Matrix Autoregressive Model for Hyperspectral Anomaly Detection

Jingxuan Wang, Jinqiu Sun, Yu Zhu, Yong Xia, *Member, IEEE*, and Yanning Zhang, *Senior Member, IEEE*

***Abstract*—For anomaly detection in hyperspectral imagery, the scene can be treated as a combination of the background and the anomalies. Once a pure background hyperspectral image (HSI) is obtained, the anomalies can be easily located. In this article, we detect the anomalies via a matrix autoregressive model (MARM) to reconstruct the background HSI. Specifically, some informative and discriminative bands are first selected and come into a new HSI with less bands. Second, the new HSI can be treated as a collection of profiles in the row direction. Based on this, the background can be regularly reconstructed via the MARM. The regressive model not only respects the original matrix structure in the row profiles but also utilizes both the spatial and spectral correlations for the detection process. Finally, the classical Reed Xiaoli detector is applied to the difference cube between the band-selected HSI and the HSI reconstructed by MARM, achieving a final detection map with higher accuracy. Experimental results and data analysis on four different sensors captured datasets with different resolutions have validated the effectiveness of the proposed method.**

***Index Terms*—Anomaly detection, hyperspectral image (HSI), matrix autoregressive.**

I. Introduction

HYPERSPECTRAL imagery captures both the spatial and the spectral information of the scene simultaneously, achieving a 3-D image cube [1]. The 3-D hyperspectral images (HSIs) are characterized by their rich spectral information, which can be utilized to identify the materials by their unique reflective spectra [2]. In this way, the HSIs have gained wide applications in addressing the issue of geological application, military rescue, and mineral exploration [3]. Among these applications, anomaly detection seeks to discriminate the abnormal observations without the prior information about the target. Anomalous targets have not been exactly defined and are generally treated as the pixels which derivate from the background clutter distributions. According to different application scenarios, abnormal targets may exhibit different entities, such as the infected trees in the jungles, the wounded soldiers in the forests, the spilling oil on the ocean, rare minerals in geological applications, and some man-made vehicles and airplanes for aerial searches. In this way, hyperspectral anomaly detection has drawn extensive attention in both the military and the civilian aspects [4].

The anomalous locations in the scene are specified as the following two main differences: the spectral difference and the spatial difference. On the one hand, spectral curves of the anomalous locations are severely different from those of the background locations. Meanwhile, the anomalous locations usually appear in a format of several pixels (or even sub-pixels), which are embedded in a local homogeneous region [5]. Based on these two differences, many works have been proposed to determine the anomalies in the HSIs [6]. The most classical one is the Reed–Xiaoli (RX) detector, which assumes that all the background pixels follow the Gaussian distribution [7]. In this article, the classical RX detector is also named the Global-RX, which is different from the other RX related detectors. For the Global-RX detector, the distribution of the background pixels is formulated by calculating the mean and the covariance matrix. By measuring the Mahalanobis distance between the pixels and the background, the resulting generalized likelihood ratio can be obtained. Considering that the backgrounds may also consist of several different types of materials, all the background pixels follow a uniform distribution deviated from the real scene. In this way, the local RX detector is proposed, which applies the RX detector in a local region by sliding the window [8]. The kernel RX detector maps the original HSI into a feature space with higher dimension, and the RX detector is applied to the higher feature, which overcomes the high false alarm rate brought by the unappropriated model [9], [10]. In addition, many other preprocessing techniques have also been exploited to make an improvement of the detection accuracy, such as the fractional Fourier entropy (FrFT) [11], spectral derivatives (Deriv) [13], and some others [12]. With the combination of the RX detector, these preprocessed detection methods are abbreviated as the FrFT-RX, Deriv-RX, and some others.

Considering that the original HSI can be treated as a combination of the background and anomalies, many detectors have been proposed to reconstruct the background HSI [14]. Obviously, the key to success for this type of methods is an accurate estimation of the background HSI [15]. During the reconstruction of the background HSI, low rankness [16] and sparse representation [17], [18] are the most commonly used

Manuscript received 12 July 2022; revised 29 August 2022; accepted 20 September 2022. Date of publication 26 September 2022; date of current version 14 October 2022. This work was supported by the National Natural Science Foundation of China under Grant 61901384, Grant 61971356, and Grant U19B2037. *(Corresponding authors: Jingxuan Wang; Yanning Zhang.)*

Jingxuan Wang, Yu Zhu, Yong Xia, and Yanning Zhang are with the School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China, also with the National Engineering Laboratory for Integrated Aerospace-Ground-Ocean Big Data Application Technology, Northwestern Polytechnical University, Xi'an 710072, China, and also with the Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: jxwang0130@163.com; yuzhu@nwpu.edu.cn; yxia@nwpu.edu.cn; ynzhang@nwpu.edu.cn).

Jinqiu Sun is with the School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: sunjinqiu@nwpu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2022.3209204

tools. The representative background bases are utilized to build the overcomplete dictionary [19]. In this way, any background pixel in the HSI can be expressed as a sparse linear combination of the prelearned dictionary, while the anomalies cannot. Difference between the original HSI and the reconstructed HSI is incorporated to locate the anomalies [20]. To be specific, this kind of methods depends on that the original HSI can be treated as a combination of a low-rank background matrix and a sparse anomaly component or not. Based on this, Li suggested that the sparse component can be treated as a mixture of Gaussian noises (LSDM-MoG). In this way, the anomalies can be detected by the Manhattan distance [21]. Cheng and Wang [22] applied the graph regularization and total variation regularization into the low-rank representation of the background (GTVLRR), which preserves the local structure of the HSI. Considering the noise is also sparse, some works attempt to separate the noise from the sparse space and to find the true anomalies, which is formulated in the classical GoDec method [23]. Similar to this idea, a component decomposition analysis (CDA) method is proposed to treat the HSI as a combination of the principal components, independent components, and a noise component [24].

Besides the low rankness of the dictionary matrix, considering the intrinsically 3-D characteristics of the HSIs, the three-order tensor is also widely applied to denote the HSIs [25], [26]. The tensor decomposition-based detector (TenB) is applied to eliminate the background and highlights the anomalies [27]. The TenB detector abandons the first principal component in the three dimensions from the original HSI, achieving an acceptable detection accuracy. Based on the premise that most pixels of the scene belong to the background, it is convenient to locate some background pixels. With the located background pixels, the background pixels can be reconstructed via a tensor completing method (TCD) [28] and the anomalies can be detected. Considering the group sparse prior in the HSI, a prior-based tensor approximation (PTA) method is proposed for detection, which combines low-rank, sparse, and piecewise smooth with the advantages of tensor representation [29].

Inspired by the great success of the deep learning method in computer vision, many works have also been proposed based on the deep learning to detect the anomalies [30]. Depending on whether sample support is needed, the deep learning based methods can be classified into the supervised and unsupervised categories. Typically unsupervised detection methods include the autoencoder networks and generative adversarial networks [31], which aims at achieving the low-dimensional expression of each pixel. An encoder–decoder long short-term memory-based anomaly detector has been proposed to reconstruct the HSI [32]. The Mahalanobis distance is utilized to detect the anomalies for the dimension reduced HSI. The fully convolutional autoencoder network is also applied to reconstruct the background, in which the anomalies appear as reconstruction errors [33]. To tackle the anomalies being mixed in the training process and to preserve the geometric structure among samples during the learning process, a robust graph autoencoder (RGAE) detector with a $l_{2,1}$-norm is proposed [34]. The supervised deep learning based methods usually transfer the detection problem into some other tasks, such as classification [35] or abundance matrix extraction [36].

All the mentioned detectors dealt with the HSIs either from the spatial (width–height) aspect or from the spatial-spectral (three-order tensor) aspect. However, seen from the spatial-spectral (height–depth) aspect, the profiles in the row direction are also highly correlated. In this way, the row profiles (patches in the height–depth domain) can be simulated by a matrix autoregressive model (MARM) to formulate the background HSI and to highlight the anomalies in the scene. Detection on the highlighted scene is supposed to be more accurate than that on the original scene.

The main contributions of the proposed MARM method can be summarized as follows.

