

Hyperspectral Anomaly Detection Method Based on Adaptive Background Extraction

MIN LI^{1,2}, PUHUANG LI¹, AND HAIYAN XU¹

¹College of Internet of Things Engineering, Hohai University, Changzhou 213022, China

²Division of Translational Imaging, Department of Psychiatry, Columbia University, New York, NY 10032, USA

Corresponding author: Haiyan Xu (hyxu@hhu.edu.cn)

This work was supported in part by the Nature Funding of China under Grant 61701169 and Grant 61671202, and in part by the National Key Research and Development Plan under Grant 2016YFC0401606.


ABSTRACT Anomaly detection based on clustering is a classic method that supplies a simplified manner to describe a cluttered background. However, traditional clustering methods need to know the number of clusters in advance and attempt to classify all the background pixels at one time. In addition, compared with large background clusters, small clusters are hard to discriminate due to their small populations. In this paper, an anomaly detection method based on adaptive background extraction is proposed. We apply an unsupervised clustering method to determine the cluster centers according to only the similarity of the spectral signature. To reduce the influence of the population, we propose to extract background clusters iteratively. Every iteration, we only cluster the larger clusters and extract them from the data-set. In the next iteration, the remaining pixels are clustered again. Without interference from the larger clusters, the centers of smaller clusters will appear obviously. The clustering process stop when the number of remaining pixels nears the appearance probability of anomaly (generally approximately 10%~20%). Then, only anomalies and few background pixels remain to test. Finally, every extracted background cluster, as a viewer, is applied to measure the anomaly salience of the test pixels. In addition, a weighted summation is proposed to fuse the different salience values from different viewers. Simulation experiments on two sets of real data are presented to demonstrate the superiority of the proposed method.

INDEX TERMS Hyperspectral, target detection, anomaly detection, clustering, adaptively.

I. INTRODUCTION

Hyperspectral technology provides almost continuous spectral information of the ground surface and makes it possible to detect and extract features and targets on the ground [1], [2]. An anomaly is a kind of target that has obviously special spectral signature compared with the surrounding background. Without atmospheric correction and radiation calibration, anomaly detection technology discriminates the interested target or area based on its special spectral characteristics [3]. As an unsupervised target detection technology, anomaly detectors have been widely applied in environmental protection, precision agriculture, geological exploration, national defense security, and search and rescue in public accident [4]–[6].

Background model-based anomaly detection has been the most popular method for about thirty years, and it is still

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy .

a major and effective solution for anomaly detection today. The basic idea is that an anomaly can be separated if the background or non-anomaly can be described [7]. For this, accurate background description has become the key point in these methods. The Gaussian model and multiple Gaussian mixture model are the typical background description models because of their ease of calculation.

In 1990, Reed L. proposed local model-based anomaly detector RXAD, which describes the background by a dual window [8]. RXAD constructs a local background model, based on the statistic feature of an external window. Then, it measures the anomaly saliency of the test pixel in an internal window. Local background modeling provides a simplified method for background description. However, it is easy to trap into a local optimum. In addition, the size of the sliding dual window is hard to choose, while too large a window will induce contamination from anomalies, and too small a window cannot ensure the accuracy of the background description. For this, different local background models were

proposed to decrease the influence of the windows size, such as the single local area [9], [10] and multi-local area (MSAD) [11].

Compared with the local model, the global model is aimed to obtain a homogenous description of the background [12], [13]. The global model constructs a uniform model which is based the whole background features. However, the diversity of ground objects in the background increases the difficulty of the global background description. Many solutions have been proposed to simplify the complexity of the global model, such as the multivariate normal-based anomaly detector (MVN) [14], the cluster-based anomaly detector (CBAD) [15], [16] and subspace model-based anomaly detector (SSM) [11]. The MVN hypothesizes that the background consists of different materials, and the spectral character of every material obeys the Gaussian distribution. The cluster-based background description is a compromise method, which can effectively avoid the contamination problem and obtain a global background model. CBAD clusters the background into multiple clusters, and the spectral signature of every cluster is simple. In addition, the whole background is characterized by mixture Gaussian model. However, CBAD needs to know the number of clusters in advance. In addition it is hard to know how many kinds of objects exist in a cluttered background. Another kind of model simplification method depends on spatial transformation, such as nonlinear kernel function mapping. For example, Kernel-RXAD simplifies background features by nonlinear kernel transform [17], [18]. Then, the cluttered background can be described by a Gaussian distribution in the high-dimensional feature space. However, the appropriate scale parameter is a bottle neck problem in these methods. A small scale will induce discontinuity, and large scale also has the problem of over-smoothing. The random selection anomaly detector (RSAD) [19] constructs multiple background model, and calculates the anomaly saliency of test pixels by different background models. Finally, different detection results are fused to obtain the final result. A cluster kernel RX algorithm (CKRX) [20], which induces cluster process to compute kernel distance and has fewer computations than traditional Kernel RXAD, is proposed as a generalization of kernel RX(KRX).

With an accurate background model, anomalies are easy to separate due to their obvious spectral difference. However, pure background pixels are difficult to extract due to the limited spatial resolution and complex spectrum of ground objects. In addition, the statistical model cannot address the cluttered background. For this, there are some remedy methods to improve the background model-based detector. A background purification method is proposed to obtain a purified background set, which improves the performance of LRXD, GRXD and KRXD [21]. It is also a good method for estimating a weight matrix and using the detected result of any AD method. A large weight value means a high probability of being an anomaly [22].

