Hyperspectral Anomaly Detection Using Collaborative Representation With Outlier Removal

Hongjun Su, Senior Member, IEEE, Zhaoyue Wu, Qian Du, Fellow, IEEE, and Peijun Du, Senior Member, IEEE

Abstract—Recently, collaborative representation detector (CRD) has been popularly used for hyperspectral anomaly detection. For the original CRD, the least squares solution becomes more unstable when more classes, i.e., samples for anomaly detection are involved, and the detection error is likely to happen if the test pixel is an anomalous pixel and several samples from background are similar anomalous. In this paper, we propose a hyperspectral anomaly detection method that uses CRD with principal component analysis (PCA) for removing outlier (PCARoCRD). According to the different background modeling methods, global and local versions are proposed. In the proposed algorithm, the spatial-domain PCA is adopted to extract main pixel information of global/local background that will be used as samples for collaborative representation, and simultaneously the information of abnormal pixels in global/local background can be removed. Fewer useful samples can also keep the detection result stable. Experimental results indicate that the PCARoCRD outperforms the original CRD, kernel version of CRD, advanced CRD (CRDBORAD), morphology-based CRD, Global Reed–Xiaoli (RX) algorithm, and the Local RX.

Index Terms—Anomaly detection, collaborative representation (CR), hyperspectral imagery, PCA, target detection.

I. INTRODUCTION

HYPERSPECTRAL remote sensing is an advanced technology, which can obtain many very narrow spectral continuous bands in the range of visible, near infrared, infrared, and thermal infrared spectrum [1]–[3]. Target detection aims to distinguish the target from other objects and judge its existence in each pixel [4]. Due to the spectral diagnostic ability, more accurate target detection is possible by using hyperspectral imagery. Hyperspectral target detection plays an important role in military and civil applications, such as military reconnaissance, recognition of camouflaged target, rare mineral discovery in geology, and precision agriculture [5], [6].

Generally, target detection is divided into two categories: target detection and anomaly detection. The former knows target information, while the latter does not. The observed material spectrum in an actual scene may not truly represent deterministic shapes, and even the spectral signature of the same ground object may not be necessarily the same. The main factors for this phenomenon are atmospheric conditions, sensor noise, light intensity, material composition, and location, etc. In this case, anomaly detection has obvious advantages over the traditional matched filter-based target detection. Without prior information about target, anomaly detection detects anomalous pixels different from most pixels in the image scene [7]–[11].

Anomaly detection has attracted extensive attention. The Reed–Xiaoli (RX) algorithm proposed by Reed and Xiaoli Yu is one of the most widely used anomaly detection algorithms [12]–[15], and it is a constant false alarm detection algorithm with maximum likelihood detection. There are two versions of the RX: global and local. It is based on a hypothesis that the background pixels around an anomalous pixels obey the multivariate Gaussian model, and the parameters of the model are obtained by estimating the mean and covariance of the background. Therefore, accurate global/local background statistics is crucial. In fact, this assumption may not be correct for hyperspectral imagery, leading to a higher false alarm rate. Chang and Ren proposed a series of improved algorithms [16]–[21], which can improve the efficiency of anomaly detection. Kwon has proposed a nonlinear RX algorithm (KRX) [22], which can project linearly nonseparable data into a high-dimensional feature space, where anomalies and background become more separable. Moreover, a nonlinear detector called support vector data description (SVDD) [23]–[25] is also proposed to deal with linear inseparability problems.

With the development of the compressed sensing theory and sparse coding [26], [27], sparse representation (SR) has become a hot topic [28]–[31]. It assumes that most natural signals can be represented simply by several atoms in an over-complete dictionary, with little loss of information [32]. However, in practical applications, it may be difficult to obtain the over-complete dictionary, resulting in inaccurate representation. To this end, Zhang et al. proposed collaborative representation (CR) [33], [34], where it was argued that the collaboration of samples improves the accuracy of the algorithm rather than the sparsity.
$l_2$-norm minimization can be used to find the solution. In [33], this viewpoint has been confirmed. Simultaneously, Zhang et al. also pointed out a disadvantage of CR, i.e., more classes (and thus more samples) for CR will make the solution more unstable.

