

Article

Anomaly Detection in Hyperspectral Imagery Based on Low-Rank Representation Incorporating a Spatial Constraint

Kun Tan ^{1,2,3,†} , Zengfu Hou ^{3,†}, Donglei Ma ³, Yu Chen ^{3,*} and Qian Du ⁴

¹ Key Laboratory of Geographic Information (Ministry of Education), East China Normal University, Shanghai 200241, China

² School of Geographic Sciences, East China Normal University, Shanghai 200241, China

³ Key Laboratory for Land Environment and Disaster Monitoring of NASG, China University of Mining and Technology, Xuzhou 221116, China

⁴ Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762, USA

* Correspondence: chenyu@cumt.edu.cn; Tel.: +86-021-5434-1227

† These authors contributed equally to this work.

Received: 7 May 2019; Accepted: 27 June 2019; Published: 3 July 2019



Abstract: Hyperspectral imagery contains abundant spectral information. Each band contains some specific characteristics closely related to target objects. Therefore, using these characteristics, hyperspectral imagery can be used for anomaly detection. Recently, with the development of compressed sensing, low-rank-representation-based methods have been applied to hyperspectral anomaly detection. In this study, novel low-rank representation methods were developed for anomaly detection from hyperspectral images based on the assumption that hyperspectral pixels can be effectively decomposed into a low-rank component (for background) and a sparse component (for anomalies). In order to improve detection performance, we imposed a spatial constraint on the low-rank representation coefficients, and single or multiple local window strategies was applied to smooth the coefficients. Experiments on both simulated and real hyperspectral datasets demonstrated that the proposed approaches can effectively improve hyperspectral anomaly detection performance.

Keywords: anomaly detection; hyperspectral; low-rank representation; local window; spatial constraint

1. Introduction

Due to the high spectral resolution and abundant spectral information in a large number of spectral bands, hyperspectral image (HSI) data have been widely used to distinguish different targets on the ground. According to the availability of target information, target detection algorithms can be divided into two categories: with or without a priori target knowledge [1]. Since target information is often difficult to obtain in practice [2], anomaly detection does not require any a priori information of targets and is of great importance in real applications [3,4].

Anomaly detection involves modeling the background and using the difference between the pixels and the background to detect anomalous pixels. Many different anomaly detection algorithms have been proposed. The well-known Reed–Xiaoli (RX) algorithm, considered to be a classical detection algorithm [5], is a second-order matched filtering algorithm, for which the similarity between the test pixel and the background is calculated by Mahalanobis distance. When the entire image is considered for background modeling, this is known as the global RX detector (GRXD). If the RX detector estimates the background using local statistics, it is referred to as the local RX detector (LRXD) [6]. However, neither of these two algorithms can exclude the influence of anomalous features on covariance. Until recently, many researchers have made contributions to anomaly detection. Billor et al. proposed the

blocked adaptive computationally efficient outlier nominator (BACON) [7] algorithm, which tries to select the growing subset of outlier-free points and obtain robust background statistics to weaken the contamination by anomalies. Imani proposed the median–mean line RX detector (MML-RXD) [8], in which the benefits of the median–mean metric are used to provide more reliable background samples. Yao et al. were inspired by the bilateral filter in digital image processing and proposed the bilateral-filter-based anomaly detector (BFDA) [9], which substitutes the weighted surrounding neighbors for centric pixels to remove anomalous points. Some enhanced RX-based algorithms such as the weighted-RXD (WRXD) [10] and the linear filter-based RXD [10] were also proposed. They aim at increasing the probability of anomaly detection by improving the estimation of background statistics to exclude the influence of anomalous features on covariance.

Generally speaking, in real hyperspectral imagery, the background information is very complicated and cannot be described with just the covariance matrix and mean vector. To avoid obtaining the accurate covariance matrix of the background, some linear-representation-based methods have been successfully applied to hyperspectral anomaly detection. These include the unsupervised nearest regularized subspace (UNRS) [11] detector proposed by Li et al. and the collaborative-representation-based detector (CRD) [12] for anomaly detection, both of which are based on the assumption that the spectrum of the central pixel is similar to that of the surrounding pixel and can be linearly represented by the surrounding pixels. To make full use of the spatial information of neighboring pixels, the two detectors adopt a dual window strategy, which leads to the detection accuracy being affected by the size of the internal and external windows.

Unlike UNRS and CRD, with the development of compressed sensing, sparse-representation-based methods have been extensively researched. By using a small number of similar signals to represent signals, noise and anomalous signal information can be effectively suppressed. Some sparse-representation-based methods have been applied to hyperspectral image anomaly detection. For instance, the sparse representation detector (SRD) was proposed based on the assumption that hyperspectral pixels can be well represented by only a few pixels [13,14]. Sparse representation theory has a good application in the processing and representation of one-dimensional information, but for two-dimensional matrix data, sparsity can only represent the structural information of the dataset.

Due to the redundancy and diversity of the observed data, two-dimensional matrix data are presented with low-rank characteristics. By performing a low-rank constraint on the matrix, the global structure information of the two-dimensional matrix can be better described. Low-rank representation (LRR) can be considered as a special case of sparse representation, extending from one-dimensional sparsity to two-dimensional low rank.

Some low-rank representation models have also been applied to hyperspectral anomaly detection, such as the classical robust principal component analysis (RPCA) [15], which decomposes the matrix into low-rank and sparse matrices. Chen et al. [16] proposed the RPCA-RX anomaly detection method, with the classic RX detector being applied to the sparse matrix. Unlike RPCA, the LRR [1,17] model assumes that the data are drawn from multiple subspaces, which is better suited for hyperspectral images due to the complex background features of real data. Because of the correlation between the representation coefficients, the low-rank constraint of the matrix can effectively describe the global structure of the dataset. Therefore, Xu et al. [18] proposed the low-rank representation sum-to-one (LRRSTO) anomaly detection model. In order to make the representation coefficients robust, a sum-to-one constraint was added.

In recent years, sparse subspace clustering (SCC) theory has attracted considerable attention and become a research hotspot due to its superiority. Zhai et al. first introduced the SSC [19] model to HSIs based on the assumption that pixels belonging to the same land-cover class approximately lie in the same subspace. The SSC algorithm generally calculates the affinity matrix by solving the sparse representation coefficients and uses the spectral clustering algorithm to segment the data into different subspaces.

However, directly applying the SSC algorithm to HSIs fails to take advantage of the rich spectral and spatial information of HSIs. In light of this, the SSC incorporating spatial information (SSC-S) [20] model was proposed. Based on the hypothesis that the background pixels in the neighborhood window have the same background as the central testing pixel, Zhang et al. hypothesized that the representation coefficients of the neighborhood pixels in the local window should also be similar to the central testing pixel. In the SSC-S model, they eliminated the influence of noise and other errors by constraining the space representation coefficient to enhance the clustering effect.

In this study, in order to make full use of spectral similarity in a local spatial neighborhood of HSI, we were inspired by the SSC-S algorithm to introduce a spectral constraint term to the LRRSTO model. We then developed two novel anomaly detection methods based on single or multiple local windows and the low-rank representation sum-to-one model (SLW_LRRSTO/MLW_LRRSTO).

