Residual-Driven Band Selection for Hyperspectral Anomaly Detection

Xiaodi Shang, *Graduate Student Member, IEEE*, Meiping Song, *Member, IEEE*, Yulei Wang, *Member, IEEE*, and Haoyang Yu, *Member, IEEE*

***Abstract*— This letter proposes an unsupervised band selection (BS) algorithm named residual driven BS (RDBS) to address the lack of *a priori* information about anomalies, obtain a band subset with high representation capability of anomalies, and finally improve the anomaly detection (AD). First, an anomaly and background modeling framework (ABMF) is developed via density peak clustering (DPC) to pre-determine the prior knowledge of the anomalies and background. Then, the DPC-based constraints are applied to R-Anomaly Detector (RAD), and three band prioritization (BP) criteria are derived to obtain the representative band subset for anomalies. Experiments on two datasets show the superiority of RDBS over other BS algorithms and verify that the obtained band subsets are strongly representative of anomalies.**

***Index Terms*— Density peak clustering (DPC), hyperspectral anomaly detection (AD), R-Anomaly Detector (RAD), residual-driven band selection (RDBS).**

I. Introduction

HYPERSPECTRAL anomaly detection (AD) which relies on the difference of statistical distribution between the target and background in the absence of any prior information, has more practical application value and thus has attracted extensive attention [1]. The classic model, Reed–Xiaoli Detector (RXD) [2] based on Mahalanobis distance, is to handle the signal patterns with non-negligible and unknown intensities in several optical bands, which has a strong theoretical basis and can obtain stable performance. Subsequently, other detectors are committed to improving RXD, including R-Anomaly Detector (RAD) [3], kernel RX [4], and so on. These full-band detectors achieve good detection. However, there are inevitably a certain amount of interference bands with some undesirable statistical and geometric properties in hyperspectral image (HSI), such as noisy bands or meaningless bands, which seriously affect the detection accuracy of full-band anomaly detectors.

In fact, each ground object has its unique and significant spectral features that differ from the background [5]. This is also true for anomalies. That is, finding the aforementioned feature set specified by anomaly will help the anomaly detector. For this purpose, band selection (BS) is effective. It not only maintains the physical significance of the ground objects but also retains the abnormal spectral information to the greatest extent [6]. Unfortunately, the lack of anomaly spectral information poses more challenges for BS of AD. Currently, most BS methods are developed for classification or supervised target detection (e.g., [5]–[7]) and are not relevant to AD. Even if BS for AD is available, it is mostly designed based on the amount of image information. For example, He *et al.* [8] proposed a recursive SAM-based BS (RSAM-BBS) and developed an unsupervised anomaly detector for demonstrating the utility of RSAM-BBS. Inspired by RXD, Huber-Lerner *et al.* [9] implemented BS for whitened images based on the bands' Gaussianity level. Furthermore, by fully exploiting the potential physical features that favor AD to constrain the unsupervised network, Xie *et al.* [10] established a selection criterion to adaptively select a band subset containing the discriminative and informative features between the anomaly and background. To retain the key information of abnormal targets in the band, Chang *et al.* [11] proposed a new subspace-selection-based discriminative forest (SSDF) method. Finally, Andika *et al.* [12] proposed a BS algorithm that uses entropy and histogram counts to select effective band subset. These band subsets are difficult to characterize specific anomalies and obtain high detection accuracy.

Therefore, in view of the great significance but limited research on this area, this letter proposes a residual-driven BS (RDBS) algorithm for AD to cope with the absence of prior information and enhance the discrimination ability of band subset for anomalies. First, an anomaly and background modeling framework (ABMF) proposed in RDBS divides images into anomaly set and background set through density peak clustering (DPC) [13] and treats them as prior knowledge to constraint band prioritization (BP) criteria. Then, inspired by RAD, RDBS designs three RDBP criteria depending on the above constraints, namely, minimum signal residual (MinSR), maximum background residual (MaxBR), and minimum signal to background residual ratio (MinSBR). The major contributions are emphasized as follows.

(1) DPC, which can effectively model the anomalies and backgrounds, is adopted to solve the dilemma of difficulty in designing a suitable BP criterion for AD due to the lack of anomaly information, and it is finally used to guide the BS of AD.