Li and Du proposed a hyperspectral anomaly detection algorithm based on collaborative representation (CRD) [35]. There is also a kernel version of CRD (KCRD). According to Li and Du [35], CRD algorithm is simpler and more efficient than other detectors. However, the performance of CRD is not optimal. On the one hand, the defect of CR makes CRD unstable; on the other hand, if the test pixel is an abnormal pixel and several pixels from surrounding pixels are similar anomalies, the judge error may occur. For the second limitation, Vafadar and Ghassemian [36] proposed CR with outlier removal anomaly detector (CRBORAD) method to improve the accuracy. This method removes outlier by disregarding the small probability pixels in the Gaussian distribution. However, the actual situation does not necessarily conform to the statistical law, the improvement of detection accuracy may be limited. Recently, the morphology-based collaborative representation detector (MCRD) was proposed in [37], which can make better use of spatial information. However, it does not consider the pollution of abnormal pixels and instability of CR, and there are many parameters to be tuned.

Principal component analysis (PCA) is often used to reduce dimension to extract useful spectral information [38]–[40]. More often, there are thousands of pixels in an image, but they only have a few types of objects. So pixel-information is redundant. In this paper, we propose an algorithm that uses CRD with PCA for removing outlier (PCArOCR). The PCA can extract background information, thus eliminating the influence of outliers on the detection accuracy and making the result stable. Two versions of the PCArOCR are proposed: global and local. For the global PCArOCR detector, the PCA of the whole image is conducted, and then the extracted pixel information is used to represent each test pixel. For the local PCArOCR detector, a sliding dual window is adopted, and then PCA is used to extract the major background pixel information as samples to represent the test pixel. If the test pixel can be well represented, it is a normal pixel; otherwise it is claimed to be an abnormal pixel.

The main advantages of the proposed PCArOCR methods are as follows.

1) They are more practical because they do not make any assumptions (such as background pixels obey Gaussian distribution) or require any prior information.

2) Using PCA to extract few and useful information to make the result stable.

3) The proposed detectors remove the effect of abnormal pixels in background modeling and produce more accurate detection.

The remainder of this paper is organized as follows. Section II introduces the related work on CR and CRD. Section III presents the proposed global and local PCArOCR detectors. Section IV describes the experimental result of the proposed PCArOCR and several related detectors. The conclusion is drawn in Section V.

II. BACKGROUND

A. CR

CR assumes that the collaboration among atoms improves representation accuracy. For each pixel $y$ of size $b \times 1$, where $b$ is the number of bands, all the atoms from dictionary $B$ are involved in the representation so that $y \approx B\alpha$ where $\alpha$ is a weight vector. $l_2$-norm is adopted to constrain the $\alpha$, i.e., all the atoms in $B$ are assigned with similar small coefficients. The CR of $y$ by $B$ can be formulated as

$$\hat{\alpha} = \arg \min_{\alpha} \| y - B\alpha \|_2^2 + \lambda \| \alpha \|_2^2$$

where $\lambda$ is the regularization parameter. When the atoms in $B$ are too many, the solution may become unstable.

B. CRD

CRD algorithm supposed that each normal pixel $y$ can be well collaboratively represented via the surrounding pixels $X_s \in \mathbb{R}^{b \times s}$ ($s$ is the number of samples), while the abnormal pixel cannot. It can be formulated as

$$\arg \min_{\alpha} \| y - X_s\alpha \|_2^2 + \lambda \| \Gamma_y \alpha \|_2^2$$

where $\Gamma_y$ is the distance-weighted Tikhonov regularization. The reconstruction error $r_1$ can be obtained as

$$r_1 = \| y - \hat{y} \|_2 = \| y - X_s\hat{\alpha} \|_2.$$  

Consider that if $y$ is an abnormal pixel and several pixels from $X_s$ are also abnormal and similar to $y$, and then the corresponding coefficients in $\alpha$ are larger than normal pixels, the judge error is likely to happen.

C. CRBORAD

CRBORAD takes into account the effect of anomalous pixels from $X_s$ in CRD. After selecting the background pixels, the possible abnormal pixels whose value is greater than the threshold_{max} or smaller than the threshold_{min} are removed from the following representation. The threshold can be estimated as

$$\text{threshold}_{\text{max}} = \mu + 2 \times \sigma$$

$$\text{threshold}_{\text{min}} = \mu - 2 \times \sigma$$

where $\mu$ and $\sigma$ are the mean and standard deviation of samples in CRD.

