Article

Spectra-Based Selective Searching for Hyperspectral Anomaly Detection

Chensong Yin [1,2], Chengshan Han [1,*], Xucheng Xue [1] and Liang Huang [1,2]

[1] Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China; yinchensong16@mails.ucas.ac.cn (C.Y.); xuexucheng@ciomp.ac.cn (X.X.); huangliang@ciomp.ac.cn (L.H.)
[2] University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: hancs@ciomp.ac.cn; Tel.: +86-0431-86708135

Abstract: The research on hyperspectral anomaly detection algorithms has become a hotspot, driven by a lot of practical applications, such as mineral exploration, environmental monitoring and the national defense force. However, most existing hyperspectral anomaly detectors are designed with a single pixel as unit, which may not make full use of the spatial and spectral information in the hyperspectral image to detect anomalies. In this paper, to fully combine and utilize the spatial and spectral information of hyperspectral images, we propose a novel spectral-based selective searching method for hyperspectral anomaly detection, which firstly combines adjacent pixels with the same spectral characteristics into regions with adaptive shape and size and then treats those regions as one processing unit. Then, by fusing adjacent regions with similar spectral characteristics, the anomaly can be successfully distinguished from background. Two standard hyperspectral datasets are introduced to verify the feasibility and effectiveness of the proposed method. The detection performance is depicted by intuitive detection images, receiver operating characteristic curves and area under curve values. Comparing the results of the proposed method with five popular and state-of-the-art methods proves that the spectral-based selective searching method is an accurate and effective method to detect anomalies.

Keywords: selective searching; hyperspectral imagery; anomaly detection; oversegmentation

1. Introduction

Hyperspectral imaging (HSI) endowed with abundant spectral information is a powerful tool for target detection in remote sensing because a variety of ground objects, such as various natural land covers and artificialities, have distinct spectral signatures [1–4]. Target detection seeks to distinguish the specific target pixels from various backgrounds via prior target spectra or anomalous properties [5–7]. In the past few decades, hyperspectral target detection has attracted a lot of attention in view of its importance in a growing number of applications [8], such as mineral exploration [9], environmental monitoring [10] and the national defense force [11]. Usually, target detection without prior knowledge of the target is called anomaly detection [11–15]. This technique can directly identify targets with spatial or spectral characteristics that are quite different from the surroundings [16,17].

In general, statistical and geometrical strategies are widely used for current anomaly detection algorithms. As the benchmark of many statistical-based anomaly detection algorithms, the Reed–Xiaoli (RX) algorithm [18] assumes that the background of HSI is a multivariate Gaussian distribution and statistics such as the mean and covariance values of the background are computed. Then, the Mahalanobis distance between each pixel and the background is calculated to score the probability that the pixel is an anomaly. The RX algorithm has two common versions depending on the methods of background description [19]. One is the global RX (GRX), in which all pixels in the image are viewed as background and used to calculate background statistics. The other is the local RX (LRX) [20–22], in which

a sliding fixed-size dual-window is employed to select those surrounding pixels as the background pixels of each tested pixel to calculate local background statistics. However, the multivariate Gaussian distribution is too simple to accurately characterize the complicated background of real-world hyperspectral scenes. As a result, a series of modified RX-based anomaly detectors was designed [23–27]. For example, the covariance matrix of an image was standardized in a regularized-RX detector [25], which could avoid the pathological state of matrix inversion. The kernel-RX (KRX) detector [26] is generally regarded as a nonlinear version of the RX, in which the origin HSI data is mapped to a high-dimensional feature space via the kernel theory to strengthen the differences between anomaly target and background. The cluster-based anomaly detection (CBAD) [27] subdivides the entire image into several clusters, and the RX detector is implemented to detect the anomaly target in each cluster. However, background statistics that are more or less contaminated by anomaly pixels and noise may not effectively characterize the background information. In order to further settle this problem, spectral transform such as fractional Fourier transform (FrFT) [28] were employed for hyperspectral anomaly detection to handle nonstationary noise introduced during the imaging process from satellite or airborne sensors and to extract the intrinsic features of the hyperspectral image. In Ref. [29], the FrFT was applied to process the hyperspectral data, and the original data were transformed to the intermediate domain, including the reflection spectrum and its Fourier transformation information, so that the anomaly target could be clearly distinguished from the background and noise. Then, the RX method achieved better detection performance in this optimal fractional domain. This method is call fractional Fourier entropy (FFE). In Ref. [30], low rank and sparse matrix decomposition (LRaSMD) was utilized to achieve background purification by decomposing the image data into background, noise and anomaly after extracting intrinsic features from the hyperspectral image via FrFT. Then background atoms were selected to build the covariance matrix for Mahalanobis distance-based RX anomaly detection, named feature extraction and background purification anomaly detector (FEBPAD). In addition to the above Gaussian model-based detector, some robust statistic modeling detectors have been developed, including the discriminative metric learning-based anomaly detector [31], random-selection-based model [32], and nonlinear learning-based detection method [33]. These anomaly detectors construct background statistics through different strategies, thereby avoiding contamination of background statistics by abnormal pixels. However, it is still difficult to secure satisfactory anomaly detection performance when processing images with complex backgrounds.