(2) Inspired by RAD, this letter proposes a new unsupervised BS algorithm based on residual theory, RDBS, which is a new attempt and breakthrough in the field of AD. The results show that the proposed method can obviously reduce band redundancy and seek the bands more significant to the anomalies, thus improving the detection performance.

Manuscript received June 27, 2021; revised September 1, 2021 and October 19, 2021; accepted November 10, 2021. Date of publication November 12, 2021; date of current version January 10, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61971082 and Grant 61890964 and in part by the Fundamental Research Funds for the Central Universities under Grant 3132019341. *(Corresponding author: Meiping Song.)*

The authors are with the Center for Hyperspectral Imaging in Remote Sensing (CHIRS), Information and Technology College, Dalian Maritime University, Dalian 116026, China (e-mail: shangxd329@dlmu.edu.cn; smping@163.com; 7340912@qq.com; yuhy@dlmu.edu.cn).

Digital Object Identifier 10.1109/LGRS.2021.3128013

II. Anomaly Detector

RXD is an effective and most widely used anomaly detector. By replacing the sample covariance matrix $\mathbf{K}$ with the correlation matrix $\mathbf{R}$, an alternative detector for AD, RAD, is further developed. Assuming that $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^{N}$ represents the HSI composed of N sample vectors, where $\mathbf{x}_i = (x_i^1, \ldots, x_i^L)^T$ is the ith sample with L spectral bands, two types of anomaly detectors, RXD and RAD, can be briefly described as follows.

A. Reed–Xiaoli Detector

RXD based on the generalized likelihood ratio test under binary hypothesis is a constant false alarm rate (CFAR) detection detector. Let the detected pixel be $\mathbf{x}$, the RXD can be denoted as

$$\delta_{\text{RXD}}(\mathbf{x}) = (\mathbf{x} - \mu)^T \mathbf{K}^{-1} (\mathbf{x} - \mu) \tag{1}$$

where μ and $\mathbf{K}$ are the sample mean and the covariance matrix, respectively. Mathematically, RXD can be regarded as the inverse process of principal component analysis (PCA). PCA can find the main direction of information distribution according to the eigenvectors corresponding to the largest eigenvalues of $\mathbf{K}$. However, anomalies with low probability in the image will not be included in these main components, because their corresponding energy is very small. So, the smaller the eigenvalue, the larger the $\delta_{\text{RXD}}(\mathbf{x})$, which is the reason why RXD can be effectively applied to hyperspectral AD.

B. R-Anomaly Detector

Replacing $(\mathbf{x} - \mu)$ and $\mathbf{K}$ with the sample vector $\mathbf{x}$ and the correlation matrix $\mathbf{R}$, $\mathbf{R} = (1/N)\mathbf{X}\mathbf{X}^T$, RAD can be defined as

$$\delta_{\text{RAD}}(\mathbf{x}) = \mathbf{x}^T \mathbf{R}^{-1} \mathbf{x}. \tag{2}$$

Compared with RXD, which only deals with the second-order statistics by $\mathbf{K}^{-1}$, RAD can both handle the first-order and second-order statistics using $\mathbf{R}^{-1}$.

III. Residual-Driven Band Selection

This section establishes an unsupervised BS approach for AD named RDBS. Two keys of RDBS are the ABMF to provide priori knowledge and RDBP criteria to determine band priority, which are described below.

A. Anomaly and Background Modeling Framework

It is difficult to formulate a suitable criterion to evaluate the band priority without any prior information, which hinders the development of BS for AD. The density-based clustering methods can identify the outliers that do not belong to any cluster and are very suitable for detecting anomalies [14]. Besides, compared with k-means which usually achieves better clustering only on datasets with spherical distribution, such density-based clustering method can be applied to datasets with various complex shapes. DPC proposed in [13] can classify the pixels in the image based on density properties and has been successfully applied to HSI processing [15]. In detail, anomalies are usually isolated and can be viewed as a low-density pixel by DPC, while backgrounds are widely distributed in HSI with high density. Given a data $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^{N}$ to be clustered, the density ρ_i of any samples $\mathbf{x}_i$ can be obtained by DPC via (3)

$$\rho_i = \sum_{j \neq i} e^{-\left(\frac{d_{ij}}{d_c}\right)^2} \tag{3}$$

where $d_{ij} = ||\mathbf{x}_i - \mathbf{x}_j||_2^2$ is denoted as the Euclidean distance between samples $\mathbf{x}_i$ and $\mathbf{x}_j$. d_c represents the cut-off distance to determine the radius of the search region. Due to the definition of density, anomalies can be considered as low-density pixels, while other pixels are considered as higher density backgrounds. DPC can well separate the anomalies from backgrounds so that the constraint vectors formulated by the DPC-based anomaly set and background set can be viewed as the priori information to develop AD-driven BS algorithms. ABMF is implemented as follows.