III. PROPOSED METHOD

A. CRD With Global PCA Remove Outlier

Global PCArOCR exploits PCA to extract main background information as the samples for CR. The three-dimensional (3-D) hyperspectral cube with $n$ pixels is resized into a 2-D matrix $Y \in \mathbb{R}^{n \times b}$, $Y = [y_1, y_2, \ldots, y_n]^T$. The spatial-domain PCA transformation can be expressed as

$$X = W^T Y$$

where $W$ is the PCA projection matrix and $X$ is the transformed data, $X = [x_1, x_2, \ldots, x_n]^T$. The main components in $X$
contain most of the background information, and the information of anomalous pixels is contained in the minor components. We choose the first $m$ principal components in $X$ as the samples for CR, i.e., $X_m = [x_1, x_2, \ldots, x_m]$ of size $b \times m$. In this way, fewer samples are used which contain most of the information of the original data in the spatial domain. For each test pixel $y$, the objective is to find weight vector $\alpha$ such that $\|y - X_m \alpha\|^2$ is minimized under the constraint that $\|\alpha\|^2$ is also minimized. Therefore, the objective function is

$$\text{arg min}_{\alpha} \|y - X_m \alpha\|^2 + \lambda \|\alpha\|^2$$

where $\lambda$ is a regularization parameter. The solution is

$$\hat{\alpha} = (X_m^T X_m + \lambda I)^{-1} X_m^T y.$$  

(8)

If the samples used to compute $\alpha$ are different from the test pixel, then the corresponding coefficients of the representation are different. The higher the similarity between the sample and the test pixel, the greater the coefficient of the sample. The distance-weighted Tikhonov regularization [41], [42] is employed to adjust the weight vector $\Gamma_y$

$$\Gamma_y = \begin{bmatrix} \|y - x_1\|_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \|y - x_m\|_2 \end{bmatrix}$$

where $x_1, x_2, \ldots, x_m$ are the columns of $X_m$. The new optimization problem becomes

$$\text{arg min}_{\alpha} \|y - X_m \alpha\|^2 + \lambda \|\Gamma_y \alpha\|^2$$

(9)

which means that the Euclidean distance is crucial to the penalty values. Actually, if the sample and the test pixel are not similar and the distance between them is large, then the coefficient needs to be small; otherwise, the coefficient is large. Moreover, Tikhonov regularization reinforces the capability or incapability of samples to represent the pixel under test on the PCA basis. Similarly, the solution to (10) is

$$\hat{\alpha} = (X_m^T X_m + \lambda \Gamma_y) \Gamma_y^{-1} X_m^T y.$$  

(11)

Then, the reconstruction residual is

$$r_1 = \|y - \hat{y}\|_2 = \|y - X_m \hat{\alpha}\|_2.$$  

(12)

If $r_1$ is larger than a threshold, then the test pixel $y$ is claimed to an anomalous pixel. The overall description of the Global PCAroCRD algorithm is given as Algorithm 1.

**Algorithm 1:** The Global PCAroCRD Algorithm.

**Input:** 3-D hyperspectral cube, regularization parameter $\lambda$, and the number of principal components $m$

**Compute:**
1) data $X$ is transformed by (6);
2) the sample matrix $X_m$ is obtained.

**for all pixels do**
1) $\Gamma_y$ is calculated by (9);
2) the weight vector $\alpha$ is calculated by (11);
3) the residual is calculated by (12) and compared with the threshold.

**end for**

**Output:** Anomaly detection map.

**Algorithm 2:** The Local PCAroCRD Algorithm.

**Input:** 3-D hyperspectral cube, regularization parameter $\lambda$, window size $(w_{out}, w_{in})$, and the number of principal components $m$

**for all pixels do**
1) For each test pixel $y$, a matrix $Y$ is collected based on the dual window;
2) the transformed data $X$ is calculated by (6);
3) $\Gamma_y$ and $X_m$ are obtained;
4) the weight vector $\alpha$ is calculated by (11);
5) the Euclidean distance is calculated by (12) and compared with the threshold.

**end for**

**Output:** Anomaly detection map.

The overall description of the local PCAroCRD algorithm is given as Algorithm 2, and its flowchart is shown in Fig. 1.

