Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Probabilistic anomaly detector for remotely sensed hyperspectral data

Lianru Gao
Qiandong Guo
Antonio Plaza
Jun Li
Bing Zhang

SPIE.

Probabilistic anomaly detector for remotely sensed hyperspectral data

Lianru Gao,[a,*] Qiandong Guo,[b] Antonio Plaza,[c] Jun Li,[d] and Bing Zhang[a]

[a]Chinese Academy of Sciences, Institute of Remote Sensing and Digital Earth, Key Laboratory of Digital Earth Science, No. 9 Dengzhuang South Road, Beijing 100094, China

[b]University of South Florida, School of Geosciences, 4202 East Fowler Avenue, Tampa, Florida 33620, United States

[c]University of Extremadura, Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Avda. de la Universidad S/N, Cáceres 10071, Spain

[d]Sun Yat-sen University, School of Geography and Planning, No. 135 Xingang Xi Road, Guangzhou 510275, China

Abstract. Anomaly detection is an important technique for remotely sensed hyperspectral data exploitation. In the last decades, several algorithms have been developed for detecting anomalies in hyperspectral images. The Reed-Xiaoli detector (RXD) is one of the most widely used approaches for this purpose. Since the RXD assumes that the distribution of the background is Gaussian, it generally suffers from a high false alarm rate. In order to address this issue, we introduce an unsupervised probabilistic anomaly detector (PAD) based on estimating the difference between the probabilities of the anomalies and the background. The proposed PAD takes advantage of the results provided by the RXD to estimate statistical information for the targets and background, respectively, and then uses an automatic strategy to find the most suitable threshold for the separation of targets from the background. The proposed technique is validated using a synthetic data set and two real hyperspectral data sets with ground-truth information. Our experimental results indicate that the proposed method achieves good detection ratios with adequate computational complexity as compared with other widely used anomaly detectors.
 [DOI: 10.1117/1.JRS.8.083538]

Keywords: hyperspectral imaging; anomaly detection; multivariate normal distribution; Reed-Xiaoli detector; automatic thresholding.

Paper 14170 received Mar. 24, 2014; revised manuscript received Sep. 12, 2014; accepted for publication Sep. 25, 2014; published online Nov. 3, 2014.

1 Introduction

Target detection is concerned with the identification of targets with interest that appear with low probabilities in a given scene.[1] Based on the availability of *a priori* information, we can roughly categorize available techniques into two main classes: (1) unsupervised techniques[2,3] and (2) supervised techniques, among which we can distinguish between signature matching-based techniques, in which the target of interest is defined by a reference spectrum,[4,5] and subspace matching-based techniques, in which the target of interest is represented by a set of basis vectors that account for the target signal variation.[6,7]

A representative of signature matching-based methods is constrained energy minimization,[4,5] a powerful method to address signature matching-based detection problems. Another example is the adaptive cosine estimate,[8] an adaptable filter specifically designed to preserve the target signal based on a generalized likelihood ratio test. The target-constrained interference minimized filter can be used for situations in which several target spectra are known *a priori*.[6,7] This algorithm designs a finite impulse response filter w to minimize the output energy while retaining target information.

*Address all correspondence to: Lianru Gao, E-mail: gaolr@radi.ac.cn

0091-3286/2014/$25.00 © 2014 SPIE

As opposed to supervised methods, the Reed-Xiaoli detector (RXD)[9–11] is a widely used method for unsupervised target detection. It relies on two main assumptions: (1) anomalies are the pixels whose signatures are spectrally distinct from their surroundings and (2) anomalies occur with low probabilities in the scene.[3] Based on these characteristics, the RXD adopts a multivariate normal distribution model to represent the hyperspectral data and treats anomalies as outliers in probabilistic sense. This model enables us to calculate the probability of a pixel under test (PUT) as belonging to the background, with the resulting value providing an idea about the degree of anomaly at the pixel. This algorithm is also a constant false-alarm rate (FAR) adaptive anomaly detector.[12] However, the algorithm may suffer from a high FAR. This is due to several reasons. First and foremost, the assumption that the background follows a multivariate normal distribution is sometimes impractical since in many scenarios, the scene contains a variety of objects that are too complex.[13] This assumption increases the difficulty in separating targets from background due to shadows, illumination effects, atmospheric interferers, or spectral similarity between different classes.[14] In addition, some anomalous targets may be subpixel in nature,[4] and these targets tend to be less distinctive from the background in spectral space, which complicates their detection. Last but not least, hyperspectral images exhibit spectral correlation which leads to ill-conditioned covariance matrices, and the calculation of the inverse may suffer from large accumulated rounding errors that ultimately result in inaccuracies in the RXD algorithm. All of these factors complicate anomaly detection and increase the possibility of detecting background and outlier pixels as targets, thus increasing false alarms. Since these false alarms negatively impact the efficiency of detectors, how to reduce them becomes a critical issue in anomaly detection.

Several strategies have been proposed to improve the performance of the RXD. Some available approaches have focused on increasing the separability between targets and background through feature selection or feature extraction, including subspace RXD (SSRXD),[2,15,16] kernel RXD (KRXD),[12,17–19] and random projection RXD.[20] SSRXD is suitable under the assumption that the spectral characteristics of the target are well represented in a subspace. KRXD uses a strategy based on mapping the original data to a high-dimensional space by means of a kernel function. Random projection RXD is deployed on orthogonally projected data.[20] Another strategy for improving the RXD has been focused on exploiting *a priori* information directly obtained from the scene. This is the case of the segment-RXD (SRXD),[21] which uses clustering techniques to calculate the covariance matrix and the mean vector, with the ultimate goal of obtaining refined information for the background. Other methods seek to refine the background by removing anomalies, with the ultimate goal of obtaining a more accurate estimation of the background. Examples include the locally adaptable iterative RX (LAIRX) in Ref. 22 and the local-RXD (LRXD) in Ref. 21. LAIRX allows the RXD to iteratively refine background estimation using first- and second-order statistics. In turn, LRXD estimates the statistics for the background by using a sliding window around each PUT. The window size is normally determined by the size of the anomalies of interest. Finally, other algorithms for improvement include the blocked adaptive computationally efficient outlier nominators (BACON)[23] or the random selection-based anomaly detector (RSAD),[24] which both use iterative approaches in order to refine background samples and obtain a more accurate estimation of the background information.

In this paper, we develop a new probabilistic anomaly detection (PAD) algorithm which characterizes the target and background by segmenting the image into a target set and a background set. These sets are then used as *a priori* information in order to calculate the probabilities of a given PUT to be declared as target (anomaly) or background. This additional information is used to overcome the lack of knowledge about the target and background signals, thus making the method fully automatic. Once the targets signals are removed from the background by means of the proposed PAD, the background is much closer to a multivariate normal distribution. This property allows us to reduce the FAR without decreasing the probability of detection when applying the RXD algorithm.

The reminder of this paper is organized as follows. Section 2 provides an overview of the classic RXD algorithm and its variations. Section 3 introduces the proposed PAD method for suppressing the false alarms. In Sec. 4, we evaluate the proposed method by using a synthetic data set. Section 5 presents extensive experiments with real hyperspectral data sets and

comparisons with other methods. Section 6 draws some conclusions and provides hints at plausible future research lines.

2 Multivariate Normal Model and the RXD Algorithm

Anomaly detection is one of the problems encountered without prior target information.[1,25] Thus, the target signal is unknown which is a main barrier for the detection. However, two features of targets can help targets be detected. The spectra of targets are supposed to be significantly different from the spectra of the background. Besides, targets occur in the image with relatively low probabilities.[2] The RXD algorithm takes advantage of these two features to detect the anomalies. This method is based on the multivariate normal model, which can be simply described as follows. Let H_1 denote the target signal and H_0 denote the background signal. With these notations in mind, the detection problem can be written as

$$H_0: \mathbf{x} = \mathbf{b}, \tag{1}$$

$$H_1: \mathbf{x} = \mathbf{s} + \mathbf{b}, \tag{2}$$

where $\mathbf{x}$ is a sample pixel vector, $\mathbf{s}$ is the target signal, and $\mathbf{b}$ is the background clutter for which we assume a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, i.e., $\mathbf{b} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Therefore, we have $\mathbf{x}|H_0 \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbf{x}|H_1 \sim N(\boldsymbol{\mu} + \mathbf{s}, \boldsymbol{\Sigma})$. With these definitions in mind, we can obtain the probability density function of the background and target, respectively, as

$$p(\mathbf{x}|H_0) = \frac{1}{(2\pi)^{K/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \tag{3}$$

and

$$p(\mathbf{x}|H_1) = \frac{1}{(2\pi)^{K/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}-\mathbf{s})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}-\mathbf{s})}, \tag{4}$$

where K is the number of bands of the original hyperspectral image. Since an anomaly $\mathbf{x}_s$ is expected to be significantly different from the background, $p(\mathbf{x}_s|H_0)$ should be very small for an anomalous pixel. Therefore, for a given background, as $1/[(2\pi)^{K/2}|\Sigma|^{1/2}]$ is fixed, $(\mathbf{x}_s - \boldsymbol{\mu})^T\Sigma^{-1}(\mathbf{x}_s - \boldsymbol{\mu})$ should be larger for an anomalous pixel than for a background pixel. Based on this observation, RXD implements a filter specified by

$$D_{\mathrm{RXD}}(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}). \tag{5}$$

The result of the RX detector in Eq. (5) can be thresholded to perform anomaly detection. Such a procedure is actually equivalent to thresholding the background probability density function.