Algorithm 1 ABMF

Input:	Hyperspectral image $\mathbf{X}$, the cut-off distance d_c, and the density threshold $\tilde{\rho}$.
Step 1:	Obtain the density ρ of all the samples in $\mathbf{X}$ by (3).
Step 2:	Divide all the pixels into the anomaly set $\mathbf{D}$ and the background set $\mathbf{B}$. That is, if $\rho_i < \tilde{\rho}$, $\mathbf{x}_i$ is regarded as anomaly and $\mathbf{x}_i \in \mathbf{D}$. Otherwise, suppose $\mathbf{x}_i$ belongs to background and $\mathbf{x}_i \in \mathbf{B}$.
Step 3:	Generate the anomaly-based constraint vector $\mathbf{c}$ and the background-based constraint vector $\tilde{\mathbf{c}}$ according to $\mathbf{D}$ and $\mathbf{B}$. If $\mathbf{x}_i \in \mathbf{D}$, then $c_i = 1$. If not, $c_i = 0$. On the contrary, If $\mathbf{x}_i \in \mathbf{B}$, then $\tilde{c}_i = 1$. Otherwise, $\tilde{c}_i = 0$.
Output:	Constraint vectors $\mathbf{c}$ and $\tilde{\mathbf{c}}$

B. Residual-Driven BP Criteria

This sub-section designs three BP criteria of RDBS, which make full use of the prior information provided by ABMF and can select more representative band subset for anomalies.

Due to the anomaly detector $\delta_{\text{RAD}}(\mathbf{x})$ which performs pixel by pixel, if we consider the RAD detection on the whole, then $\delta_{\text{RAD}}(\mathbf{x})$ can be rewritten as $\delta_{\text{RAD}}(\mathbf{X})$ given by

$$\delta_{\text{RAD}}(\mathbf{X}) = \mathbf{X}^T \mathbf{R}^{-1} \mathbf{X}. \tag{4}$$

Then, the ABMF developed in Section III-A is adopted to generate the anomaly and background constraint vectors. Generally speaking, the anomaly constraint vector $\mathbf{c} = (1, 1, 0, 0, \ldots, 1, \ldots, 0)^T$ is defined according to the DPC clustering map where the corresponding n signal index positions in DPC clustering map are defined as 1, and the others are defined as 0. Therefore, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$ actually becomes $\mathbf{X} = [\mathbf{s}_1, \mathbf{s}_2, \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{s}_n, \ldots, \mathbf{b}_{N-n}]$ with n target signal vectors $\mathbf{s}$ and $(N-n)$ background signal vectors $\mathbf{b}$. Finally, with the constraint $\mathbf{c}$, the total energy of $\mathbf{X}$ can be expressed as

$$E(\mathbf{X}) = \mathbf{c}^T \left(\mathbf{X}^T \mathbf{R}^{-1} \mathbf{X}\right) \mathbf{c} = (1/N) \mathbf{c}^T \left[\mathbf{X}^T \left(\mathbf{X}\mathbf{X}^T\right)^{-1} \mathbf{X}\right] \mathbf{c}. \tag{5}$$

Let $\mathbf{U} = \mathbf{X}^T$, the orthogonal subspace of $\mathbf{U}$ is $\mathbf{P}_{\mathbf{U}}^{\perp} = \mathbf{I} - \mathbf{U}\mathbf{U}^{\#}$ with $\mathbf{U}^{\#} = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T$. Then, (5) can be further

