

# An Improved Hyperspectral Image Anomaly Detection Algorithm using Low-Rank Representation

Dongying Bai

Air defense and missile defense academy, Airforce  
Engineering University, Xi'an, China  
277726615@qq.com

Hailin Tian

Air defense and missile defense academy, Airforce  
Engineering University, Xi'an, China  
458429355@qq.com

Dongli Tang\*

Air defense and missile defense academy, Airforce  
Engineering University, Xi'an, China

Zhaozhan Li

Air defense and missile defense academy, Airforce  
Engineering University, Xi'an, China  
751932965@qq.com

## ABSTRACT

Anomaly detection in hyperspectral images has drawn much attention in recent years. In order to provide a high-quality background dictionary for low-rank representation-based anomaly detector, from the perspective of dictionary learning, an anomaly detection method based on low-rank representation with an online-learned double sparse dictionary is proposed. Firstly, the double sparsity structure is adopted to the dictionary learning model to enhance the adaptivity. Next, to improve the dictionary training efficiency, the double sparse dictionary structure is modified and a corresponding online dictionary learning algorithm is proposed. The experimental results on five real-world hyperspectral datasets show that our method can achieve a reliable anomaly detection result and the background suppression performance is satisfying.

## CCS CONCEPTS

• Applied computing; • Physical sciences and engineering;

## KEYWORDS

Hyperspectral imagery, Anomaly detection, Low-rank representation, Dictionary learning

## ACM Reference Format:

Dongying Bai, Dongli Tang, Hailin Tian, and Zhaozhan Li. 2021. An Improved Hyperspectral Image Anomaly Detection Algorithm using Low-Rank Representation. In *The 5th International Conference on Computer Science and Application Engineering (CSAE 2021), October 19–21, 2021, Sanya, China*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3487075.3487170>

## 1 INTRODUCTION

Hyperspectral images (HSIs) usually contain hundreds or even thousands of narrow spectral bands that are barely about wide [1].

\*Correspondence: mnzx116@163.com

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
CSAE 2021, October 19–21, 2021, Sanya, China

© 2021 Association for Computing Machinery.  
ACM ISBN 978-1-4503-8985-3/21/10...\$15.00  
<https://doi.org/10.1145/3487075.3487170>

Abundant spatial and spectral information with high resolution in hyperspectral images make it possible to recognize different ground objects precisely. Such process is referred as target detection and it can be categorized into two types in terms of the availability of prior information: supervised and unsupervised. Since it is difficult to obtain the spectral information for targets of interest, unsupervised target detection is more frequently studied in practical which is known as anomaly detection.

In hyperspectral images, anomalies occupy only a few pixels and have evidently different spectral characteristics from the background. In order to effectively extract anomalies, plenty of methods have been proposed. The Reed-Xiaoli anomaly detector [2] is the benchmark anomaly detection method in this type. It measures the Mahalanobis distance between the test pixels and the estimated background which is under the assumption of multivariate normal distribution. Since the results of RX are easily interfered by anomalies in the background and the calculation of the covariance matrix is time-cost, various modified RX based methods have been proposed to overcome these drawbacks. The regularized RX method [3] aims to solve the ill conditioning problem of the matrix inversion by imposing a regularization for the background covariance matrix. The weighted RX method [4] tries to attenuate the anomaly contamination by re-weighting the background with the Gaussian probability. The subspace-based RX uses the significant eigen vectors of the covariance matrix to estimate the background so that the anomaly interference can be eliminated. Kwon et al. [5] projected the RX detector into a higher space so that the anomalies can be more effectively separated from the background. Kernel adaptive subspace detector [6] employs the RBF kernel to represent the data so that the translation invariance. Johnson et al. [7] proposed an independent-component-analysis-based anomaly detection method. They use the adaptive Wiener filter to suppress the background while retaining the anomalies. Chiang et al. [8] located the anomalies by projection pursuit to search anomaly pixels with evolution algorithm. Huck et al. [9] proposed a constant-false-alarm-rate detection method via combining projection pursuit and binary test hypothesis. Sparse-representation (SR) based anomaly detection methods have been hot topics recently. Li et al. [10] aimed to improve the performance of SR based anomaly detector by adaptively estimating local regions so that the potential anomaly contamination can be eliminated. Zhu et al. [11] established a global background dictionary with extracted

background-related endmembers in the hyperspectral images. In [12], the discriminative power of the dictionary is enhanced with an atom re-weighting strategy which estimates the usage frequency of an atom as its weight. Low-rank-representation (LRR) based methods also play important roles in HSI anomaly detection. The work in [13] introduced sparsity constraint for LRR model so that a more robust detection performance can be achieved. Yang et al. [14] provided a dictionary with pure background information for LRR model. As an intrinsic property of HSI, Tan et al. [15] analyzed the similarity among pixels in the local regions and used it as a spatial constraint for LRR based anomaly detector.

In this paper, in order to obtain a more efficient detection result, we propose a novel HSI anomaly detection based on low-rank representation model with an online-learned double sparse background dictionary. The main contributions of this paper can be summarized as follows.

1. An online dictionary learning algorithm specially designed for double sparse background dictionary is proposed so that a discriminative background dictionary can be formed with less training samples.
2. Based on the background dictionary, an LRR based detection method is proposed.

