in electrical engineering, from University of Maryland-Baltimore County (UMBC), Baltimore, in 1999 and 2001 respectively.

From 1992 to 1996, he was a Software Engineer in Taipei. He was a Research Assistant in the Anesthesiology Research Laboratory, UMBC, from 1996 to 1998, and was a Research Assistant in the Remote Sensing Signal and Image Processing Laboratory at UMBC from 1999 to 2001. Since 2001, he has been an Assistant Professor with the Department of Electrical Engineering, Lunghwa University of Science and Technology, Taoyuan, Taiwan. His research interests include signal processing, communications, pattern recognition, and remote sensing image processing.

Dr. Chiang is a member of SPIE.