1) As far as we know, it is the first time that the MARM has been applied to the hyperspectral anomaly detection field for the background reconstruction. The MARM innovatively utilized the spatial-spectral information conveyed by the row profiles for reconstructing the background. Meanwhile, an optimal neighborhood reconstruction strategy is first applied to eliminate some redundant or noisy bands and to reduce the computational cost.
2) There are four HSIs that are captured by three different sensors, and the scenes in the HSIs and the spatial resolution of the HSIs are different. Row profiles in different HSIs exhibit different correlations. In this way, the number of profiles exploited for the regression model is varied for different HSIs.
3) In addition, different from the traditional vector autoregressive model that neglects the close height–width correlation in the row profiles, the matrix regressive model keeps the original structure in the matrix, and it is more interpretable.

II. Related Works

A. Spectral-Spatial Modeling of the Background

Since there is no information known for the anomaly detection process, spectral-spatial modeling of the background information is a direct way for the detecting. The main idea is to reconstruct the pixels in the background using a certain dictionary. The low-rankness and sparsity are the common tools for the regularization process. The reconstruction model can be expressed as

$$\arg\min \operatorname{rank}(L) + \lambda ||S||_0, s.t.\ X = L + S \tag{1}$$

where X is the HSI to be detected, and L and S are the low-rankness matrix and the sparse matrix, respectively. $||\cdot||_0$ is the l_0 norm, which represents the number of nonzeros elements in the matrix. λ is the weighting factor between the low rankness and sparsity. This is the classical robust principal comment analysis formulation [37], [38]. Solving for the loss function with a l_0 norm is an NP-hard problem. In this way, the l_0 norm is usually degenerated into the l_1 norm or the l_2 norm. The $\operatorname{rank}()$ is usually transferred into the nuclear norm [39]. To make finer discrimination between the outlier and the noise, a robust subspace learning detector has been proposed via $\arg\min \operatorname{rank}(L + OT) + \lambda ||S||_0$, where the outliers are treated as additive sparse corruptions [40]. In order to separate the noises

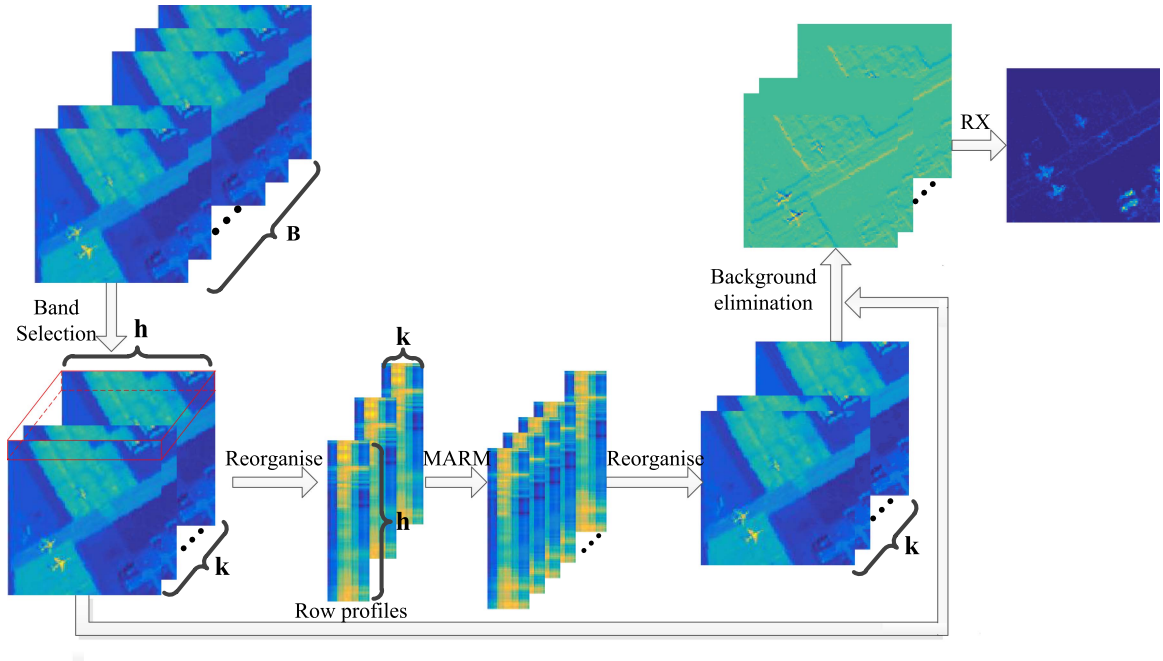


Fig. 1. Architecture of the proposed MARM detector.

from the anomalies, some works also import the noises with sparsity in the loss function, which can be formulated as

$$\arg\min \operatorname{rank}(L) + \lambda_1 ||S||_0 + \lambda_2 ||N||_0, \text{s.t.} X = L + S + N. \tag{2}$$

When solving for this equation, it also should be degenerated into the convex format. A CDA that represents a data space as a three-component mixture model in orthogonal decomposition has been proposed [41].

These mentioned detectors reconstruct the background of the HSIs either from the spatial (width–height) aspect or from the spatial-spectral (three order tensor) aspect. However, seen from the height–depth aspect, the profiles in the row direction are also highly correlated, which can be specially utilized for reconstructing the background.

B. Autoregressive Model

The vector autoregressive model was first applied to model the multivariate time series with temporal dependence [42]. As the time series data were changed from the single variable into the multivariate ones, the traditional autoregressive model is too simple to formulate the process [43]. The vector autoregressive model consists of not only the serial dependence within each marginal series but also the interdependence across different marginal series. In this way, it has been utilized in the statical, economic, and even signal processing aspects [44]. It assumes that there is a linear relationship between the output variables and the previous variables. The traditional vector autoregressive model can be modeled as

$$y_t = \sum_{k=1}^{d} \alpha_k y_{t-k} + \in_t, t = d+1, \ldots, T \tag{3}$$

where α_k represents the regression coefficient. Here, d is the order of the model, and it can also be regarded as the number of previous series, which are correlated with the current one. The aim of the loss function is solving for the α_k. $\in_t$ is usually treated as the Gaussian noise. The least square method is utilized for the solving process. Inspired by the effectiveness of VAR in predicting the time series, many works have been proposed to make a further improvement of the VAR, such as the sparse VAR (sVAR). The sVAR constrains that most of the coefficients are zero, which suits for the situation that only a few previous series make a significant influence of the current one [45].

III. Proposed MARM

The proposal contains three parts: band selection, background reconstruction via MARM, and final detection. The detailed description about these three parts has been presented in the following sections.

Fig. 1 exhibits the schematic diagram of the proposed method, which consists of the following steps. First, the bands in the HSI are selected via an optimal neighboring reconstruction strategy, an operation neglects the redundant bands or noisy bands. Then, the previous p observed row profiles are utilized to formulate an MARM. The model can be utilized to predict the following row profile, and then a background HSI can be obtained. Finally, the classical RX detector is applied to the difference cube between the band-selected HSI and the HSI reconstructed by MARM, and achieving the final detection map.

To facilitate discussion, some frequently used terms are expressed here. Let $\bar{X} \in R^{w \times h \times B}$ represent the original HSI with B bands, in which w and h denote the width and the height of the scene. $X \in R^{w \times h \times k}$ denotes the selected HSI with k bands,

and it can be reorganized into a row cube $\hat{X} \in R^{h\times k\times w}$, which can be viewed as a collection of row profiles. Given a specific row number i (i ranges from 1 to w), it comes into a matrix $\hat{X}_i \in R^{h\times k}$, and we call it a row profile. For the HSIs, the pixels in the rows are highly correlated in the spatial domain. Suppose a series of row profiles have been observed, they can be denoted as a sequence $[\hat{X}_{i+1}\ \hat{X}_{i+2}\ \hat{X}_{i+3}\ \cdots\ \hat{X}_{i+p}] \in R^{h\times k\times p}$ and be utilized to infer the following band $\hat{X}_{i+p+1}$.

A. Band Selection

There are hundreds of bands in the HSI that come into their rich spectral discrimination ability. However, these bands not only burden the computational complexity but also convey some redundant or noisy information, which hinders the accuracy of their applications [46].