The remainder of this paper is organized as follows. In Section 2, the proposed method is demonstrated. In Section 3, experiments are conducted and the results are analyzed. In Section 4, the conclusions are drawn.

## 2 THE PROPOSED METHOD

### 2.1 Low-Rank Representation Model

In order to cope with the drawbacks of principal component analysis, John Wright et al. [16] proposed robust principal component analysis (RPCA): for a given high-dimensional dataset  $\mathbf{X}$ , it can be separated as a linear combination of a low-rank matrix  $\mathbf{L}$  and a sparse matrix  $\mathbf{S}$ , and the aim of RPCA is to reconstruct the original dataset with  $\mathbf{L}$  and  $\mathbf{S}$ . The corresponding model can be expressed as follows,

$$\min_{\mathbf{L}, \mathbf{S}} \text{rank}(\mathbf{L}) + \lambda \|\mathbf{S}\|_0 \quad \text{s.t.} \quad \mathbf{X} = \mathbf{L} + \mathbf{S} \quad (1)$$

where  $\lambda$  is a balancing parameter,  $\|\cdot\|_0$  is the  $l_0$ -norm that counts the non-zero entries in the matrix. Solving the above optimal problem is NP-hard and it is often relaxed into the following convex problem:

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathbf{L}\|_* + \lambda \|\mathbf{S}\|_1 \quad \text{s.t.} \quad \mathbf{X} = \mathbf{L} + \mathbf{S} \quad (2)$$

where  $\|\cdot\|_*$  is the nuclear norm of the matrix which is the sum of all singular values, and  $\|\cdot\|_1$  represents the  $l_1$ -norm that is calculated by summing the absolute values of entries in the matrix.

In an HSI, since the bands are highly correlated and the background is generally homogeneous, the background can be regarded to have low-rank property. Furthermore, the background usually consists of several materials and the corresponding spectral vectors are linear combinations of endmembers. Therefore, the background data can be assumed drawn from multiple data subspace. This assumption corresponds with LRR model [17], which can be expressed as

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathbf{Z}\|_* + \lambda \|\mathbf{S}\|_{2,1} \quad \text{s.t.} \quad \mathbf{X} = \mathbf{DZ} + \mathbf{S} \quad (3)$$

where  $\|\cdot\|_{2,1}$  denotes the sum of  $l_2$ -norm of all columns,  $\mathbf{D}$  is the background dictionary and the atoms span the whole background data space. The sparse matrix  $\mathbf{S}$  contains anomaly signals in its column space. To obtain the sparse matrix  $\mathbf{S}$ , Augmented Lagrange Method is often utilized. Accordingly, the problem in (3) can be solved by optimizing the following Lagrange cost function:

$$L = \|\mathbf{J}\|_* + \lambda \|\mathbf{S}\|_{2,1} + \text{tr}(\mathbf{Y}_1^T (\mathbf{X} - \mathbf{DZ} - \mathbf{S})) + \text{tr}(\mathbf{Y}_2^T (\mathbf{Z} - \mathbf{J})) + \frac{\mu}{2} (\|\mathbf{X} - \mathbf{DZ} - \mathbf{S}\|_F^2 + \|\mathbf{Z} - \mathbf{J}\|_F^2) \quad (4)$$

where  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$  is the Lagrange multipliers,  $\mu > 0$  is the penalty parameter. Firstly,  $\mathbf{J}$  is obtained by fixing the other variables:

$$\mathbf{J} = \arg \min \|\mathbf{J}\|_* + \text{tr}(\mathbf{Y}_2^T (\mathbf{Z} - \mathbf{J})) + \frac{\mu}{2} (\|\mathbf{X} - \mathbf{DZ} - \mathbf{S}\|_F^2 + \|\mathbf{Z} - \mathbf{J}\|_F^2) \quad (5)$$

the analytic solution of  $\mathbf{J}$  can be obtained by

$$\mathbf{J} = \Theta_{1/\mu}(\mathbf{Z} + \mathbf{Y}_2/\mu) \quad (6)$$

where  $\Theta_{1/\mu}$  denotes the singular value thresholding operator which can be expressed as

$$\Theta_{1/\mu}[\mathbf{R}] = \mathbf{U}(\text{diag}(\{\sigma_j - 1/\mu\}_+))\mathbf{V}^* \quad (7)$$

Next,  $\mathbf{Z}$  can be obtained by

$$\mathbf{Z} = (\mathbf{I} + \mathbf{D}^T \mathbf{D})^{-1} (\mathbf{D}^T (\mathbf{X} - \mathbf{S}) + \mathbf{J} + (\mathbf{D}^T \mathbf{Y}_1 - \mathbf{Y}_2)/\mu) \quad (8)$$

Then  $\mathbf{S}$  can be derived by optimizing the following problem

$$\mathbf{S} = \arg \min \lambda \|\mathbf{S}\|_{2,1} + \text{tr}(\mathbf{Y}_1^T (\mathbf{X} - \mathbf{DZ} - \mathbf{S})) + \frac{\mu}{2} (\|\mathbf{X} - \mathbf{DZ} - \mathbf{S}\|_F^2 + \|\mathbf{Z} - \mathbf{J}\|_F^2) \quad (9)$$