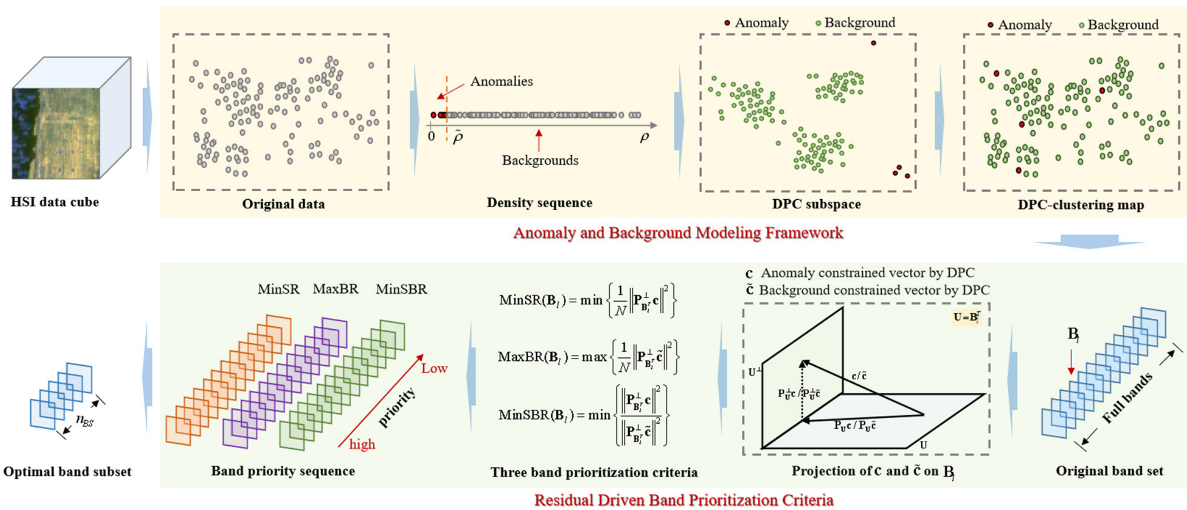


Fig. 1. Implementation of RDBS.

expressed as

$$E(\mathbf{X}) = (1/N)\mathbf{c}^T\left[\mathbf{U}\left(\mathbf{U}^T\mathbf{U}\right)^{-1}\mathbf{U}^T\right]\mathbf{c} = (1/N)\mathbf{c}^T\left(\mathbf{I}-\mathbf{P}_{\mathbf{U}}^{\perp}\right)\mathbf{c}. \tag{6}$$

Due to $\mathbf{P}_{\mathbf{U}}^{\perp} = (\mathbf{P}_{\mathbf{U}}^{\perp})^T(\mathbf{P}_{\mathbf{U}}^{\perp})$, (6) can also be written as

$$E(\mathbf{X}) = (1/N)\mathbf{c}^T\left(\mathbf{I}-\mathbf{P}_{\mathbf{U}}^{\perp}\right)\mathbf{c} = (1/N)\left\{||\mathbf{c}||^2 - \left|\left|\mathbf{P}_{\mathbf{U}}^{\perp}\mathbf{c}\right|\right|^2\right\}. \tag{7}$$

Regarding $\mathbf{c}$ as a binary coded vector of anomalies and backgrounds, $\mathbf{P}_{\mathbf{U}}^{\perp}\mathbf{c}$ can be denoted as the projection of $\mathbf{c}$ to the orthogonal subspace $\mathbf{U}^{\perp}$, that is, the residual vector of $\mathbf{c}$ on $\mathbf{U}$ (i.e., $\mathbf{X}^T$). So $||\mathbf{P}_{\mathbf{U}}^{\perp}\mathbf{c}||^2$ can be defined as the residual value of $\mathbf{c}$ after projection to $\mathbf{U}^{\perp}$. It should be noted that the smaller the $||\mathbf{P}_{\mathbf{U}}^{\perp}\mathbf{c}||^2$, the higher the similarity between $\mathbf{c}$ and $\mathbf{U}$ (i.e., $\mathbf{X}^T$), i.e., the stronger the representation ability of $\mathbf{X}$ to $\mathbf{c}$. So, the larger the $E(\mathbf{X})$, the better the characterization ability of $\mathbf{X}$ to $\mathbf{c}$.