There are several ways to obtain background samples for estimating $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$. The global-RXD (GRXD) and LRXD use different strategies. GRXD is given by

$$D_{\mathrm{GRXD}}(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu}_G)^T\boldsymbol{\Sigma}_G^{-1}(\mathbf{x} - \boldsymbol{\mu}_G), \tag{6}$$

where $\boldsymbol{\mu}_G$ and $\boldsymbol{\Sigma}_G$ are the mean vector and covariance matrix of all pixels in the image. For a given pixel, let $\boldsymbol{\mu}_8$ be the mean value of its eight surrounding neighbors. Then, LRXD is given by

$$D_{\mathrm{LRXD}}(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu}_8)^T\boldsymbol{\Sigma}_G^{-1}(\mathbf{x} - \boldsymbol{\mu}_8). \tag{7}$$

It is noticeable that LRXD generally uses $\boldsymbol{\Sigma}_G$ as the covariance matrix instead of $\boldsymbol{\Sigma}_8$. This is because, in hyperspectral data, the number of spectral bands is much higher than 8 and, in this case, $\boldsymbol{\Sigma}_8$ is a singular matrix and could not be inversed.[21] Although RXD is very popular for solving anomaly detection problems, it could suffer from a high FAR which may hinder the detection of real targets as outliers. Improved RXD algorithms, such as SRXD, KRXD and LAIRX, can reduce the FAR. SRXD requires the number of pixels in the cluster to be equal

to or greater than the number of spectral bands in order to be able to invert the covariance matrix, which may not be possible in some scenarios.[26] LAIRX requires an empirical threshold value that is quite critical for its performance as it is defined using a manual procedure. KRXD suffers from computational complexity and is sensitive to target interference. In the following section, we develop a new improved version of the RXD that follows a different optimization strategy.

3 PAD Algorithm

As mentioned in previous sections, the main obstacle for detecting anomalies by using RXD and other algorithms is the lack of *a priori* knowledge of the targets or background. On the other hand, the fundamental assumption of many anomaly detectors is that the distribution of the background is Gaussian.[26,27] This assumption is a very common model applied for anomaly detection problems so that the background can be described in a statistical model. In practice, this assumption cannot be met in many scenarios, which leads to a high FAR.[22] One point that violates the assumption is that the background samples often contain target information as they are intrinsically mixed. In this section, we develop a new unsupervised PAD which is based on estimating the difference between the probabilities of the anomalies and the background. The proposed PAD takes advantage of the results provided by the RXD to estimate statistical information for the targets and background, respectively, and then uses an automatic strategy to find the most suitable threshold for the separation of targets from the background. The algorithm can be summarized by the following steps:

Step 1. In the first step, we apply the classic RXD algorithm and obtain the initial probabilities associated with each PUT.

Step 2. In the second step, we segment the image into targets and background by means of an automatic thresholding method, which is calculated by seeking a specific feature point of the histogram curve.

Step 3. In the third step, we calculate the probabilities of the PUT to be declared as targets or as the background. Then we subtract the probability of the PUT to be declared as the background from the probability of the PUT being part of the target set.

At this point, it is important to emphasize that, in the second step, a threshold is calculated to separate targets from the background. This is a crucial process in our anomaly detection scheme. Specifically, we aim at setting a confidence coefficient γ in order to determine a rejection probability $P(\alpha_0) = \gamma$ so that, if $\delta_{\mathrm{RXD}}(\mathbf{x}) > \alpha_0$, then the PUT (say, $\mathbf{x}$) will be categorized as an anomaly.[3] However, in previous works, γ was generally set in supervised manner. We emphasize that, in most cases, it is expected that the number of anomalies will be low compared with the number of background pixels. As a result, a peak in the histogram distribution of RXD values can generally be found in the vicinity of the boundary between targets and background. The peak generally appears as a feature point in the histogram curve, that is, the change ratio of the feature point will be a minimum in the interval. Based on this principle, we automatically obtain the threshold value as follows:

1. First, we set two confidence coefficients γ_{Lower} and γ_{Upper}. Based on $P(\alpha_0) = \gamma$, we obtain the interval $[\mathrm{TH}_{\mathrm{Lower}}, \mathrm{TH}_{\mathrm{Upper}}]$ so that the threshold will be located in this interval.
2. Then, we divide this section into N equal-length small sections. The number of RXD values in section i—$[\mathrm{TH}_i, \mathrm{TH}_{i+1}]$ is defined as T_i. Then, the threshold is selected as $\mathrm{TH} = \mathrm{TH}_{\hat{i}+1}$, where $\hat{i} = \arg\min_i\{T_{i+1}/T_i\}$.

At this point, it is important to emphasize that, opposite to other conventional strategies that set a threshold by considering a rejection probability that is decided manually, our approach expands the point into an interval instead and then seeks for a possible threshold in that interval. As a result, the way to calculate the threshold is more realistic and adaptive to different image data sets than the threshold decided by a single point.

For illustrative purposes, Fig. 1 presents an example of the histogram of RXD values in the ideal situation, in which targets and the background are chi-square distributed. The threshold TH is calculated by the proposed method. From Fig. 1, we can see that the proposed threshold is very close to the point in which the background and target curves cross, and therefore appears to

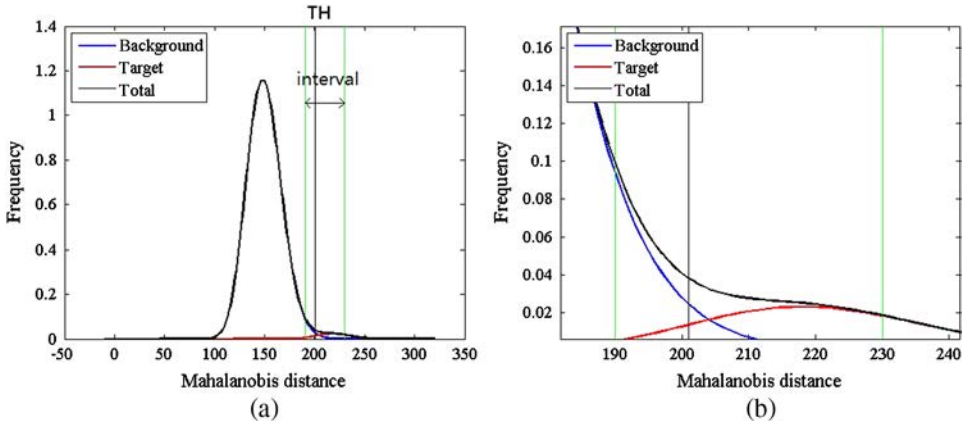

Fig. 1 (a) An example of the interval of values and threshold selected by the proposed PAD method when determining the threshold for a real hyperspectral scene. In this example, we set the ratio of the number of target pixels to the number of background pixels to 1:50. (b) Magnified version in which the specific threshold selected and its relation to the target and background distributions is displayed.

represent a good compromise. If the number of anomalies is too small, the peak in the histogram will not be very obvious. However, the feature point whose change ratio is a minimum can still be found in the interval if it is set appropriately. Compared with other conventional thresholding methods, our presented thresholding approach intends to be more adaptive and feasible in different scenarios. In fact, thresholding RXD results are tricky for at least two reasons. One is because the number of anomalies is uncertain, so their probability distribution function is not reliable. Another reason is the signals of anomalies are variable; they cannot be described by one Gaussian distribution. These facts will reduce the accuracy of separating anomalies from background. Thus, if the number of anomalies is too small or there are several kinds of targets in the image, the anomalies cannot be totally distinguished from the image by a thresholding method. However, the preprocessing of thresholding can still refine the estimation of the background and provide samples for the estimation of targets, which are the main advantages of our proposed method.

The proposed PAD has a similar theoretical foundation to RXD. In the RXD model, the signals of the target and the background are represented as indicated in Eqs. (1) and (2). As mentioned previously, in many scenarios, the background data contain anomalous signals to violate the Gaussian assumption. The proposed method which is based on the Gaussian assumption separates potential targets from background for a more accurate estimation. Specifically, the RXD algorithm assumes that the target signal is a constant vector $\mathbf{s}$. Therefore, the covariance matrices of the background and targets are the same. Unfortunately, in most images, we have many different types of targets, which results in the fact that the assumption of a constant vector $\mathbf{s}$ as the representation of the target signal is inappropriate. Different from RXD method, in our interpretation, it is more reasonable that the target signals $\mathbf{s}$ are expressed as variable vectors and have independent covariance matrices, which are different from the matrix of the background. We noted $\mathbf{s} \sim N(\boldsymbol{\mu}_\mathbf{s}, \boldsymbol{\Sigma}_\mathbf{s})$. As target signals have much more energy than those of background,[3] $\boldsymbol{\Sigma}_\mathbf{s}$ is dominant to $\boldsymbol{\Sigma}$. $\mathbf{x}|H_1$ can be considered as a multivariate normal distribution whose covariance matrix is very close to $\boldsymbol{\Sigma}_\mathbf{s}$. We noted that $\mathbf{x}|H_1 \sim N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$. Thus, in PAD, $p(\mathbf{x}|H_0)$ is the same as the corresponding expression in RXD, and $p(\mathbf{x}|H_1)$ is given by

$$p(\mathbf{x}|H_1) = \frac{1}{(2\pi)^{K/2}|\boldsymbol{\Sigma}_1|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_1)}. \tag{8}$$