It can be solved by using the  $l_{1/2}$  minimization operator addressed as

$$\mathbf{S} = \Omega_{1/\mu}(\mathbf{X} - \mathbf{DZ} + \mathbf{Y}_1/\mu) \quad (10)$$

$$\mathbf{S}(:, i) = \begin{cases} \frac{\|\mathbf{Q}(:, i)\|_2 - 1/\mu}{\|\mathbf{Q}(:, i)\|_2} \mathbf{Q}(:, i), & \|\mathbf{Q}(:, i)\|_2 > 1/\mu \\ 0 & \text{others} \end{cases} \quad (11)$$

### 2.2 Online-Learned Double Sparse Background Dictionary

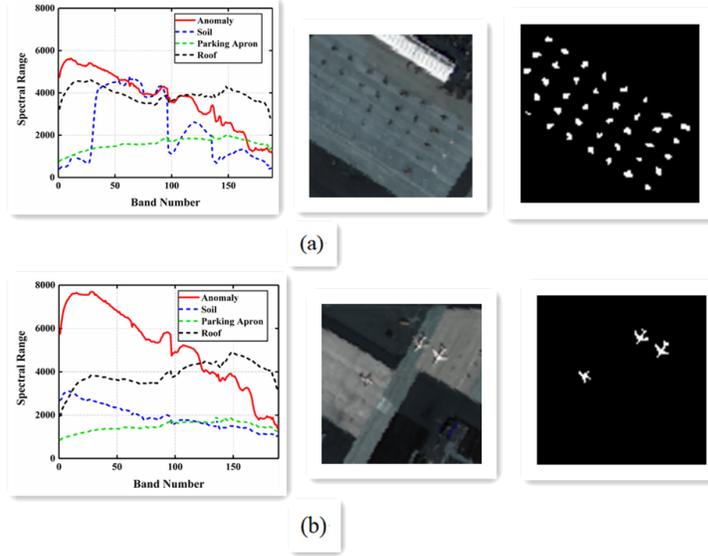
The effectiveness of separation between anomalies and backgrounds for LRR model significantly depends on the background dictionary. Generally, randomly selecting samples from the HSI and dictionary learning are two main method to construct the background dictionary. In this paper, we construct the background dictionary with dictionary learning. Furthermore, traditional dictionary learning requires extensive computational resource and can be time-cost, so that an efficient dictionary learning algorithm is demanded. In order to address this issue, the double sparsity model is applied to the background dictionary and a corresponding online dictionary learning algorithm is proposed in this section.

Traditional dictionary learning process can be addressed as follows.

$$\min_{\mathbf{D}, \mathbf{X}} \frac{1}{2} \|\mathbf{Y} - \mathbf{DX}\|_F^2 \quad \text{s.t.} \quad \|\mathbf{x}_i\|_0 \leq p, \quad i = 1, 2, \dots, l \quad (12)$$

where  $\mathbf{X}$  is the coefficient matrix,  $\mathbf{Y}$  is the training sample set. We then apply the double sparsity model in the dictionary learning process which can be transformed as follows:

$$\min_{\mathbf{A}, \mathbf{X}} \frac{1}{2} \|\mathbf{Y} - \Phi \mathbf{A} \mathbf{X}\|_F^2, \quad \text{s.t.} \quad \begin{cases} \|\mathbf{a}_i\|_0 \leq s, & i = 1, 2, \dots, m \\ \|\mathbf{x}_j\|_0 \leq p, & j = 1, 2, \dots, l \end{cases} \quad (13)$$



**Figure 1: Spectrum Curves, Pseudo Images and Ground Truth Maps of Two San Diego Airport HSI Datasets: (a) San Diego Airport 1 Dataset; (b) San Diego Airport 2 Dataset.**

where  $\Phi$  is the base dictionary and  $\mathbf{A}$  is the sparse dictionary. In the dictionary update process, the base dictionary is fixed while the sparse matrix is updated. We use the Kronecker product of two Discrete Cosine Transform (DCT) matrices as the base matrix, expressed as follows

$$\Phi = \Phi_S \otimes \Phi_S \quad (14)$$

where  $\Phi_S$  is the DCT matrix. Inspired by the works in [18] and [19], in order to save computational resource and time, we employ the online dictionary learning algorithm and modify it to fit the double sparsity model. Then, the sparse dictionary  $\mathbf{A}$  can be updated by the following equation:

$$\mathbf{A}^{t+1} = \text{H}_k [\mathbf{A}^t + \eta^t \nabla f(\mathbf{A}^t)] \quad (15)$$

where  $\mathbf{A}^t$  is the estimation of  $\mathbf{A}$  in  $t$  th iteration,  $\text{H}_k$  is the hard-thresholding operator and  $\nabla f(\mathbf{A}^t)$  is the current gradient at the  $t$  th iteration.  $\eta$  is the step factor and can be calculated as follows:

$$\eta^t = \frac{\|\nabla f(\mathbf{A}_S^t)\|_F}{\|\Phi \nabla f(\mathbf{A}_S^t) \mathbf{X}_S^t\|_F} \quad (16)$$

where  $S$  is the current support set. Aiming to provide a more efficient structure for online dictionary learning, we modify the double sparse dictionary learning cost function into the following form:

$$\|\mathbf{Y} - \Phi \mathbf{A} \mathbf{X}\|_F^2 = \|\Phi^T (\mathbf{Y} - \Phi \mathbf{A} \mathbf{X})\|_F^2 = \|\Phi^T \mathbf{Y} - \mathbf{A} \mathbf{X}\|_F^2 \quad (17)$$

let  $\Phi^T \mathbf{Y} = \mathbf{C}$ , the dictionary learning process can be expressed as:

$$\min_{\mathbf{A}, \mathbf{X}} \frac{1}{2} \|\mathbf{C} - \mathbf{A} \mathbf{X}\|_F^2 \text{ s.t. } \begin{cases} \|\mathbf{a}_i\|_0 \leq s, & i = 1, 2, \dots, m \\ \|\mathbf{x}_j\|_0 \leq p, & j = 1, 2, \dots, l \end{cases} \quad (18)$$