Apart from the aforementioned statistical strategy-based detection methods, geometrical strategy-based detection methods have recently gained much attention. This strategy usually considers that the background pixels can be approximated as a certain combination of a group of primary spectra extracted from HSI data, while the anomaly target is exceptional and cannot be reconstructed with those primary spectra. For instance, in the subspace-based detector [34], a low-dimension feature space considered as a collection of background spectra is constructed by a linear transformation from the HSI data cube. The representation residual of test pixels can be calculated by projecting all the pixels onto the background orthogonal subspace and then used for anomaly detection. Sparse representation and collaborative representation [35–37] have been recently used for anomaly detection of hyperspectral images. In the low-rank and sparse representation (LRASR) detector [38], the low-rank representation is applied to model the background, while a sparsity-induced regularization term is appended to the representation coefficients matrix which contains the background information. K-means is used to select the background dictionary's atoms from HSI data cube. The principle of the collaborative representation-based detector (CRD) [39] is that each background pixel can be approximately represented by its spatial surrounding pixels in a sliding size-fixed local dual-window with the form of linear combinations, while anomalies cannot. The ℓ_2-norm minimization of the representation weight vector is employed to enhance the collaboration among these pixels. By subtracting the predicted background result from the origin HSI data, a residual image can be obtained

to reveal the anomaly pixels. However, regardless of the sparse representation-based detector or collaborative representation-based detector, the performance of those detectors highly depends on the background dictionary that is chosen. If the background dictionary is inexact or contaminated by anomalies, the detection performance may deteriorate. This means that the background dictionary extracted in sparse representation needs to be as accurate and complete as possible, and the size of the dual windows is one of the most critical factors in collaborative representation. In Ref. [40], the various local spatial distribution information of the neighboring pixels of a test pixel were considered by adding a summation strategy to the local window in the CRD algorithm, which improved the accuracy of linear representation and the detection performance, but this came at the cost of increased computational complexity and time, called local summation anomaly detection based on collaborative representation and inverse distance weight (LSAD-CR-IDW).

In this paper, to fully combine and utilize the spatial and spectral information of hyperspectral images and to implement hyperspectral anomaly detection, a spectra-based selective searching (Triple-S) hyperspectral anomaly detection algorithm is proposed. Unlike most existing algorithms, the adaptive regions composed of adjacent pixels with the same spectral characteristics are referred to as the processing unit rather than a single pixel. In addition to the differences between the anomalies and the background, similar properties among the anomaly pixels and among the background pixels are exploited in the proposed detector via two phases: oversegmentation and region merging. First, a graph-based image segmentation method [41] is applied to HSIs via a complementary pixel-grouping criterion, which can dynamically adjust the weight of the shape and amplitude in the measurement of pixel spectral similarity. This method can quickly and effectively obtain a set of initial regions, each of which is spectrally uniform and does not contain anomalous and background pixels simultaneously. In addition, the segmentation criterion is adaptive in capturing anomalies in all possible locations and scales without exhaustive searching. Then, the similarity of each pair of adjacent regions after the initial oversegmentation is calculated based on the spectral mean of each region and recorded in the initial similarity matrix. The measure of similarity is the linear combination of the spectral angle and the correlation coefficient between each pair of adjacent regions. Based on the similarity matrix, region merging via the greedy algorithm is imposed to merge similar regions successively and finally leave dissimilar regions. Further, anomalies are the regions that are distinguished from the background of the HSIs. Thanks to the above spectral-based selective search process, the proposed algorithm is effective and handy without assuming a statistical background distribution and constructing a background or target dictionary. The detection performance can be assessed by intuitive detection images, receiver operating characteristic curves and area under curve values. We evaluated and compared the proposed method using two widely-used real hyperspectral datasets. The results showed that the proposed method possessed higher detection accuracies and lower false alarm rate than several popular and state-of-the-art methods.