To avoid the extraction of background pixels, some methods utilize characteristics of low-rank and a sparse background to detect anomalies. With the assumption of similarity of spectral features and spatial features of the background, many of the latest anomaly detection (AD) attempt to learn a dictionary and detected anomaly by reconstruction error [23]–[27]. Other methods apply slowly varying signal analysis [28], matrix decomposition [29], [30] or optimal filters [31], [32] to detected targets.

Without the spectrum signature of the anomaly target, it is hard for the algorithm to determine the background and anomaly accurately. Anomaly detectors based on a background model need pure background pixels to construct a background statistic model. Many works have been performed to reduce the purity requirements of background pixels or enhance the robustness of the model to anomaly contamination. Recently, clustering and iterative calculation have become increasingly popular in anomaly detectors. Unsupervised clustering, as a type of self-organized classification method, only extracts most background pixels which supply a relatively accurate measurement for anomaly saliency [22] and spatial-spectral similarity [25]. It also supplies a relatively accurate data-set to construct a background dictionary [24], [26], [27]. Iterative computations is another effective method for solving this kind of problem. Because of the lack of a specific quantity of anomaly appearance probability, anomaly detection (AD) is according to the relative rareness of the spatial distribution and difference in spectral features. That means we do not know the correct answer, but we obtain a closer possible answer by iterative calculation.

In this paper, we propose an anomaly detection method that extracts background pixels iteratively by an adaptive clustering process. It has three contributions as follow:

- (1) We apply an adaptive clustering method to cluster background pixels. Cluster centers can be identified quickly and adaptively according to only the similarity of the spectral signature.
- (2) To reduce the influence of larger clusters, an iterative clustering process is proposed to extract background clusters step by step. After eliminating larger clusters, centers of small clusters will appear in the density-distance chart.
- (3) Considering the confidence level of different clusters, a weight fusion is proposed to fuse the detected results from different viewers. The weight of every cluster is calculated by a weight function.

This paper is arranged as follows. Section 2 introduces the main idea and processing steps of our method. Section 3 evaluates the effectiveness of our method by experiment and analysis. Section 4 discusses the selection of parameters. The final section conclude.

II. METHOD

The proposed adaptive background extraction-based anomaly detection method contains two parts: iterative clustering and multi-viewer anomaly detection.

First, an adaptive clustering method is applied to extract background clusters iteratively. An innovative clustering method, as an unsupervised clustering method, can adaptively and quickly determine the cluster centers [33]. Referring to the spectrum similarity of pixels, we construct a density-distance chart of hyperspectral data. In the chart, cluster centers are far away from the majority. Considering the influence of larger cluster on small clusters, we only cluster and extract 2~4 larger clusters in one iteration. Then, we draw the density-distance chart of the remaining pixels and clusters. The centers of smaller clusters will appear without interference from larger background clusters. Most background pixels are extracted by clustering until the number of the remaining pixels near the appearance probability of anomaly (generally approximately 10%).

Second, we use all the extracted background clusters as viewers to measure the anomaly salience of the remaining pixels. For every pixel, there are different anomaly salience values measured by different viewers. In addition, the final result is fused by a weighted summation.

A. ITERATIVE CLUSTERING

Just like most AD methods, spectral similarity is the basic criterion for discriminating target and background. In this paper, we apply an innovative clustering method for extracting background pixels according to the similarity of the spectrum signature. It constructs two measure parameters: density and distance. Density is the local population around the point. Distance is the other parameter, which is decided by density. This clustering method assigns the cluster centers as the points which have higher population and are far away from each other. Based on the local density and relative distance, cluster centers can be found by competition between different clusters.

Hyperspectral data \mathbf{X} has N pixels $\mathbf{X} = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_N\}$, $\mathbf{X} \in R^{N \times K}$ and every pixel \vec{x}_i has K bands. The local density of pixel \vec{x}_i is defined as the number of pixels \vec{x}_i with a similar spectrum signature in a local neighborhood.

$$\rho_i = \sum_{j=1}^N \chi(d_{ij} - T_d) \tag{1}$$

where $d_{i,j} = \|\vec{x}_i - \vec{x}_j\|^2$. It measures the similarity of the spectrum signature between the i th pixel \vec{x}_i and the j th pixel \vec{x}_j . $\chi(\cdot)$ is a step function $\chi(d_{ij} - T_d) = 1$, if $d_{ij} < T_d$, otherwise $\chi(d_{ij} - T_d) = 0$. T_d is the threshold. The area of neighborhood is confined by cutoff radius T_d .

Pixels with higher local density ρ_i have a higher possibility to be a cluster center. However, many pixels may have higher density in the same cluster. There is other knowledge that cluster centers should be far away from each other. Then, relative distance δ_i of pixel \vec{x}_i is defined according to the value of ρ_i and ρ_j .

$$\delta_i = \begin{cases} \min_{j:\rho_j > \rho_i} (d_{ij}), & \rho_i < \rho_j \\ \max_{j=1,2,\dots,N} (d_{ij}), & \rho_i > \rho_j \end{cases} \tag{2}$$