Correspondingly, the gradient at current iteration and the step factor can be re-written as follows.

$$\begin{aligned} \nabla f(\mathbf{A}^t) &= \mathbf{A}^t \mathbf{X}^t (\mathbf{X}^t)^T - \mathbf{C} (\mathbf{X}^t)^T \\ \eta^t &= \frac{\|\nabla f(\mathbf{A}_S^t)\|_F}{\|\nabla f(\mathbf{A}_S^t) \mathbf{X}_S^t\|_F} \end{aligned} \quad (19)$$

The sparse coding phase is functioned with the orthogonal matching pursuit (OMP) algorithm. After the sparse dictionary  $\mathbf{A}$  is obtained, the background dictionary can be calculated by  $\mathbf{D} = \Phi \mathbf{A}$ , and it can be used in LRR model to extract the sparse matrix.

### 3 EXPERIMENTS AND ANALYSIS

In this section, we evaluate the effectiveness by conducting experiments on five real-world HSI datasets. The experiments are conducted on a laptop with a CPU Intel I5-7400hq, 16 GB RAM and an operating system of Windows 10.

#### 3.1 Data Description

The first two real-world HSI dataset are from the San Diego Airport HSI dataset captured by Airborne Visible /Infrared Imaging Spectrometer (AVIRIS). The raw data is consist of 224 bands and ranging from  $0.37 \sim 2.50 \mu\text{m}$ . In this experiment, two subregions with the size of  $100 \times 100$  from the whole scene are selected. After removing the low SNR and water-absorption bands, there are 189 bands left. The background mainly contains roof, soil and parking apron. The anomaly targets in two datasets are aircrafts in the parking apron. The spectrum curves, the pseudo images and the ground truth maps of two AVIRIS San Diego Airport datasets are depicted in (1).

The third to fifth datasets are selected from the Air-Beach-Urban (ABU) HSI dataset. The HSIs in this dataset were mainly captured by AVIRIS and Reflective Optics System Imaging Spectrometer. Three datasets have the same size of  $100 \times 100$ . The third dataset is from

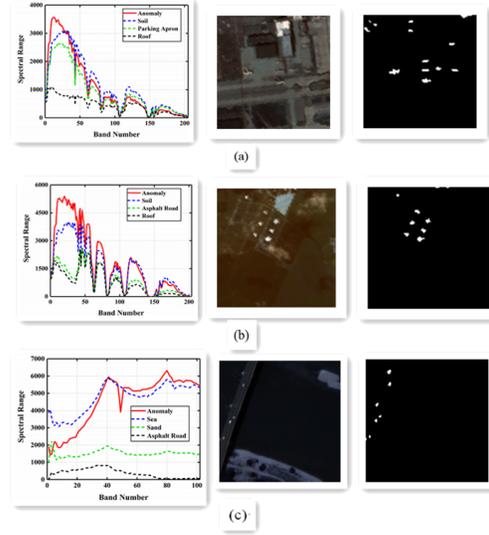


Figure 2: Spectrum Curves, Pseudo Images and Ground Truth Maps of Three Abu Hsi Datasets: (a) ABU-Airport Dataset; (b) ABU-Urban Dataset; (c) ABU-Beach Dataset.

Table 1: The Parameters Settings

Methods	Parameters
GKRX	Number of centroids: 600, kernel width: 40
LRASR	$K = 12, P = 10, \lambda = 0.01, \beta = 0.1$
LSMAD	$r = 4, k = 0.004, r = 3$
LSUNR	$win_{in} = 5, win_{out} = 9, \lambda = 0.1$
SSEAD	$\lambda_1 = 0.1, \lambda_2 = 0.01$
Ours	$\lambda = 0.1, \mu = 10^{-6}, \mu_{max} = 10^6, \rho = 1.1, \epsilon = 10^{-5}$

the ABU Airport dataset and contains 205 bands. The anomaly targets in this scene are several aircrafts in the parking apron, and the background mainly includes soil, parking apron and roof. The fourth dataset is from the ABU Urban dataset and contains 204 bands. The anomaly targets in this HSI are airplanes lying on the ground, and the background mainly contains soil, asphalt road and roof. The fifth dataset is from the ABU beach dataset and contains 102 bands. The anomaly targets in this scenario are several vehicles, and the background mainly includes sea, sand and asphalt road. The spectrum curves, the pseudo images and the ground maps of three datasets are shown in (2).