In this article, we incorporate the simple but effective optimal neighboring reconstruction method to select some representative bands. Suppose there are k bands selected from the original HSI $\bar{X}$, the optimization function is formulated as

$$\arg\min_{k,W} \mathcal{L}(D) \quad \text{s.t.} f(\bar{X}) = f([\bar{X}_{s1}\bar{X}_{s2}\ldots\bar{X}_{sk}])W + D \quad (4)$$

in which $W = [w_1, w_2, \ldots, w_k] \in R^{k\times B}$ and $D = [d_1, d_2, \ldots, d_k] \in R^{M\times B}$ denote weight matrix and error matrix, respectively. $f()$ is a function that reorganizes the three-order tensor or two-order matrix into their previous order format, such as the two-order matrix or one-order vector. $s1, s2, \ldots, sk$ are the indices of the selected bands in the $\bar{X}$. The loss function $\mathcal{L}()$ is expressed as

$$\mathcal{L}(D) = \sum_{j=1}^{k} g_\tau(||d_j||_2) \quad (5)$$

where $g_\tau()$ is a noise reducer which limits the upper bound of the error. It is formulated as

$$g_\tau(e) = \begin{cases} e & e \le \tau \\ \tau & e > \tau \end{cases}. \quad (6)$$

With the design of $g_\tau()$, one single inexactly reconstructed band will not greatly influence the search process. Optimization function of expression (4) can be converted into two subproblems and be optimized by the augmented Lagrangian method [47].

B. Matrix Autoregressive Model

Once obtained the band-selected HSI X, it can be reorganized into a new HSI $\hat{X}$, which can be treated as a collection of w row profiles. Different from the spectral difference between neighboring bands, it is normal that responses at different wavelengths are different, which comes into the spectral discrimination. The neighboring row profiles are supposed to be consistent in the background scene. It should be noted that the autoregressive model is a typically statistical model in the time series analysis, in which the current series is highly correlated with the previous series [48]. This process can be described as

$$x_t = \sum_i A_i x_{t-hi} + \varepsilon_t \quad (7)$$

in which $x_t \in R^{M\times 1}$ is the time series, $\mathrm{H} = \{h_1, \ldots, h_d\}$ is a set of time lags, A is a coefficients matrix, and ε_t represents the Gaussian noise.

In this article, we borrow it for analyzing the HSI row profiles. Different from the time series which usually presents in a format of vector, the row profiles are matrices. In this way, the autoregressive model is named MARM. Let $\hat{X}_i$ denote the matrix with size of $h \times k$ at the ith row, and the matrix-valued row series under the autoregressive framework takes the form

$$\hat{X}_i = A\hat{X}_{i-1}B^T + E_i \quad (8)$$

in which both $A \in R^{h\times h}$ and $B \in R^{k\times k}$ are square matrices, which are the coefficient matrices of the MARM. Here, h represents the height of the HSI, and k is the number of bands. It means the current row profile X_i can be linearly represented by its former profile X_{i-1} in the MARM. $E_t \in R^{h\times k}$ denotes the regression error.

To solve for the two coefficient matrices A and B in the object function, we have minimized the Frobenius norm of the regression error. In this way, the optimization function can be formulated as follows:

$$\arg\min_{A,B} \frac{1}{2}\sum_{i=2}^{p} ||\hat{X}_i - A\hat{X}_{i-1}B^T||_F^2. \quad (9)$$

It should be noted that both the coefficient $\frac{1}{2}$ and square operation of the Frobenius norm are utilized for the convenience of optimization while making no effect on the optimal variables of the functions.

Taking partial derivatives of (9) with respect to the entries of A and B, respectively, the gradient condition for the alternating least square method can be obtained as

$$\sum_i AX_{i-1}B^TBX_{i-1}^T - \sum_i X_iBX_{i-1}^T = 0$$
$$\sum_i BX_{i-1}^TA^TAX_{i-1} - \sum_i X_i^TAX_{i-1} = 0. \quad (10)$$

The expression (9) is guaranteed to have at least one global minimum, so the solutions of expression (10) are guaranteed to exist. In this way, the optimization of A and B can be expressed as

$$A^l = \left(\sum_{i=2}^{p} \hat{X}_i B^{l-1}\hat{X}_{i-1}^T\right)\left(\sum_{i=2}^{p} \hat{X}_{i-1}(B^{l-1})^T B^{l-1}\hat{X}_{i-1}^T\right)^{-1}$$
$$B^l = \left(\sum_{i=2}^{p} \hat{X}_i^T A^l \hat{X}_{i-1}\right)\left(\sum_{i=2}^{p} \hat{X}_{i-1}^T (A^l)^T A^l \hat{X}_{i-1}\right)^{-1} \quad (11)$$

in which l ranges from 1 to the maximum iteration number. The maximum iteration number is empirically set as 100 in the proposed method. When solving for A^1, there is a B^0 involved, which is randomly initialized.

Once the optimal A and B are achieved, the $p+1$th row profiles can be reconstructed. Let i iterates from 1 to $w-p$, $w-p$ subsets $([\hat{X}_1,\ \hat{X}_{i+2},\ \hat{X}_{i+3},\ \cdots, \hat{X}_p]\ , \ldots,$ $[\hat{X}_{w-p},\ \hat{X}_{w-p+1},\ \hat{X}_{w-p+2},\ \cdots, \hat{X}_{w-1}])$ can be obtained, thus

TABLE I
INFORMATION OF THE FOUR TEST HSIS

name	sensor	spatial resolution	size	anomalous object	number of anomalous pixels
SanDiego	AVIRIS	3.5m	$100 \times 100 \times 186$	airplanes	57
HYDICE	HYDICE	1.56m	$80 \times 100 \times 175$	roofs and cars	21
airport-2	AVIRIS	7.1m	$100 \times 100 \times 205$	airplanes	87
beach-4	ROSIS	1.3m	$150 \times 150 \times 102$	vehicles on the bridge	68

Algorithm 1: Pseudocode of the MARM Algorithm.

Input: The original HSI, $\bar{X}$; The number of bands to be selected, k; The number of rows to utilized for regression, p;
Output: The detection map;
1: Extracting k bands from the $\bar{X}$ via exp 4, and achieving the HSI X;
2: Reorganize the X into $\hat{X}$;
3: **for** $i = 0; i < w - p; i++$ **do**
4: The continuous p row profiles $(\hat{X}_{i+1}, \hat{X}_{i+2}, \ldots, \hat{X}_{i+p})$ in $\hat{X}$ are utilized for solving A and B;
5: Achieving the $\hat{H}_{i+p+1}$;
6: **end for**
7: reorganize the $\hat{H}$ into H;
8: The Global-RX applied to X-H;
9: **return** Detection map;

reconstructing the row profiles $(\hat{H}_{p+1}, \ldots, \hat{H}_w)$. The former p profiles are inherited from $\hat{X}$. The reconstructed background HSI $\hat{H} \in R^{h \times k \times w}$ can be transposed into $H \in R^{w \times h \times k}$, whose first p rows are the same with those of the band-selected HSI X.

C. Final Detection via the Global-RX Detector

The difference between the band-selected HSI X and the background HSI H can be denoted via a simple subtraction operation

$$C = X - H \tag{12}$$

in which H is another format of the row profiles reconstructed by A, X, and B. In this way, C can be viewed as the reconstruction error of the X and H, which are supposed to follow the Gaussian distribution. In this way, the classical Global-RX detector is applied to C achieving the final detection map. The proposed MARM can be summarized in Algorithm 1.

IV. EXPERIMENTS

A. Datasets

To verify the effectiveness of our proposed MARM for hyperspectral anomaly detection, four HSIs captured by three different sensors with different resolutions are utilized, which are described in detail as follows.

1) HSIs Captured by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) Sensor: There are two HSIs captured by the AVIRIS being applied to the experiment. The first one depicts the airport region over the San Diego airport, CA, USA, in 1998, whose spatial resolution is 3.5 m per pixel. There are 224 bands ranging from the ultraviolet 370 nm to the near-infrared 2510 nm. After removing the low-SNR and water absorption bands, 189 bands are left for the experiment, whose spatial size is 100×100. The three aircrafts are marked as anomalies, with a number of 58 pixels. The second HSI is the airport-2 HSI, which were captured during a flight over the Los Angeles region in 2011, whose spatial resolution is 7.1 m per pixel [50]. After removing the bands with low quality, the size of the HSI utilized for the experiment is $100 \times 100 \times 205$, among which 87 pixels are marked as anomalies.

2) HSI Captured by the Hyperspectral Digital Imagery Collection Experiment (HYDICE) Sensor: There is an HSI captured by the HYDICE sensor being utilized for the experiment. It is collected over an urban area, CA, USA, whose size is $80 \times 100 \times 175$. There are 21 pixels being marked as anomalies, which are cars and roofs. The spectrum ranges from 400 to 2500 nm. Its spatial resolution is 1.56 m per pixel and the spectral resolution is 10 nm per band. The background land-cover types include parking lot and some others.