Unlike traditional LRRSTO models, we added a spatial constraint to the LRRSTO model in our SLW_LRRSTO methods. Based on the hypothesis that the background pixels in the neighborhood window share the same background as the central testing pixel, we used a single local window to smooth the LRR coefficients matrix. For the MLW_LRRSTO methods, we adopted multiple local windows to enhance the utilization of spatial information. However, in the LRRSTO model, the overcomplete dictionary composed of data itself contains a large number of redundant dictionary atoms, which greatly slows down the speed of the algorithm. In order to accelerate the operation, in the proposed SLW_LRRSTO/MLW_LRRSTO models, we adopted the random selection method to construct the initial dictionary.

2. Related Work

2.1. LRRSTO Detection Algorithm

LRR can effectively describe the global correlation of the observed HSI. Based on the LRR, the HSI can be decomposed into a background part and a sparse part. The anomaly information can then be detected in the sparse part. Xu et al. [18] applied the sum-to-one constraint to enhance the robustness of the representation coefficient in the LRRSTO method, which is formulated as

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_{2,1} \quad s.t. \quad X = DZ + E, \quad Z^T \mathbf{1}_{d \times 1} = \mathbf{1}_{MN \times 1} \quad (1)$$

where $X \in R^{p \times MN}$ is the 2D HSI data matrix with p bands and MN pixels, $D \in R^{p \times d}$ is the dictionary matrix, $Z \in R^{d \times MN}$ is the LRR coefficient matrix, $Z^T \mathbf{1}_{d \times 1} = \mathbf{1}_{MN \times 1}$ is the sum-to-one constraint, and E is the sparse matrix.

2.2. SSC-S Clustering Algorithm

The SSC model [21] can be expressed as

$$\operatorname{argmin}_{Z,E} \|Z\|_1 + \frac{\lambda}{2} \|E\|_F^2 \quad s.t. \quad X = XZ + E, \quad \operatorname{diag}(C) = 0, \quad Z^T \mathbf{1} = \mathbf{1} \quad (2)$$

where $Z \in R^{MN \times MN}$ is the representation coefficient matrix, $E \in R^{p \times MN}$ is the representation error term, and λ is the tradeoff parameter. Here, $\operatorname{diag}(C) = 0$ is used to eliminate trivial solutions.

According to Tobler's first law, neighborhood pixels within a local window usually belong to the same class as the central pixel under test. They also have similar sparse coefficients to the central pixel [22]. If the mean representation coefficient vectors of the neighborhood pixels are different, the central pixel under test may be noise. Zhang et al. [20] proposed a novel SSC model incorporating spatial neighborhood information by utilizing a mean constraint for sparse representation coefficients. In this model, \tilde{Z} is utilized to regularize Z , and $\frac{\alpha}{2} \|Z - \tilde{Z}\|_F^2 < \varepsilon$ is the spatial constraint, which limits

the difference between the test pixel and the mean of its neighborhood pixels. The representation coefficient matrix can then be obtained by solving the following optimization problem:

$$\min_{\mathbf{Z}, \tilde{\mathbf{Z}}, \mathbf{E}} \|\mathbf{Z}\|_1 + \frac{\alpha}{2} \|\mathbf{Z} - \tilde{\mathbf{Z}}\|_F^2 + \frac{\lambda}{2} \|\mathbf{E}\|_F^2 \quad s.t. \quad \mathbf{X} = \mathbf{XZ} + \mathbf{E}, \text{diag}(\mathbf{C}) = 0, \mathbf{Z}^T \mathbf{1} = \mathbf{1} \quad (3)$$

where $\frac{\alpha}{2} \|\mathbf{Z} - \tilde{\mathbf{Z}}\|_F^2 < \varepsilon$ is the spatial constraint, and α and λ are the regularization parameters.

2.3. The Local Summation Anomaly Detection (LSAD) Algorithm

A circular window is utilized to calculate the covariance matrix and the mean vector for the pixels that are between the outer and inner windows in the LRRXD [23]. However, if there are anomalous pixels in the frame, the local background distribution is not the best representation for the testing pixel. Du et al. [24] proposed a novel local summation anomaly detection (LSAD) method which exploits a second-order Mahalanobis distance statistical feature and a multiple local window filter to establish a local summation anomaly detection strategy.

3. Proposed Methods

3.1. Single Local Window

In the LRRSTO model, the spatial similarity between adjacent background pixels is not considered; that is, the neighborhood pixels of local windows have the same category as the central pixel under test. In order to make full use of the correlation information between local background pixels, we constrained the representation coefficients of central pixels by using mean filtering. Based on the LRRSTO model, the spatial constraint of the low-rank representation coefficient $\|\mathbf{Z} - \tilde{\mathbf{Z}}\|_F^2 < \varepsilon$ was used. The local distributions of the single window are obtained by the single local window filter, but it does not include the pixel under test in the detection statistics. Therefore, we proposed a novel single local window anomaly detection method based on the low-rank representation sum-to-one model (SLW_LRRSTO):

$$\operatorname{argmin}_{\mathbf{Z}, \mathbf{E}} \|\mathbf{Z}\|_* + \eta \|\mathbf{Z} - \tilde{\mathbf{Z}}\|_F^2 + \lambda \|\mathbf{E}\|_{2,1} \quad s.t. \quad \mathbf{X} = \mathbf{DZ} + \mathbf{E}, \mathbf{Z}^T \mathbf{1} = \mathbf{1} \quad (4)$$

where $\mathbf{Z}^T \mathbf{1} = \mathbf{1}$ is the sum-to-one constraint, $\|\mathbf{Z} - \tilde{\mathbf{Z}}\|_F^2$ is the spatial constraint in the single local window, \mathbf{Z} is the representation coefficient matrix, $\tilde{\mathbf{Z}}$ is the smoothing filtering result of \mathbf{Z} , and η and λ are the tradeoff parameters.

3.2. Multiple Local Background Statistics

The abovementioned SLW_LRRSTO algorithm only utilizes a single window filter of a specified size. Ideally, the local neighborhood pixels of the single local window do not contain anomalous pixels. However, when the local neighborhood contains anomalous pixels, the detector cannot obtain ideal detection performance. To solve this problem, we proposed using multiple local windows.

In Figure 1, the purple squares represent anomalous pixels, and the white squares represent the same category background pixels. Figure 1a is the ideal distribution of neighboring pixels, where the center pixel is the anomaly and the neighborhood pixels are the background. The mean value of the neighborhood pixels differs from that of the center test pixel value, which can improve the detection accuracy. Figure 1b shows that the center pixel is the anomaly, and the neighborhood contains anomalous pixels. With the mean filter window (red box), the mean value of the neighborhood cannot detect the center pixel as the anomaly. However, with the bottom-right local window filter (green box), the neighborhood pixels do not contain the anomalous pixel, and the center pixel can be detected as an anomaly. Figure 1c shows that the center pixel is a background pixel and the neighborhood pixels include an anomalous pixel in the single local window. For the single mean filter window (red box),

the mean value of the neighborhood indicates the center pixel is an anomaly. Using the bottom-right local window mean filter (green box), the neighborhood pixels do not contain the anomaly, which can prevent the center pixel from being identified as an abnormal target.