IV. EXPERIMENTS

In order to illustrate the performance of the proposed PCAroCRD, three real datasets experiments were conducted. Seven typical detectors were selected in the experiment, such as global RX, local RX, CRD, KCRD, CRBORAD, MCRD, global and local PCAroCRD. The detection performance of each algorithm is displayed through three real datasets. The parameter setting of the proposed algorithm was discussed, and two real datasets were used to verify that the proposed algorithm can resist the effect of abnormal pixels on the detection result, and the detection performance of PCAroCRD more stable when the number of samples is large, while the CRD and CRBORAD cannot.

A. Hyperspectral Datasets

The first dataset was obtained from the Airborne Visible/Infrared Imaging Spectrometer covering the Moffett Field, CA, USA, at the southern end of the San Francisco Bay on August 20, 1992. The spatial resolution is approximately 20 m. The scene consists of 512 × 512 pixels with 224 bands spanning the wavelength interval of 0.4–2.5 $\mu$m. In the experiment, a subimage of size $80 \times 80$ is selected, including the anomalies.
Fig. 1. Schematic illustration of local PCAroCRD for hyperspectral anomaly detection.

Fig. 2. (a) Pseudocolor image of the Moffett Field scene. (b) Ground-truth map of anomalous pixels.

Fig. 3. (a) Pseudocolor image of the HYDICE urban scene. (b) Ground-truth map of anomalous pixels.

Fig. 4. (a) Pseudocolor image of the HyMap scene. (b) Ground-truth map of anomalous pixels.

The second dataset was collected by HYDICE airborne sensor covering the urban area. This scene consists of 80 × 100 pixels with 175 bands after removal of water-absorption bands. The spatial resolution is approximate 1 m. There are 21 anomalous pixels, representing cars and roof. The scene and the ground truth map of anomalies are illustrated in Fig. 3.

The third dataset was acquired by the HyMap airborne hyperspectral imaging sensor covering one area of Cooke City, MT, USA on July 4, 2006, with the spatial size of 200 × 800 and 126 spectral bands spanning the wavelength interval of 0.4–2.5 μm. The spatial resolution is approximately 3 m. This image has seven types of targets that are four fabric panel targets and three vehicle targets. In the experiment, a subimage of 68 × 238 was cropped, including all the targets as presented in Fig. 4.

B. Experimental Setting

The performance of detection can be evaluated with receiver operating characteristic (ROC) curves and the area under the ROC curve (AUC) [44]. The larger the area under the ROC curve, the better the algorithm. AUC (%) is adopted to measure the detection accuracy. In order to compare the maximum detection performance of algorithms, the optimized parameters such as λ and window size (w_out, w_in) are adopted. They were obtained when other parameters remain unchanged and the AUC is the maximum. In the experiment, the parameters of seven algorithms for three real datasets are set as Tables I–III. And the best detection results in the Tables I–III have been shown in bold, as in the other tables.

C. Detection Performance

According to Tables I–III, the ROC curves of the seven algorithms for three real datasets are displayed in Figs. 5–7. For
TABLE I
PARAMETERS SETTING FOR MOFFETT FIELD WHEN THE DETECTORS PERFORM BEST

<table>
<thead>
<tr>
<th>Detectors</th>
<th>Window size</th>
<th>Parameter $\lambda$</th>
<th>AUC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local RX</td>
<td>(15,13)</td>
<td></td>
<td>79.12</td>
</tr>
<tr>
<td>CRD</td>
<td>(13,11)</td>
<td>1e-4</td>
<td>94.46</td>
</tr>
<tr>
<td>KCRD</td>
<td>(13,11)</td>
<td>1e-4</td>
<td>94.40</td>
</tr>
<tr>
<td>CRBORAD</td>
<td>(13,11)</td>
<td>1e-4</td>
<td>96.44</td>
</tr>
<tr>
<td>MCRD</td>
<td>(13,11)</td>
<td>1e-6</td>
<td>98.76</td>
</tr>
<tr>
<td>Global PCAroCRD</td>
<td></td>
<td>1e-4</td>
<td>99.18</td>
</tr>
<tr>
<td>Local PCAroCRD</td>
<td>(15,7)</td>
<td>1e-1</td>
<td>99.32</td>
</tr>
</tbody>
</table>