As indicated earlier, the proposed method estimates the covariance matrix and the mean vector of the target and background sets by using different samples. We noted the target set as V_1 and the background set as V_0. The number of pixels in V_1 and V_0 are denoted as N_1 and N_0, respectively. The background set V_0 contains purer background information and the samples of target

set V_1 are very different from the background. Thus, the covariance matrix $\mathbf{\Sigma}_m$ and the mean vector $\boldsymbol{\mu}_m$ will be more accurate if evaluated by V_m, $(m = 0,1)$. The proposed method assumes that the background and target signals follow a multivariate normal distribution, therefore, the posterior probabilities $p(\mathbf{x}_m|\mathbf{x})$ are obtained as follows:

$$p(\mathbf{x}|H_m) = \frac{1}{(2\pi)^{K/2}|\mathbf{\Sigma}_m|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_m)^T\mathbf{\Sigma}_m^{-1}(\mathbf{x}-\boldsymbol{\mu}_m)} \qquad (m = 0,1), \tag{9}$$

$$p(H_m) = \frac{N_m}{N_0 + N_1}, \tag{10}$$

$$p(H_m|\mathbf{x}) = p(\mathbf{x}|H_m) \times p(H_m)/p(\mathbf{x}). \tag{11}$$

In this case, we can use the log-likelihood ratio:

$$\lg[p(H_m|\mathbf{x})] = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_m)^T\mathbf{\Sigma}_m^{-1}(\mathbf{x}-\boldsymbol{\mu}_m) - \frac{K}{2}\lg(2\pi) - \frac{1}{2}\lg(|\mathbf{\Sigma_m}|) + \lg[p(H_\mathbf{m})] - \lg[p(\mathbf{x})], \tag{12}$$

and then subtracting $\lg[p(H_0|\mathbf{x})]$ from $\lg[p(H_1|\mathbf{x})]$, we obtain $\mathrm{PAD}(\mathbf{x})$ as follows:

$$\begin{aligned}\mathrm{PAD}(\mathbf{x}) = {} & (\mathbf{x}-\boldsymbol{\mu}_0)^T\mathbf{\Sigma}_0^{-1}(\mathbf{x}-\boldsymbol{\mu}_0) - (\mathbf{x}-\boldsymbol{\mu}_1)^T\mathbf{\Sigma}_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_1) \\ & + \lg(|\mathbf{\Sigma}_0|) + \lg[p(H_1)] - \lg(|\mathbf{\Sigma}_1|) - \lg[p(H_0)].\end{aligned} \tag{13}$$

As $\lg(|\mathbf{\Sigma}_0|) + \lg[p(H_1)] - \lg(|\mathbf{\Sigma}_1|) - \lg[p(H_0)]$ is constant for a considered image, $\mathrm{PAD}(\mathbf{x})$ can be simplified as

$$\mathrm{PAD}(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu}_0)^T\mathbf{\Sigma}_0^{-1}(\mathbf{x}-\boldsymbol{\mu}_0) - (\mathbf{x}-\boldsymbol{\mu}_1)^T\mathbf{\Sigma}_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_1) = \mathrm{RXD}_0(\mathbf{x}) - \mathrm{RXD}_1(\mathbf{x}). \tag{14}$$

Resulting from the previous expression, $\mathrm{RXD}_0(\mathbf{x})$ is the distance between the PUT $\mathbf{x}$ and the background, and $\mathrm{RXD}_1(\mathbf{x})$ represents the distance from $\mathbf{x}$ to the target set. In this context, the different value of these two probabilities is used as a criterion by the proposed PAD method to evaluate the feasibility of a pixel being considered as an anomaly, while the RXD just adopts the background probability. This is a significant difference of our newly proposed PAD with regards to the RXD. An important reason why the proposed PAD exhibits good anomaly detection capability is that it highlights the anomaly twice, i.e., by the distance from the background set and by the distance from the target set. In Eq. (14), the factor $\mathrm{RXD}_0(\mathbf{x})$ withholds the background signals and highlights the anomalies as outliers. The function $\mathrm{RXD}_1(\mathbf{x})$ underscores the targets which appear with higher probability in the targets set. As a result, some minor signals (which are probably detected as targets by RXD, including noise and some background pixels) are withheld by the proposed PAD. As a result, the presented method is expected to improve its anomaly detection accuracy by reducing false alarms caused by noise and background.

There are several other algorithms applying the idea of refining the estimation of background for anomaly detection, i.e., LAIRX, BACON, and RSAD. The following parts will introduce these methods with detailed descriptions.

3.1 *LAIRX*

The LAIRX algorithm is an unsupervised anomaly detector for hyperspectral imagery based on data characteristics in the image.[22] The RXD algorithm is prone to a high FAR because the Gaussian assumption is inaccurate.[28] The background estimation of RXD contains target signals and noise which contaminate the homogeneity of the background. Although there is lack of prior knowledge about targets and background for anomaly detection problems, LAIRX is used to refine the estimation of the background by analyzing data characteristics. The goal of LAIRX is reducing target signals and noise from the estimation of the background. The steps of LAIRX are as follows:

Step 1. Apply the classic RXD algorithm and obtain the RXD scores associated with each PUT.
Step 2. Threshold anomalous pixels from background. Those RX scores that exceed $\chi^2_{\alpha,K}$ will be referred to as anomalies.
Step 3. Iterate Step 2. If the anomalous set in Step 2 is identical to the set in the previous iteration, then go to Step 4; otherwise, return to Step 2.
Step 4. Nominate the pixels in the final set as anomalies.

3.2 *BACON*

BACON is a very popular algorithm to identify outliers in multivariate and regression data. Although many anomaly detectors customarily assume that one given hyperspectral imagery is homogeneous, in fact, it often contains outliers such as targets or noisy pixels. On the other hand, all multiple outlier detectors in the past have suffered from a computational cost that increased rapidly with the sample size. BACON can be computed quickly, and often requires less than five evaluations of the model to fit the data, regardless of the sample size. The BACON algorithm consists of the following steps:

Step 1. Compute the RXD score for each PUT. Select m smallest values of RXD scores as the background subset. The condition of $m > K$ should be met, where K is the number of bands.
Step 2. Obtain the square root of RXD scores based on the new background subset.
Step 3. Those pixels whose square root values of RXD scores are smaller than $c_{nKr}\chi_{K,\alpha}$ will be selected as the new background set. c_{nKr} can be computed as follows:

$$c_{nKr} = c_{nK} + c_{hr}, \tag{15}$$

$$c_{hr} = \max\{0, (h-r)/(h+r)\}, \qquad h = [(n+K+1)/2], \tag{16}$$

$$h = (n+K+1)/2, \tag{17}$$

$$c_{nK} = 1 + (K+1)/(n-K) + 1/(n-h-K), \tag{18}$$

where n is the total number of pixels and r is the number of pixels in the current background subset.
Step 4. Iterate Step 2 and Step 3 until the size of the basic subset no longer changes.
Step 5. Map anomalies to the image space.

3.3 *RSAD*

The RSAD algorithm randomly selects representative background samples from the image each time to estimate background statistical information, identifies anomalies by statistical differences, and finally fuses all the detection results. The steps of RSAD are described as follows:

Step 1. Randomly select background pixels as the initial basic subset of the observed pixels from the hyperspectral imagery.
Step 2. Compute the square root of RXD value of each pixel vector based on the initial basic subset.
Step 3. Those pixels whose square root values of RXD scores are smaller than $c_{nKr}\chi_{K,\alpha}$ will be selected as the new background set. The procedures to compute c_{nKr} are same as those of BACON.
Step 4. Iterate Step 2 and Step 3 until the basic subset no longer changes.
Step 5. Nominate the pixels excluded by the final basic subset as anomalies.

All of these three methods threshold target pixels from the background for a more accurate estimation of the background. Similar to these methods, the presented algorithm keeps the

background more homogeneous by analyzing data characteristics. In the following section, we evaluate the performance of the PAD method using synthetic data.

4 Synthetic Data Experiments

This section performs experiments with a synthetically generated data set that was used to evaluate the performance of the proposed PAD method. The anomalies in this simulated data set were generated following a simple target implantation strategy.[29–31] The performance metrics are based on the receiver operating characteristic (ROC);[1,32] specifically, we analyze the area under the curve (AUC)[30] as a baseline metric for analyzing the performance of different detectors. In the following, we first describe the procedure used for synthetic data generation and then present the experiments that were conducted.

4.1 *Synthetic Data*

The synthetic data were generated from the target detection blind-test scenes provided by Rochester Institute of Technology (RIT).[33] These data were collected by a HyMap instrument over Cook City in Montana, on July 2006. The selected portion (illustrated in Fig. 2) has 280×800 pixels in size and 126 spectral bands. The spatial resolution of the data is $\sim$3 m.