The contents of the manuscript are arranged as follows. Section 2 first introduces the common measurement of spectral similarity and then presents a detailed description of the proposed Triple-S hyperspectral anomaly algorithm. The employed evaluation indicators of anomaly detection performance and the data set used in the experiment are also described in this section. Section 3 shows the experimental detection results of the proposed algorithm implemented on two real hyperspectral datasets. Section 4 introduces the parameter setting of our Triple-S and comparable algorithms. Anomaly detection performance of the proposed algorithm is also discussed qualitatively and quantitatively. The competitive detection performance of the proposed algorithm is demonstrated by comparing it with several representative methods. The summary and outlook of the algorithm and performance are provided in Section 5.

2. Materials and Methods

In this section, the common measurement of spectral similarity is firstly introduced. Then the detailed description of the proposed method is presented. Figure 1 shows a schematic of the proposed spectra-based selective searching algorithm for anomaly detection in hyperspectral images. This algorithm contains two main phases: oversegmentation and region merging. First, a graph-based image segmentation technique was applied to the HSI to divide the image into a series of initial regions as the processing unit. The similarities between all neighboring regions were then calculated and recorded. Subsequently, the greedy algorithm was applied to iteratively merge the regions until the separation of the anomalies from the background was achieved. The widely-used criteria to evaluate the anomaly detection performance and two common datasets are introduced briefly.
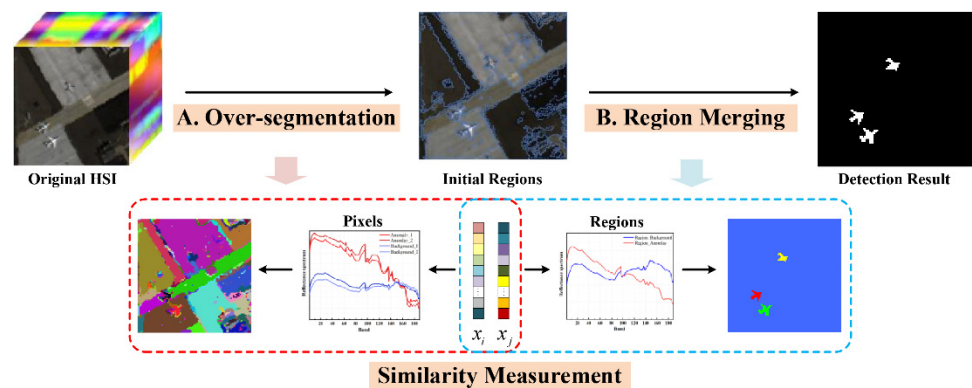
Figure 1. Flowchart of the proposed algorithm.

2.1. Measurement of Spectral Similarity

Consider a 3-D hyperspectral cube by $X = \{x_i\}_{i=1}^{n}$ in $\mathbb{R}^d$, where n denotes the number of pixels in the HSI, d represents the number of spectral bands, and x_i represents the spectrum of each pixel of the HSI, which can be regarded as a vector quantity with each entry of the vector representing the reflectance value at each spectral band. Therefore, each pixel spectral curve is an efficient tool for describing the characteristics of the corresponding ground object. The spectral similarity criterion demonstrates the similarity degree between two given spectral curves. Generally, there are two types of spectral similarity criteria. One is a numerical index that mainly describes the difference in radiant luminance of ground objects, such as Euclidean distance (ED) and Lance distance (LD, also known as the Canberra distance) [42]. The other is the shape index, which mainly describes the difference in waveform among spectral curves, such as the spectral angle measure (SAM) and correlation coefficient (*r*). Neither the numerical index nor the shape index can fully reflect the degree of similarity between spectral vectors. In this study, the numerical index and shape index are combined to accurately describe the similarity between spectra.