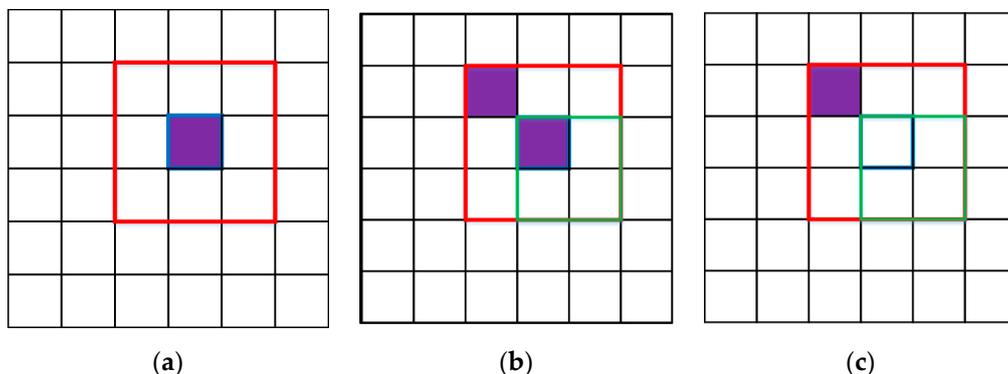


Figure 1. Analysis between the center test pixel and neighboring pixels in a single local window. The purple squares represent anomalous pixels, and the white squares represent the same category background pixels: (a) ideal distribution of neighboring pixels; (b) the neighborhood contains anomalous pixels; (c) the center pixel is a background pixel.

Based on the SLW_LRRSTO model, the representation coefficient spatial constraint of the single window is extended to the spatial constraint of multiple local windows. As shown in Figure 2, the size of the red single window is $N \times N$ (where N is odd). We used a small green window, the size of which is $L = (N + 1)/2$ to process the data in the red single window from left to right and top to bottom. $M = (N - L + 1) \times (N - L + 1)$ is the number of multiple local windows.

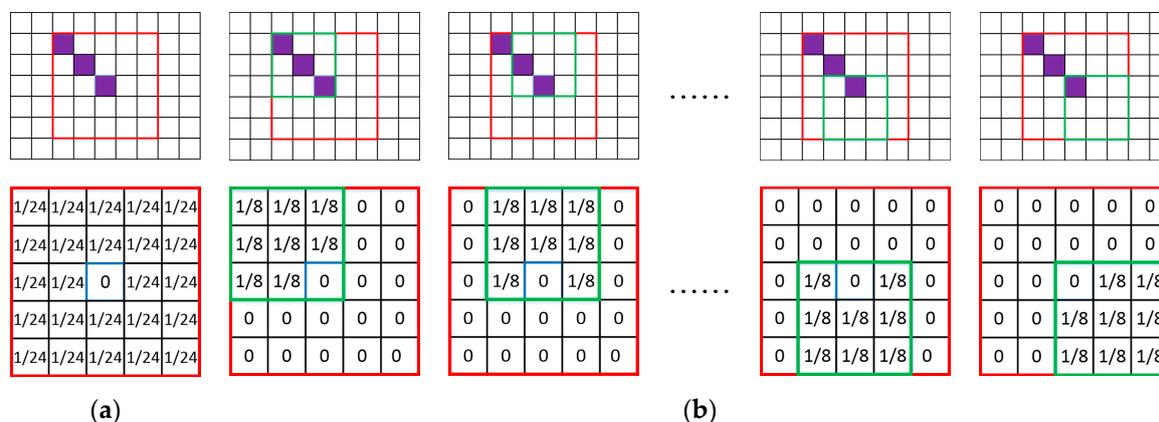


Figure 2. Schematic diagram of a single local window and multiple local windows and the corresponding filters ($N = 5$): (a) single local window; (b) multiple local windows.

Figure 2a shows a single window of size 5×5 which is expanded to nine local windows. The corresponding filter is shown in Figure 2b, where four of nine local windows contain anomalous pixels. This differs from the single local window, in which the neighborhood pixels include two anomalous pixels. On the one hand, the multiple local window strategy can effectively reduce the number of abnormal pixels in the neighborhood. On the other hand, the situation of the local window containing abnormal pixels rarely occurs.

In Formula (4), the spatial constraint of the single local window $\eta \|Z - \tilde{Z}\|_F^2$ is extended to $\|Z - \tilde{Z}_1\|_F^2 + \|Z - \tilde{Z}_2\|_F^2 + \dots + \|Z - \tilde{Z}_M\|_F^2 < \varepsilon$, because $\left\| \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix} * Z - \begin{bmatrix} \tilde{Z}_1 \\ \vdots \\ \tilde{Z}_M \end{bmatrix} \right\|_F^2 < \varepsilon$, where M is the

number of multiple local windows and I is the identity matrix. We defined $\tilde{I} = [I, I, \dots, I]^T$ and $\tilde{Z} = [\tilde{Z}_1, \tilde{Z}_2, \dots, \tilde{Z}_M]^T$. The objective function is then

$$\operatorname{argmin}_{Z,E} \|Z\|_* + \eta \|\tilde{I} * Z - \tilde{Z}_F^2\| + \|E\|_{2,1} \quad \text{s.t. } X = DZ + E, Z^T \mathbf{1} = \mathbf{1}. \tag{5}$$

We used the inexact augmented Lagrange multiplier (IALM) algorithm [1,17,25] to solve the MLW_LRRSTO model and add auxiliary variable J . The formula can be written as

$$\min_{Z,E} \|J\|_* + \eta \|\tilde{I} * Z - \tilde{J}\|_F^2 + \lambda \|E\|_{2,1} \quad \text{s.t. } X = DZ + E, Z = J, \mathbf{1}_d^T * Z = \mathbf{1}_n^T \tag{6}$$

where $\tilde{J} = \begin{bmatrix} \tilde{J}_1 \\ \vdots \\ \tilde{J}_M \end{bmatrix} = \begin{bmatrix} \tilde{Z}_1 \\ \vdots \\ \tilde{Z}_M \end{bmatrix}$ is the auxiliary variable. The augmented Lagrange function is

$$\begin{aligned} L(Z, J, E, Y^1, Y^2, Y^3, \mu) &= \|J\|_* + \beta \|\tilde{I} * Z - \tilde{J}\|_F^2 + \lambda \|E\|_{2,1} + \langle Y^1, X - AZ - E \rangle + \langle Y^2, Z - J \rangle \\ &+ \langle Y^3, \mathbf{1}_d^T * Z - \mathbf{1}_n^T \rangle + \frac{\mu}{2} (\|X - AZ - E\|_F^2 + \|Z - J\|_F^2 + \|\mathbf{1}_d^T * Z - \mathbf{1}_n^T\|_F^2) = \|J\|_* + \beta \|\tilde{I} * Z - \tilde{J}\| \\ &+ \lambda \|E\|_{2,1} + \frac{\mu}{2} (\|X - AZ - E + \frac{Y^1}{\mu}\|_F^2 + \|Z - J + \frac{Y^2}{\mu}\|_F^2 + \|\mathbf{1}_d^T * Z - \mathbf{1}_n^T + \frac{Y^3}{\mu}\|_F^2) \\ &- \frac{1}{2\mu} (\|Y^1\|_F^2 + \|Y^2\|_F^2 + \|Y^3\|_F^2) \end{aligned} \tag{7}$$

where (Y^1, Y^2, Y^3) are the Lagrange multipliers and $\mu > 0$ is the penalty parameter.