In this work, we use a target implantation method to simulate a set of anomalous targets in the considered hyperspectral data set over Cook City in Montana. The advantage of using a target implantation method is that we can evaluate the performance of the detectors in a totally controlled environment.[29] Specifically, our method generates the anomalous targets using a synthetic spectral signature $\mathbf{z}$ with a specified abundance fraction f from a desired target $\mathbf{t}$, contaminated by a background signature $\mathbf{b}$, in the spatial position in which the target of interest is simulated.[30,31] This means that the implanted targets are all sub-pixel in nature. For the simulations, we use a simple linear mixture model as follows:

$$\mathbf{z} = f \cdot \mathbf{t} + (1 - f) \cdot \mathbf{b}. \tag{19}$$

Figure 3 shows the spectral signatures of the six targets (corresponding to different man-made materials present in different locations of the scene, as provided by the target detection blind-test scenes provided by RIT) that have been implanted in the hyperspectral scene, together with two considered background signatures. The image portion in which the targets have been implanted is denoted as ROI-1 (see Fig. 2). This portion corresponds to an open vegetation region with dimensions $100 \times 100 \times 126$ containing few anomalous pixels. Vegetation can be considered as the main homogeneous background in this region. In order to evaluate the presented method, we use six kinds of targets that have been implanted into ROI-1 using Eq. (19), where the characteristics of the targets are summarized in Table 1. We have synthetically generated six different test images, where each image contains only one type of target

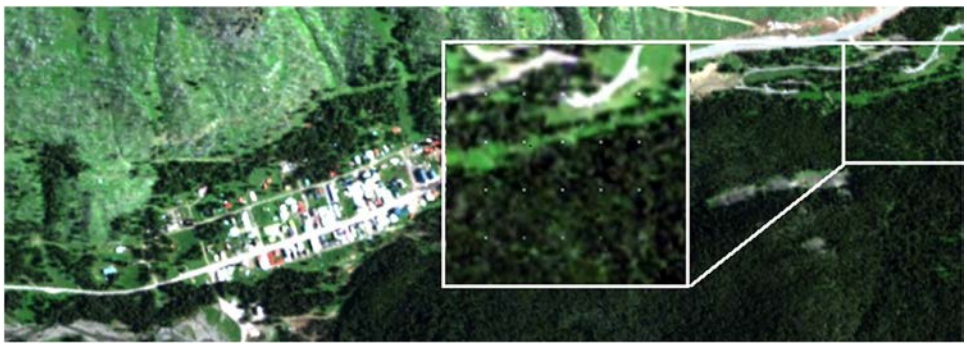

Fig. 2 HyMap hyperspectral image over Cook City, Montana, with 20 embedded targets in a region called ROI-1.

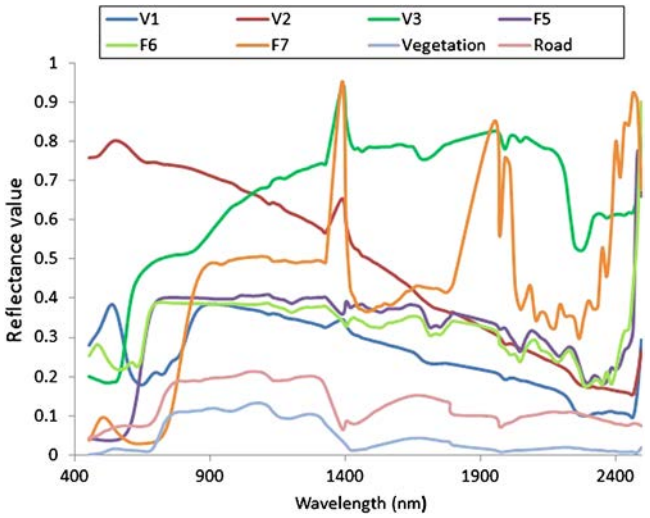

Fig. 3 Spectral signatures of the six targets implanted in the image and two types of backgrounds.

Table 1 Characteristics of the implanted target spectra in our synthetic data experiments.

Name	Type
V1	1993 Chevy Blazer
V2	1997 Toyota T100
V3	1985 Subaru GL Wagon
F5	Maroon Nylon Target
F6	Gray Nylon Target
F7	Green Cotton Target

with different fractions. In each test image, 4×5 targets have been implanted, where the abundance fractions f follow an arithmetic progression with a difference of 0.02 from left to right and top to bottom, and where the maximum value is 0.4 on the top-left target, and the minimum value is 0.02 on the bottom-rightmost target. The main reason why we set 0.4 as the maximum value of f is that, when f is higher than 0.4, there are no false alarms when the target is detected. We decided to set the common difference as 0.02 since we would like to make a complete evaluation of the capability of the detectors with a sufficiently small difference of the target abundance f.

4.2 *Preliminary Evaluation*

This subsection describes the results obtained by the proposed PAD method using the synthetic data set described in the previous subsection. Specifically, we use global [denoted as PAD (Global)] and local [denoted as PAD(Local)] models to estimate the background information. PAD(Global) uses the GRXD results to perform the segmentation, and PAD(Local) uses the LRXD results. The window size of the local methods was empirically set to 3×3 since it was experimentally observed that this parameter leads to the best results in our experiments. This section also provides a detailed comparison between our proposed methods and other widely used methods such as LAIRX2, BACON, and RSAD. Our main reason for using LAIRX2 instead of LAIRX is that our proposed PAD calculates the RX after one segmentation, and LAIRX2 performs two RX-type iterations.[22] Therefore, a comparison of PAD to LAIRX2 is

more natural. Similarly, we refer to BACON2 and RSAD2 in our experiments to indicate that these methods are also configured to perform two iterations in the experiments.

Figure 4 presents the binary images (separating targets and background) obtained by the considered algorithms after using an optimally selected threshold. As shown in Fig. 4, PAD and LAIRX2 can detect more implanted targets than the classic RXD, regardless of whether a local or global model is used. Specifically, both PAD(Global) and LAIRX2(Global) can detect at least 17 out of 20 targets, whereas GRXD can only detect 12. For instance, if we consider only the simulated targets of the F7 type, PAD(Global), PAD(Local), and LAIRX2(Local) can all detect the targets with f equal or higher than 0.04. LAIRX2(Global) can detect the targets with an f equal or higher than 0.06. However, both GRXD and LRXD can only detect the targets whose f is not <0.18. This implies that the proposed method can detect subpixel anomalies of less abundance from the target signal. Further, BACON and RSAD can detect more implanted targets than other algorithms in the considered experiments with synthetic data. These two methods are iterative and generally more time consuming than the proposed method, as indicated by their processing times measured on a desktop PC equipped with two Intel(R) Core(TM) Duo processors at 3.00 GHz with 2 GB of RAM memory.

4.3 *ROC Analysis*

ROC curves provide a widely used strategy to evaluate anomaly detectors. It establishes the correspondence between the detection probability and the FAR. Figure 5 presents the ROC curves obtained for the different methods considered in this experiment with synthetic data. As Fig. 5 shows, compared with the conventional RXD, our proposed methods, PAD(Global) and PAD(Local), increase the probability of detection while reducing the FAR. Table 2 reports the AUC scores, which were used for quantitative evaluation of the detection performance (the larger the AUC, the better the detection results). It is noticeable that PAD results in a lower FAR than LAIRX2 with the same probability of detection, regardless of whether a global or local model is considered. On the other hand, local-based detectors such as LRXD, PAD(Local) and LAIRX2(Local) outperformed the corresponding global-based detectors GRXD, PAD(Global), and LAIRX2(Global). The main reason is that all targets in the synthetic image are subpixel in size, hence the local context around each pixel provides a better estimation for background information. Although BACON and RSAD achieve the best performance in terms of AUC for the synthetic image, they are also more computationally expensive than the presented methods. If we take F7-type targets again, we can see that PAD(Global) provides slightly lower AUC scores than BACON and RSAD. However, the execution time of BACON is 2.79 times higher than the one measured for PAD(Global) and the time of RSAD is 2.15 times higher than the one measured for PAD(Global). With these issues in mind, we can generally conclude that the proposed methods provide a better balance between detection performance and time consumption than other algorithms considered in our experiments.

4.4 *Performances of PAD with Various Thresholds*

Admittedly, PAD seeks a threshold from an interval, which is set experientially. It is necessary to present the performances of PAD with various thresholds. In this additional experiment, we demonstrate whether or not a fluctuation of the threshold has a great influence on the detection results. In the thresholding process, except for the threshold application in PAD, the three other thresholds have been set to $\gamma = 0.97$, $\gamma = 0.95$, and $\gamma = 0.90$, respectively. We apply each threshold on PAD(Global) and get ROC curves and AUC for all six target implanted images.