For arbitrary band $\mathbf{B}_l$ instead of $\mathbf{X}$, a new BP criterion called minimum signal residual (MinSR) can be defined as

$$\mathrm{MinSR}(\mathbf{B}_l) = \min\left\{(1/N)\left|\left|\mathbf{P}_{\mathbf{B}_l^T}^{\perp}\mathbf{c}\right|\right|^2\right\}. \tag{8}$$

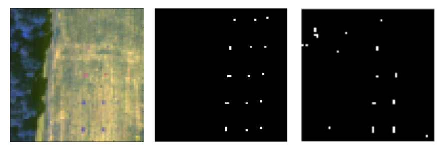
Similarly, if the values of 1 and 0 in $\mathbf{c}$ are exchanged, the background constraint vector can be defined as $\tilde{\mathbf{c}} = (0, 0, 1, 1, \ldots, 0, \ldots, 1)^T$, and another BP criterion named the maximum background residual (MaxBR) can be developed as

$$\mathrm{MaxBR}(\mathbf{B}_l) = \max\left\{(1/N)\left|\left|\mathbf{P}_{\mathbf{B}_l^T}^{\perp}\tilde{\mathbf{c}}\right|\right|^2\right\}. \tag{9}$$

MaxBR mainly considers the similarity between $\mathbf{B}_l$ and the background which accounts for the main proportion of HSI. It considers that the larger the $\mathrm{MaxBR}(\mathbf{B}_l)$, the weaker the response of $\mathbf{B}_l$ to the background, that is, the band can better express the abnormal information.

Finally, integrating MinSR and MaxBR, the third BP criterion, namely, minimum signal to background residual ratio (MinSBR), is proposed, which takes the information of anomalies and backgrounds into comprehensive consideration

$$\mathrm{MinSBR}(\mathbf{B}_l) = \min\left\{\left|\left|\mathbf{P}_{\mathbf{B}_l^T}^{\perp}\mathbf{c}\right|\right|^2 \Big/ \left|\left|\mathbf{P}_{\mathbf{B}_l^T}^{\perp}\tilde{\mathbf{c}}\right|\right|^2\right\}. \tag{10}$$



(a) (b) (c)

Fig. 2. HYDICE dataset. (a) False-color image. (b) Reference map. (c) DPC clustering map with $d_c = 6\%$.

(a) (b) (c)

Fig. 3. Urban dataset. (a) False-color image. (b) Reference map. (c) DPC clustering map with $d_c = 4\%$.

If $\mathbf{B}_l$ has good representation for anomalies, it should have strong response to the anomalies and weak expression to the backgrounds. Finally, RDBS is illustrated in Fig. 1.

IV. Experiments

In our experiments, two widely used HSIs are available to verify the effectiveness of RDBS for hyperspectral AD. Also, the experimental results are compared and analyzed in detail.

A. Datasets

1) Hyperspectral Digital Imagery Collection Experiment (HYDICE) Scene: It was collected by the airborne HYDICE sensor in 1995. Its size is of 64 × 64 with 169 spectral bands, and the corresponding false-color image is shown in Fig. 2(a). Also, there are 15 panels containing 19 pixels to be detected as shown in Fig. 2(b). The panel pixels in each row are the same materials.

2) Urban Scene: The second is an urban scene which was also acquired by HYDICE sensor. After removing the water absorption and noise bands, the image cube with size of 80 × 100 × 162 is finally used in the experiment. Fig. 3(a) and (b) shows the false color map and its reference

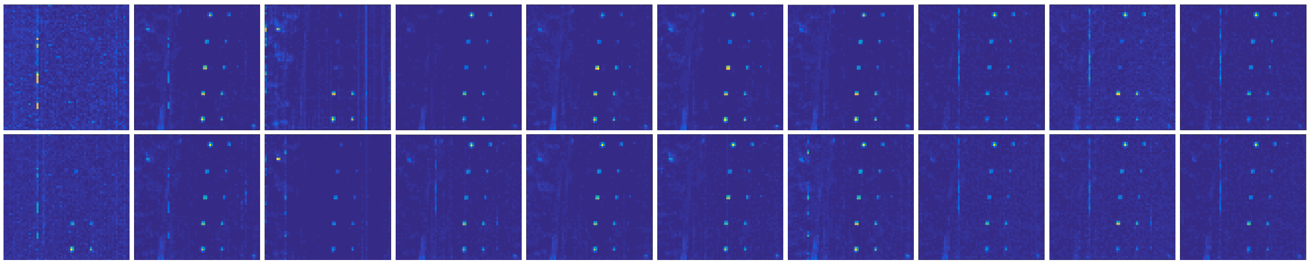


Fig. 4. RAD detection maps with band subset selected by ten types of BS methods for the HYDICE dataset. It should be noted that n_{BS} is 13 and 26 in the first and second rows, respectively. (a) SQ. (b) UBS. (c) MVPCA. (d) E-FDPC. (e) ONR. (f) OCF. (g) FNGBS. (h) MinSR. (i) MaxBR. (j) MinSBR.

map, respectively. More specifically, there are 19 abnormal pixels different from the background.