To determine the appropriate similarity criterion according to the specific situation of the spectral curve, the information entropy of the spectral difference curve is introduced. First, as the name suggests, the spectral difference curve is obtained by subtracting the two spectra and taking the absolute value. When the spectral curves of two pixels are similar in shape but different in the overall amplitude, the spectral difference curve can be approximated as a straight line with a slope of zero. When there is a significant difference in the shapes of the spectral curves of the two pixels, the spectral difference curve is a curve with obvious fluctuations. The information entropy of the spectral difference curve can be used to describe the nature of the spectral curve difference to dynamically determine the similarity criterion. The information entropy [43] of the spectral difference curve is defined as

$$Ent = -\sum_{t=1}^{d} P_t \lg(P_t) \tag{1}$$

where $P_t = x_t / \sum\limits_{t=1}^{d} x_t$ is the probability of the tth value in the spectral difference vector x.
The larger the Ent, the more obvious the difference between the shapes of the spectral curves. When $Ent = 0$, the shapes of the spectral curves are completely consistent. Lance distance [42] is selected as the numerical index of spectral similarity in the proposed Triple-S algorithm because it is not sensitive to singular values and can suppress noise very well. In addition, the LD value ranges from zero to one, which is convenient for combining with the shape index. The Lance distance can be expressed as

$$LD = \frac{1}{d}\sum_{t=1}^{d} \frac{|x_{it} - x_{jt}|}{x_{it} + x_{jt}} \tag{2}$$

where x_{it} and x_{jt} are the entries of the two spectral vectors in the tth band, respectively.

SAM [39] is a widely used criterion to evaluate the shape similarity of two spectral vectors by calculating the angle between two spectral vectors x_i and x_j, and the angle can be calculated by

$$sa = \arccos\left(\frac{x_i \cdot x_j}{\|x_i\| \cdot \|x_j\|}\right) \tag{3}$$

The obtained angle intuitively shows the similarity of the spectral curves in shape. The value range of Equation (3) is from 0 to π. The smaller the angle, the more similar the two spectral curves are.

The correlation coefficient [39] is another common criterion to describe the similarity of two spectral vector shapes. The correlation coefficient can be calculated by

$$r = \frac{\sum\limits_{t=1}^{d}\left(x_{it} - \overline{x_i}\right) \cdot \sum\limits_{t=1}^{d}\left(x_{jt} - \overline{x_j}\right)}{\sqrt{\sum\limits_{t=1}^{d}\left(x_{it} - \overline{x_i}\right)^2 \cdot \sum\limits_{t=1}^{d}\left(x_{jt} - \overline{x_j}\right)^2}} \tag{4}$$

where $\overline{x_i}$ and $\overline{x_j}$ are the means of the two spectral vectors in all bands. The value range of Equation (4) is from -1 to 1. The closer the correlation coefficient is to one, the more similar are the shapes of the two spectral vectors.

In order to accurately describe the similarity between spectral vectors, it is necessary to adopt a comprehensive similarity criterion. In this study, the information entropy of the spectral difference curve is combined with the correlation coefficient and the Lance distance to form a dynamic similarity operator to realize accurate initial segmentation of hyperspectral images.

2.2. Proposed Anomaly Detection Algorithm

2.2.1. Oversegmentation of HSI

As regions can yield richer information than a single pixel, we regard regions instead of pixels as the basic unit to more precisely represent background and anomaly characteristics. Therefore, we adapt the graph-based image segmentation algorithm [41] to hyperspectral images to integrate neighboring pixels with high similarity into a series of regions.

In general, a hyperspectral image can be represented as $HSI = (X, E)$, which is composed of pixel spectra $x \in X$ and edges $(x_i, x_j) \in E$ representing the connection of the pairs of neighboring pixels. Each edge $(x_i, x_j) \in E$ has a corresponding weight $w(x_i, x_j)$ which is the spectral similarity between the two neighboring pixels x_i and x_j connected by that edge. The segmentation of the HSI is a subdivision of the pixels into a series of regions such that each region corresponds to a connected component C and the pixels in each component are homogeneous. This requires that edges between two pixels in the same region should have relatively low weights, meaning that they are more similar. In order to effectively separate anomaly pixels from background pixels, we define the weight

$w(x_i, x_j)$ as the following dynamic similarity operator according to the measurement of spectral similarity.

$$w = Ent \times (1 - r) + LD \tag{5}$$

The smaller the weight, the more similar are the two pixels, and the sooner the two pixels will be combined into a region.