From Fig. 6 and Table 3, although we can observe that the thresholds that we set fluctuate in a given range, the detecting performances of PAD(Global) still remain on a very high level. The improvements in GRXD from PAD are obvious with various thresholds. The results of this experiment give powerful support to the adaptive nature of our presented algorithm. Moreover, PAD (in combination with the presented thresholding method) obtains the highest AUC value in four out of six images. This fact indicates that our thresholding method could

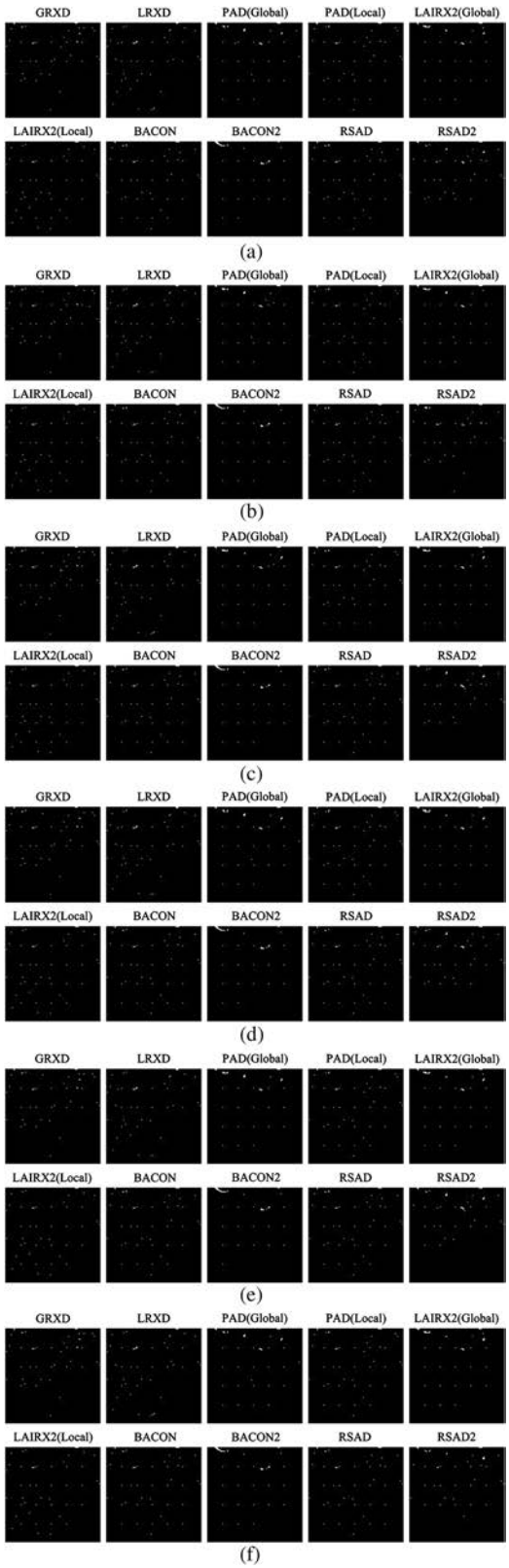

Fig. 4 Detection results obtained by the different methods with synthetic data generated using different targets. In all cases, the rejection probability γ was set to 99.5%. (a) V1, (b) V2, (c) V3, (d) F5, (e) F6, and (f) F7.

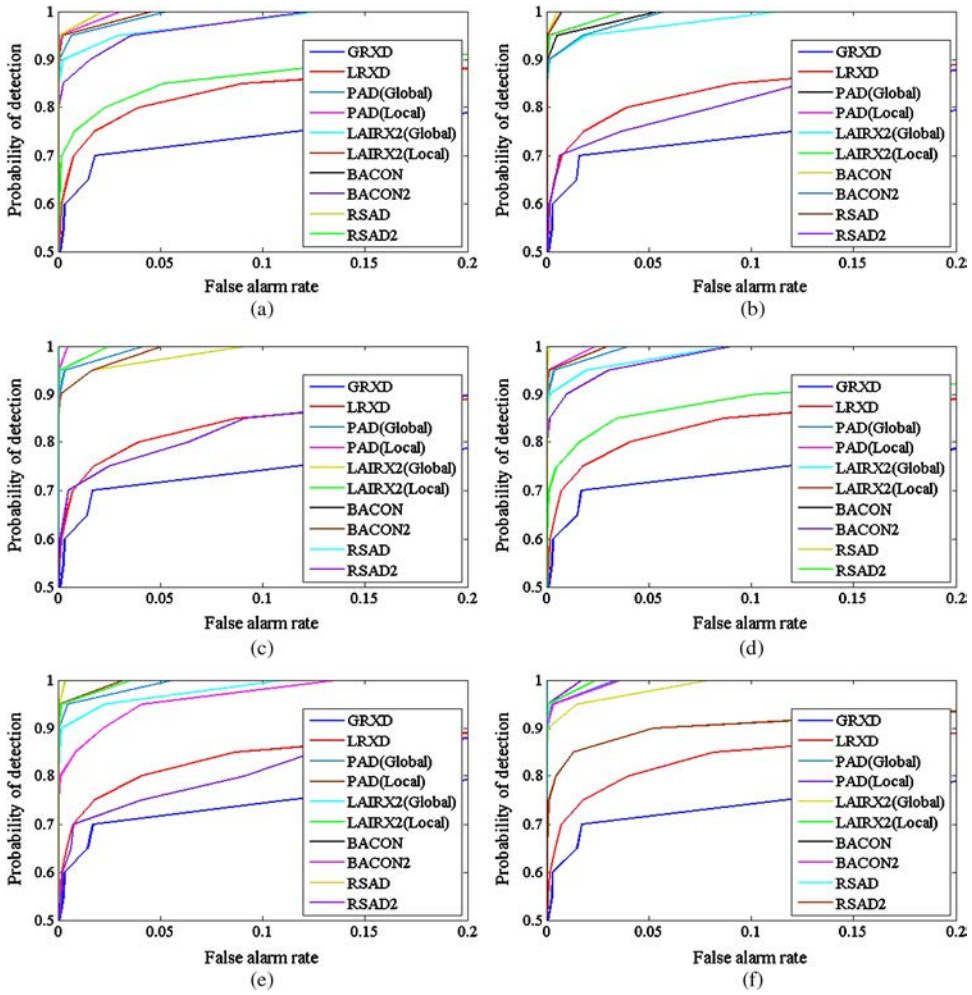

Fig. 5 Receiver operating characteristic (ROC) curves corresponding to the detection results reported in Fig. 4. (a) V1, (b) V2, (c) V3, (d) F5, (e) F6, and (f) F7.

assist PAD(Global) to get the best detecting performance in most scenarios, as compared with other thresholds that were applied in PAD(Global).

5 Real Image Experiments

In this section, two real hyperspectral data have been used to evaluate the performance of the proposed anomaly detectors. In the following, we describe the data sets used and the results obtained by different anomaly detectors when applied to these two widely used data sets.

5.1 *Hyperspectral Data Sets*

5.1.1 *World Trade Center*

This data set was collected by the Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS), operated by NASA's Jet Propulsion Laboratory, over the World Trade Center (WTC) area in New York, on September 16, 2001 (just 5 days after the terrorist attacks that collapsed the two main towers in the WTC complex).[34] A portion of 200×200 pixels (with 224 spectral bands between 0.4 and 2.5 μm) was selected for the experiments. This area covered the hot spots corresponding to latent fires at the WTC, which can be considered as anomalies. Figure 7(a) shows a false color representation of the portion selected for experiments, whereas Fig. 7(b) displays a ground-truth data set, which comprises the spatial location of the hot spots provided by the United States Geological Survey (USGS).

Table 2 Area under the curve (AUC) and processing times (s) for the detectors reported on Figs. 4 and 5.

Target	Algorithm	AUC	Times (s)	Algorithm	AUC	Times (s)	Algorithm	AUC	Times (s)
V1	GRXD	0.9079	0.706	PAD(Global)	0.9983	1.673	LAIRX2(Global)	0.9952	1.092
	LRXD	0.9422	0.883	PAD(Local)	0.9992	1.800	LAIRX2(Local)	0.9987	1.769
	BACON	0.9994	5.708	RSAD	0.9994	3.079			
	BACON2	0.9943	1.280	RASD2	0.9538	0.826			
V2	GRXD	0.9098	0.592	PAD(Global)	0.9983	1.568	LAIRX2(Global)	0.9960	1.102
	LRXD	0.9420	0.918	PAD(Local)	0.9990	1.816	LAIRX2(Local)	0.9989	1.772
	BACON	0.9998	4.546	RSAD	0.9998	1.939			
	BACON2	0.9976	1.284	RASD2	0.9356	0.842			
V3	GRXD	0.9114	0.899	PAD(Global)	0.9987	1.883	LAIRX2(Global)	0.9968	1.123
	LRXD	0.9455	0.896	PAD(Local)	0.9998	1.808	LAIRX2(Local)	0.9993	1.769
	BACON	1.0000	4.112	RSAD	1.0000	2.546			
	BACON2	0.9979	1.300	RASD2	0.9458	0.833			
F5	GRXD	0.9100	0.594	PAD(Global)	0.9987	1.534	LAIRX2(Global)	0.9967	1.129
	LRXD	0.9446	0.891	PAD(Local)	0.9993	1.804	LAIRX2(Local)	0.9991	1.789
	BACON	0.9999	4.140	RSAD	0.9999	2.020			
	BACON2	0.9957	1.322	RASD2	0.9613	0.880			
F6	GRXD	0.9105	0.580	PAD(Global)	0.9983	1.516	LAIRX2(Global)	0.9960	1.104
	LRXD	0.9446	0.905	PAD(Local)	0.9991	1.801	LAIRX2(Local)	0.9990	1.769
	BACON	0.9999	4.153	RSAD	0.9999	2.596			
	BACON2	0.9930	1.314	RASD2	0.9358	0.860			
F7	GRXD	0.9127	0.569	PAD(Global)	0.9990	1.473	LAIRX2(Global)	0.9972	1.143
	LRXD	0.9438	0.896	PAD(Local)	0.9995	1.854	LAIRX2(Local)	0.9993	1.788
	BACON	1.0000	4.111	RSAD	1.0000	3.161			
	BACON2	0.9989	1.300	RASD2	0.9734	0.854			

5.1.2 *SpecTIR data*

These data were collected in the framework of the SpecTIR Hyperspectral Airborne Rochester Experiment (SHARE).[35] The data set was collected on July 29, 2010 by the ProSpecTIR-VS2 sensor containing 360 bands from 390 to 2450 nm with a 5-nm spectral resolution. The ground resolution is ∼1 m. In the image, road and vegetation are the main backgrounds and red and blue fabrics (sized 9, 4, and 0.25 m^2) were purposely placed as targets. We selected an area of 180×180 pixels that contains fabric targets, as displayed in Fig. 8(a), for the experiments. Figure 8(b) displays the ground-truth location of the targets in the considered experiment.