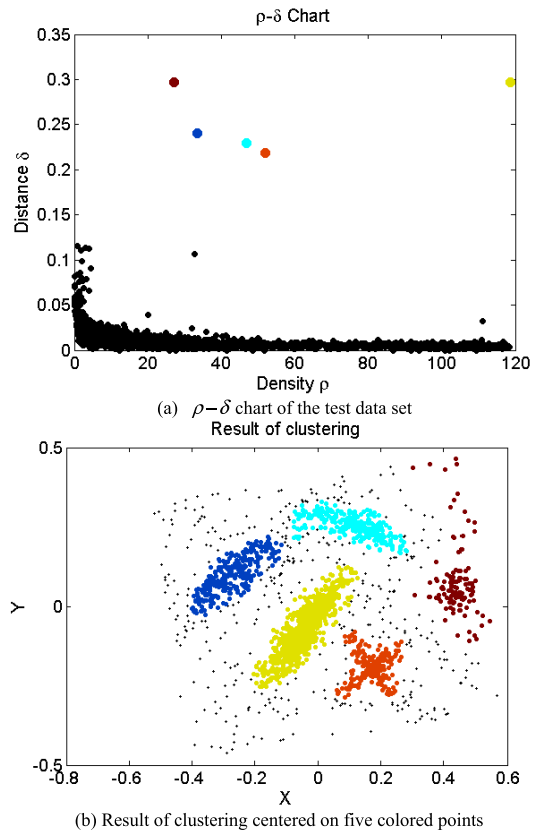


FIGURE 1. Clustering of a random test data-set.

where δ_i is the maximum distance between \vec{x}_i and $\vec{x}_j (i \neq j)$, when $\rho_i > \rho_j$. That means pixel \vec{x}_i is a cluster center, and relative distance δ_i confines the maximum radius of this cluster. δ_i is the minimum distance between \vec{x}_i and \vec{x}_j , when $\rho_i < \rho_j$. That means \vec{x}_i is a follower of a cluster center at pixel \vec{x}_j . Relative distance δ_i is the distance from pixel \vec{x}_j to its center. A density-distance chart ($\rho - \delta$ chart) is drawn to determine cluster centers adaptively. In $\rho - \delta$ chart, cluster centers stand out because the centers have larger density ρ_i and a larger distance δ_i simultaneously. Then, the centers of the clusters are chosen artificially.

Take a random data-set as an example and draw its $\rho - \delta$ chart to show the clustering process. In Fig. 1(a), five colored points, which have both higher density and higher distance, are far from the majority. Considering that density implies the size of the cluster, the cluster centered at the yellow point has the highest population, and the cluster centered at the brown points has a lower population. In Fig. 1(a), there are some black points that depart from the majority due to their larger δ and ρ . However, they are not remarkable compared with those colored points. Therefore, in the same chart, the centers of the small clusters are inhibited by those larger clusters. Taking these colored points as cluster centers, we obtain five clusters (showed in Fig. 1(b)) which are colored the same as their center. In addition, other black points do not belong to any cluster.

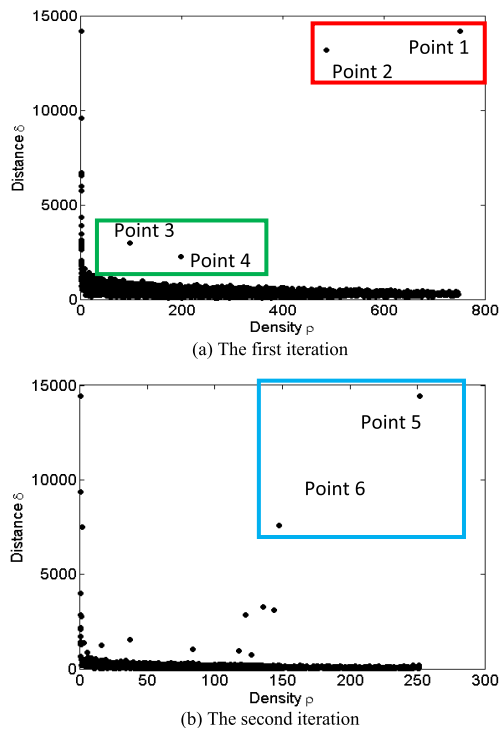


FIGURE 2. $\rho - \delta$ chart of every step of the iteration.

Considering the complexity of hyperspectral data, the size of the background clusters is different. To extract as many background clusters as possible, especially small clusters, we propose to cluster the background iteratively. In each iteration, we cluster the larger clusters and extracted it from the data set, and then cluster the remaining data. Without the influence of the large clusters, the centers of the smaller clusters will emerge in the new $\rho - \delta$ chart. The iterative process stops when the number of remaining pixels nears a lower limit. Fig. 2 shows the iterative clustering process of a random data set.

In Fig. 2 (a), there are two pixels far from the majority, which can be considered as cluster centers. The clusters centered at “point 1” and “point 2” have larger sizes. The other two points labeled as “point 3” and “point 4” are not remarkable compared with “point 1” and “point 2”. We extract two clusters centered at “point 5” and “point 6” from the dataset and redraw the $\rho - \delta$ chart for the remaining data (Fig. 2(b)). Many points are released from the majority without the inhibition of the largest cluster. Another “point 1” and “point 2” are removed from the majority. This time we extract two background clusters centered at these two points. Clustering processing continues iteratively until the number of remaining pixels is lower than the threshold. Considering the appearance probability of anomalies is 10%, we set the threshold to 20% to avoid loss.