The IALM [1,17,25] is a multiple-variable optimization problem, which can be solved by alternately updating one variable by the minimizing function with the other variables fixed. The problem can be divided into the following subproblems:

Step 1: Fix (J, Z, E) and update J . The objective function can then be written as follows:

$$J = \operatorname{argmin}_J \|J\|_* + \frac{\mu}{2} \|J - \left(Z + \frac{Y^2}{\mu} \right)\|_F^2 \tag{8}$$

Step 2: Update \tilde{J}_i in each local window and use these auxiliary variables to combine \tilde{J} :

$$\tilde{J}_{k+1} = [\tilde{J}_1, \tilde{J}_2, \dots, \tilde{J}_M]^T \tag{9}$$

Step 3: Fix (J, Z, E) and update Z . The objective function can then be written as follows:

$$Z = \operatorname{argmin}_Z \beta \|\tilde{I} * Z - \tilde{J}\|_F^2 + \frac{\mu}{2} \|X - AZ - E + \frac{Y^1}{\mu}\|_F^2 + \frac{\mu}{2} \|Z - J + \frac{Y^2}{\mu}\|_F^2 + \frac{\mu}{2} \|\mathbf{1}_d^T * Z - \mathbf{1}_n^T + \frac{Y^3}{\mu}\|_F^2 \tag{10}$$

Step 4: Fix (J, Z, E) and update E . The objective function can then be written as follows:

$$E = \operatorname{argmin}_E \lambda \|E\|_{2,1} + \frac{\mu}{2} \left\| X - AZ - E + \frac{Y^1}{\mu} \right\|_F^2 \tag{11}$$

The procedure of the proposed method is summarized as follows Algorithm 1, in which Θ, Ω , and $\lambda_{2,1}$ are the singular value thresholding and $\lambda_{2,1}$ is the minimization operator.

Algorithm 1. Inexact ALM algorithm for MLW_LRRSTO**Input:** data matrix X , size N of single local window, parameters $\beta > 0$ and $\lambda > 0$ **Initialize:** $Z_0 = J_0 = \tilde{J}_0$, $E_0 = 0$, $Y_0^1 = Y_0^2 = Y_0^3 = 0$, $\mu_0 = 10^{-2}$, $\mu_{max} = 10^{10}$, $\rho = 1.1$,
 $\varepsilon_1 = 10^{-6}$, $\varepsilon_2 = 10^{-3}$ **While** $\|X - AZ_k - E_k\|_F / \|X\|_F \geq \varepsilon_1$ **or** $\mu_k \max(\|Z_k - Z_{k-1}\|_F, \|E_k - E_{k-1}\|_F, \|Z_k - J_k\|_F) / \|X\|_F \geq \varepsilon_2$ **Continue**(1) update variable J_{k+1}

$$J_{k+1} = \Theta_{\mu_k^{-1}} \left(Z_k + \frac{Y_k^2}{\mu_k} \right)$$

(2) update \tilde{J}_i , compose \tilde{J}_{k+1}

$$\tilde{J}_{k+1} = [\tilde{J}_1, \tilde{J}_2 \cdots \tilde{J}_M]^T$$

(3) update variable Z_{k+1}

$$Z = \left(2\beta \tilde{I} \tilde{I}^T + \mu A^T A + \mu I + \mu \mathbf{1}_d \mathbf{1}_d^T \right)^{-1} \left(2\beta \tilde{I} \tilde{J} + \mu A^T \left(X - E + \frac{Y^1}{\mu} \right) + \mu \left(J - \frac{Y^2}{\mu} \right) + \mu \mathbf{1}_d \left(\mathbf{1}_n^T - \frac{Y^3}{\mu} \right) \right)$$

(4) update variable E_{k+1}

$$E_{k+1} = \Omega_{\lambda \mu_k^{-1}} \left(X - AZ_{k+1} + \frac{Y_k^1}{\mu_k} \right)$$

(5) update Lagrange multipliers $Y_{k+1}^1, Y_{k+1}^2, Y_{k+1}^3$

$$Y_{k+1}^1 = Y_k^1 + \mu_k (X - AZ_{k+1} - E_{k+1});$$

$$Y_{k+1}^2 = Y_k^2 + \mu_k (Z_{k+1} - J_{k+1});$$

$$Y_{k+1}^3 = Y_k^3 + \mu_k (\mathbf{1}_d^T Z_{k+1} - \mathbf{1}_n^T);$$

(6) update penalty parameter μ_{k+1}

$$\mu_{k+1} = \min(u_{max}, \rho \mu_k) \text{ where}$$

$$\rho = \begin{cases} \rho_0, & \text{if } \mu_k \max(\|Z_{k+1} - Z_k\|, \|E_{k+1} - E_k\|, \|Z_{k+1} - J_k\|) / \|X\|_F \leq \varepsilon_2 \\ 1, & \text{otherwise} \end{cases}$$

End while**Output:** the optimal solution (J, Z, E)

4. Experiments and Analysis

In order to evaluate the effectiveness of the proposed methods, we conducted our experiments on three hyperspectral images. The first experiment was used to analyze the propriety of the proposed methods, and the other two experiments were used to demonstrate its effectiveness. It is worth noting that these three datasets were obtained after preprocessing including atmospheric correction.

4.1. Data Description

4.1.1. Simulated Hyperspectral Image

The hyperspectral data used in the simulated experiment [26,27] were collected by the airborne visible/infrared imaging spectrometer (AVIRIS) over Salinas Valley, CA, USA, which can be downloaded from the GIC website [28]. The original image comprising vegetables, bare soil, and vineyard fields had a spatial size of 512×217 pixels and was made up of 224 spectral bands in the wavelength range of 370–2510 nm. In total, 202 bands were used in the experiment after the noise and water absorption bands had been removed. The image scene in pseudocolor is shown in Figure 3a. A region with a size of 120×120 was selected to generate the simulated data and is shown in Figure 3b. The 25 specific anomalous pixels were randomly selected from the whole image, with different spectra from the subimages, and are shown in Figure 3c. Figure 3d is the corresponding anomaly location map. The

anomalous pixels were simulated by the target implantation method [29]. A simple linear mixture model was adopted for the implanted pixels as

$$z = f \times t + (1 - f) \times b \quad (12)$$

where f is the abundance fraction (ranging from 0.04 to 1), t is the specific anomaly spectrum, and b is the background spectrum.

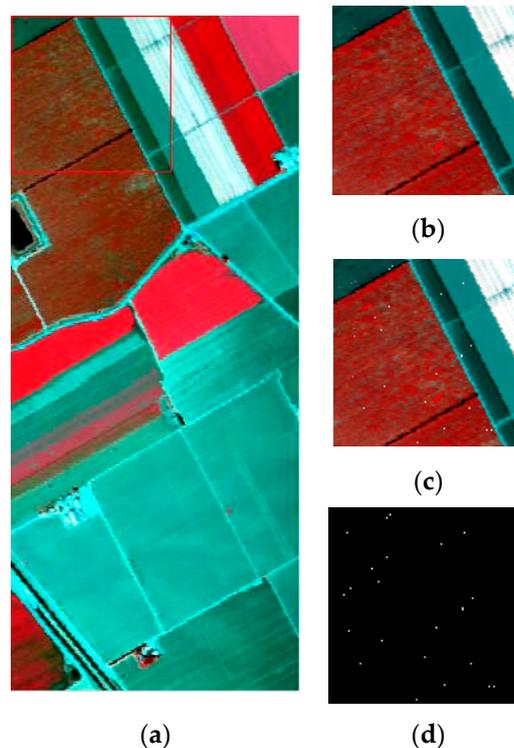


Figure 3. AVIRSI_Salinas dataset: (a) false-color image of the whole scene (R:65, G:33, B:21); (b) false-color image of the subarea; (c) false-color image of the synthetic; (d) the anomaly location map.