TABLE II
PARAMETERS SETTING FOR HYDICE URBAN WHEN THE DETECTORS PERFORM BEST

<table>
<thead>
<tr>
<th>Detectors</th>
<th>Window size</th>
<th>Parameter $\lambda$</th>
<th>AUC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global RX</td>
<td>(9.7)</td>
<td></td>
<td>98.57</td>
</tr>
<tr>
<td>Local RX</td>
<td>(9.7)</td>
<td>1e-5</td>
<td>99.65</td>
</tr>
<tr>
<td>CRD</td>
<td>(9.7)</td>
<td>1e-6</td>
<td>99.84</td>
</tr>
<tr>
<td>KCRD</td>
<td>(9.7)</td>
<td>1e-6</td>
<td>99.82</td>
</tr>
<tr>
<td>CRBORAD</td>
<td>(11.7)</td>
<td>1e-6</td>
<td>99.83</td>
</tr>
<tr>
<td>MCRD</td>
<td>(9.7)</td>
<td>1e-6</td>
<td>99.82</td>
</tr>
<tr>
<td>Global PCAroCRD</td>
<td></td>
<td>1e-4</td>
<td>99.55</td>
</tr>
<tr>
<td>Local PCAroCRD</td>
<td>(9.7)</td>
<td>1e-5</td>
<td>99.81</td>
</tr>
</tbody>
</table>

TABLE III
PARAMETERS SETTING FOR HYMAP WHEN THE DETECTORS PERFORM BEST

<table>
<thead>
<tr>
<th>Detectors</th>
<th>Window size</th>
<th>Parameter $\lambda$</th>
<th>AUC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global RX</td>
<td>(11.3)</td>
<td></td>
<td>64.58</td>
</tr>
<tr>
<td>Local RX</td>
<td>(11.3)</td>
<td></td>
<td>67.81</td>
</tr>
<tr>
<td>CRD</td>
<td>(15.11)</td>
<td>1e-7</td>
<td>75.56</td>
</tr>
<tr>
<td>KCRD</td>
<td>(13.11)</td>
<td>1e-5</td>
<td>74.09</td>
</tr>
<tr>
<td>CRBORAD</td>
<td>(15.13)</td>
<td>1e-8</td>
<td>76.30</td>
</tr>
<tr>
<td>MCRD</td>
<td>(13.11)</td>
<td>1e-6</td>
<td>78.07</td>
</tr>
<tr>
<td>Global PCAroCRD</td>
<td></td>
<td>1e-2</td>
<td>79.23</td>
</tr>
<tr>
<td>Local PCAroCRD</td>
<td>(11.9)</td>
<td>1e1</td>
<td>90.57</td>
</tr>
</tbody>
</table>

Fig. 5. For the AVIRIS Moffett Field data, detection performance of the Local RX, CRD, KCRD, CRBORAD, MCRD, Global PCAroCRD, and Local PCAroCRD.

Fig. 6. For the HYDICE urban image data, detection performance of the Global RX, Local RX, CRD, KCRD, CRBORAD, MCRD, Global PCAroCRD, and Local PCAroCRD.

Fig. 7. For the HyMap image data, detection performance of the Global RX, Local RX, CRD, KCRD, CRBORAD, MCRD, Global PCAroCRD, and Local PCAroCRD.

The first dataset, Fig. 5 shows the performance of local RX, CRD, KCRD, CRBORAD, MCRD, global PCAroCRD, and local PCAroCRD. The parameters of each algorithm are set as detailed in Table I. For the proposed algorithm, when the number of principal components changes from 1 to 20, taking the best result as the final performance of the algorithm. Fig. 5 shows that global/local PCAroCRD outperforms all other detectors which is expected because the global/local PCAroCRD can resist the defect of CR and anomalous pixels from the background. It is obvious that the proposed algorithm is better than CRBORAD at removing the outlier from the background. Compared to other algorithms, the local PCAroCRD always yields high probability of detection for all false alarm rate and the probability of detection is 1, when the false alarm rate is about 0.1.
For the second HYDICE urban dataset, anomaly detection is relatively easier. In order to observe the differences of seven algorithms, we zoom in the ROC curve and display false alarm rate values between 0 and 0.1 for the x-axis. As shown in Fig. 6, CRD, KCRD, CRBORAD, MCRD, and local PCAroCRD perform better than global RX, local RX, and global PCAroCRD. The probability of detection of CRD, KCRD, CRBORAD, and local PCAroCRD is 1 when the false alarm rate is about 0.01. The probability of detection of local RX and MCRD is 1 when the false alarm rate is 0.02–0.03. The probability of detection of global PCAroCRD is 1 when the false alarm rate is about 0.06. The probability of detection of global RX is less than 1 when the false alarm rate is 0.1. The optimal parameters of each algorithm are shown in Table II, and AUC (%) values for seven detectors (global RX, local RX, CRD, KCRD, CRBORAD, MCRD, global, and local PCAroCRD) are 98.57, 99.65, 99.84, 99.82, 99.83, 99.82, 99.55, and 99.81, respectively.