B. MULTI-VIEWER ANOMALY DETECTION

After iterative clustering, most background data have been extracted. To obtain a comprehensive and objective valuation of anomaly saliency, each background cluster is viewed as

an independent observer to measure anomaly saliency of the remaining data. For example, the Mahalanobis distance of the i th test pixel \vec{x}_i is defined based on the background cluster C_1 , where $C_1 = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_i, \dots, \vec{x}_{N_1}\}$ has N_1 pixels.

$$D_1(\vec{x}_i) = \sqrt{(\vec{x}_i - \mu) \Sigma^{-1} (\vec{x}_i - \mu)^T} \vec{x}_i \in \mathbf{X} \quad (3)$$

where Σ is the covariance matrix of background cluster C_1 ; μ is the mean value of C_1 . Mahalanobis distance $D_1(\vec{x}_i)$ measures the similarity between pixel \vec{x}_i and background cluster C_1 . A larger $D_1(\vec{x}_i)$ indicates a larger difference between and background cluster C_1 and equals higher anomaly saliency. If the background \mathbf{U} has three clusters $\mathbf{U} = \{C_1, C_2, C_3\}$, every pixel \vec{x}_i will have three anomaly saliency value with different background clusters.

$$\begin{aligned} D_1(\vec{x}_i), & \left\{ \vec{x}_i, C_1 \right\} \Big| \vec{x}_i \in \mathbf{X} \\ D_2(\vec{x}_i), & \left\{ \vec{x}_i, C_2 \right\} \Big| \vec{x}_i \in \mathbf{X} \\ D_3(\vec{x}_i), & \left\{ \vec{x}_i, C_3 \right\} \Big| \vec{x}_i \in \mathbf{X} \end{aligned} \quad (4)$$

To obtain the final anomaly saliency of pixel \vec{x}_i , we need to fuse all the saliency values $D_1(\vec{x}_i)$.

Considering the larger background cluster occupied most of the background area, anomaly saliency obtained by this cluster should have higher credibility. In contrast, small background clusters only occupied a small part of the background. And we cannot exclude the probability that small clusters are anomaly or target clusters. Therefore, the anomaly saliency measured by small clusters will be adopted cautiously. So, a weighted summation is proposed to fuse the results from different viewers. The weighted summation fusion process is denoted as follow.

$$\mathbf{M}(\vec{x}_i) = \sum_{c=1}^C \gamma_c (1 - D_c(\vec{x}_i)) \{C_c \in \mathbf{U} | \mathbf{U} = C_1, C_2, \dots, C_C\} \quad (5)$$

where γ_c is the weight of cluster C_c . The value of γ_c is calculated by the normal weight function according to the size of the background cluster C_c . A larger cluster has a larger weight, a smaller cluster has a smaller weight.

$$\gamma_c = \begin{cases} 0 & N_c < 0 \\ 1 - \exp\left(-\left(\frac{N_c}{N}\right)^2\right) & 0 \leq N_c \leq N \end{cases} \quad (6)$$

where $C_c \in \{C_1, C_2, \dots, C_C\}$, N_c is the pixel number of cluster C_c , N is the total number of pixels in data-set. The proposed method is showing as follow.

III. EXPERIMENT

A. DATA SET DESCRIPTION

Two kinds of hyperspectral data are applied to evaluate the performance of our proposed method.

Adaptive Background Extraction Algorithm

Input: hyperspectral data matrix $\mathbf{X}_{N \times K}$, anomalies percent $p = 20\%$.

Output: final detection result of every pixel $M(\vec{x}_i)$.

- Step 1: Normalized the hyperspectral data $\mathbf{X}_{N \times K}$;
 - Step 2: Calculate density ρ_i and distance δ_i of every pixel \vec{x}_i in data set \mathbf{X} ;
 - Step 2.1: Identify centers of larger clusters by the $\rho - \delta$ chart;
 - Step 2.2: Cluster the larger clusters and extract them from the data set.
 - Step 2.3: Redraw the $\rho - \delta$ chart for the remaining data;
 - Step 2.4: if the number of remaining pixels is more than 20% of the total number, then go to Step 2.1.
 - Step 3: Calculate the anomaly salience $D_c(\vec{x}_i)$ of the pixel \vec{x}_i with the background cluster \mathbf{C}_c $c = 1, 2, \dots, C$
 - Step 4: Obtain weight γ_c of cluster \mathbf{C}_c by the normal weight function according to the size of cluster \mathbf{C}_c ;
- $$\gamma_c = 1 - \exp\left(-\left(\frac{N_c}{N}\right)^2\right) \quad 0 \leq N_c \leq N$$
- Step 5: All of the anomaly salience value of the pixel \vec{x}_i are fused by weighted summation;
- $$M(\vec{x}_i) = \sum_{c=1}^C \gamma_c (1 - D_c(\vec{x}_i)) \{ \mathbf{C}_c \in \mathbf{U} | \mathbf{U} = \mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_C \}$$

The first group of experimental data, the HYDICE Urban data-set is download from the U.S. Army Engineer Research and Development Center website. The HYDICE data includes 210 spectral bands ranging from 400 nm to 2,500 nm, with a spatial resolution of 3 m. After removing the bands of water vapor absorption and low SNR (1-4, 76, 87, 101-111,136-153, 198-210), 162 effective bands remained. The size of the original data is $307 \times 307 \times 210$. In this paper, image blocks with a size of 80×100 are cropped for the simulation experiment. This area contains vegetation, buildings, asphalt, motor vehicles and other ground objects, as shown in Fig. 3(a). These data also provide the ground truth, as shown in Fig. 3(c), where the bright target is the anomaly, and the size of each anomaly target is between 1×2 and 2×2 pixels.