The experiment selects the anomaly detector RAD for AD and compares ten different BS algorithms including sequential BS (SQ), uniform BS (UBS), MVPCA [16], E-FDPC [17], ONR [18], OCF [7], FNGBS [19], and three proposed BS algorithms MinSR, MaxBR, and MinSBR. Specifically, SQ selects the desired n_{BS} bands sequentially. UBS selects the band subset from the total bands at equal intervals. Besides, the DPC clustering maps obtained by ABMF of two images are shown in Figs. 2(c) and 3(c). Here, arrange the distance values between all the pixels in ascending order. By experimental analysis, we take the value at 6% and 4% of the sequence as d_c for these two datasets. Since the anomaly is usually isolated with very low content, so the density threshold $\tilde{\rho}$ is assumed as 1.

For the HYDICE and Urban datasets, n_{BS} estimated by virtual dimensionality (VD) [20] with false alarm rate of 10^{-5} is 13 and 11, respectively. However, it turns out that the VD-determined n_{BS} is generally under-estimated, and thus n_{VD} and $n_{2\mathrm{VD}}$ are selected as n_{BS}. Besides, the RAD detection map and the 3-D receiver operating characteristic (3D ROC) curve [21] are used to evaluate the performance of BS. The 3D ROC curve is a function of a threshold parameter τ, detection probability ($\mathrm{P_D}$), and false alarm probability ($\mathrm{P_F}$). It can be divided into three 2-D ROC curves of ($\mathrm{P_D}, \mathrm{P_F}$), ($\mathrm{P_D}$, τ), and ($\mathrm{P_F}$, τ) and further obtain the corresponding area under the curve (AUC) to measure the overall detection performance, target detection capability, and background suppression ability of the detector. The higher the AUC($\mathrm{P_D}$,$\mathrm{P_F}$) and AUC($\mathrm{P_D}$, τ), the better the detection performance. The smaller the AUC($\mathrm{P_F}$, τ), the better the background suppression.

B. Results and Analysis

First, for the HYDICE dataset, Fig. 4 shows the detection maps on the band subset composed of n_{VD} and $n_{2\mathrm{VD}}$ bands selected by ten different BS algorithms. Notably, the results of FNGBS are the best results for five runs at $n_{\mathrm{BS}} = 13$. As shown in the first row of Fig. 4, RAD hardly detects any anomaly on the band subset selected by SQ. The detection result on the band subset selected by MVPCA is also unsatisfactory, with only the anomalies in the fourth and fifth rows showing strong responses. In contrast, the detection results on the band subsets selected by the rest are much better. Not only are the anomalies detected prominently but also the background is also well-suppressed. However, the UBS and ONR still have a stronger response to the background. Finally, when n_{BS} increases to $n_{2\mathrm{VD}}$, the detection results on different band subsets are improved to some extent.

TABLE I

AUC Values for the HYDICE Dataset

n_{BS}	$n_{\mathrm{VD}} = 13$			$n_{2\mathrm{VD}} = 26$		
AUC	($\mathrm{P_D}$, $\mathrm{P_F}$)	($\mathrm{P_D}$, τ)	($\mathrm{P_F}$, τ)	($\mathrm{P_D}$, $\mathrm{P_F}$)	($\mathrm{P_D}$, τ)	($\mathrm{P_F}$, τ)
SQ	0.4822	0.0349	0.0411	0.7749	0.2008	0.0306
UBS	0.9868	0.3952	0.0148	0.9853	0.3356	0.0160
MVPCA	0.9228	0.2694	0.0223	0.9701	0.1455	0.0119
E-FDPC	0.9873	0.2616	0.0090	**0.9931**	0.3058	0.0152
ONR	0.9684	0.3921	0.0144	0.9839	0.3595	0.0150
OCF	0.9829	0.4339	0.0155	**0.9915**	0.3522	0.0163
FNGBS	0.9861	0.3362	0.0129	0.9893	0.3869	0.0180
MinSR	**0.9898**	0.2113	0.0163	0.9898	0.2648	0.0195
MaxBR	0.9813	0.3033	0.0261	**0.9942**	0.2927	0.0226
MinSBR	0.9895	0.2346	0.0164	**0.9908**	0.2556	0.0186
Full bands	0.9900	0.3584	0.0435			