To evaluate the statistical significance of differences with the proposed methods, the nonparametric McNemar test was employed to HYDICE dataset. For two methods to be compared, let \( d_{11} \) denote the number of targets that both methods can correctly detected, \( d_{22} \) is the number of targets that both cannot, \( d_{12} \) is the number of targets detected by method 1 but not method 2, and \( d_{21} \) is the number of targets detected by method 2 but not method 1. Then, the McNemar’s test statistic for these two methods can be defined as

\[
\frac{d_{12} - d_{21}}{\sqrt{d_{12} + d_{21}}}.
\]

When \(|z| > 1.96\), the difference between the two methods are considered to be statistically significant (5% level of significance). If the local PCAroCRD is detector 1 and CRD is detector 2, the \( z = 0 \) when false alarm rate is 0.01, and the \( z = -1 \) when false alarm rate is 0.001. Therefore, we think the performance of the two algorithms is similar.

For the third dataset, the best detection performance of seven algorithms is illustrated in Fig. 7. The optimal parameters for the seven detectors are shown in Table III. In Fig. 7, the AUC of the local PCAroCRD is the greatest. The global PCAroCRD exhibits higher probability of detection only when the false alarm is about 0.17 to 0.6. The global and local RX perform poorly, and the CRD, KCRD, and CRBORAD perform similarly. The MCRD outperforms CRD, KCRD, and CRBORAD when the false alarm is about 0.1–0.5. According to Table III, the AUC(%) values for seven detectors (global RX, local RX, CRD, KCRD, CRBORAD, MCRD, and global and local PCAroCRD) are 64.58, 67.81, 75.56, 74.09, 76.30, 78.07, 79.23, and 90.57, respectively. Obviously, the performance of local PCAroCRD is the best, and the global PCAroCRD is better than other algorithms.

D. Parameter Analysis

The proposed algorithm has four parameters, i.e., window size \((w_{\text{out}}, w_{\text{in}})\), regularization parameter \(\lambda\) and the number of principal components \(m\). Therefore, it is necessary to discuss the effect of each parameter on the accuracy of the algorithm.

For the proposed algorithm, in order to verify the ability of removing the abnormal pixels in background and making the detection stable when the number of samples is larger, dual windows of different sizes \((w_{\text{out}} = 15, w_{\text{in}} = 3, 5, 7, 9, 11, 13; w_{\text{out}} = 13, w_{\text{in}} = 3, 5, 7, 9, 11)\) for Moffett Field and HYDICE urban were selected in the experiment. In the experiment, the parameter \(\lambda\) is optimal. As shown in Fig. 8 (a) and (b), with the increase of the background pixels, the detection performance of CRD and CRBORAD is greatly affected, while the local PCAroCRD is very small. For the detection accuracy, local PCAroCRD is better than CRBORAD and CRD except the size of dual window is \((13, 9)\). This is mainly related to the distribution of background pixels and abnormal pixels.
However, in general, the local PCARoCRD has high stability and accuracy.

The AUC was computed to evaluate the performance of the proposed local PCARoCRD and CRD with varying window size \((w_{\text{out}}, w_{\text{in}})\) and \(\lambda\). For the first dataset, as shown in Tables IV and V, with the change of parameters, the AUC \((\%)\) of local PCARoCRD is between 95.97 and 99.32, while the AUC \((\%)\) of CRD is between 71.03 and 94.46. When the dual window size is constant, the local PCARoCRD is a little more sensitive to the parameter \(\lambda\) than the CRD. But the PCARoCRD is not sensitive to the dual window size. For the second and third datasets, we observed the same phenomenon from Tables VI to IX that the AUC \((\%)\) of local PCARoCRD is only slightly changed with the change of parameters, and the local PCARoCRD is a little sensitive to the parameter \(\lambda\) but insensitive to the dual window size. The AUC \((\%)\) of local PCARoCRD is always high for all the datasets. This is because it not only uses the spatial-domain PCA to extract several intrinsic background spectral signatures to maintain the stability, but also removes the abnormal pixels in the background. For the second and third data, the AUC \((\%)\) of global PCARoCRD is calculated when the parameter \(\lambda\) is changed. Table X is the HYDICE urban data and Table XI is the HyMap image data. It can be seen from the two tables that the global PCARoCRD is sensitive to the parameter \(\lambda\).