The second group of experimental data is HyMap data. These data is taken by the HyMap imaging spectrometer for Cook city, Montana. The spatial resolution of the HyMap imager is 3 m, and it provides 126 spectrum bands ranging from 450 nm to 2,500 nm. The size of the original data image is $280 \times 800 \times 126$. We only take part of the data as the simulation experiment data with a size of 100×100 pixels, as shown in Fig. 3(b). There are six abnormal targets in this region, including building areas and vehicles. The corresponding ground truth shows the location of the anomaly target, and

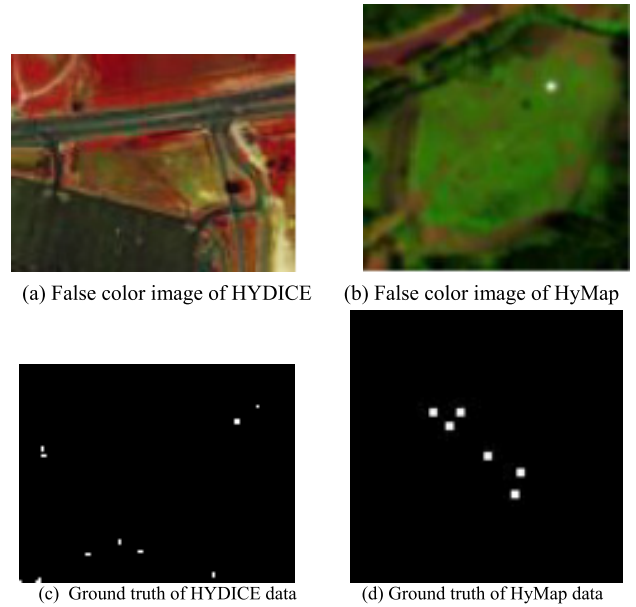


FIGURE 3. Two groups of experimental data and their ground truth.

the size of anomaly target is approximately 2×2 to 3×3 , as shown in Fig. 3(d).

B. DETECTION PERFORMANCE

To verify the performance of the proposed algorithm, CBAD, LRX, WSCF, 2DCAD and other algorithms are selected as comparison methods to detect anomalies. CBAD is the first detection method which introduces clustering into the background description. LRX is the most classic anomaly detector, and WSCF has similar accuracy as LRX. 2DCAD is a brand new method which is different from the idea of background description. It detects an anomaly by the edge feature of targets. The performances of the detection methods are evaluated by visual display and criteria quantitative parameters. The classical receiver operating characteristic (ROC) and area under the curve (AUC) are used as evaluation criteria.

The background in HYDICE includes roads, water bodies, airports, green land and other ground objects. It can be seen in Fig. 3(a) that all background ground objects are evenly distributed and have similar sizes. The anomalies in these data are cars on the road, which are small and sparse. According to the size of targets, the sliding windows of the comparison algorithms are chosen as internal windows is 3×3 and external windows is 7×7 . CBAD needs to set the number of clusters in advance. Here, it is set to 8, the same as paper in [15]. The WSCF algorithm set the parameter $\alpha = 0.2$. In the proposed method, we set the threshold of the iterative process is to $p = 0.2$, which indicates that the extraction process of the background cluster will stop when the number of remaining pixels is 20% of the total quantity of hyperspectral data. Finally, we apply the normal weight function to compute the weight of different background clusters.

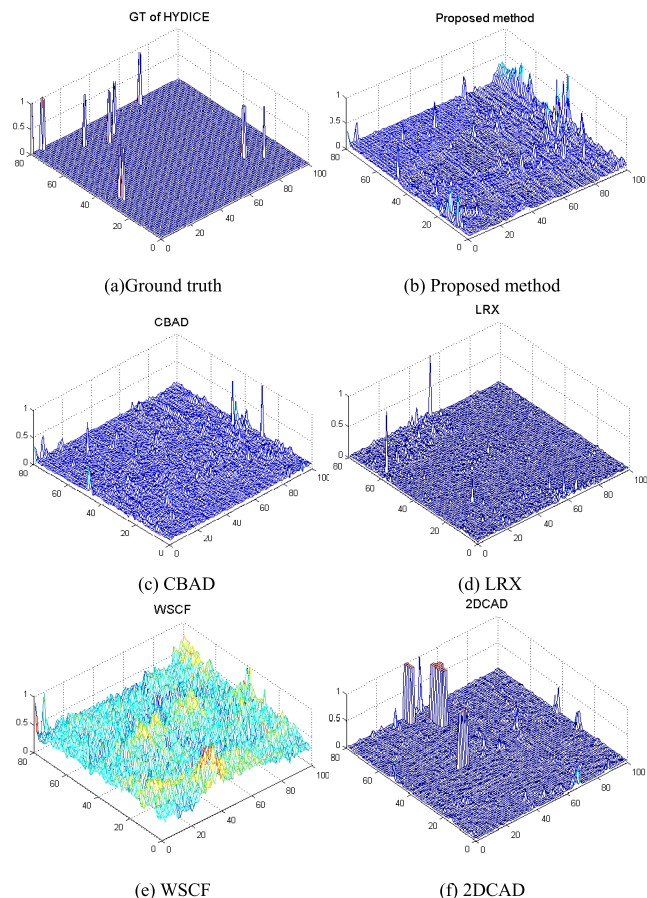


FIGURE 4. 3-D illustrations of anomaly saliency for HYDICE data.

For HYDICE data, 7 background clusters are extracted in total. In addition, for HyMap data, 8 background clusters are extracted. In the simulated experiment, the centers of the cluster are chosen artificially. Every time, we chose at least 2 centers.

Anomaly saliency of the proposed method and the comparison algorithms are shown in Fig. 4(a)-(f) in 3-D. Fig. 4(a) is the ground truth (GT) of the HYDICE data. It can be seen that most anomaly targets appear on the left side of Fig. 4(a). In Fig. 4(b), the location and number of anomalous targets are mostly coherent with GT, and the distribution structure of the ground object is more obvious than others. We cannot see the significant geographical structure in the other four saliency maps, Fig.4(c)-Fig. 4(f).