TABLE II

AUC Values for the Urban Dataset

n_{BS}	$n_{\mathrm{VD}} = 11$			$n_{2\mathrm{VD}} = 22$		
AUC	($\mathrm{P_D}$, $\mathrm{P_F}$)	($\mathrm{P_D}$, τ)	($\mathrm{P_F}$, τ)	($\mathrm{P_D}$, $\mathrm{P_F}$)	($\mathrm{P_D}$, τ)	($\mathrm{P_F}$, τ)
SQ	0.9768	0.5271	0.0940	0.9622	0.5930	0.1161
UBS	**0.9883**	0.1776	0.0128	**0.9912**	0.1522	0.0128
MVPCA	0.8824	0.0406	0.0139	0.9204	0.0490	0.0157
E-FDPC	**0.9928**	0.3841	0.0280	**0.9873**	0.1726	0.0152
ONR	**0.9916**	0.1893	0.0109	**0.9920**	0.1628	0.0146
OCF	**0.9896**	0.2472	0.0146	**0.9908**	0.1249	0.0120
FNGBS	**0.9916**	0.3712	0.0255	**0.9882**	0.1773	0.0158
MinSR	**0.9848**	0.1769	0.0235	**0.9887**	0.1231	0.0193
MaxBR	**0.9942**	0.1567	0.0137	**0.9948**	0.1752	0.0193
MinSBR	**0.9878**	0.0940	0.0132	**0.9875**	0.1354	0.0184
Full bands	0.9822	0.2041	0.0342			

Tables I and II record the AUC values of the detection results of seven comparison algorithms and the proposed algorithms on the two datasets, respectively. Of course, the detection on full-band set is attached to the last row of the table. Since FNGBS differs each time, the average of the five runs is recorded. In all the methods, the AUC($\mathrm{P_D}$,$\mathrm{P_F}$) value will be bold if it is higher than the AUC($\mathrm{P_D}$,$\mathrm{P_F}$) value of the full-band set, while the highest AUC($\mathrm{P_D}$,$\mathrm{P_F}$) will be marked in red. For the HYDICE dataset, when $n_{\mathrm{BS}} = 13$, SQ and MinSR

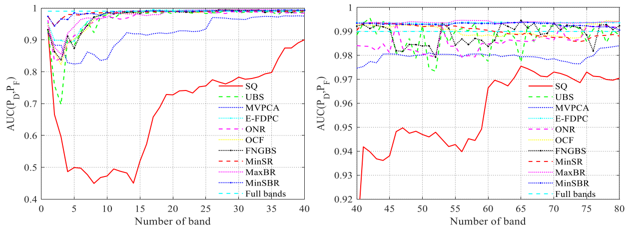


Fig. 5. AUC values of the progressive band subset for HYDICE dataset.

are the worst and the best. AUC(P_D, P_F) of MinSBR, E-FDPC, UBS, FNGBS, OCF, and MaxBR is slightly lower than that of MinSR. Their results are comparable to the full-band set. Moreover, ONR ranks third from the bottom. The AUC of MVPCA is only higher than that of SQ. Similarly, when $n_{BS} = 26$, MaxBR, E-FDPC, OCF, and MinSBR all obtain more accurate results. MinSR performs slightly worse than the full-band set, but still leads the rest of the BS algorithms, which further proves the superiority of RDBS. The above indicates that the amount of information contained in the band subset with n_{2VD} bands can more adequately express and identify the anomalies.

Likewise, for the Urban dataset, it can be observed that MaxBR obtains the highest AUC(P_D,P_F) at both 11 bands and 22 bands. MinSR and MinSBR have a moderate performance, but the band subsets they selected are still helpful for AD compared with the full-band set. The latest algorithms also obtained higher AUC values, but worse than MaxBR. Moreover, MVPCA performs the worst. In conclusion, MaxBR based on a background prior with a large amount of guidance information obtains a more stable and superior performance.