5.2 *Experimental Results*

This section reports the experimental comparison conducted using different detectors and the considered real hyperspectral scenes. As the size of the targets in these two real images is usually more than one pixel, the local methods in the experiments have been implemented using a dual

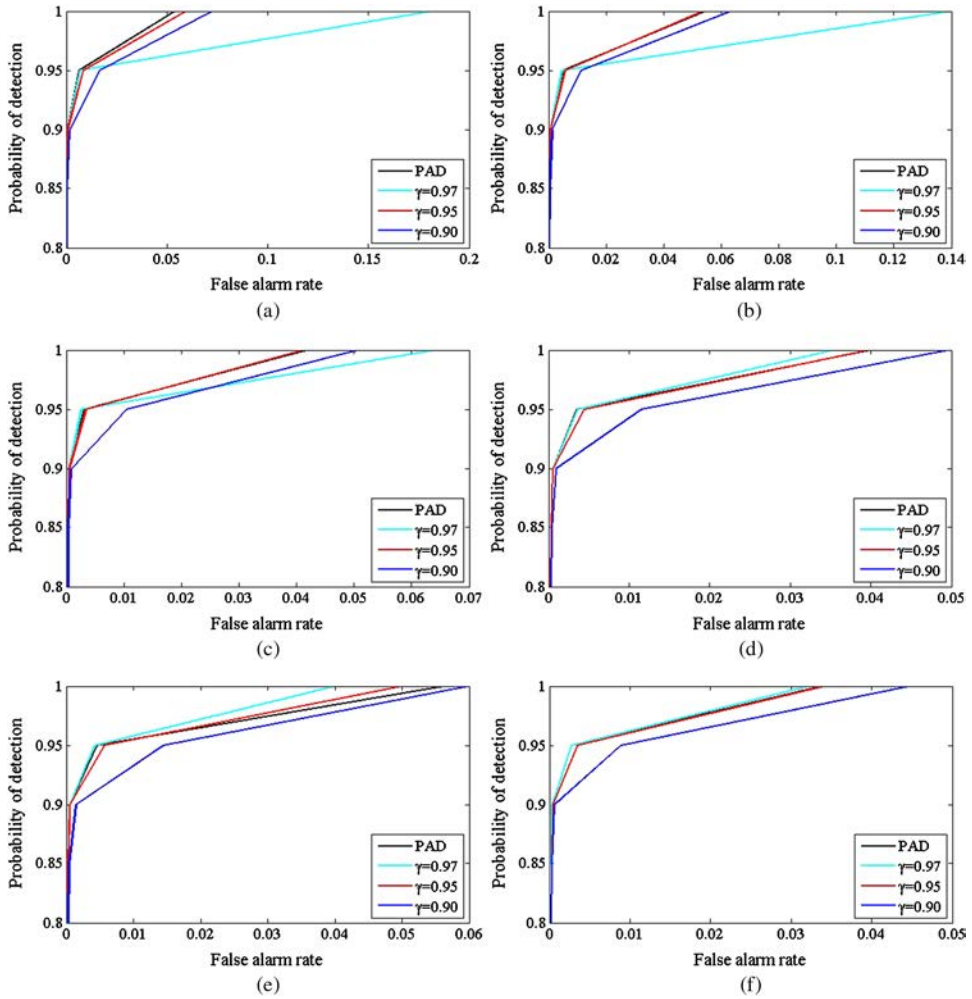

Fig. 6 ROC curves of different thresholds on image (a) V1, (b) V2, (c) V3, (d) F5, (e) F6, and (f) F7.

window approach[36] to better estimate the background information. In our experiments, the inner and outer window sizes for the dual windows were set to 5×5 and 15×15 pixels, respectively, after considering the size of the targets.

5.2.1 *World Trade Center*

The detection results provided by the proposed methods for the WTC scene are presented in Fig. 9. These results indicate that the proposed PAD methods can highlight the targets more effectively than the conventional RXD, regardless of whether a global or local method is

Table 3 AUC for PAD with various thresholds reported on Fig. 6.

Target	PAD	$\gamma = 0.97$	$\gamma = 0.95$	$\gamma = 0.90$
V1	0.9983	0.9951	0.9980	0.9972
V2	0.9983	0.9962	0.9983	0.9977
V3	0.9987	0.9982	0.9987	0.9981
F5	0.9987	0.9989	0.9981	0.9987
F6	0.9983	0.9987	0.9984	0.9976
F7	0.9990	0.9990	0.9989	0.9984

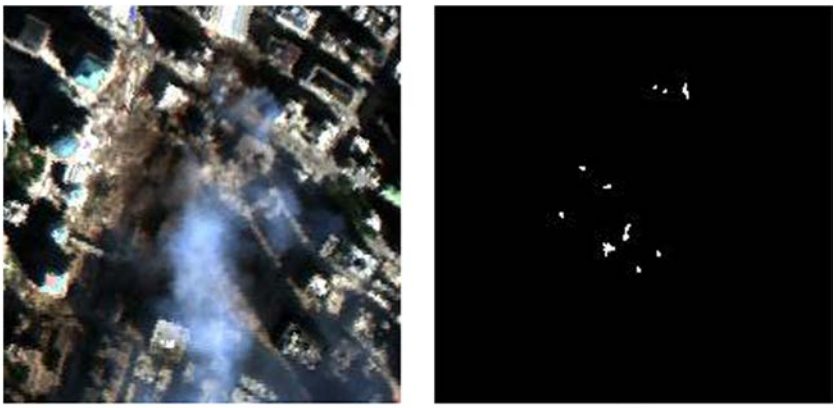

Fig. 7 (a) AVIRIS image over the World Trade Center in New York City. (b) Ground-truth map indicating the spatial location of hot spot fires, available from the United States Geological Survey.

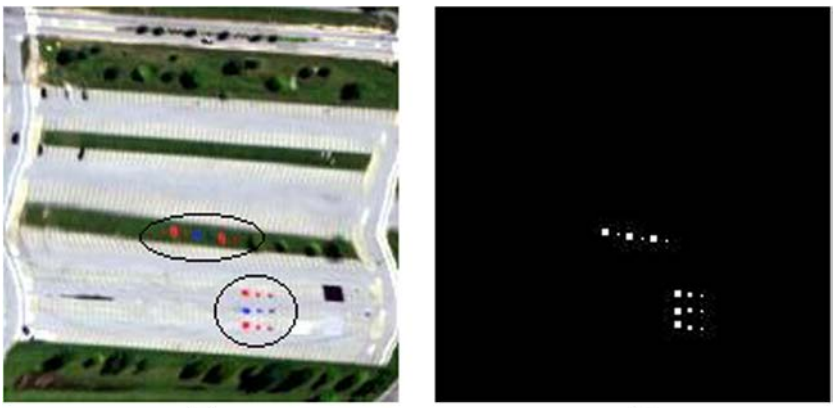

Fig. 8 (a) SpecTIR hyperspectral image with the targets highlighted by black ellipses. (b) Ground-truth information.

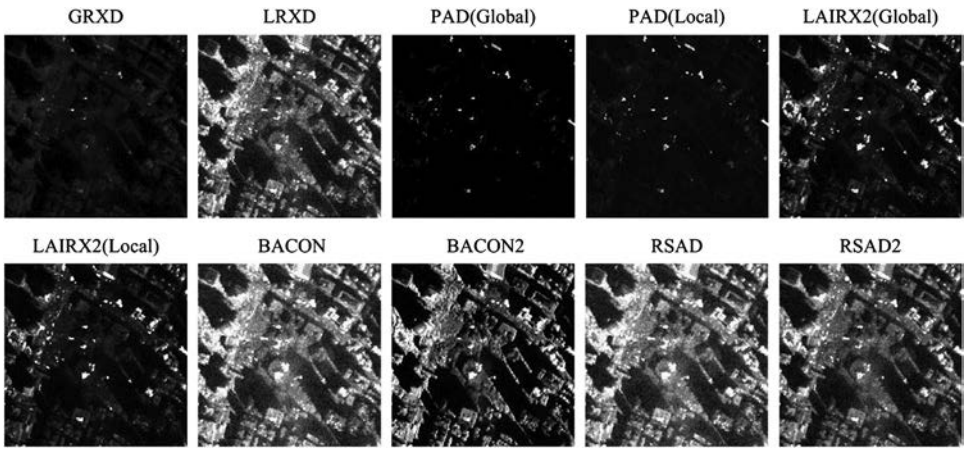

Fig. 9 Detection results obtained by different algorithms for the AVIRIS World Trade Center data.

used. The corresponding ROC curves and AUC are presented in Fig. 10 and Table 4, which indicate that the proposed methods exhibit better detection results than the classic RXD. It is remarkable that BACON outperforms other algorithms for this data set, however, it is more computationally expensive than LAIRX2 and RSAD. On the other hand, BACON2 and RSAD2 obtain a lower probability of detection than BACON and RSAD, although they

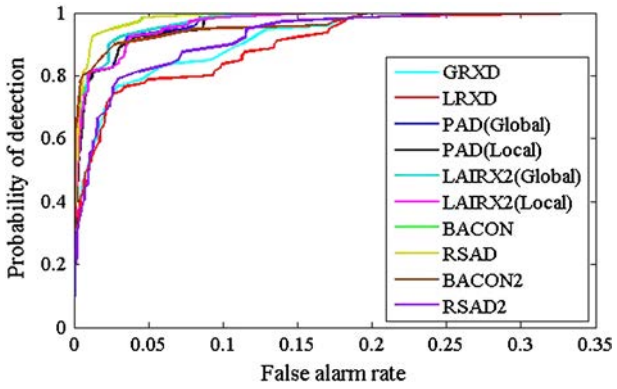

Fig. 10 ROC curves corresponding to the detection results described in Fig. 9.