3) HSI Captured by the Reflective Optics System Imaging Spectrometer (ROSIS) Sensor: The beach-4 was captured by the ROSIS, which depicts the Pavia region. Its size is $150 \times 150 \times 102$, whose spatial resolution is 1.3 m per pixel. The vehicles on the bridge are marked as the anomalies. The background includes the bridge, river, bare soil, and buildings.

The corresponding detail of these four test datasets has been listed in Table I.

TABLE II
ABLATION STUDY OF THE MODEL

	MARM-RX$^\circ$		MARM-RX*		MARM-RX	
	BS ($\surd$)	MAR ($\times$)	BS ($\times$)	MAR ($\surd$)	BS ($\surd$)	MAR ($\surd$)
SanDiego	0.9653		0.9606		0.9656	
HYDICE	0.9895		0.9895		0.9944	
airport-2	0.9342		0.9464		0.9643	
beach-4	0.9676		0.9759		0.9767	
Average	0.9642		0.9681		0.9753	

B. Experimental Setup

In the experiments, comparison is made with several other methods, which also preprocessed the original HSI via some typical methods, including the Deriv-RX, the FrFT-RX, and the classical Global-RX. Meanwhile, two typical matrix decomposition methods are also applied for the comparison, including the LSDM-MoG and PTA. The parameters are set according to the suggestion in the corresponding paper, such as the initial rank and number of mixture Gaussian noise in the LSDM-MoG, the truncated low-rank, and some other penalty parameters. In addition, comparison has also been made with two deep learning

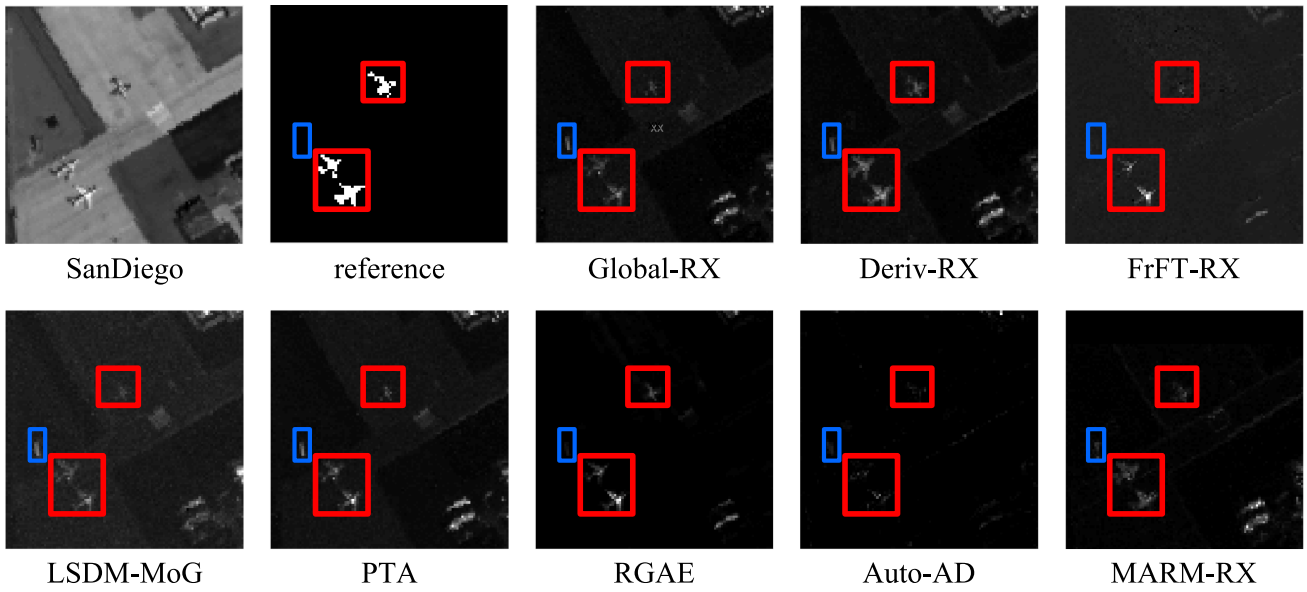


Fig. 2. Visual difference of different methods on the SanDiego.

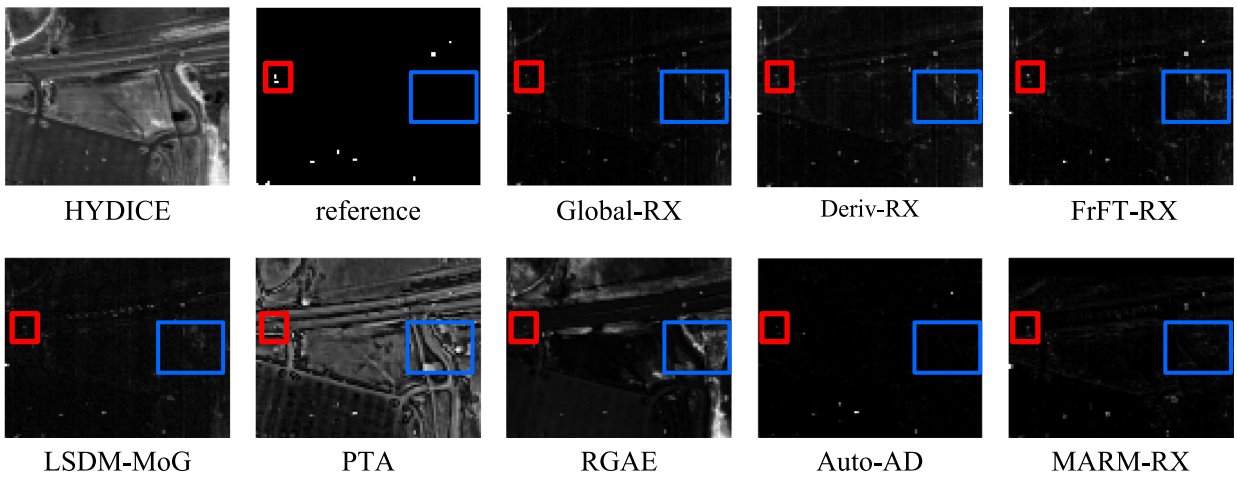


Fig. 3. Visual difference of different methods on the HYDICE.

based anomaly detection method RGAE and the fully convolutional autoencoder (Auto-AD). For the RGAE, three parameters are adjusted according to the paper to find the best performance. It should be noted that there are two parameters involved in the MARM, including the number of bands being selected (k) and the number of row profiles (p) exploited in the MARM. The k is empirically iterated from 20 to 30 via a stride 1, and the p is iterated from 5 to 15 to obtain the optimal detection performance. The 3-D receiver operating characteristics (ROC) analysis [49] is incorporated to make a measurement of the proposed method, which includes three area under the curves (AUCs). The three AUCs are the $\mathrm{AUC}_{(D,F)}$, $\mathrm{AUC}_{(D,\tau)}$, and $\mathrm{AUC}_{(F,\tau)}$. These three measurements, respectively, measure the effectiveness of the detector, the target detection ability of the detector, and the background suppression ability of the detector. For the first two measurements, the higher, the better. The $\mathrm{AUC}_{(F,\tau)}$ is on the contrary. Moreover, visual detection maps achieved by different detectors have also been exhibited to make a visual assessment.

C. Ablation Study

Experiments have been conducted to illustrate the effectiveness of the model parameters, including the band selection and the MARM model. The corresponding AUCs of the four HSIs have been listed in Table II, in which "BS" and "MAR" represent the band selection module and the matrix autoregressive module, respectively. The first column data illustrate the AUC values of the Global-RX detector directly applied to the band-selected HSI, which has been denoted by a superscript°. Performance comparison between the Global-RX and the MARM-RX° has validated the effectiveness of the band selection strategy. Meanwhile, comparison between the MARM-RX° and the MARM-RX has further validated the effectiveness of the MARM.