Also, we use 3-D to display the anomaly saliency of different detection methods for HyMap data, which is shown in Fig. 5(a)-(f). The first figure is the ground truth of this data; the targets are at the center part of this area. In Fig. 5(b), anomaly saliency of most background are zero. In addition, at the center part, the saliency value is larger than the background.

For other methods, shown in Fig. 5(c)-(e), the background also has higher anomaly saliency. In 2DCAD, the background and target can be separated by anomalous saliency.

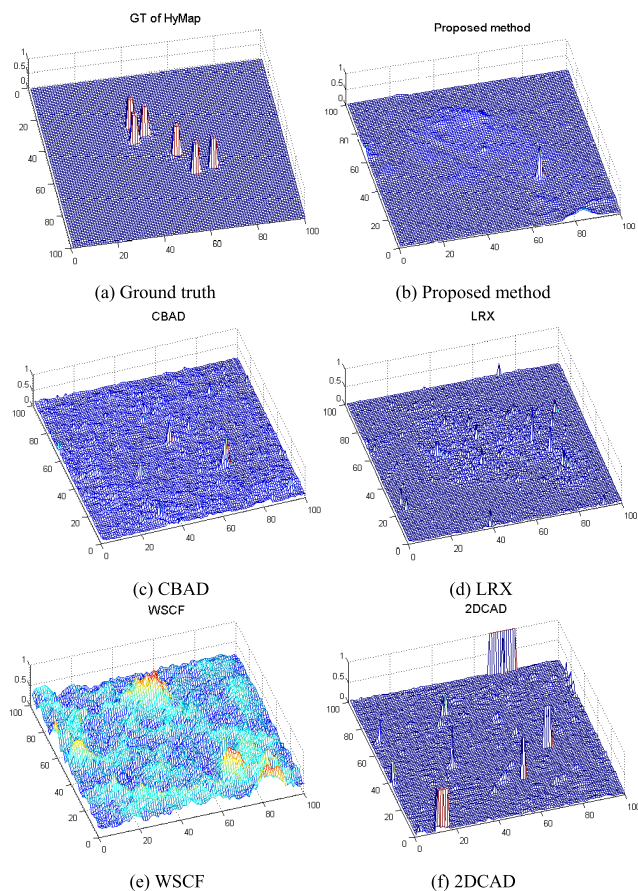


FIGURE 5. 3-D illustrations of anomaly saliency for HyMap data.

However, the false alarm rate is higher than other methods. The proposed method is effective in inhibiting background and highlighting anomalies.

According to the anomalous saliency value, an ROC curve is applied to evaluate the abnormal detection effect of each algorithm on the HYDICE and HyMap data. Fig. 6 shows the hit rate of the proposed method. Compared with other algorithms, the proposed method is higher at the same false alarm rate obviously. Additionally, the AUC is larger than that of the other algorithms. Fig. 7 is a comparison of the ROC curve and the AUC for HyMap data. It also shows that the proposed method has a better performance than that of the other methods.

IV. DISCUSSION

CBAD determines that the number of anomalies is too small to form a single cluster, and its spectral distribution is obviously different from the background. Based on this assumption, hyperspectral data is clustered into a specified number of clusters, and the model attempts to obtain a global background model. The proposed method applied a clustering method to extract the background. However, it is a coarse method compared with CBAD. Without the number of clusters, the clustering process stops when the most

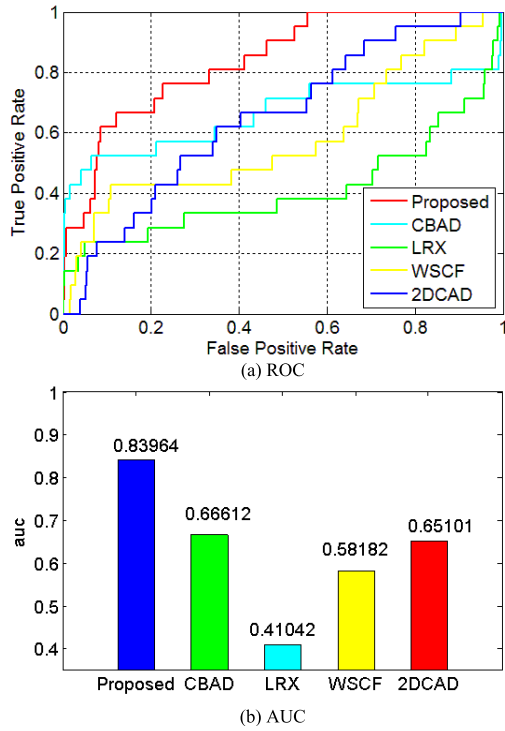


FIGURE 6. Quantitative comparison of anomaly detection results for the HYDICE data.