Table 4 AUC and processing times (s) for the detectors reported on Figs. 9 and 10.

Algorithm	AUC	Times (s)	Algorithm	AUC	Times (s)	Algorithm	AUC	Times (s)
GRXD	0.9689	6.917	PAD(Global)	0.9886	20.914	LAIRX2(Global)	0.9905	13.526
LRXD	0.9638	57.469	PAD(Local)	0.9878	70.604	LAIRX2(Local)	0.9888	100.571
BACON	0.9947	87.835	BACON2	0.9861	14.005			
RSAD	0.9946	52.384	RSAD2	0.9722	9.800			

are less time-consuming. From the results in Fig. 10, we can also conclude that the global detectors perform better than the corresponding local detectors. This is due to the fact that the real targets are bigger than one pixel in size. As a result, the local estimation is expected to contain more target signals.

5.2.2 *SpecTIR data*

The detection results obtained in our experiments for the SpecTIR image are reported on Fig. 11. The binary images obtained after thresholding the detection results with empirically selected optimal thresholds are reported in Fig. 12. The binary images indicate that PAD(Global) and PAD(Local) can detect a higher number of targets than the classic RXD-based algorithms (GRXD and LRXD), particularly when the targets are subpixel in nature. More specifically, PAD(Global) and PAD(Local) can detect five subpixel targets (0.25 m^2) out of six and LAIRX2(Global) can detect four subpixel targets, whereas GRXD and LRXD cannot detect

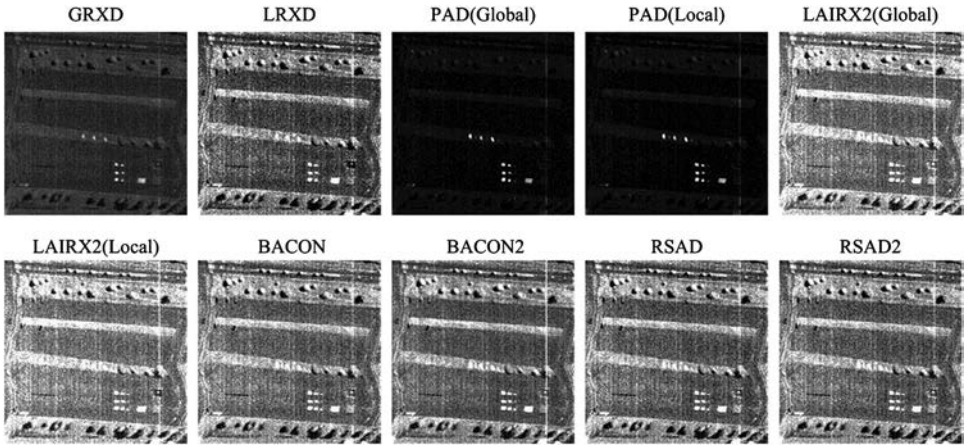

Fig. 11 Detection results obtained by different algorithms for the SpecTIR data.

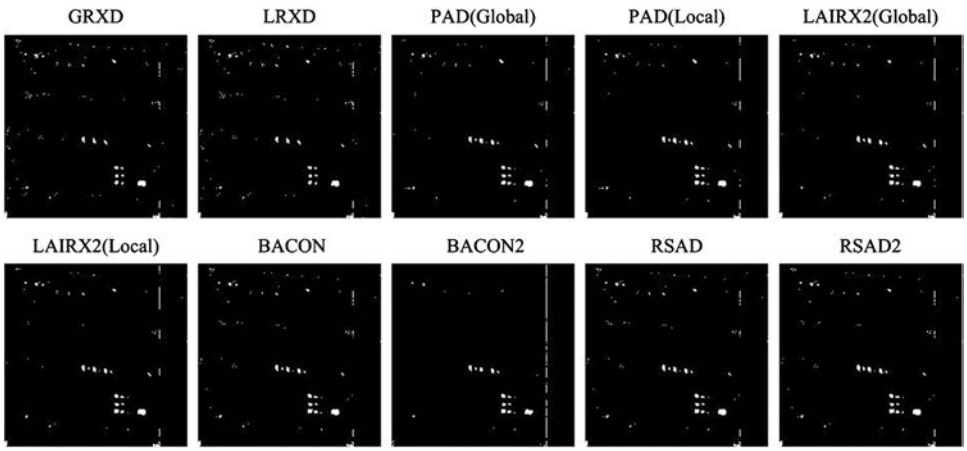

Fig. 12 Binary images obtained after thresholding the results in Fig. 11 using empirically selected optimal thresholds. The ratio of anomaly was set as 1%.

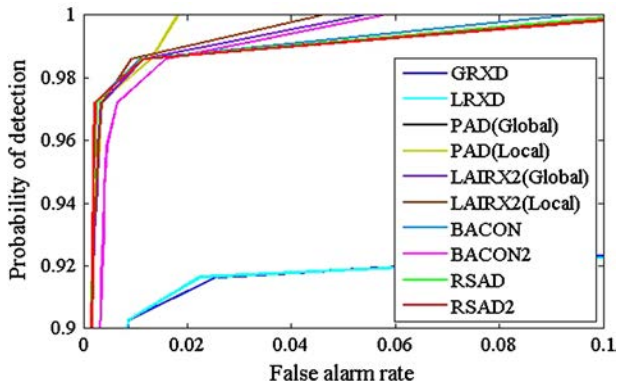

Fig. 13 ROC curves obtained for different methods on the SpecTIR data.

any of them. BACON and RSAD also perform similarly in terms of detection accuracy when compared with PAD. On the other hand, the proposed methods resulted in a lower number of false alarms, as indicated by Fig. 12. This demonstrates that the proposed methods achieved a generally good performance in the task of suppressing background and noise information.

On the other hand, the ROC curves obtained for different detectors for the SpecTIR data are reported in Fig. 13. From these curves and the corresponding AUC scores reported in Table 5, we can conclude that PAD(Global) and PAD(Local) obtained the best performances for the considered data set. Specifically, PAD(Global) and PAD(Local) obtained significant improvements when compared with GRXD and LRXD and also obtained a slight performance increase in terms of AUC when compared to BACON, RSAD, and LAIRX2. A general observation is that local methods are more computationally expensive than their corresponding global methods, mainly due to the fact that we are using a dual window-based strategy in order to enhance their detection results. It is also noticeable from Table 5 that the execution time of PAD(Global) is

Table 5 AUC and processing times (seconds) for the detectors reported on Figs. 11–13.

Algorithm	AUC	Times (s)	Algorithm	AUC	Times (s)	Algorithm	AUC	Times (s)
GRXD	0.9683	12.283	PAD(Global)	0.9993	31.508	LAIRX2(Global)	0.9990	23.425
LRXD	0.9651	67.802	PAD(Local)	0.9993	87.006	LAIRX2(Local)	0.9991	131.820
BACON	0.9988	77.982	BACON2	0.9985	27.963			
RSAD	0.9987	54.584	RSAD2	0.9987	17.598			

lower than the execution times of BACON and RSAD. In fact, BACON and RSAD can detect more targets than BACON2 and RSAD2, although they are more computationally expensive due to their iterative nature. Since PAD considers the probability of the PUT to be considered as targets, it can withhold noise and underscore major targets (fabrics have at least 72 pixels) in the image. This also explains why PAD achieved the best performance in the task of detecting the fabrics while reducing noise and background for the SpecTIR data.

Summarizing, the experiments conducted using both synthetic and real data sets indicate the satisfactory performance of all the detectors presented in this paper. Although LAIRX2 slightly outperformed PAD in the WTC data set, our proposed methods are superior to LAIRX2 on the synthetic and SpecTIR data sets. On the other hand, our methods provide an automatic threshold as opposed to LAIRX2 in which the optimal threshold value should be empirically selected. As a result, the threshold selection method implemented in PAD is quite flexible and adaptive to various scenarios. Our results also revealed that PAD(Global) outperformed BACON and RSAD for the SpecTIR data set. For the WTC image, however, their performances are reversed. As BACON and RSAD are iterative methods, they are more time-consuming than PAD(Global). This is particularly the case for large hyperspectral images. In addition, PAD has the capacity to exploit its adaptive threshold strategy in order to obtain a more accurate estimation of the background.