The influence of the principal component selection on the detection results was also studied. In Fig. 9, the detection accuracy of the local PCARoCRD is almost constant after the number of principal components is greater than 2, 4, and 6 in the Moffett Field, HYDICE, and HyMap datasets, respectively. In Fig. 10, the detection accuracy of the global PCARoCRD is almost constant for HYDICE urban dataset when the number of principal components is 4, and it remains stable for all the number of principal components for Moffett Field and HyMap experiments.
TABLE IX
AUC PERFORMANCE (%) OF LOCAL PCAroCRD WITH VARYING WINDOW SIZE ($w_{\text{win}}, w_{\text{out}}$) AS WELL AS $\lambda$ FOR THE HYMap IMAGE DATA

<table>
<thead>
<tr>
<th>Window size</th>
<th>$10^{4}$</th>
<th>$10^{5}$</th>
<th>$10^{6}$</th>
<th>$10^{7}$</th>
<th>$10^{8}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9,7)</td>
<td>89.11</td>
<td>89.55</td>
<td>87.34</td>
<td>82.85</td>
<td>78.54</td>
</tr>
<tr>
<td>(11,7)</td>
<td>85.62</td>
<td>85.79</td>
<td>87.05</td>
<td>84.07</td>
<td>79.28</td>
</tr>
<tr>
<td>(11,9)</td>
<td>89.65</td>
<td>90.57</td>
<td>89.77</td>
<td>84.64</td>
<td>79.50</td>
</tr>
<tr>
<td>(13,7)</td>
<td>82.30</td>
<td>82.19</td>
<td>84.00</td>
<td>84.91</td>
<td>80.41</td>
</tr>
<tr>
<td>(13,9)</td>
<td>86.51</td>
<td>87.13</td>
<td>88.03</td>
<td>85.18</td>
<td>80.47</td>
</tr>
<tr>
<td>(13,11)</td>
<td>88.39</td>
<td>88.70</td>
<td>88.25</td>
<td>84.64</td>
<td>80.17</td>
</tr>
<tr>
<td>(15,7)</td>
<td>80.26</td>
<td>77.29</td>
<td>76.51</td>
<td>79.29</td>
<td>78.33</td>
</tr>
<tr>
<td>(15,9)</td>
<td>80.77</td>
<td>77.46</td>
<td>78.83</td>
<td>81.60</td>
<td>79.08</td>
</tr>
<tr>
<td>(15,11)</td>
<td>83.08</td>
<td>83.63</td>
<td>85.64</td>
<td>85.44</td>
<td>80.95</td>
</tr>
<tr>
<td>(15,13)</td>
<td>87.89</td>
<td>88.26</td>
<td>88.79</td>
<td>86.15</td>
<td>80.76</td>
</tr>
</tbody>
</table>

TABLE X
AUC PERFORMANCE (%) OF GLOBAL PCAroCRD WITH VARYING $\lambda$ FOR THE HYMap IMAGE DATA

<table>
<thead>
<tr>
<th>Parameter $\lambda$</th>
<th>$10^{-5}$</th>
<th>$10^{-4}$</th>
<th>$10^{-3}$</th>
<th>$10^{-2}$</th>
<th>$10^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global PCAroCRD</td>
<td>99.55</td>
<td>99.55</td>
<td>99.54</td>
<td>98.72</td>
<td>89.47</td>
</tr>
</tbody>
</table>

TABLE XI
AUC PERFORMANCE (%) OF GLOBAL PCAroCRD WITH VARYING $\lambda$ FOR THE HYMap IMAGE DATA

<table>
<thead>
<tr>
<th>Parameter $\lambda$</th>
<th>$10^{-5}$</th>
<th>$10^{-4}$</th>
<th>$10^{-3}$</th>
<th>$10^{-2}$</th>
<th>$10^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global PCAroCRD</td>
<td>79.01</td>
<td>79.23</td>
<td>79.26</td>
<td>79.15</td>
<td>77.74</td>
</tr>
</tbody>
</table>

Fig. 9. Performance of the local PCAroCRD for three datasets when the number of principal components is in 1–20.