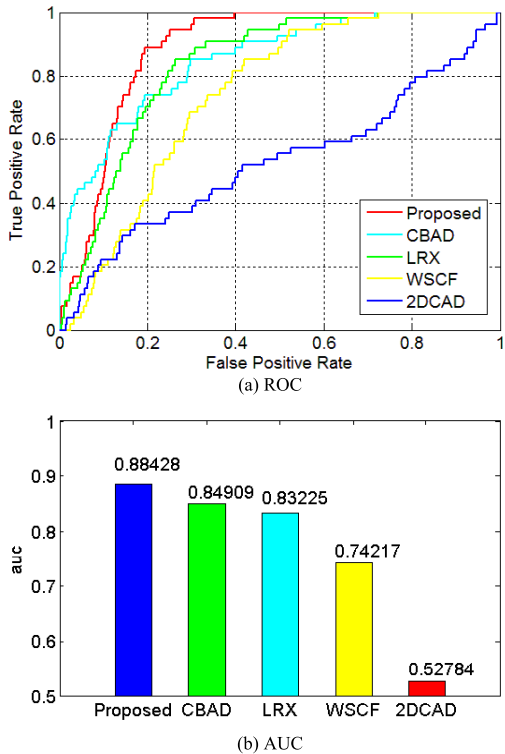


FIGURE 7. Quantitative comparison of anomaly detection results for the HyMap data.

background pixels are extracted, which avoids contamination from anomalies, and reduces the limitation of prior information. Paper [20] proposed a simple filter which expect to

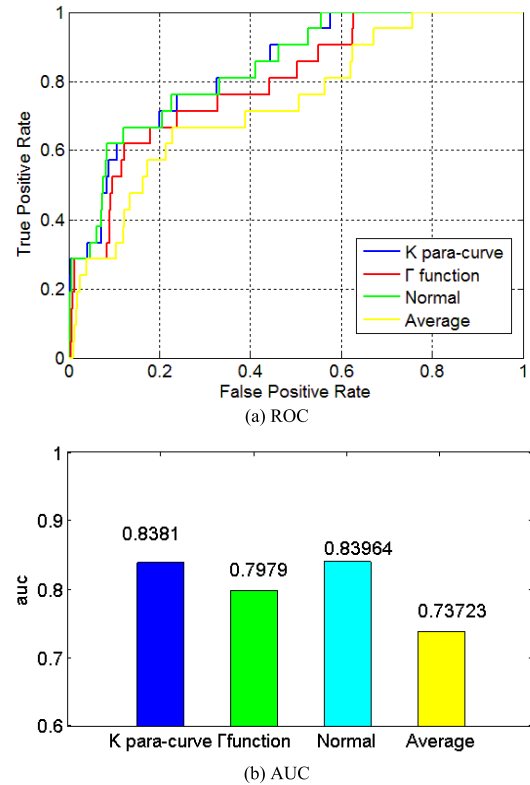


FIGURE 8. Quantitative comparison of anomaly detection results with different weight function.

improve the anomaly detection performance. In the future study, we will improve our method with this filter.

LRX and WSCF are based on local background modeling. Their performances are related to the size of the double windows and fitness between the background and model. 2DCAD detects anomalies by the mutations in the edge of the target. It also implies that the background should be smooth and simple. A simple background is more beneficial for LRX, WSCF and 2DCAD.

A. COMPARISON OF DIFFERENT WEIGHT FUNCTIONS

In this paper, we use the normal function to estimate the weight of every background cluster. In addition to this, there are some other weight functions, such as parabolic types, the Γ function and the average method. To check if different weight functions will affect the accuracy of the final result, four weight functions are applied to calculate the weight of every background cluster. Fig. 8 compares the detection results with different weights. The ROC curve (Fig. 8 (a)) and AUC value (Fig. 8 (b)) shows that parabolic type function and Γ function have lower AUC values. However, there is little difference between different weight functions.

Without knowledge of the numbers and sizes of the clusters, the proposed method iteratively and adaptively extracts background pixels according to the spectral similarity. There are still some limitations, which need to be further studied. First, in every iteration, we need to choose more than one

TABLE 1. Comparison of running times (s).

Data	Proposed method	Method				
		CBAD	2DCAD	WSCF	LRX	SVDD
HyMAP	194.65	0.28	37.37	0.63	76.59	1486.82
HYDICE	268.87	0.23	47.10	0.62	60.95	2023.31

center to cluster, because the clustering method obtains cluster centers by competition between different centers. Second, we use an empirical value as the iteration termination condition in the background extraction process. However, it is not the optimal value in complex situations. Third, cluster centers are chose artificially during the iterative clustering process. These questions will be addressed in our next study.

B. COMPARISON OF RUNNING TIMES

The computational cost is another important fact, which affects the anomaly detectors efficiency. We compare the running time of the proposed method and other classical methods in TABLE 1. Timings are taken on an Intel Core i5-2520M, 2.5 GHz CPU with 8 GB RAM. All the experiments are implemented on the MATLAB R12a platform. In this paper, we calculate the spectral distance of pixels and save it as a.mat file in advance. The calculation time does not include this part of the calculation time. Our method needs more time than other classical methods. Iterative computation always has high computational costs. Acceleration of the proposed method will be an interesting future research topic.

V. CONCLUSION

For anomaly detection of hyperspectral data, the accuracy of background extraction is directly related to the accuracy of subsequent detection. With the cluttered background, background pixels are difficult to identify because of the contamination by anomalies and the limited spatial resolution. In this paper, an adaptive background extraction method is applied to obtain the most background pixels. Without knowing the specific number of clusters, background pixels are extracted iteratively by data-driven cluster processing. Every background cluster is viewed as an observer to measure the anomaly salience of the tested pixels. In addition, a weighted summation is applied to fuse the different results from different viewers. Experimental results demonstrate that the proposed method has advantages of adaptiveness, robustness and efficiency under cluttered backgrounds. Nice performance also been showed by the ROC curve and AUC index compared with the other classical methods. However, the computational cost of the proposed method is higher, and the selection of cluster centers needs human assistance. These will be points for improvement in further research.