4.1.2. Parameters Analysis

The initial choices of different parameters were important for the proposed methods, which involve two regularization parameters: η and λ . For this simulated dataset, we adopted the target embedding method to fully control the generating environment of subpixel anomalous targets. Therefore, in order to show the superiority and fully excavate the potential of the proposed methods, we found the optimal parameters by a trial-and-error method. Figure 4 shows how the detection performance changed as the parameters were changed. The range of parameters η and λ was set to 0–2, and the specific data results are shown in Table 1. The data results in Figure 4 and Table 1 show that MLW_LRRSTO was sensitive to λ . For the simulated dataset, it achieved low area under the curve (AUC) values when λ was larger than 0.5. In this range, the average AUC value for the AVIRIS_Salinas data was 0.9815. Therefore, we empirically set $\eta = 1$ and $\lambda = 0.1$ for all the datasets in our experiments.

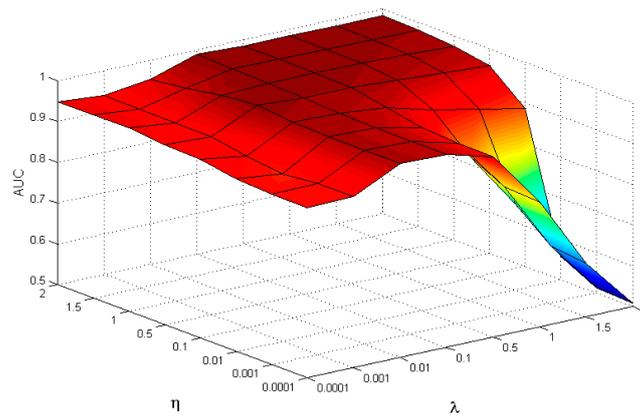


Figure 4. Results with different parameters for AVIRIS_Salinas dataset.

Table 1. Results with different parameters for the AVIRIS_Salinas data.

$\eta \backslash \lambda$	0.0001	0.001	0.01	0.1	0.5	1	1.5	2
0.0001	0.9693	0.9606	0.9516	0.9623	0.9623	0.9622	0.9620	0.9617
0.001	0.9694	0.9612	0.9517	0.9623	0.9623	0.9622	0.9620	0.9617
0.01	0.9698	0.9666	0.9530	0.9625	0.9623	0.9623	0.9620	0.9618
0.1	0.9476	0.9633	0.9640	0.9685	0.9635	0.9627	0.9623	0.9620
0.5	0.9482	0.9482	0.9804	0.9774	0.9693	0.9661	0.9645	0.9636
1	0.9486	0.9469	0.9783	0.9794	0.9737	0.9694	0.9673	0.9659
1.5	0.9487	0.9463	0.9670	0.9801	0.9760	0.9718	0.9693	0.9677
2	0.9487	0.9460	0.9624	0.9813	0.9777	0.9737	0.9710	0.9692

4.1.3. Real Hyperspectral Images

The first real experimental image [1,4,30] was obtained by the HYDICE sensor over an urban area, which can be downloaded from the website [31]. The image had a spectral resolution of 10 nm and a spatial resolution of 1 m. The whole image had a spatial size of 307×307 pixels. It comprised 210 spectral bands, and 162 bands were used in the experiment after the noise and water absorption bands had been removed. The image scene in pseudocolor is shown in Figure 5a. A region with a size of 80×100 pixels was selected as the test data in the upper-rightmost area of the scene and is shown in Figure 5b. The ground-truth map is shown in Figure 5c. Twenty-one pixels were anomalies, which were cars and roofs, because they had spectra that differed from the background.

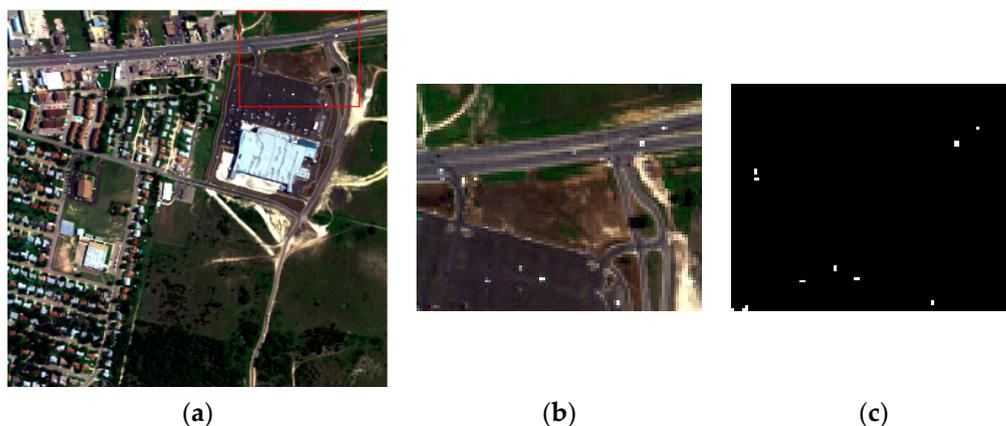


Figure 5. HYDICE_Urban dataset: (a) false-color image of the whole scene; (b) false-color image of the detection area; (c) the ground-truth map.

The second real dataset [32–34] was acquired by the AVIRIS sensor over San Diego, CA, USA. This image had a spatial resolution of 3.5 m and the whole image had a spatial size of 400×400 pixels. It comprised 224 spectral channels, and 186 channels were used in the experiment after the noise and water absorption bands had been removed. The image scene in pseudocolor is shown in Figure 6a. The top-left 100×100 of the scene was chosen as the test image, as shown in Figure 6b. The ground-truth map is shown in Figure 6c. Fifty-eight pixels were anomalies from the subimage. The airplanes were considered as anomalies, the main bodies of which were pure pixels and edges were mixed pixels.

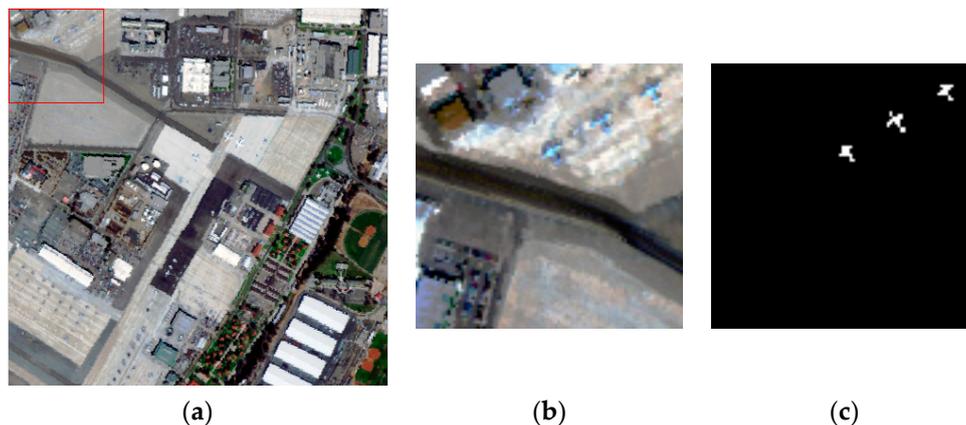


Figure 6. AVIRIS_SanDiego dataset: (a) false-color image of the whole scene; (b) false-color image of the detection area; (c) the ground-truth map.