Fig. 10. Performance of the global PCAroCRD for two datasets when the number of principal components is in 1–20.

E. Computing Time

The computational complexity of the seven detectors run in a personal computer with 2.7 GHz CPU and 8.0 GB memory was reported in Table XII. All the experiments were carried out in MATLAB software. We can see that the global PCAroCRD and MCRD are more time-consuming than other detectors as expected. Global RX detectors can save significant amount of time. CRD, CRBORAD, and local PCAroCRD took about the same amount of time.

<table>
<thead>
<tr>
<th>Detectors</th>
<th>Experiment data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moffett Field</td>
<td>- 0.16 0.25</td>
</tr>
<tr>
<td>HYMap image</td>
<td>23.91 21.33</td>
</tr>
<tr>
<td>HYDICE Urban</td>
<td>32.28 3.20</td>
</tr>
<tr>
<td>Global RX</td>
<td>22.83 12.46</td>
</tr>
<tr>
<td>Local RX</td>
<td>5.55 3.28</td>
</tr>
<tr>
<td>CRD</td>
<td>22.83 12.46</td>
</tr>
<tr>
<td>MCRD</td>
<td>64.30 49.52</td>
</tr>
<tr>
<td>Global PCAroCRD</td>
<td>176.35 92.15</td>
</tr>
<tr>
<td>Local PCAroCRD</td>
<td>20.22 3.92</td>
</tr>
</tbody>
</table>

TABLE XII
COMPUTING TIME FOR ALL OF THE EXPERIMENT DATA SETS (IN SECONDS)

V. CONCLUSION

In this paper, an improved CR approach for hyperspectral anomaly detection is proposed. This algorithm uses CR detectors with PCA remove outlier (PCAroCRD) and has two versions: global and local PCAroCRD. The proposed algorithm can greatly reduce the errors of judgment when the test pixel is anomalous and several samples from background are similarly anomalous, and the proposed algorithm can remain stable when the number of samples for representation becomes large. The experimental results showed that the performance of the proposed algorithm is better than several widely used detectors.

REFERENCES


Zhaoyue Wu received the B.E. degree in surveying and mapping engineering from the School of Earth Sciences and Engineering, Hohai University, Nanjing, China, in 2018. She is currently working toward the M.E. degree in photogrammetry and remote sensing at the School of Earth Sciences and Engineering, Hohai University, Nanjing, China. Her research interests include hyperspectral remote sensing imagery, collaborative representation, and machine learning in hyperspectral remote sensing imagery target detection.

Qian Du (S’98–M’00–SM’05–F’18) received the Ph.D. degree in electrical engineering from the University of Maryland, Baltimore, MD, USA, in 2000. She is currently the Bobby Shackouls Professor with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA. She is also an Adjunct Professor with the College of Surveying and Geo-informatics, Tongji University, Shanghai, China. Her research interests include hyperspectral remote sensing image analysis and applications, pattern classification, data compression, and neural networks.

Peijun Du (M’07–SM’12) received the Ph.D. degree in geodesy and survey engineering from the China University of Mining and Technology, Xuzhou, China, in 2001. He is currently a Professor in remote sensing and geographical information science with Nanjing University, Nanjing, China. He has authored more than 70 articles in international peer-reviewed journals, and more than 100 papers in international conferences and Chinese journals. His research interests include remote sensing image processing and pattern recognition, hyperspectral remote sensing, and applications of geospatial information technologies.

Dr. Du is a Fellow of the SPIE-International Society for Optics and Photonics. She was the recipient of the 2010 Best Reviewer Award from the IEEE Geoscience and Remote Sensing Society. She was a Co-Chair of the Data Fusion Technical Committee of the IEEE Geoscience and Remote Sensing Society from 2009 to 2013, and the Chair of the Remote Sensing and Mapping Technical Committee of the International Association for Pattern Recognition from 2010 to 2014. She has served as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, the Journal of Applied Remote Sensing, and the IEEE SIGNAL PROCESSING LETTERS. Since 2016, she has been the Editor-in-Chief for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. She was the General Chair of the fourth IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Shanghai, in 2012.