In addition to the definition of edge weights, the criterion for determining whether two components can be the same component is critical. In general, when the weight between two connected components is less than the maximum value of the internal weights for the two components, they can be considered as the same component. In math, this criterion can be expressed as

$$Con(C_i, C_j) = \begin{cases} true & if\ Dif(C_i, C_j) \leq MIDC(C_i, C_j) \\ false & otherwise \end{cases} \tag{6}$$

where $Con(C_i, C_j)$ represents whether two components (C_i, C_j) should be connected and become the same component, and $Dif(C_i, C_j)$ describes the difference between two components and is defined as the minimum weight connecting the two components, as follows:

$$Dif(C_i, C_j) = \min_{x_1 \in C_i, x_2 \in C_j, (x_1, x_2) \in E} w(x_1, x_2) \tag{7}$$

$MIDC(C_i, C_j)$ is the minimum internal difference, which is defined as

$$MIDC(C_i, C_j) = \min(IDC(C_i) + \tau(C_i), IDC(C_j) + \tau(C_j)) \tag{8}$$

where $IDC(C)$ is the largest weight of the minimum spanning tree (MST) of the component $MST(C, E)$ and expresses the internal difference of a component, given by

$$IDC = \max_{e \in MST(C,E)} w(e) \tag{9}$$

For small components, $IDC(C)$ cannot accurately estimate the local characteristics of the component. For example, $IDC(C) = 0$ when the component is a single pixel. Therefore, the threshold function $\tau(C)$ is necessary for segmentation to control the degree of difference between two different components to be greater than their respective internal differences.

$$\tau(C) = \frac{K}{[C]} \tag{10}$$

where, $[C]$ denotes the number of pixels in component C and K is a constant parameter that controls the size of the formed components, therefore setting a scale of oversegmentation. A larger K creates a preference for larger components. Moreover, K is not the minimum component size. This threshold function can make the segmentation criterion adaptive, which can preserve detail in the information of high-variability regions while ignoring detail in low-variability regions. Anomalies can occur at any location and in any shape within the HSI. This oversegmentation method is completely data-driven and divides the image into a series of initial regions with freewill shape and size. Anomaly targets of different sizes or shapes can form some initial regions according to the image data. In addition, the oversegmentation method has high efficiency and its running time is related to the number of image pixels.

2.2.2. Region Merging

By setting an appropriate segmentation parameter K and using a dynamic similarity criterion among the anomalous pixels and background pixels, the applicable oversegmentation of hyperspectral images is achieved via the graph-based image segmentation method. It is important to further coalesce the background regions so that anomaly pixels can be

distinguished. The principle of region merging is that the background and anomaly regions are dissimilar. Therefore, the similarity computation between all neighboring regions is an important basis for judging whether the regions belong to the same class or object [44–46]. The average spectrum of all pixels in each region is utilized to represent the spectral characteristics of regions, thereby calculating the similarities of adjacent regions. Only the shape similarity of the regional mean spectrum is considered here, and the similarity is computed by the linear combination of the spectral angle and the correlation coefficient, given by

$$simi = \left(1 - \frac{sa}{\pi}\right) + \frac{(1+r)}{2} \tag{11}$$

The value range of Equation (11) is from 0 to 2. The larger *simi* is, the more similar the two neighboring regions are, and the sooner the two regions will be merged into one region. The similarities of all pairs of initial neighboring regions are calculated and recorded in a similarity matrix. The greedy algorithm is used to iteratively merge the two adjacent regions with the highest similarity. A threshold T for constraining the degree of region consolidation is preset, whose range is determined by the minimum and the maximum of the similarity matrix. Then, the maximum value of the similarity matrix is extracted and compared with the threshold. When the maximum is larger than the threshold, the corresponding two contiguous regions are merged into a new region. The similarity values related to the two merged regions in the original similarity matrix are deleted. The similarity matrix is updated by computing the similarities between the new region and its neighboring regions. The regions stop merging until the iterative loop condition is not satisfied, that is, the maximum value in the similarity matrix is less than the threshold. Finally, those regions that fail to merge with most of the pixels in the image are anomalous targets to be detected. The flow of the proposed spectra-based selective searching algorithm is summarized in Algorithm 1.