The third real dataset was acquired during the “Viareggio 2013 trial” hyperspectral data collection campaign [35], which can be downloaded from the website [36]. Datasets in this campaign were acquired by a pushbroom hyperspectral “Sistema Iperspettrale Modulare Galileo Avionica” sensor mounted on an ultralight aircraft. In this paper, the subset of D1F12H1 was utilized. The image had a spatial size of 375×450 pixels and comprised 511 spectral bands ranging from 400 to 1000 nm. The spatial resolution and spectral resolution were 0.6 m and the 2.3 nm, respectively. The image scene in true color is shown in Figure 7a. It covers a parking lot, several sports facility buildings, and a football field in a suburban vegetated area in Viareggio, Italy. In the ground-truth map (Figure 7b), there are three vehicles, four panels, and two reference calibration tarps with 135 anomalous pixels to be detected in the scene.

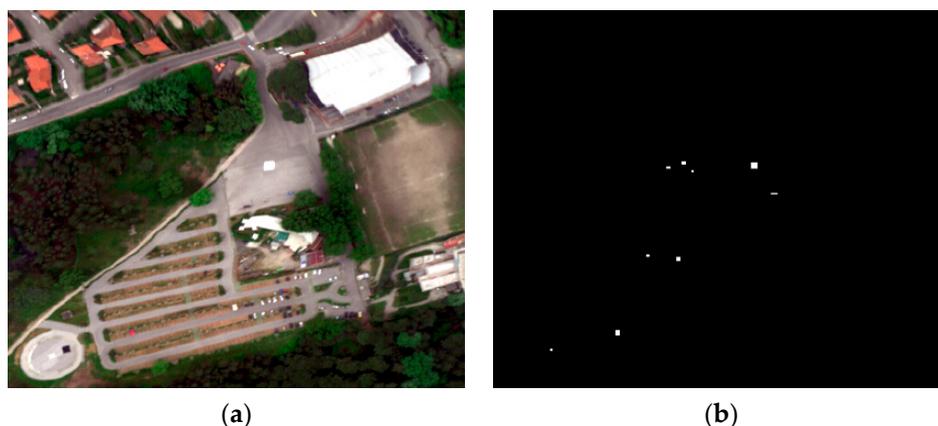


Figure 7. Viareggio dataset: (a) RGB image of the scene (R:219, G:144, B:66); (b) the ground-truth map.

4.2. Detection Performance

The proposed methods (MLW_LRRSTO and SLW_LRRSTO) were compared to the GRXD, LRXD, WRXD, RPCA-RX, and LRRSTO. In different datasets, the window size parameters were set differently

for LRXD, and the parameters of other methods remained the same. The regularization parameters for LRXD, RPCA-RX, and LRRSTO were empirical parameters in the experiments. In the LRXD, the covariance matrix size depends on the number of bands, and in order to have an accurate sample covariance matrix, the number of captured pixels should be no less than band number + 1. Therefore, in AVIRIS_Salinas, HYDICE_Urban, AVIRIS_SanDiego, and Viareggio datasets, the inner and outer window parameters (W_{out} , W_{in}) were set to (17, 9), (19, 5), (19, 9), and (25, 7), respectively, after extensive searching. Parameter λ of RPCA was set to 0.01 and parameter λ of LRRSTO was set to 0.1. The size of the single local window was set to 3 for MLW_LRRSTO and SLW_LRRSTO, and the tradeoff parameters were set as $\eta = 1$ and $\lambda = 0.1$ for MLW_LRRSTO and SLW_LRRSTO. It is worth noting that the detection results of all the methods in this paper are the average value of 10 repeated experiments. The parameters of each detection method were set as shown in Table 2.

Table 2. Parameters of the different detection methods.

Method	Parameters
GRXD	—
LRXD	(W_{out}, W_{in})
WRXD	—
RPCA_RX	$\lambda = 0.01$
LRRSTO	$\lambda = 0.1$
SLW_LRRSTO	$W_{Local} = 3, \eta = 1, \lambda = 0.1$
MLW_LRRSTO	$W_{Local} = 3, \eta = 1, \lambda = 0.1$

4.2.1. AVIRIS_Salinas Experiment

Figure 8 shows the receiver operating characteristic (ROC) curves [37]. An important observation from Figure 8 is that the performance of RPCA was clearly better than that of GRXD, LRXD, and WRXD. Compared with LRRSTO, the proposed SLW_LRRSTO and MLW_LRRSTO had a higher probability of detection in the condition of a low false alarm rate. MLW_LRRSTO was the best method in terms of the overall detection performance. Furthermore, we also computed the AUC of the ROC [37] to evaluate the performance of these methods. The results are shown in Table 3, where the proposed MLW_LRRSTO achieved the highest score, which was higher than the second-highest score from the SLW_LRRSTO. Table 3 also lists the execution time of the different methods. The GRXD and WRXD had shorter computational time, but the detection rates were lower than RPCA_RX and LRRSTO. The RPCA_RX and LRRSTO took less time than MLW_LRRSTO, while compared with it, the SLW_LRRSTO and MLW_LRRSTO had higher detection rates. The SLW_LRRSTO and MLW_LRRSTO had less computational time than LRXD, in which the inverse function of local background requires heavy computational time.

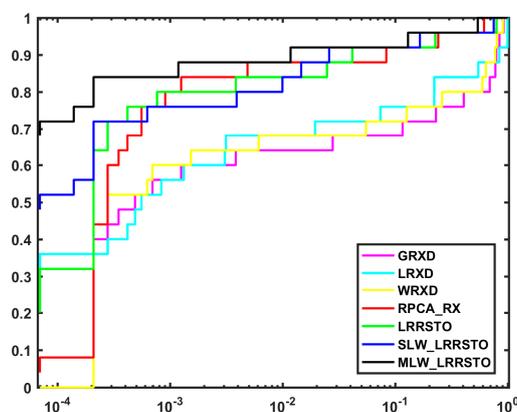


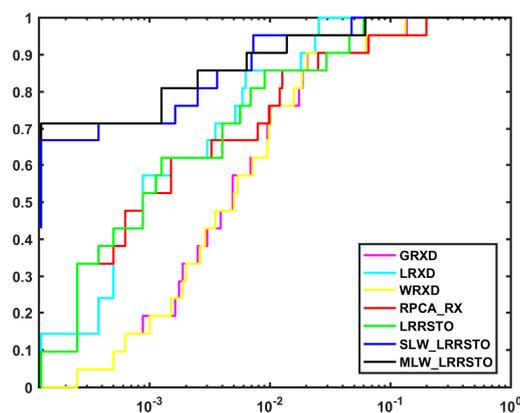
Figure 8. Receiver operating characteristic (ROC) evaluation of the different methods for the AVIRIS_Salinas data.

Table 3. Area under the curve (AUC) values for the different detectors with the AVIRIS_Salinas dataset.

Method	GRXD	LRXD	WRXD	RPCA_RX	LRRSTO	SLW_LRRSTO	MLW_LRRSTO
AUC	0.8073	0.8306	0.8332	0.9619	0.9750	0.9840	0.9854
Time/s	0.12	77.81	0.45	10.03	45.38	37.67	72.74

4.2.2. HYDICE_Urban Experiment

The ROC curves and AUC values are shown in Figure 9 and Table 4, respectively. The proposed MLW_LRRSTO and SLW_LRRSTO generated higher probability of detection when the false alarm rate was low, and they had similar ROC curves. SLW_LRRSTO achieved the highest AUC score among all the detectors, and MLW_LRRSTO obtained the second-highest score. Therefore, it can be concluded that MLW_LRRSTO and SLW_LRRSTO were more effective methods than GRXD, LRXD, WRXD, RPCA, and LRRSTO for the detection of anomalous pixels in the HYDICE_urban dataset. As shown in Table 3, according to the computational time of the different methods, we can see that GRXD, WRXD, and RPCA_RX had relatively fast operation speed and LRXD took longer to calculate. The MLW_LRRSTO took the longest time to calculate, in which the use of multiple local windows requires heavy computational time, but the proposed MLW_LRRSTO and SLW_LRRSTO generated higher detection rates than the others.