6 Conclusion and Future Lines

Anomaly detection is a very active topic of research in hyperspectral imagery. In anomaly detection scenarios, generally no prior information is available. Although some classic detectors (like RXD and its variations) have been very popular in the literature, they generally suffer from a high FAR. In order to mitigate this issue, this paper has presented a new method for anomaly detection called PAD, which aims at decreasing the FAR without sacrificing the anomaly detection accuracy. A key aspect of PAD is that it mines knowledge from RXD results by separating the information from the target and background sets for improved estimation purposes. In particular, the segmentation can refine the estimation of the background statistics, leading to a background estimation that is closer to a multivariate normal distribution. Our experimental results, conducted with both synthetic and real hyperspectral data sets, indicate that PAD has a good performance in the task of underscoring the major targets in the scene and reducing noise-induced false alarms. The proposed method also incorporates an automatic procedure for adaptively estimating a threshold value. Although the proposed approach is shown to be quite effective in computational terms, our future research work will be focused on accelerating the performance of the method even more by resorting to high performance computing architectures such as multicore processors and commodity graphics processing units.

Acknowledgments

The authors would like to express appreciation to the Digital Imaging and Remote Sensing Group Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, for providing the Target Blind Detection Test data sets. This research was supported by the Key Research Program of the Chinese Academy of Sciences under Grant No. KZZD-EW-TZ-18, and by the National Natural Science Foundation of China under Grants No. 41325004, and by the Interdisciplinary and Collaborative S&T Innovation Research Team on Advance Earth Observation System, CAS.

References

1. M. T. Eismann, *Hyperspectral Remote Sensing*, SPIE Press, Bellingham, Washington (2012).
2. D. W. J. Stein et al., "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.* **19**(1), 58–69 (2002).
3. C. I. Chang and S. S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **40**(6), 1314–1325 (2002).
4. C. I. Chang and D. C. Heinz, "Constrained subpixel target detection for remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.* **38**(3), 1144–1159 (2000).

5. W. H. Farrand and J. C. Harsanyi, "Mapping the distribution of mine tailings in the Coeur d'Alene River Valley, Idaho, through the use of a constrained energy minimization technique," *Remote Sens. Environ* **59**(1), 64–76 (1997).
6. H. Ren and C. I. Chang, "Target-constrained interference-minimized approach to subpixel target detection for hyperspectral images," *Opt. Eng.* **39**(12), 3138–3145 (2000).
7. Q. Du and C. I. Chang, "A signal-decomposed and interference-annihilated approach to hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.* **42**(4), 892–906 (2004).
8. S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.* **47**(9), 2538–2541 (1999).
9. I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoustics Speech Signal Process.* **38**(10), 1760–1770 (1990).
10. X. L. Yu, I. S. Reed, and A. D. Stocker, "Comparative performance analysis of adaptive multispectral detectors," *IEEE Trans. Signal Process.* **41**(8), 2639–2656 (1993).
11. X. L. Yu et al., "Automatic target detection and recognition in multiband imagery: a unified ML detection and estimation approach," *IEEE Trans. Image Process.* **6**(1), 143–156 (1997).
12. H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **43**(2), 388–397 (2005).
13. A. Huck and M. Guillaume, "Asymptotically CFAR-unsupervised target detection and discrimination in hyperspectral images with anomalous-component pursuit," *IEEE Trans. Geosci. Remote Sens.* **48**(11), 3980–3991 (2010).
14. J. Meola et al., "Detecting changes in hyperspectral imagery using a model-based approach," *IEEE Trans. Geosci. Remote Sens.* **49**(7), 2647–2661 (2011).
15. E. Lo, "Maximized subspace model for hyperspectral anomaly detection," *Pattern Anal. Appl.* **15**(3), 225–235 (2012).
16. S. Matteoli et al., "An automatic approach to adaptive local background estimation and suppression in hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.* **49**(2), 790–800 (2011).
17. H. Goldberg, H. Kwon, and N. M. Nasrabadi, "Kernel eigenspace separation transform for subspace anomaly detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.* **4**(4), 581–585 (2007).
18. A. Mehmood and N. M. Nasrabadi, "Kernel wavelet-Reed-Xiaoli: an anomaly detection for forward-looking infrared imagery," *Appl. Opt.* **50**(17), 2744–2751 (2011).
19. L. Capobianco, A. Garzelli, and G. Camps-Valls, "Target detection with semisupervised kernel orthogonal subspace projection," *IEEE Trans. Geosci. Remote Sens.* **47**(11), 3822–3833 (2009).
20. J. E. Fowler and Q. Du, "Anomaly detection and reconstruction from random projections," *IEEE Trans. Image Process.* **21**(1), 184–195 (2012).
21. N. Gorelnik, H. Yehudai, and S. R. Rotman, "Anomaly detection in non-stationary backgrounds," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pp. 1–4, CDROM, Reykjavík, Iceland (2010).
22. Y. Taitano, B. Geier, and K. Bauer, "A locally adaptable iterative RX detector," *EURASIP J. Adv. Signal Process.* **2010**(1), 341908 (2010).
23. N. Billor, A. S. Hadi, and P. F. Velleman, "BACON: blocked adaptive computationally efficient outlier nominators," *Comput. Statist. Data Anal.* **34**(3), 279–298 (2000).
24. B. Du and L. P. Zhang, "Random-selection-based anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **49**(5), 1578–1589 (2011).
25. P. C. Hytla et al., "Anomaly detection in hyperspectral imagery: comparison of methods using diurnal and seasonal data," *J. Appl. Remote Sens.* **3**(1), 033546 (2009).
26. C. E. Caefer et al., "Improved covariance matrices for point target detection in hyperspectral data," *Opt. Eng.* **47**(7), 076402 (2008).
27. S. Matteoli, M. Diani, and G. Corsini, "Improved estimation of local background covariance matrix for anomaly detection in hyperspectral images," *Opt. Eng.* **49**(4), 046201 (2010).
28. A. Banerjee, P. Burlina, and R. Meth, "Fast hyperspectral anomaly detection via SVDD," in *IEEE Int. Conf. Image Processing*, Vol. 4, IEEE, San Antonio, TX (2007).

29. S. M. Schweizer and J. M. F. Moura, "Efficient detection in hyperspectral imagery," *IEEE Trans. Image Process.* **10**(4), 584–597 (2001).
30. S. Khazai et al., "Anomaly detection in hyperspectral images based on an adaptive support vector method," *IEEE Geosci. Remote Sens. Lett.* **8**(4), 646–650 (2011).
31. M. S. Stefanou and J. P. Kerekes, "A method for assessing spectral image utility," *IEEE Trans. Geosci. Remote Sens.* **47**(6), 1698–1706 (2009).
32. J. M. Molero et al., "Anomaly detection based on a parallel kernel RX algorithm for multicore platforms," *J. Appl. Remote Sens.* **6**(1), 061503 (2012).
33. D. Snyder et al., "Development of a Web-based application to evaluate target finding algorithms," in *IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, Vol. 2, pp. 915–918, IEEE, Boston, MA (2008).
34. J. M. Molero et al., "Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **6**(2), 801–814 (2013).
35. J. A. Herweg et al., "SpecTIR hyperspectral airborne rochester experiment data collection campaign," *Proc. SPIE* **8390**, 839028 (2012).
36. S. Kwon, S. Z. Der, and N. M. Nasrabadi, "Adaptive anomaly detection using subspace separation for hyperspectral imagery," *Opt. Eng.* **42**(11), 3342–3351 (2003).

Lianru Gao received his PhD degree in cartography and geographic information system from Institute of Remote Sensing Applications, Chinese Academy of Sciences (CAS), Beijing, China, in 2007. He is currently an associate professor with the Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, CAS. His research interests include spectral feature analysis, hyperspectral image processing, target detection, and image simulation.

Qiandong Guo is currently working toward his PhD degree in the School of Geosciences, University of South Florida, Florida, USA. He received his MS degree in cartography and geographic information system from Institute of Remote Sensing and Digital Earth, CAS, Beijing, China, in 2014. His research interests include hyperspectral image processing, target detection, and artificial intelligence.

Antonio Plaza received his PhD degrees in computer engineering from the University of Extremadura, Caceres, Spain, in 2002. He is currently an associate professor (with accreditation for full professor) with the Department of Technology of Computers and Communications, University of Extremadura, where he is the head of the Hyperspectral Computing Laboratory (HyperComp). He is currently serving as the editor-in-chief of the *IEEE Transactions on Geoscience and Remote Sensing Journal*.

Jun Li received her European doctorate for signal processing (SIGNAL) under the joint supervision of Prof. José M. Bioucas-Dias and Prof. Antonio Plaza in 2010. She is currently a professor at the School of Geography and Planning, Sun Yat-sen University, Guangzhou, China. Her research interests include hyperspectral image classification and segmentation, spectral unmixing, signal processing, and remote sensing.

Bing Zhang received his PhD degree in remote sensing from Institute of Remote Sensing Applications, CAS, Beijing, China. He is currently a professor and the deputy director of the Institute of Remote Sensing and Digital Earth, CAS. His research interests include development of physics-based models and image processing software for the use of hyperspectral remote sensing data in solving problems in geology, hydrology, ecology, and botany.