### 3.2 Evaluation Criteria and Parameters Settings

We use RX [1], Global KRX [4], LRASR [20], LSMAD [21], LSUNR [22], and SSEAD [23] as comparison methods, and we name our method as Ours. The parameter settings of our method and comparison methods are listed in Table 1. In the experiments we draw 1/5 samples from the test dataset to form the training sample set.

The quantitative evaluation criteria we use in the experiments are receiver operator curve (ROC) and background-anomaly separability maps. The detection probability and the false alarm rate

are respectively defined as follows.

$$P_d = \frac{N_{cd}}{N_t} \quad (20)$$

$$F_a = \frac{N_{fd}}{N} \quad (21)$$

where  $N_{cd}$  is the number of correctly detected anomaly pixels,  $N_t$  is the number of true anomaly pixels,  $N_{fd}$  denotes the false detected pixels and  $N$  denotes the total number of pixels in the test image.

### 3.3 Results and Analysis

The experimental results on five real-world HSI datasets are shown in (3). The first column shows the ground truth maps as references. The detection results of two San Diego Airport, ABU-Airport, ABU-Urban and ABU-Beach are depicted in the first to the fifth row, respectively. It can be seen from the first row that the detection performance of our method significantly outperforms the others. Compared to RX and LRASR, our method can identify the anomalies more prominently. From the results of SSEAD, GKRX and LSMAD, it can be observed that the roof on the top right of the scene severely interferes the detection performances of these methods. The results depicted in the second row illustrates that our method can effectively recognize the anomalies and suppress the background. The results of RX, GKRX, LRASR and LSMAD are all interfered by false alarms. It can be also observed from third row to fifth row that our method can significantly separate the anomalies from the background on three ABU datasets. For the comparison methods, LSUNR and SSEAD can achieve comparable detection performances while our method has evidently stronger background suppression ability. The results of RX show that it can identify anomalies in ABU-Airport and ABU-Urban while it suffers from heavy false alarms. As for the results of GKRX, LRASR and LSMAD, they can detect most of the anomalies in three ABU datasets, yet there exist large amounts of pixels that have high response

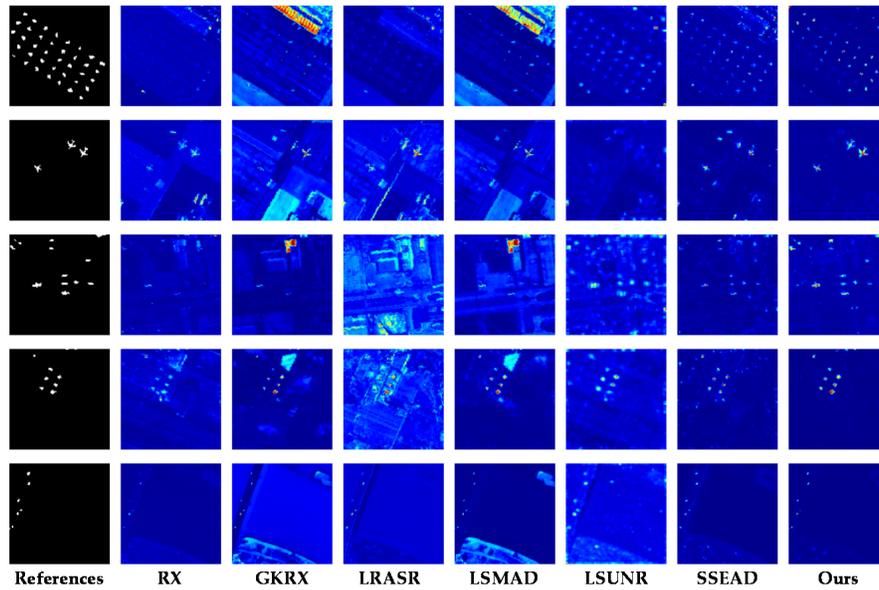


Figure 3: Ground Truths and Color Detection Maps.