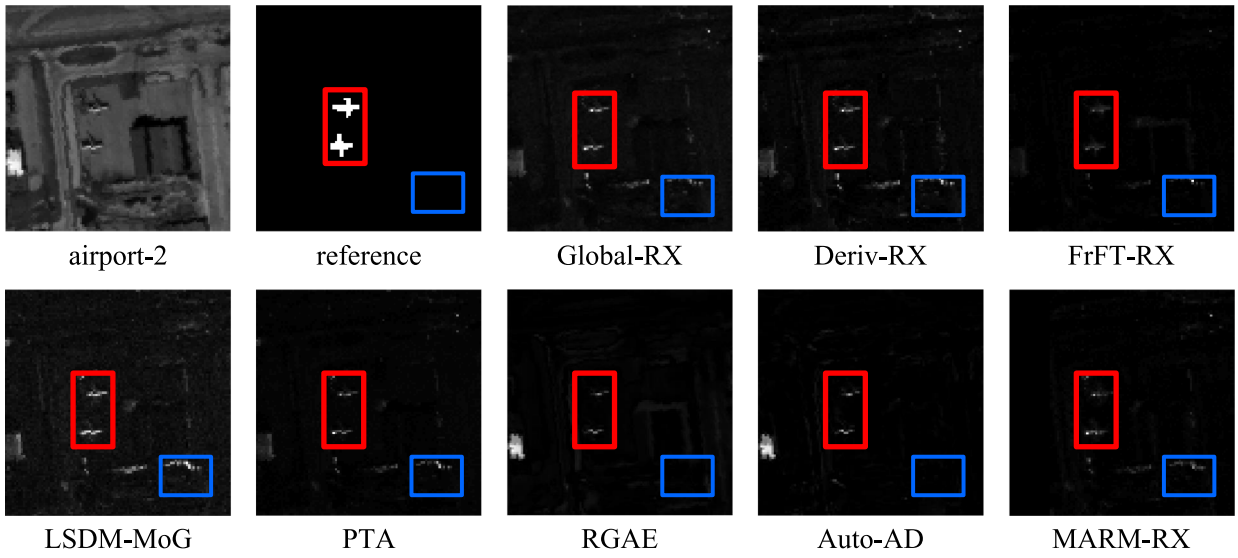


Fig. 4. Visual difference of different methods on the airport-2.

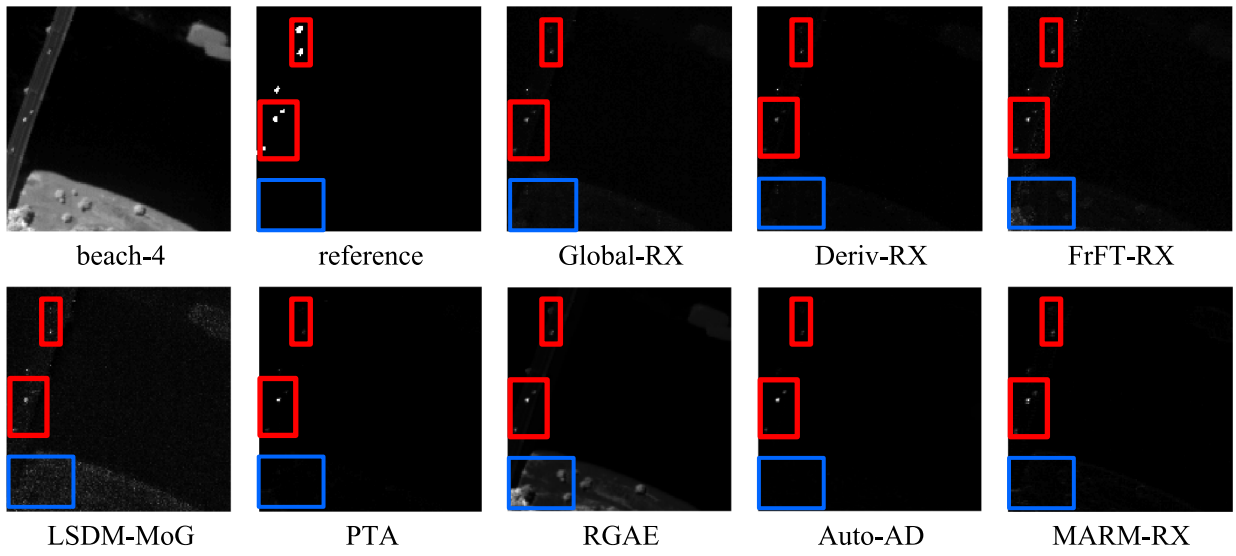


Fig. 5. Visual difference of different methods on the beach-4.

In addition, the second column denotes the whole band being sent into the MARM, which has been marked by a superscript*. Comparison between the MARM-RX* and MARM-RX has demonstrated the effectiveness of the band selection.

D. Data Analysis

Figs. 2–5 and have exhibited the visual detection maps of the proposed MARM and its competitors, in which the blue rectangles denote the regions which do not contain the anomalies, and the red rectangles represent the regions which contain the anomalies. The first column denotes the 100th band of the test HSIs, and the second column is the referenced detection map. All the rest of columns represent the maps achieved by the other detectors.

The corresponding objective measurements have been listed in Table III, in which the optimal and the suboptimum $\text{AUC}_{(D,F)}$ have been highlighted by the bold and underline format, respectively. The $\text{AUC}_{(D,F)}$ is utilized to measure the effectiveness of the detector, the $\text{AUC}_{(D,\tau)}$ is utilized to detect the target detection ability of the detector, and $\text{AUC}_{(F,\tau)}$ is utilized to measure the background suppression ability of the detector. It is noted that the MARM always effectively suppresses the background and achieves the acceptable detection performance.

When it comes to the visual exhibition of the SanDiego in Fig. 2, it is noticed that both the RGAE and the Auto-AD detectors exhibit comparatively low responses to the backgrounds, which is consistent to the $\text{AUC}_{(F,\tau)}$ of these two detectors in Table III. However, the Auto-AD detector also exhibits a low response to the anomalies, resulting its low detection rate.

TABLE III
MEASUREMENTS OF THE PROPOSED METHOD AND THE COMPETITORS ON THE FOUR HSIS

		Global-RX	Deriv-RX	FrFT-RX	LSDM-MoG	PTA	RGAE	Auto-AD	MARM-RX
SanDiego	$\mathbf{AUC}_{(D,F)}$	0.9403	<u>0.9788</u>	0.7790	0.9318	0.9294	**0.9920**	0.9455	0.9656
	$\mathbf{AUC}_{(D,\tau)}$	0.1778	0.2481	0.2260	0.2332	0.2040	0.1814	0.0552	0.1452
	$\mathbf{AUC}_{(F,\tau)}$	0.0589	0.0532	0.1184	0.1280	0.0701	0.0080	0.0043	0.0225
HYDICE	$\mathbf{AUC}_{(D,F)}$	0.9858	0.9875	<u>0.9930</u>	0.9727	0.7963	0.7810	0.9751	**0.9944**
	$\mathbf{AUC}_{(D,\tau)}$	0.2404	0.3088	0.4388	0.2429	0.4386	0.1847	0.2343	0.2387
	$\mathbf{AUC}_{(F,\tau)}$	0.0350	0.0465	0.0395	0.0479	0.2270	0.0905	0.0049	0.0212
airport-2	$\mathbf{AUC}_{(D,F)}$	0.8404	0.8840	<u>0.9604</u>	0.8419	0.8428	0.7484	0.8623	**0.9643**
	$\mathbf{AUC}_{(D,\tau)}$	0.1841	0.1650	0.1013	0.2680	0.1223	0.0775	0.0855	0.1264
	$\mathbf{AUC}_{(F,\tau)}$	0.0516	0.0394	0.0147	0.0759	0.0275	0.0196	0.0106	0.0156
beach-4	$\mathbf{AUC}_{(D,F)}$	0.9531	0.9589	0.9222	0.8801	0.9597	0.9187	<u>0.9742</u>	**0.9767**
	$\mathbf{AUC}_{(D,\tau)}$	0.1214	0.0921	0.1359	0.1788	0.0742	0.1296	0.0909	0.0960
	$\mathbf{AUC}_{(F,\tau)}$	0.0233	0.0122	0.0254	0.0540	0.0061	0.0233	0.0009	0.0097
Average	$\mathbf{AUC}_{(D,F)}$	0.9299	<u>0.9523</u>	0.9141	0.9066	0.8821	0.8600	0.9393	**0.9752**
	$\mathbf{AUC}_{(D,\tau)}$	0.1809	0.2035	0.2291	0.2307	0.2098	0.1433	0.1165	0.1516
	$\mathbf{AUC}_{(F,\tau)}$	0.0422	0.0378	0.0495	0.0765	0.0826	0.0353	0.0052	0.0172

*:The optimal and second optimal values have been highlighted by the bold and underline format, respectively.