Algorithm 1: Spectra-based Selective Searching for Anomaly Detection

Input 　　1). A hyperspectral image *HSI*

　　　　　　2). The parameter K for the creation of initial region

　　　　　　3). The threshold T for the region merging

Output 　A two-dimensional map recording the detection result *Img-R*

Procedure 　For the input hyperspectral image data *HSI*

　　　　　　1). Obtain initial region $R = \{r_1, \cdots, r_n\}$ using Felzenszwalb and Hutttenlocher (2004);

　　　　　　2). Initialise similarity set $Simi = \varnothing$;

　　　　　　3). For each neighboring region pair $\left(r_i, r_j\right)$ **do**

　　　　　　　　Calculate similarity $Simi\left(r_i, r_j\right)$;

　　　　　　　　$Simi = Simi \cup Simi\left(r_i, r_j\right)$;

　　　　　　　　Get highest similarity $maxSimi = \max(Simi) = Simi\left(r_i, r_j\right)$;

　　　　　　4). While $maxSimi \geq T$ **do**

　　　　　　　　Merge corresponding regions $r_t = r_i \cup r_j$;

　　　　　　　　Remove similarities regarding r_i : $Simi = Simi \backslash Simi(r_i, r^*)$;

　　　　　　　　Remove similarities regarding r_j : $Simi = Simi \backslash Simi\left(r_j, r^*\right)$;

　　　　　　　　Calculating similarity Simit between region r_t and its neighbors;

　　　　　　　　Update similarity set $Simi$: $Simi = Simi \cup Simi_t$;

　　　　　　　　Update highest similarity $maxSimi = \max(Simi) = Simi\left(r_i, r_j\right)$;

2.3. Evaluation Criteria

This study evaluates the anomaly detection performance of the proposed algorithm from both qualitative and quantitative aspects. The qualitative evaluation is mainly based on the visual effect of the detection result graph, while the quantitative evaluation is carried out by drawing the receiver operating characteristic curve (ROC) [47], calculating the area under the ROC curve (AUC) value [48] and comparing the calculation time. The ROC

curve originating from the communication field is widely used to describe detection performance, showing the relationship between the true positive rate (TPR) and the false positive rate (FPR) by varying the threshold between 0 and 1. The TPR represents the detection probability, and the FPR is also known as the false alarm rate, which can be defined as follows:

$$TPR = \frac{TP}{TP+FN}$$
$$FPR = \frac{FP}{TN+FP}$$

(12)

where *TP*, *TN*, *FP*, and *FN* stand for true positive, true negative, false positive, and false negative, respectively. In general, a higher TPR value means that the algorithm detection performance is better when the FPR value is the same. Moreover, the corresponding ROC curve will be closer to the upper left corner of the coordinate plane, and the area under the curve will be larger. However, observing the ROC curve alone cannot directly and intuitively judge and compare the effectiveness of the anomaly detectors. Thus, AUC is taken into consideration, which is the value of the area under the ROC curve. The closer the AUC value is to unity, the better the detection performance of the algorithm is.
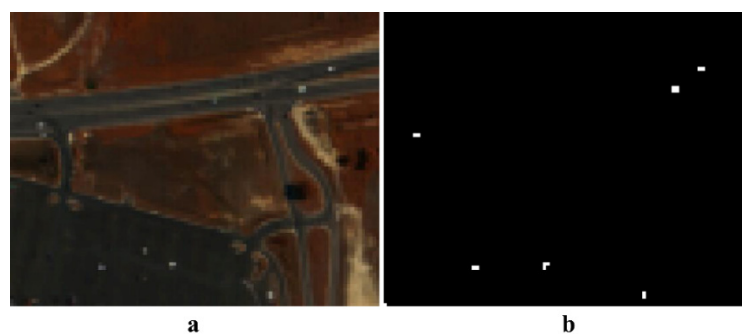
2.4. Hyperspectral Dataset