Finally, to demonstrate the high stability of the proposed algorithms, Fig. 5 shows the AUC(P_D,P_F) curves with increasing n_{BS} of HYDICE. To observe more clearly, two sets of curves are shown for n_{BS} from 1 to 40 and 41 to 80. Obviously, SQ and MVPCA perform the worse, especially poor and unstable with low n_{BS}. Besides, the AUC values of UBS, FNGBS, and MaxBR are similar in different n_{BS}. However, the curve of MaxBR is flatter and more stable than the wobbly curve of UBS and FNGBS. Although E-FDPC, ONR, and OCF also obtain fairly high values, it is still MaxBR occupies the highest value under most n_{BS}. Finally, although MinSR and MinSBR are usually under-performing when n_{BS} is small, with the increase of n_{BS} to 20, their results are generally consistent with MaxBR, and better than the full-band detection results.

The above illustrates the necessity of appropriate BP criteria and the advantages of background-based guidance for AD. The proposed methods are effective, which fully consider the band's ability of anomaly representation and background suppression, thereby can obtain a more stable AD detection result.

V. Conclusion

This letter proposes an unsupervised BS algorithm, RDBS, for AD, which somewhat alleviates the dilemma of slow development of BS for AD due to the lack of *a priori* information. The experiments show that the proposed BPs not only effectively reduce the band redundancy but also obtain band subsets with strong characterization of anomalies. In particular, MaxBR, built on the basis of background statistical information, achieves excellent detection results and the highest stability.

Acknowledgment

The authors would like to thank Prof. Chein-I Chang for providing the dataset of HYDICE panel scene.

References

[1] 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.

[2] 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.

[3] C.-I. Chang and S.-S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 6, pp. 1314–1325, Jun. 2002.

[4] J. Zhou, C. Kwan, B. Ayhan, and M. T. 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.

[5] Y. Tan, L. Lu, L. Bruzzone, R. Guan, Z. Chang, and C. Yang, "Hyperspectral band selection for lithologic discrimination and geological mapping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 471–486, 2020.

[6] X. Sun, H. Zhang, F. Xu, Y. Zhu, and X. Fu, "Constrained-target band selection with subspace partition for hyperspectral target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 9147–9161, 2021.

[7] Q. Wang, F. Zhang, and X. Li, "Optimal clustering framework for hyperspectral band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5910–5922, Oct. 2018.

[8] Y. He, D. Liu, and S. Yi, "Recursive SAM-based band selection for hyperspectral anomaly detection," in *Proc. SPIE*, Dalian, China, 2010, p. 76582K.

[9] M. Huber-Lerner, O. Hadar, S. R. Rotman, and R. Huber-Shalem, "Hyperspectral band selection for anomaly detection: The role of data gaussianity," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 732–743, Feb. 2016.

[10] 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.

[11] S. Chang, B. Du, and L. Zhang, "A subspace selection-based discriminative forest method for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 4033–4046, Jun. 2020.

[12] F. Andika, M. Rizkinia, and M. Okuda, "A hyperspectral anomaly detection algorithm based on morphological profile and attribute filter with band selection and automatic determination of maximum area," *Remote Sens.*, vol. 12, no. 20, p. 3387, Oct. 2020.

[13] A. Rodriguez and A. Laio, "Machine learning. Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[14] B. Tu, N. Li, Z. Liao, X. Ou, and G. Zhang, "Hyperspectral anomaly detection via spatial density background purification," *Remote Sens.*, vol. 11, p. 2618, Nov. 2019.

[15] B. Tu, X. Yang, N. Li, C. Zhou, and D. He, "Hyperspectral anomaly detection via density peak clustering," *Pattern Recognit. Lett.*, vol. 129, pp. 144–149, 2020.

[16] C.-I. Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 6, pp. 2631–2641, Nov. 1999.

[17] S. Jia, G. Tang, J. Zhu, and Q. Li, "A novel ranking-based clustering approach for hyperspectral band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 88–102, Jan. 2016.

[18] 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, doi: 10.1109/TGRS.2020.2987955.

[19] Q. Wang, Q. Li, and X. Li, "A fast neighborhood grouping method for hyperspectral band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 5028–5039, Jun. 2021.

[20] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.

[21] C.-I. Chang, "An effective evaluation tool for hyperspectral target detection: 3D receiver operating characteristic curve analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 59, no. 6, pp. 5131–5153, Jun. 2021.