Both the low false alarm rate and the low detection rate lead to their unsatisfaction. The Global-RX, Deriv-RX, and the PTA exhibit higher false alarm rates, especially for the runway. When compared with the Deriv-RX detector, it is visually observed that the proposed MARM-RX detector tackles the right-up background regions more precisely. In this way, the $\mathrm{AUC}_{(F,\tau)}$ achieved by the proposed method outperforms that of the Deriv-RX. The FrFT-RX achieves high responses for the region around the upper red rectangle, which is easily mistaken as the anomalies. Compared with both the reconstruction-based methods, LSDM-MoG and PTA, the proposed MARM-RX exhibits the lowest response to the blue rectangle regions. It illustrates the best background suppression ability of the proposed MARM-RX, which is further proved by the $\mathrm{AUC}_{(F,\tau)}$ in Table III. In addition, the visual detection map of the proposed MARM-RX demonstrates a comparable performance with contrast to that of the RGAE. Nevertheless, the performance of the RGAE is highly related to the training process, which can be unstable with the variation of the datasets.

To make a comprehensive evaluation of the performance, the box plots are applied to tabular the distributions of the backgrounds and the anomalies in the response maps and to improve the interpretation of the data [51]. Fig. 6 demonstrates the distribution of the SanDiego detection maps. Different indices in the x-axis represent different detectors, which are listed at the right top of the figure. The green boxes and the blue boxes denote the distribution of the background and the anomalies, respectively. According to the data in Fig. 6, it is noticed that for the Deriv-RX, the RGAE, and the proposed MARM-RX, their lower hinges of the anomalies are still larger than the extremely large responses of the background. This proves the discrimination ability of these three detectors. For the LSDM-MoG and the PTA, their median responses of the backgrounds are almost next to those of the anomalies, which is difficult to distinguish the anomalies from the backgrounds.

For the HYDICE, the detection maps of different detectors have been plotted in Fig. 3, and it is noticed that both the PTA and the RGAE exhibit poor detection performance for this data, which is consistent with the $\mathrm{AUC}_{(D,F)}$ in Table III.

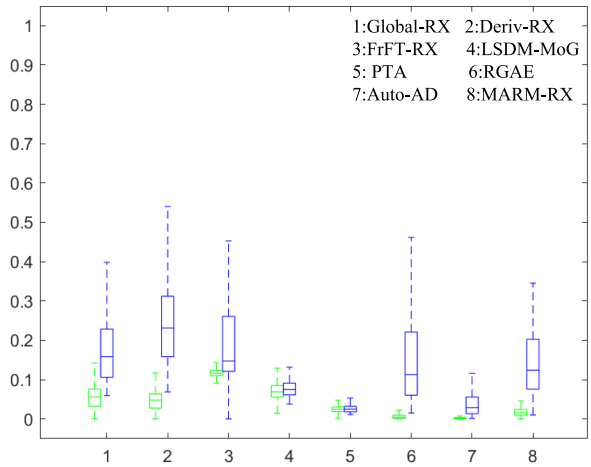


Fig. 6. Box plot of different methods on the SanDiego.

Meanwhile, the Auto-AD still exhibits the low responses to the whole scene and achieves the lowest $\mathrm{AUC}_{(F,\tau)}$. For the blue rectangle regions which are backgrounds, the proposed MARM-RX and the Auto-AD exhibit the optimal suppression performance. All the other competitors exhibit comparatively large responses to the background and are with comparatively higher false alarm rates. In addition, it is observed from all the detection maps that the Auto-AD always exhibits low responses to both the backgrounds and the anomalies. In this way, it can achieve the appealing background suppression performance but can miss some anomalies at the same time.

According to the corresponding box plot in Fig. 7, it is noticed that for the PTA and the RGAE detectors, the extremely large responses of the background are larger than the median responses of the anomalies. This phenomenon has explained the poor detection accuracy of both detectors in Table III. Compared with the FrFT-RX and Auto-AD, the anomalous box of the proposed MARM-RX is much more balanced. To be specific, the half-upper boxes (from the median value to the upper hinge) of the FrFT-RX and Auto AD are severally imbalanced from their half-lower boxes (from the lower hinge to the median

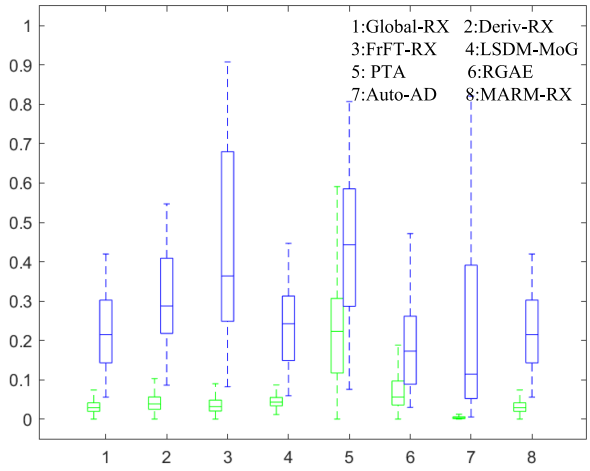


Fig. 7. Box plot of different methods on the HYDICE.

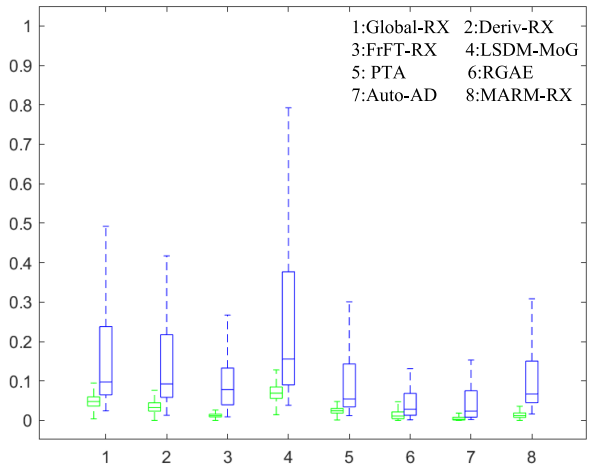


Fig. 8. Box plot of different methods on the airport-2.

value). This imbalance illustrates the distribution skewness of the anomalies in the maps.

As to the airport-2 HSI, all the Global-RX, Deriv-RX, and LSDM-MoG exhibit poor background suppression performance, especially for the blue rectangle regions. The FrFT-RX shows an overall low response neither to the anomalies nor to the backgrounds. Visually observed from Fig. 4, the PTA achieves a nearly same detection performance as the proposed MARM-RX. However, for the left-bottom background region, the proposed MARM-RX is with lower responses, which shows its better background suppression performance over the PTA. Both the deep learning based methods, RGAE and Auto-AD, mistakenly take the left buildings as the anomalies. This leads to their poor detection rate $\mathrm{AUC}_{(D,\tau)}$ in Table III.

According to the box plot in Fig. 8, it is noticed that the responses of FrFT-RX, RGAE, and Auto-AD are much smaller than the other detectors. It further proves the aforementioned low responses of these three detectors in the visual detection maps. Seen from Fig. 8, the separation between the background and the anomalies achieved by the proposed MARM-RX is the largest among the other competitors. The extremely large value

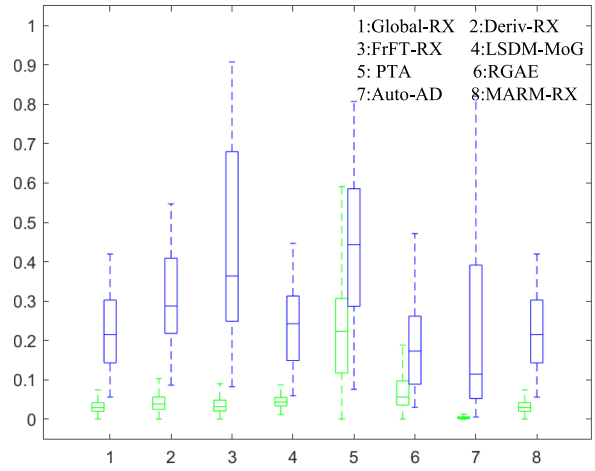


Fig. 9. Box plot of different methods on the beach-4.

of the background is still smaller than the lower hinge of the anomalies. In this way, the proposed MARM-RX achieves the best detection accuracy for the airport-2.