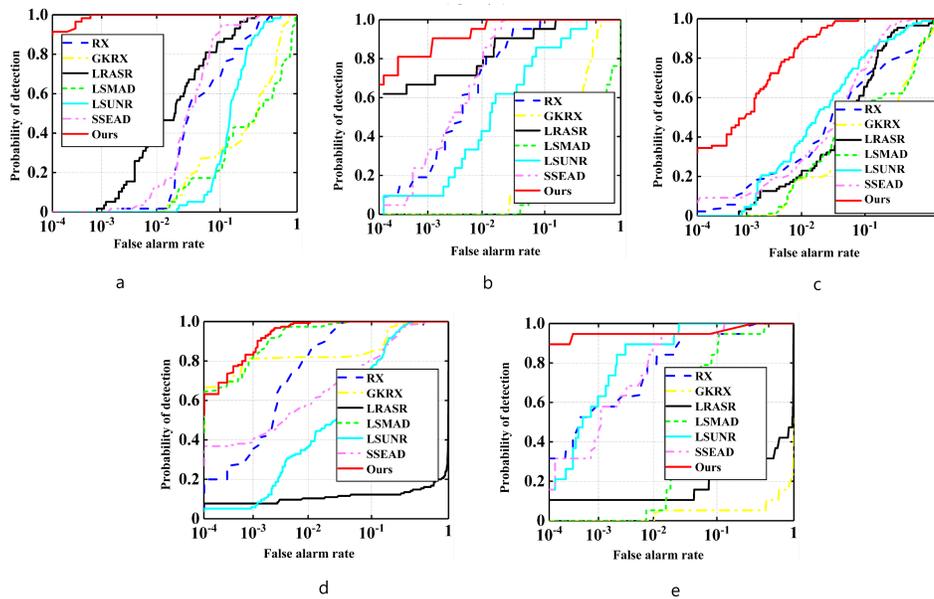


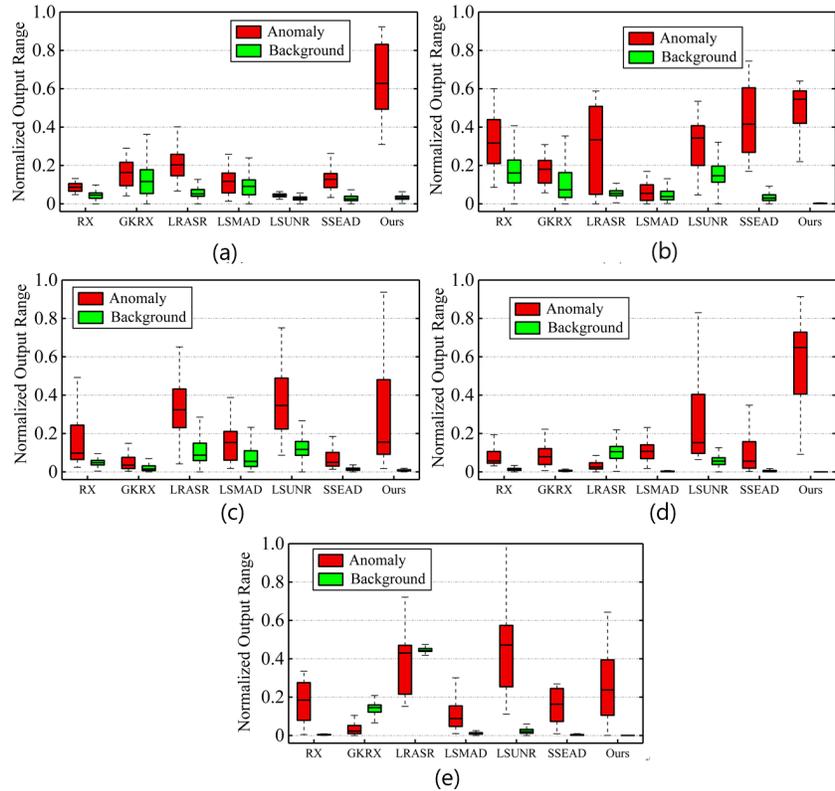
Figure 4: ROC of the Detection Results: (a) Results on San Diego Airport 1. (b) Results on San Diego Airport 2. (c) Results on ABU-Airport. (d) Results on ABU-Urban. (e) Results on ABU-Beach.

level. Compared with these methods, our method has more reliable detection results.

The ROC of all methods on five HSI datasets are shown in (4). As can be seen from Figure 4(a), the detection probability of our method reaches to around 0.9 with a false alarm rate of  $10^{-4}$  while the detection probabilities of the other methods are still zero. When the false alarm rate is  $10^{-3}$ , the other methods barely identify anomalies while the detection probability of our method achieves 1. The

superiority of our method illustrated above is also validated in Figure 4(b) and Figure 4(c). In Figure 4(d) and Figure 4(e), the areas under curve of our method are both the largest, which indicates that our method achieves the most reliable results.

To further quantitatively validate the effectiveness of our method, the separability maps of results on five HSI datasets are drawn in (5). The red box and the green box represent the output distribution range of anomalies and background, respectively. It can be observed



**Figure 5: Separability Maps of the Detection Results: (a) Results on San Diego Airport 1. (b) Results on San Diego Airport 2. (c) Results on ABU-Airport. (d) Results on ABU-Urban. (e) Results on ABU-Beach.**

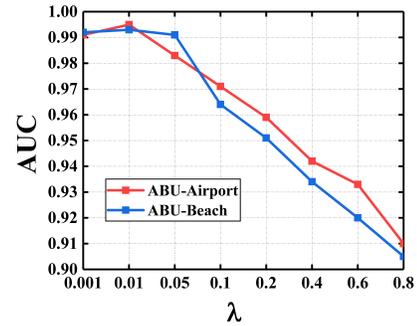
from each boxplot that the gap between two boxes of our method is the largest. This suggests that our method achieves the most effective background-anomalies separability for all five datasets. For GKRX, LRASR and LSMAD, overlaps between red box and green box exists in their output results which indicates that their detection results suffer from serious false alarms. Moreover, in these five boxplots, it can be observed that the background distribution boxes are all relatively the narrowest ones. This illustrates that our method has a more satisfied background suppression ability.

### 3.4 Parameters Analysis

In this subsection, we conduct experiments on two real-world HSI datasets to testify the parameters effects on detection results: ABU-Airport and ABU-Beach. The evaluation index we use is the area under curve (AUC). The parameters we discuss in this subsection are the balancing parameter  $\lambda$ , the proportion of samples for dictionary learning, the sparsity  $s$  of sparse dictionary  $\mathbf{A}$  and the sparsity  $p$  of the coefficient matrix  $\mathbf{X}$ .