The first hyperspectral dataset is named AVIRIS San Diego airport in this manuscript and was collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), mainly covering the urban scene where San Diego Airport is located in California, USA. The spatial resolution of the dataset was 3.5 m, and the number of spectral channels was 224, spanning the wavelength range of 370 nm to 2510 nm. After removing the corresponding water absorption regions and low signal-to-noise ratio bands (1–6, 33–35, 107–113, 153–166, and 221–224), 189 bands of the raw dataset were reserved for the experiments. The top-left 100 × 100 pixels were chosen as the test image. The land-cover type in this scene is mainly ground, road, roof, and shadow, as shown in Figure 2a. There are three aircraft which are the anomaly targets to be detected. The corresponding ground-truth map is shown in Figure 2b.

a   b
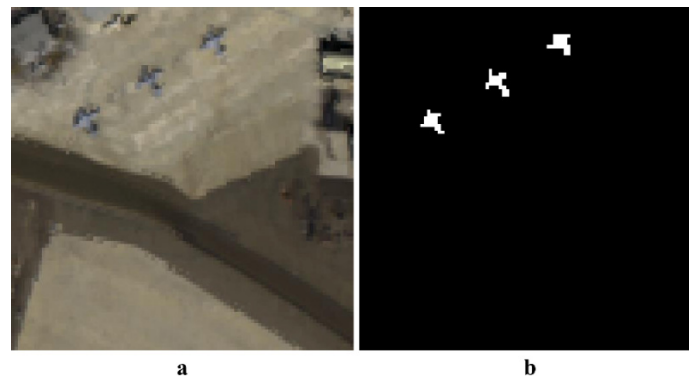
Figure 2. First real-world data set: AVIRIS San Diego airport scene. (**a**) False color image; (**b**) ground-truth map.

The second hyperspectral dataset is named HYDICE Urban in this manuscript, obtained from the Hyperspectral Digital Imagery Collection Experiment (HYDICE) remote sensor. This dataset mainly covered an urban scene consisting of some plant areas, several roofs, and some roads with vehicles. The spatial and spectral resolutions of the image are 1 m and 10 nm, respectively. This dataset contained 210 spectral channels from 400 nm to 2500 nm. Low signal-to-noise ratio and water vapor absorption bands (1–4, 76, 87, 101–111, 136–153, and 198–210) were eliminated, retaining 162 bands for the experiments. The ground truth defined that the anomalous targets were the vehicles embedded in different backgrounds. The test image considered an 80 × 100 subscene containing the anomaly pixels. Figure 3a,b show the false color image and the corresponding ground truth map of the dataset, respectively.

Figure 3. Second real-world data set: HYDICE urban scene. (**a**) False color image; (**b**) ground-truth map.
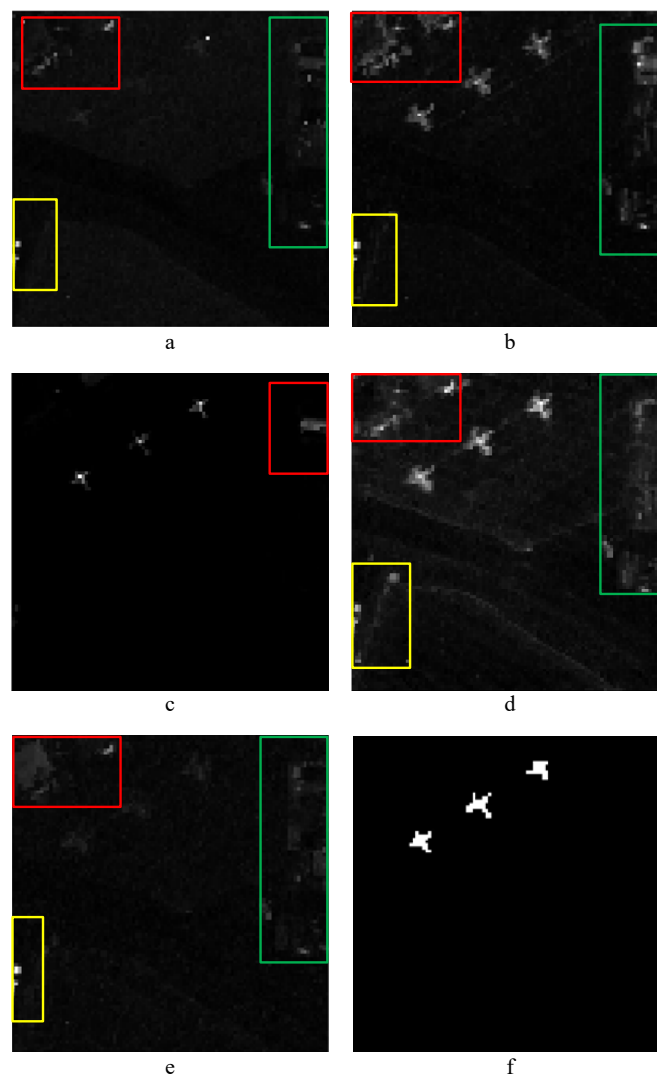
3. Results