For the beach-4 dataset, neither the Global-RX nor Deriv-RX achieves the acceptable performance for the red rectangle anomalous regions. All the FrFT-RX, LSDM-MoG, and RGAE exhibit high responses for the blue background rectangles, which result in the high false alarm rates. The rest PTA, Auto-AD, and the proposed MARM-RX exhibit the acceptable performance in the background regions. However, when detecting the upper anomalies, both the PTA and the Auto-AD are less sensitive than the proposed MARM-RX making their comparatively low detection rates.

Fig. 9 has listed the box plots of these detectors. It is noticed that the Auto-AD still exhibits the low responses to the whole scene. Meanwhile, for the PTA, the extremely large value of the background is larger than the upper hinge of the background, which indicates its poor $\mathrm{AUC}_{(D,\tau)}$ in Table III.

Furthermore, the corresponding ROC curves of the data in Table III have been plotted in Fig. 10. The detectors that belong to the same category have been marked by the same color, and different markers are utilized to make an intra differentiation. The proposed MARM-RX is plotted by red. It is noticed that the proposed MARM-RX is always with the least false alarm rate to achieve the 100% detection rate. This have demonstrated the effectiveness of the proposed detector in the anomalies-prior situation. Comparison between the matrix-decomposed LSDM-MoG and PTA has demonstrated that the proposed MARM-RX always obtains a better background suppression ability. This is achieved by exploiting the neighboring correlations between the row profiles, which proves more stable to the scene.

Computational costs of different detectors have been listed in Table IV. All the experiments have been conducted on the same platform. It is noted that the RX-based methods are much faster than the background reconstruction-based methods. For the reconstruction-based methods, including the LSDM-MoG, PTA, and the proposed method, searching for the optimal variables takes the main time of the detection process. Compared with both the LSDM-MoG and PTA, the proposed method is faster.

TABLE IV
COMPUTATIONAL COST (IN SECONDS) OF AFOREMENTIONED DETECTORS FOR THREE EXPERIMENTAL DATA

	Global-RX	Deriv-RX	FrFE-RX	LSDM-MoG	PTA	RGAE	Auto-AD	MARM-RX*	MARM-RX
SanDiego	0.16	0.13	8.79	760.44	6227.83	45087.13	40.93	2326.13	690.15
HYDICE	0.10	0.10	6.42	519.34	4508.49	40499.70	35.48	588.51	420.38
airport-2	0.15	0.14	8.42	1004.83	6899.10	46156.36	61.31	3780.99	1316.32
beach-4	0.14	0.13	12.33	1265.80	9064.33	57234.91	39.95	3496.24	988.09
Average	0.14	0.13	8.99	887.60	6674.94	47244.52	44.42	2547.96	853.74

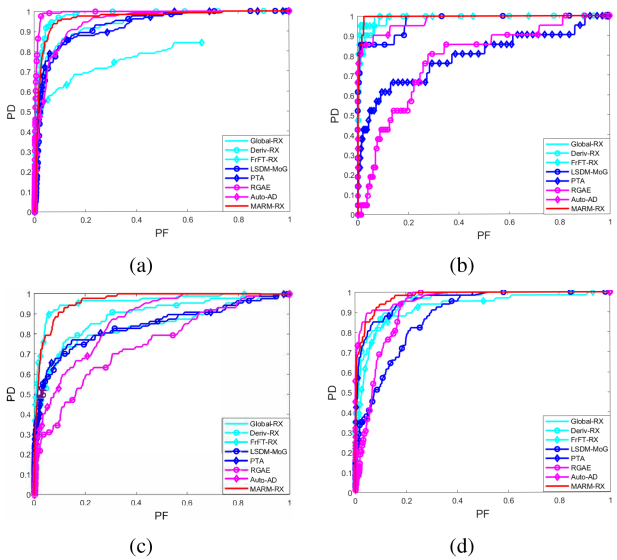


Fig. 10. 3-D ROC curves of different methods on different HSIs. (a) SanDiego. (b) HYDICE. (c) airport-2. (d) beach-4.

To be specific, for the LSDM-MoG, there are two variables to be determined. The main computational cost is caused by estimation of the low-rank component and noise component; for the PTA, there are five parameters to be determined, including the truncated low-rank r and four other hyperparameters. Among the four hyperparameters, three ones are fixed for simplification. In this way, there are two remaining variables that should be iterated for the optimal value. However, when solving for the tensor approximation, four new variables have been imported to make the loss function into an auxiliary Lagrangian function, making the detection process time-consuming; for the proposed method, there are two parameters involved. The main computational complexity is caused by the regressive representation process. However, with the band selection module, the input matrix is much smaller than the original one, which lessens the computational costs at a large degree. For the RGAE detector, there are three parameters being involved. The training is processed for all the probable parameters, making its heavy computational burden. In addition, compared with the MARM-RX* without band selection, the proposed method not only achieves the higher detection accuracy but also achieves a less computational cost.

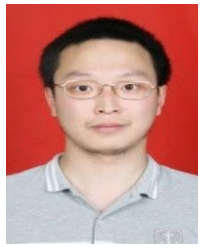
V. CONCLUSION

A novel hyperspectral anomaly detection method has been proposed via MARM in this article. It follows the original matrix structure in the row profiles, and both the spatial and spectral correlation are utilized for the detection process. Meanwhile, band selection is first applied to eliminate some noisy bands and redundant bands, and it also reduces the computational cost of the proposed methods. Experimental results and data analysis on four HSIs, which were captured by three different sensors and with different resolutions, have validated the effectiveness of the proposed method.

REFERENCES

[1] I. Racetin and A. Krtali, "Systematic review of anomaly detection in hyperspectral remote sensing applications," *Appl. Sci.*, vol. 11, no. 11, Jun. 2021, Art. no. 4878.
[2] J. H. Qu, Q. Du, Y. S. Li, L. Tian, and H. M. Xia, "Anomaly detection in hyperspectral imagery based on Gaussian mixture model," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9504–9517, Nov. 2021.
[3] M. H. Khan, Z. Saleem, M. Ahmad, A. Sohaib, H. Ayaz, and M. Mazzara, "Hyperspectral imaging for color adulteration detection in red chili," *Appl. Sci.*, vol. 10, no. 17, 2020, Art. no. 5955.
[4] C. McCann et al., "Multi-temporal mesoscale hyperspectral data of mixed agricultural and grassland regions for anomaly detection," *ISPRS J. Photo. Remote Sens.*, vol. 131, pp. 121–133, 2017.
[5] D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 58–69, Jan. 2002.
[6] H. Su, Z. Wu, H. Zhang, and Q. Du, "Hyperspectral anomaly detection: A survey," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 64–90, Mar. 2022.
[7] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.
[8] J. M. Molero, E. M. Garzón, I. García, and A. Plaza, "Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 801–814, Apr. 2013.
[9] Y. N. Gao, T. Cheng, and B. Wang, "Nonlinear anomaly detection based on spectral CSpatial composite kernel for hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 7, pp. 1269–1273, Jul. 2021.
[10] J. Zhou, C. Kwan, B. Ayhanand, and M. Eismann, "A novel cluster kernel RX algorithm for anomaly and change detection using hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6497–6504, Nov. 2016.
[11] R. Tao, X. Zhao, W. Li, H.-C. Li, and Q. Du, "Hyperspectral anomaly detection by fractional fourier entropy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 4920–4929, Dec. 2019.
[12] Y. Y. Tang, Y. Lu, and H. Yuan, "Hyperspectral image classification based on three-dimensional scattering wavelet transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2467–2480, May 2015.
[13] D. Liu and L. Han, "Spectral curve shape matching using derivatives in hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 504–508, Apr. 2017.
[14] J. Li, H. Zhang, L. Zhang, and L. Ma, "Hyperspectral anomaly detection by the use of background joint sparse representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2523–2533, Jun. 2015.
[15] N. Huyan, X. Zhang, H. Zhou, and L. Jiao, "Hyperspectral anomaly detection via background and potential anomaly dictionaries construction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2263–2276, Apr. 2019.
[16] H. Fan, Y. Chen, Y. Guo, H. Zhang, and G. Kuang, "Hyperspectral image restoration using low-rank tensor recovery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 10, pp. 4589–4604, Oct. 2017.