#### a) The parameter $\lambda$

The experimental results for the effect of parameter  $\lambda$  on two HSI datasets are depicted in (6). It can be observed that when  $\lambda$  falls in the range of [0.001, 0.01], the detection performances improve a little as  $\lambda$  increases. The AUC values dramatically fall with the



**Figure 6: Parameter Effect of  $\lambda$ .**

increasing of  $\lambda$  when  $\lambda$  is larger than 0.01. Consequently,  $\lambda$  better falls in the range of [0.001, 0.01].

#### b) The proportion of samples for dictionary learning

To validate the superiority of the proposed online double sparse dictionary learning algorithm, we design two comparison methods: 1). Remove the double sparsity structure and train the background dictionary with K-SVD algorithm. This method is named Compare A. 2). Remove the double sparsity structure and use the online dictionary learning algorithm proposed in [17]. This method is

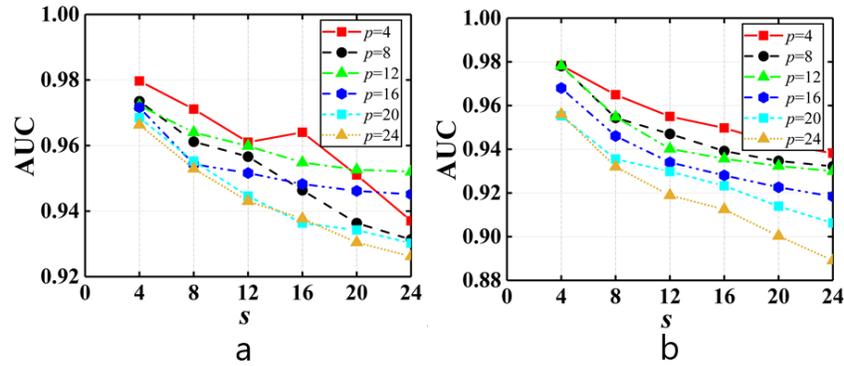


Figure 7: Parameter Effect of  $s$  and  $p$ : (a) Results on ABU-Airport. (b) Results on ABU-Beach.

Table 2: Number of Training Samples on ABU-Airport

Proportion	AUC		
	Compare A	Compare B	Ours
1/10	0.774	0.783	0.846
1/9	0.816	0.785	0.850
1/8	0.842	0.838	0.875
1/7	0.857	0.888	0.898
1/6	0.876	0.903	0.914
1/5	0.891	0.915	0.930
1/4	0.913	0.918	0.935
1/3	0.916	0.922	0.946

Table 3: Number of Training Samples on ABU-Urban

Proportion	AUC		
	Compare A	Compare B	Ours
1/10	0.778	0.775	0.845
1/9	0.807	0.796	0.867
1/8	0.816	0.813	0.883
1/7	0.837	0.838	0.905
1/6	0.853	0.844	0.921
1/5	0.867	0.866	0.938
1/4	0.875	0.874	0.941
1/3	0.879	0.881	0.943

named Compare B. The experimental results on ABU-Airport and ABU-Urban are listed in Table 2 and Table 3, respectively. From the results in Table 2, for the detection performance that our method achieves with 1/10 of the samples from the original dataset, Compare A and Compare B require about 1/8 of the original samples. The detection results that Compare A uses 1/3 of the samples for training only requires 1/7 of the samples for our method. It can be seen from the results in Table 3 that for the AUC value of around 0.84, the proportion parameter for our method is 1/8 while it is 1.3 for Compare A and Compare B. Additionally, it can also be observed from two tables that when the number of training samples is over 1/5 of the original data samples, the improvement of the

performances evidently pace down, which indicates that when we randomly draw about 1/5 of the samples from the original dataset, it could cover most of the background information.

c) The sparsity of the sparse dictionary  $\mathbf{A}$  and the coefficient matrix  $\mathbf{X}$

The experimental results for the effects of the sparsity  $s$  of  $\mathbf{A}$  and the sparsity  $p$  of  $\mathbf{X}$  are depicted in (7). As we can see that when  $p$  is fixed, the AUC values gradually decreases when  $s$  increases. This means that when the sparsity of the coefficient matrix is fixed, the increasement of the sparsity in sparse dictionary leads an attenuation to the detection performance. When  $s$  is fixed and as  $p$  increases, the detection performances gradually deteriorate. To be noticed, when  $s$  is within (4,12) and  $p$  is within (4,12), the AUC values vary in a small range. Therefore,  $s$  and  $p$  are best set in the range of (4,12).

## 4 CONCLUSIONS

In this paper, a novel anomaly detection method based on low-rank representation with an online-learned double sparse dictionary is proposed. Firstly, the double sparsity structure is applied to the dictionary learning model to improve the discriminative power for low-rank representation-based anomaly detector. Then, in order to improve the efficiency for dictionary learning, the double sparse background dictionary learning model is modified and a corresponding online dictionary learning algorithm is proposed. The experimental results show that our method can accurately identify the anomalies and effectively suppress the background simultaneously, and the proposed online double sparse dictionary learning algorithm is proved requiring less training samples compared to traditional dictionary learning algorithms.