**Figure 9.** ROC evaluation of the different methods for the HYDICE_urban dataset.**Table 4.** AUC values for the different detectors with the HYDICE_urban dataset.

Method	GRXD	LRXD	WRXD	RPCA_RX	LRRSTO	SLW_LRRSTO	MLW_LRRSTO
AUC	0.9848	0.9953	0.9851	0.9837	0.9881	0.9964	0.9957
Time/s	0.06	25.48	0.21	2.34	18.43	17.72	36.78

4.2.3. AVIRIS_SanDiego Experiment

The ROC curves of all the methods are shown in Figure 10. Compared with LRXD, the proposed MLW_LRRSTO exhibited a slightly lower probability of detection for a low false alarm rate. Compared to SLW_LRRSTO, MLW_LRRSTO exhibited a higher probability of detection in terms of the overall detection performance, and it achieved the highest probability of detection for all the false alarm rate values. The AUC values and computational time of the different methods are shown in Table 5, where the GRXD, WRXD, and RPCA_RX had shorter computational times than the others, and the MLW_LRRSTO took the longest time to calculate. However, the proposed MLW_LRRSTO yielded a score that was higher than that of the others. Compared with LRRSTO, the SLW_LRRSTO had a shorter computational time and higher detection rate, which is a significant improvement.

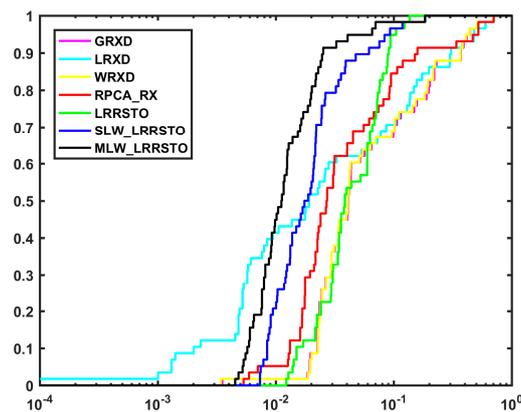


Figure 10. ROC evaluation of the different methods for the AVIRIS_SanDiego dataset.

Table 5. AUC values for the different detectors with the AVIRIS_SanDiego dataset.

Method	GRXD	LRXD	WRXD	RPCA_RX	LRRSTO	SLW_LRRSTO	MLW_LRRSTO
AUC	0.8885	0.9041	0.8901	0.9190	0.9300	0.9621	0.9780
Time/s	0.08	42.95	0.27	3.75	25.68	24.83	50.37

4.2.4. Viareggio Experiment

The ROC curves of all the methods are shown in Figure 11. Compared with LRRSTO, the proposed SLW_LRRSTO and MLW_LRRSTO exhibited a slightly lower probability of detection for a low false alarm rate. From the AUC values in Table 6, we can conclude that the GRXD and RPCA_RX had lower detection accuracy. Compared with the others, SLW_LRRSTO and MLW_LRRSTO yielded higher scores, which indicates that the two proposed methods had a significant improvement. From the computational time in Table 6, we can see that the GRXD and WRXD had shorter computational time than the others, and the LRXD had the longest time, which may have been caused by the large window size or the inverse function of local background. The computational time of SLW_LRRSTO and MLW_LRRSTO was relatively longer, but the detection accuracy was improved significantly compared with the other methods.

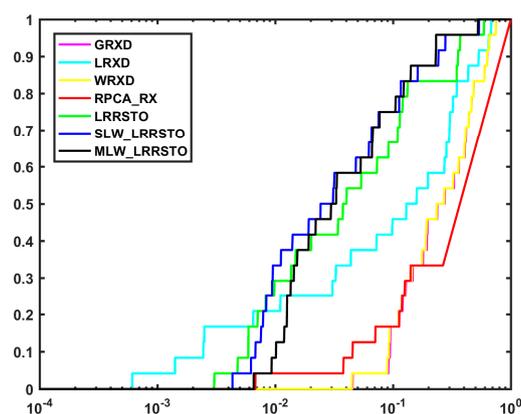


Figure 11. ROC evaluation of the different methods for the Viareggio dataset.

Table 6. AUC values for the different detectors with the Viareggio dataset.

Method	GRXD	LRXD	WRXD	RPCA_RX	LRRSTO	SLW_LRRSTO	MLW_LRRSTO
AUC	0.6884	0.7945	0.6908	0.5509	0.8825	0.9139	0.9172
Time/s	4.05	7153.51	13.43	863.73	1250.67	1936.92	3589.58

4.3. Discussion

To verify the effectiveness of the proposed methods, we conducted our experiments on three hyperspectral images which are widely used in the field of anomaly detection and compared the results with five other methods: GRXD, LRXD, WRXD, RPCA-RX, and LRRSTO.

In the proposed SLW_LRRSTO method, in order to achieve better spatial constraints, we smoothed the LRR coefficients of the central testing pixel by using a single local window. However, the spatial constraints of coefficients using a single local window were greatly affected by the anomalous pixels in the neighborhood. Therefore, in order to effectively reduce the impact of the anomalous pixels in the neighborhood, we proposed the MLW_LRRSTO algorithm, which adopts the multiple local window to constrain the spatial coefficients.

For a numerical comparison, we used the ROC curves and AUC values as the main criteria to evaluate these detection results. From the ROC curves shown in Figures 8–11, and the specific AUC values shown in Tables 3–6, it can be seen that the detection accuracy of the proposed two methods was obviously better than that of the other methods. Based on these experimental results, we can conclude that the proposed two methods showed superior detection performance compared with the others.

Although the proposed methods yielded outstanding detection performance, there still exist some shortcomings that can be further improved. The proposed method involves two regularization parameters— η and λ —which have an important influence on the detection results of images. Although the optimal parameters can be found by a trial-and-error method in synthetic data experiments, in practical applications, we can neither obtain the a priori knowledge of anomalous objects in advance nor can we find the optimal parameters by the trial-and-error method. Therefore, only empirical parameters can be used. Moreover, the matrix decomposition is very time consuming, and it is important to speed up the operation of the algorithm.

5. Conclusions

In this paper, two novel anomaly detection methods for HSIs have been proposed. The first approach is the SLW_LRRSTO anomaly detection method, which is based on LRRSTO with the combination of spectral and spatial information. SLW_LRRSTO adds a spatial constraint to the low-rank representation coefficient and smooths the low-rank representation coefficients of the central testing pixel using a single local window. The second approach is the MLW_LRRSTO method, which uses the same model as the SLW_LRRSTO method but employs a multiple local window smoothing filter strategy. To confirm the effectiveness of the proposed methods, experiments were conducted on both simulated and real hyperspectral data in comparison with other detection methods. The experimental results confirmed that SLW_LRRSTO and MLW_LRRSTO can effectively separate anomalous targets from the background, and their detection performance was clearly better than that of other detection methods. In particular, MLW_LRRSTO outperformed SLW_LRRSTO and offered the highest detection accuracy.