[17] R. Zhao, B. Du, and L. Zhang, "Hyperspectral anomaly detection via a sparsity score estimation framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3208–3222, Jun. 2017.
[18] W. Y. Xie, X. Zhang, Y. S. Li, J. Lei, J. J. Li, and Q. Du, "Weakly supervised low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 51, no. 8, pp. 3889–3990, Aug. 2021.
[19] W. Xie, X. Zhang, Y. Li, K. Wang, and Q. Du, "Background learning based on target suppression constraint for hyperspectral target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5887–5897, 2020.
[20] Y. Xu, B. Du, L. Zhang, and S. Chang, "A low-rank and sparse matrix decomposition- based dictionary reconstruction and anomaly extraction framework for hyperspectral anomaly detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 7, pp. 1248–1252, Jul. 2020.
[21] L. Li, W. Li, Q. Du, and R. Tao, "Low-rank and sparse decomposition with mixture of Gaussian for hyperspectral anomaly detection," *IEEE Trans. Cyb.*, vol. 59, no. 9, pp. 4363–4372, Sep. 2021.
[22] T. Cheng and B. Wang, "Graph and total variation regularized low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 391–406, Jan. 2020.
[23] T. Zhou and D. Tao, "GoDec: Randomized low-rank & sparse matrix decomposition in noisy case," in *Proc. 28th Int. Conf. Mach. Learn.*, Bellevue, WA, USA, 2011, pp. 33–40.
[24] C. I. Chang and J. Chen, "Hyperspectral anomaly detection by data sphering and sparsity density peaks," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5526321.
[25] S. Song, H. Zhou, L. Gu, Y. Yang, and Y. Yang, "Hyperspectral anomaly detection via tensor-based endmember extraction and low rank decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 10, pp. 1772–1776, Oct. 2020.
[26] S. Y. Sun, J. Liu, X. Chen, W. Li, and H. B. Li, "Hyperspectral anomaly detection with tensor average rank and piecewise smoothness constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, Mar. 2022, early access, doi: 10.1109/TNNLS.2022.3152252.
[27] X. Zhang, G. Wen, and W. Dai, "A tensor decomposition-based anomaly detection algorithm for hyperspectral image," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5801–5820, Oct. 2016.
[28] J. Wang, Y. Xia, and Y. Zhang, "Anomaly detection of hyperspectral image via tensor completion," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 6, pp. 1099–1103, Jun. 2021.
[29] L. Li, W. Li, Y. Qu, C. Zhao, R. Tao, and Q. Du, "Prior-based tensor approximation for anomaly detection in hyperspectral imagery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 3, pp. 1037–1050, Mar. 2022.
[30] T. Jiang, Y. Li, W. Xie, and Q. Du, "Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4666–4679, Jul. 2020.
[31] W. Xie, B. Liu, Y. Li, J. Lei, C.-I. Chang, and G. He, "Spectral adversarial feature learning for anomaly detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2352–2365, Apr. 2020.
[32] D. Zhu, B. Du, and L. Zhang, "EDLAD: An encoder-decoder long short-term memory network-based anomaly detector for hyperspectral images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2021, pp. 4412–4415, doi: 10.1109/IGARSS47720.2021.9553145.
[33] S. Wang, X. Wang, L. Zhang, and Y. Zhong, "Auto-AD: Autonomous hyperspectral anomaly detection network based on fully convolutional autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5503314.
[34] G. Fan, Y. Ma, X. Mei, F. Fan, J. Huang, and J. Ma, "Hyperspectral anomaly detection with robust graph autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art no. 5511314, doi: 10.1109/TGRS.2021.3097097.
[35] H. Luo, H. Zhu, S. Liu, Y. Liu, X. Zhu, and J. Lai, "3D auxiliary classifier GAN for hyperspectral anomaly detection via weakly supervised learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 6009805.
[36] J. Qi, Z. Gong, W. Xue, X. Liu, A. Yao, and P. Zhong, "An unmixing based network for underwater target detection from hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 5470–5487, 2021.
[37] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, pp. 1027–1063, 2009.
[38] Z. Zhou, X. Li, J. Wright, E. J. Candes, and Y. Ma, "Stable principal component pursuit," in *Proc. IEEE Int. Symp. Inf. Theory*, 2010, pp. 1518–1522.
[39] S. Jia, X. Zhang, and Q. Li, "Spectral Cspatial hyperspectral image classification using $l_{1/2}$ regularized low-rank representation and sparse representation-based graph cuts," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2473–2484, Jun. 2015.
[40] N. Vaswani, Y. Chi, and T. Bouwmans, "Rethinking PCA for modern data sets: Theory, algorithms, and applications," *Proc. IEEE*, vol. 106, no. 8, pp. 1274–1276, Aug. 2018.
[41] S. Chen, C.-I. Chang, and X. Li, "Component decomposition analysis for hyperspectral target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5516222.
[42] S. E. Heaps, "Enforcing stationarity through the prior in vector autoregressions," *J. Comput. Graph. Statist.*, to be published, doi: 10.1080/10618600.2022.2079648.
[43] A. Chen, N. Thai, and M. Stadje, "Risk management with multiple VaR constraints," *Math. Methods Oper. Res.*, vol. 88, no. 2, pp. 297–337, Oct., 2018.
[44] F. Abbasi and I. Riaz, "CO2 emissions and financial development in an emerging economy: An augmented VAR approach," *Energy Policy*, vol. 90, pp. 102–114, Mar., 2016.
[45] M. Billio, R. Casarin, and L. Rossini, "Bayesian nonparametric sparse VAR models," *J. Econ.*, vol. 212, no. 1, pp. 97–115, Sep. 2019.
[46] W. Xie, Y. Li, J. Lei, J. Yang, C.-I. Chang, and Z. Li, "Hyperspectral band selection for spectral-spatial anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3426–3436, May 2020.
[47] Q. Wang, F. Zhang, and X. Li, "Hyperspectral band selection via optimal neighborhood reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8465–8476, Dec. 2020.
[48] L. Li, X. Su, Y. Zhang, Y. Lin, and Z. Li, "Trend modeling for traffic time series analysis: An integrated study," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3430–3439, Dec. 2015.
[49] C. I. Chang, "An effective evaluation tool for hyperspectral target detection: 3D receiver operating characteristic curve analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 5131–5153, Jun. 2021.
[50] X. Kang, X. Zhang, S. Li, K. Li, J. Li, and J. A. Benediktsson, "Hyperspectral anomaly detection with attribute and edge-preserving filters," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5600–5611, Oct. 2017.
[51] M. Spitzer, J. Wildenhain, J. Rappsilber, and M. Tyers, "BoxPlotR: A web tool for generation of box plots," *Nature Methods*, vol. 11, no. 2, pp. 121–122, Feb. 2014.



Jingxuan Wang received the B.S. degree in IC design and integrated system and the M.S. degree in integrated circuit engineering from Xidian University, Xi'an, China, in 2013 and 2016, respectively. He is currently working toward the Ph.D. degree in electronics and information with the School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an, China.

His research interests include neural networks, tensor decompositions and applications, hyperspectral image processing, and parallel computing.

Jinqiu Sun received the B.S., M.S., and Ph.D. degrees in guidance navigation and control from Northwestern Polytechnical University, Xi'an, China, in 1999, 2004, and 2005, respectively.

She is currently a Professor with the School of Astronomy, Northwestern Polytechnical University. Her research work focuses on signal and image processing, computer vision, and pattern recognition.

Yu Zhu received the Ph.D. degree in computer science and technology from Northwestern Polytechnical University (NPU), Xi'an, China, in 2017.

He is currently with NPU as an Associate Researcher. His current research interests include computer vision, image super-resolution, and other image fusion topics.

Yong Xia (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in computer science and technology from the School of Computer Science and Engineering, Northwestern Polytechnical University (NPU), Xi'an, China, in 2001, 2004, and 2007, respectively.

He is currently a Professor with the School of Computer Science and Engineering, NPU. His research interests include medical image analysis, computer-aided diagnosis, pattern recognition, machine learning, and data mining.

Yanning Zhang (Senior Member, IEEE) received the B.S. degree in electronic engineering from the Dalian University of Technology, Dalian, China, in 1988, the M.S. degree in electronic engineering from Northwestern Polytechnical University, Xi'an, China, in 1993, and the Ph.D. degree from the School of Marine Engineering, Northwestern Polytechnical University, Xi'an, in 1996.

She is currently a Professor with the School of Computer Science, Northwestern Polytechnical University. She is also a Cheung Kong Professor of Ministry of Education, China. She has authored more than 200 articles. She is also the Organization Chair of the ACCV2009 and the Publicity Chair of the ICME2012. She is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. Her current research interests include remote sensing image analysis, computer vision, and pattern recognition.