## REFERENCES

- [1] Pizzolante R, Carpentieri B (2016). Multiband and Lossless Compression of Hyperspectral Images. *Algorithms*, vol. 9, no. 1:16-25.
- [2] I. S. Reed and X. Yu (1990). Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions on Acoustics Speech & Signal Processing*, vol. 38, no. 10, pp. 1760-1770.
- [3] N. M. Nasrabadi (2008). Regularization for spectral matched filter and RX anomaly detector. *Proceedings of Spie the International Society for Optical Engineering*, vol. 6966.
- [4] C. Zhao, J. Li, M. Meng, and X. Yao (2017). A Weighted Spatial-Spectral Kernel RX Algorithm and Efficient Implementation on GPUs. *Sensors*, vol. 17, no. 3, p.

- 441.
- [5] H. Kwon and N. M. Nasrabadi (2005). Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery. *IEEE Transactions on Geoscience & Remote Sensing*, vol. 43, no. 2, pp. 388-397.
  - [6] H. Kwon and N. M. Nasrabadi (2006). Kernel adaptive subspace detector for hyperspectral imagery. *IEEE Geoscience & Remote Sensing Letters*, vol. 3, no. 2, pp. 271-275.
  - [7] R. J. Johnson, J. P. Williams, and K. W. Bauer (2013). AutoGAD: An Improved ICA-Based Hyperspectral Anomaly Detection Algorithm. *IEEE Transactions on Geoscience & Remote Sensing*, vol. 51, no. 6 Part2, pp. 3492-3503.
  - [8] S. S. Chiang, C. I. Chang, and I. W. Ginsberg (2001). Unsupervised target detection in hyperspectral images using projection pursuit. *IEEE Transactions on Geoscience & Remote Sensing*, vol. 39, no. 7, pp. 1380-1391.
  - [9] A. Huck and M. Guillaume (2010). Asymptotically CFAR-Unsupervised Target Detection and Discrimination in Hyperspectral Images With Anomalous-Component Pursuit. *IEEE Transactions on Geoscience & Remote Sensing*, vol. 48, no. 11, pp. 3980-3991.
  - [10] J. Li, H. Zhang, L. Zhang, and L. Ma (2015). Hyperspectral Anomaly Detection by the Use of Background Joint Sparse Representation. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, vol. 8, no. 6, pp. 2523-2533.
  - [11] L. Zhu and G. Wen (2018). Hyperspectral Anomaly Detection via Background Estimation and Adaptive Weighted Sparse Representation. *Remote Sensing*, vol. 10, no. 2, pp. 272-285.
  - [12] R. Zhao, B. Du, and L. Zhang (2017). Hyperspectral Anomaly Detection via a Sparsity Score Estimation Framework. *IEEE Transactions on Geoscience & Remote Sensing*, vol 12, pp. 1-15.
  - [13] X. Yang, Z. Wu, J. Li, A. Plaza, and Z. Wei (2016). Anomaly Detection in Hyperspectral Images Based on Low-Rank and Sparse Representation. *IEEE Transactions on Geoscience & Remote Sensing*, vol. 54, no. 4, pp. 1990-2000.
  - [14] Y. Yang, J. Zhang, S. Song, and D. Liu (2019). Hyperspectral Anomaly Detection via Dictionary Construction-Based Low-Rank Representation and Adaptive Weighting. *Remote Sensing*, vol. 11, no. 2.
  - [15] K. Tan, Z. Hou, D. Ma, Y. Chen, and Q. Du (2019). Anomaly Detection in Hyperspectral Imagery Based on Low-Rank Representation Incorporating a Spatial Constraint. *Remote Sensing*, vol. 11, no. 13, p. 1578.
  - [16] J. Wright, A. Ganesh, S. Rao, and Y. Ma (2009). Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices.
  - [17] G. Liu, Z. Lin, and Y. Yong (2010). Robust Subspace Segmentation by Low-Rank Representation, in *International Conference on Machine Learning*.
  - [18] L. Fang and S. Zhang (2016). Double sparse dictionary learning for image super resolution, in *2016 Chinese Control and Decision Conference (CCDC)*.
  - [19] J. Sulam, B. Ophir, M. Zibulevsky, and M. Elad (2016). Trainlets: Dictionary Learning in High Dimensions. *IEEE Transactions on Signal Processing*, vol. 64, no. 12, pp. 3180-3193.
  - [20] Xu Y, Wu Z, Li J, *et al.* (2016). Anomaly Detection in Hyperspectral Images Based on Low-Rank and Sparse Representation. *IEEE Transactions on Geoscience & Remote Sensing*, vol 54, no.4:1990-2000.
  - [21] Zhang Y, Du B, Zhang L, *et al.* (2016). A Low-Rank and Sparse Matrix Decomposition-Based Mahalanobis Distance Method for Hyperspectral Anomaly Detection. *IEEE Transactions on Geoscience and Remote Sensing*, vol 54, no.3: 1-14.
  - [22] Tan, Hou, Wu, *et al.* (2019). Anomaly Detection for Hyperspectral Imagery Based on the Regularized Subspace Method and Collaborative Representation. *Remote Sensing*, vol 11, no.11:1318-1327.
  - [23] Zhao R, Du B, Zhang L (2017). Hyperspectral Anomaly Detection via A Sparsity Score Estimation Framework. *IEEE Transactions on Geoscience and Remote Sensing*, vol 55, no.6:3208-3222.