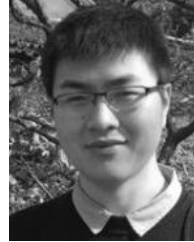
REFERENCES

- [1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [2] B. Du, Q. Wei, and R. Liu, "An improved quantum-behaved particle swarm optimization for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6003–6017, Aug. 2019.
- [3] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, 41, no. 3, pp. 1–58, 2009.
- [4] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, Sep. 2009.
- [5] M. T. Eismann, A. D. Stocker, and N. M. Nasrabadi, "Automated hyperspectral cueing for civilian search and rescue," *Proc. IEEE*, vol. 97, no. 6, pp. 1031–1055, Jun. 2009.
- [6] J. P. Kerekes and D. K. Snyder, "Unresolved target detection blind test project overview," *Proc. SPIE*, vol. 7695, May 2010, Art. no. 769521.
- [7] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 25, no. 7, pp. 5–28, Jul. 2010.
- [8] X. Yu, I. S. Reed, and A. D. Stocker, "Comparative performance analysis of adaptive multispectral detectors," *IEEE Trans. Signal Process.*, vol. 41, no. 8, pp. 2639–2656, Aug. 1993.
- [9] 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.
- [10] C. Zhao, X. Yao, and B. Huang, "Real-time anomaly detection based on a fast recursive kernel RX algorithm," *Remote Sens.*, vol. 8, no. 12, p. 1011, Dec. 2016.
- [11] A. Kanaev, E. Allman, and J. Murray-Krezan, "Reduction of false alarms caused by background boundaries in real time subspace RX anomaly detection," *Proc. SPIE*, vol. 7334, Apr. 2009, Art. no. 733405.
- [12] J. Frontera-Pons, M. A. Veganzones, F. Pascal, and J.-P. Ovarlez, "Hyperspectral anomaly detectors using robust estimators," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 720–731, Feb. 2016.
- [13] S. Matteoli, M. Diani, and J. Theiler, "An overview of background modeling for detection of targets and anomalies in hyperspectral remotely sensed imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2317–2336, Jun. 2014.
- [14] S. P. Catterall, "Anomaly detection based on the statistics of hyperspectral imagery," *Proc. SPIE*, vol. 5546, Oct. 2004, Art. no. 557161.
- [15] 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.
- [16] M. J. Carlotto, "A cluster-based approach for detecting man-made objects and changes in imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 374–387, Feb. 2005.
- [17] H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 388–397, Feb. 2005.
- [18] S. Khazai and B. Mojaradi, "A modified kernel-RX algorithm for anomaly detection in hyperspectral images," *Arabian J. Geosci.*, vol. 8, no. 3, pp. 1487–1495, Dec. 2013.
- [19] B. Du and L. Zhang, "Random-Selection-Based anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 5, pp. 1578–1589, May 2011.
- [20] 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.
- [21] Y. Zhang, Y. Fan, and M. Xu, "A Background-Purification-Based framework for anomaly target detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, to be published.
- [22] L. Zhu, G. Wen, S. Qiu, and X. Zhang, "Improving hyperspectral anomaly detection with a simple weighting strategy," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 95–99, Jan. 2019.
- [23] Q. Ling, Y. Guo, Z. Lin, and W. An, "A constrained sparse representation model for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2358–2371, Apr. 2019.
- [24] 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.

- [25] F. Li, X. Zhang, L. Zhang, D. Jiang, and Y. Zhang, "Exploiting structured sparsity for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4050–4064, Jul. 2018.
- [26] Y. Qu, W. Wang, R. Guo, B. Ayhan, C. Kwan, S. Vance, and H. Qi, "Hyperspectral anomaly detection through spectral unmixing and dictionary-based low-rank decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4391–4405, Aug. 2018.
- [27] 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.*, to be published.
- [28] R. Zhao, B. Du, L. Zhang, and L. Zhang, "Beyond background feature extraction: An anomaly detection algorithm inspired by slowly varying signal analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1757–1774, Mar. 2016.
- [29] Y. Zhang, B. Du, L. Zhang, and S. Wang, "A low-rank and sparse matrix decomposition-based mahalanobis distance method for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1376–1389, Mar. 2016.
- [30] Y. Niu and B. Wang, "Hyperspectral anomaly detection based on low-rank representation and learned dictionary," *Remote Sens.*, vol. 8, no. 4, p. 289, Mar. 2016.
- [31] S. Chang, B. Du, and L. Zhang, "BASO: A background-anomaly component projection and separation optimized filter for anomaly detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 3747–3761, Jul. 2018.
- [32] Y. Wang, L.-C. Lee, B. Xue, L. Wang, M. Song, C. Yu, S. Li, and C.-I. Chang, "A posteriori hyperspectral anomaly detection for unlabeled classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3091–3106, Jun. 2018.
- [33] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, Jun. 2014.



MIN LI received the B.S. degree in electronic information engineering and the Ph.D. degree in communication engineering from Hohai University, Changzhou, in 2005 and 2011, respectively. She is currently an Assistant Professor with the College of Internet of Things, Hohai University. Her research interests include remote sensing anomaly detection, image super resolution, and target detection.



PUHUANG LI received the B.S. degree in communication engineering from the Changzhou Institute of Technology, in 2015, and the M.S. degree in communication and information engineering from Hohai University, Changzhou, China, in 2018, where he is currently pursuing the Ph.D. degree with the College of Internet of Things. His research interests include the remote sensing anomaly detection, computer science, and optical communication technology.



HAIYAN XU received the M.S. degree in mechanical and electronic engineering from Hohai University, Changzhou, China, in 2004, and the Ph.D. degree in communication engineering from Fudan University, Shanghai, China, in 2011. She is currently an Assistant Professor with the College of Internet of Things, Hohai University. Her research interests include the optical fiber signal processing and target detection.

• • •