However, there still exist some aspects which can be further improved. The focus of the next research should be on how to automatically tune the parameters of the proposed SLW_LRRSTO and MLW_LRRSTO methods and to further improve the speed of the algorithms.

Author Contributions: K.T. and Z.H. conceived and designed the experiments; Z.H. performed the experiments; K.T., D.M., Y.C., and Q.D. analyzed the data; K.T., Z.H., D.M., Y.C., and Q.D. wrote the paper.

Funding: This research was supported in part by Natural Science Foundation of China (Nos. 41871337 and 41471356) and Shan'xi Key Research and Development Program (2018ZDXM-GY-023).

Conflicts of Interest: The authors declare no conflict of interests.

References

1. Niu, Y.; Wang, B. Hyperspectral anomaly detection based on low-rank representation and learned dictionary. *Remote Sens.* **2016**, *8*, 289. [[CrossRef](#)]
2. Li, W.; Wu, G.; Du, Q. Transferred deep learning for anomaly detection in hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 597–601. [[CrossRef](#)]
3. Nasrabadi, N.M. Hyperspectral target detection: An overview of current and future challenges. *IEEE Signal Process. Mag.* **2013**, *31*, 34–44. [[CrossRef](#)]
4. Xu, Y.; Wu, Z.; Li, J.; Plaza, A.; Wei, Z. Anomaly detection in hyperspectral images based on low-rank and sparse representation. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 1990–2000. [[CrossRef](#)]
5. Reed, I.S.; Yu, X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1760–1770. [[CrossRef](#)]
6. Matteoli, S.; Diani, M.; Corsini, G. Improved estimation of local background covariance matrix for anomaly detection in hyperspectral images. *Opt. Eng.* **2010**, *49*, 046201. [[CrossRef](#)]
7. Billor, N.; Hadi, A.S.; Velleman, P.F. BACON: Blocked adaptive computationally efficient outlier nominators. *Comput. Stat. Data Anal.* **2000**, *34*, 279–298. [[CrossRef](#)]
8. Imani, M. RX Anomaly detector with rectified background. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1313–1317. [[CrossRef](#)]
9. Yao, X.; Zhao, C. Hyperspectral anomaly detection based on the bilateral filter. *Infrared Phys. Technol.* **2018**, *92*, 144–153. [[CrossRef](#)]
10. Guo, Q.; Zhang, B.; Ran, Q.; Gao, L.; Li, J.; Plaza, A. Weighted-RXD and linear filter-based RXD: Improving background statistics estimation for anomaly detection in hyperspectral imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2351–2366. [[CrossRef](#)]
11. Li, W.; Du, Q. Unsupervised nearest regularized subspace for anomaly detection in hyperspectral imagery. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium—IGARSS, Melbourne, VIC, Australia, 21–26 July 2013; pp. 1055–1058.
12. Li, W.; Du, Q. Collaborative representation for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1463–1474. [[CrossRef](#)]
13. Philpott, S.; Hankins, C. Simultaneous joint sparsity model for target detection in hyperspectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 676–680.
14. Chen, Y.; Nasrabadi, N.M.; Tran, T.D. Hyperspectral image classification via kernel sparse representation. In Proceedings of the IEEE International Conference on Image Processing, Brussels, Belgium, 11–14 September 2011; pp. 1233–1236.
15. Wright, J.; Ganesh, A.; Rao, S.; Peng, Y.; Ma, Y. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 7–10 December 2009; pp. 2080–2088.
16. Chen, S.Y.; Yang, S.; Kalpakis, K.; Chang, C.I. Low-rank decomposition-based anomaly detection. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX*; SPIE Press: Baltimore, MD, USA, 2013; p. 87430N.
17. Liu, G.; Lin, Z.; Yan, S.; Sun, J.; Yu, Y.; Ma, Y. Robust recovery of subspace structures by low-rank representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 171–184. [[CrossRef](#)] [[PubMed](#)]
18. Xu, Y.; Wu, Z.; Wei, Z.; Liu, H.; Xu, X. A novel hyperspectral image anomaly detection method based on low rank representation. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 4444–4447.
19. Zhai, H.; Zhang, H.; Zhang, L.; Li, P.; Plaza, A. A new sparse subspace clustering algorithm for hyperspectral remote sensing imagery. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 43–47. [[CrossRef](#)]
20. Zhang, H.; Zhai, H.; Zhang, L.; Li, P. Spectral–spatial sparse subspace clustering for hyperspectral remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3672–3684. [[CrossRef](#)]
21. Elhamifar, E.; Vidal, R. Sparse subspace clustering: Algorithm, theory, and applications. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2765–2781. [[CrossRef](#)]
22. Fauvel, M.; Benediktsson, J.A.; Chanussot, J.; Sveinsson, J.R. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 3804–3814. [[CrossRef](#)]

23. Borghys, D.; Kåsen, I.; Achard, V. Comparative evaluation of hyperspectral anomaly detectors in different types of background. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*; The International Society for Optical Engineering: Baltimore, MD, USA, 2012; p. 83902J.
24. Du, B.; Zhao, R.; Zhang, L.; Zhang, L. A spectral-spatial based local summation anomaly detection method for hyperspectral images. *Signal Process.* **2016**, *124*, 115–131. [[CrossRef](#)]
25. Lin, Z.; Chen, M.; Ma, Y. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv* **2010**, arXiv:1009.5055.
26. Yuan, Y.; Wang, Q.; Zhu, G. Fast hyperspectral anomaly detection via high-order 2-D crossing filter. *IEEE Trans. Geosci. Remote Sens.* **2014**, *53*, 620–630. [[CrossRef](#)]
27. Ma, D.; Yuan, Y.; Wang, Q. A sparse dictionary learning method for hyperspectral anomaly Detection with capped norm. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 648–651.
28. The Synthetic AVIRIS Salinas Hyperspectral Dataset. Available online: http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes (accessed on 13 June 2019).
29. Stefanou, M.S.; Kerekes, J.P. A method for assessing spectral image utility. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1698–1706. [[CrossRef](#)]
30. Li, J.; Zhang, H.; Zhang, L.; Ma, L. Hyperspectral anomaly detection by the use of background joint sparse representation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 2523–2533. [[CrossRef](#)]
31. The Real HYDICE Urban Hyperspectral Dataset. Available online: <http://lesun.weebly.com/hyperspectral-data-set.html> (accessed on 13 June 2019).
32. Zhu, L.; Wen, G. Hyperspectral anomaly detection via background estimation and adaptive weighted sparse representation. *Remote Sens.* **2018**, *10*, 272.
33. Zhao, R.; Du, B.; Zhang, L. Hyperspectral anomaly detection via a sparsity score estimation framework. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3208–3222. [[CrossRef](#)]
34. Zhang, L.; Zhao, C. Hyperspectral anomaly detection based on spectral–spatial background joint sparse representation. *Eur. J. Remote Sens.* **2017**, *50*, 362–376. [[CrossRef](#)]
35. Acito, N.; Matteoli, S.; Rossi, A.; Diani, M.; Corsini, G. Hyperspectral airborne “Viareggio 2013 Trial” data collection for detection algorithm assessment. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2365–2376. [[CrossRef](#)]
36. The “Viareggio 2013 trial” Hyperspectral Dataset. Available online: <http://rsipg.dii.unipi.it/> (accessed on 13 June 2019).
37. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **1982**, *143*, 29. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).