Verification experiments on two real typical hyperspectral datasets were conducted in this section to demonstrate the anomaly detection performance of the proposed spectra-based selective searching (Triple-S) algorithm. Five previous excellent anomaly detection methods, such as GRX [19], FEE [29], FEBPAD [30], LSAD-CR-IDW [40] and LRSRD [38] were used to verify the effectiveness of the Triple-S algorithm proposed in this paper. The validity of the proposed Triple-S method and these representative methods for comparison on these two real hyperspectral datasets are displayed, including detection results, receiver operating characteristic (ROC) curve [47], the area under the ROC curve (AUC) values [48] and the running time. All the algorithms were implemented in MATLAB on a Dell XPS13 personal laptop with an Intel Core i5-8250U CPU with 8 GB of RAM.

To assess our Triple-S anomaly detection algorithm both qualitatively and quantitatively, an experiment was conducted to evaluate the anomaly detection performance of the proposed Triple-S algorithm. Meanwhile, five typical anomaly detection algorithms, GRX [19], FEE [29], FEBPAD [30], LSAD-CR-IDW [40] and LRSRD [38], were employed for comparison and the corresponding detection results were obtained from the source code shared on GitHub.

As described in Section 2.2, the proposed Triple-S anomaly detector consists of two steps, oversegmentation and region merging. The results of our proposed Triple-S algorithm are shown in Figure 4a,b for AVIRIS San Diego airport dataset and Figure 4c,d for HYDICE Urban dataset. Specifically, Figure 4a,c are separately the results of oversegmentation for two test hyperspectral images. Figure 4b,d are separately the results of region merging for two test hyperspectral images. The implementation of these five algorithms used for comparison are according to the source codes from GitHub and the principles of the respective algorithms have been briefly introduced in the Introduction section and were described in detail in the corresponding references. The detection results of five comparable algorithms as well as the proposed Triple-S algorithm are shown in Figures 5a–f and 6a–f for these two test hyperspectral images. In order to assess the detection accuracy, the ROC curves of different anomaly detectors for the quantitative comparison of detection performance are shown in Figure 7 for AVIRIS San Diego airport dataset and Figure 8 for HYDICE Urban dataset.
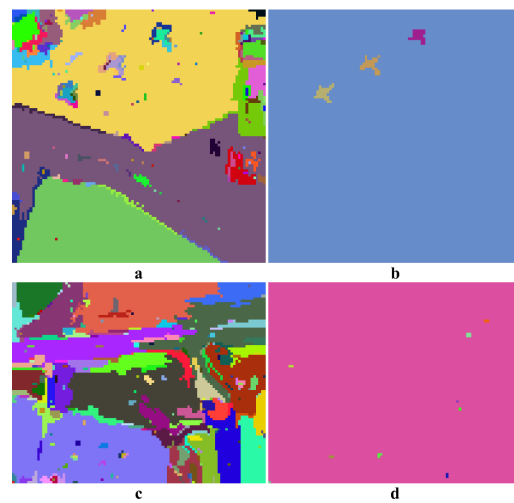
Figure 4. Results of oversegmentation and region merging of (**a**,**b**) AVIRIS San Diego airport dataset; (**c**,**d**) HYDICE urban dataset. (**a**) The result of oversegmentation for the San Diego dataset, (**b**) the result of region merging for the San Diego dataset, (**c**) the result of oversegmentation for the Urban dataset, (**d**) the result of region merging for the Urban dataset.

Figure 5. Detection results of (**a**) GRX, (**b**) FEE, (**c**) FEBPAD, (**d**) LSAD-CR-IDW, (**e**) LRSRD, and (**f**) Triple-S for AVIRIS San Diego airport dataset.
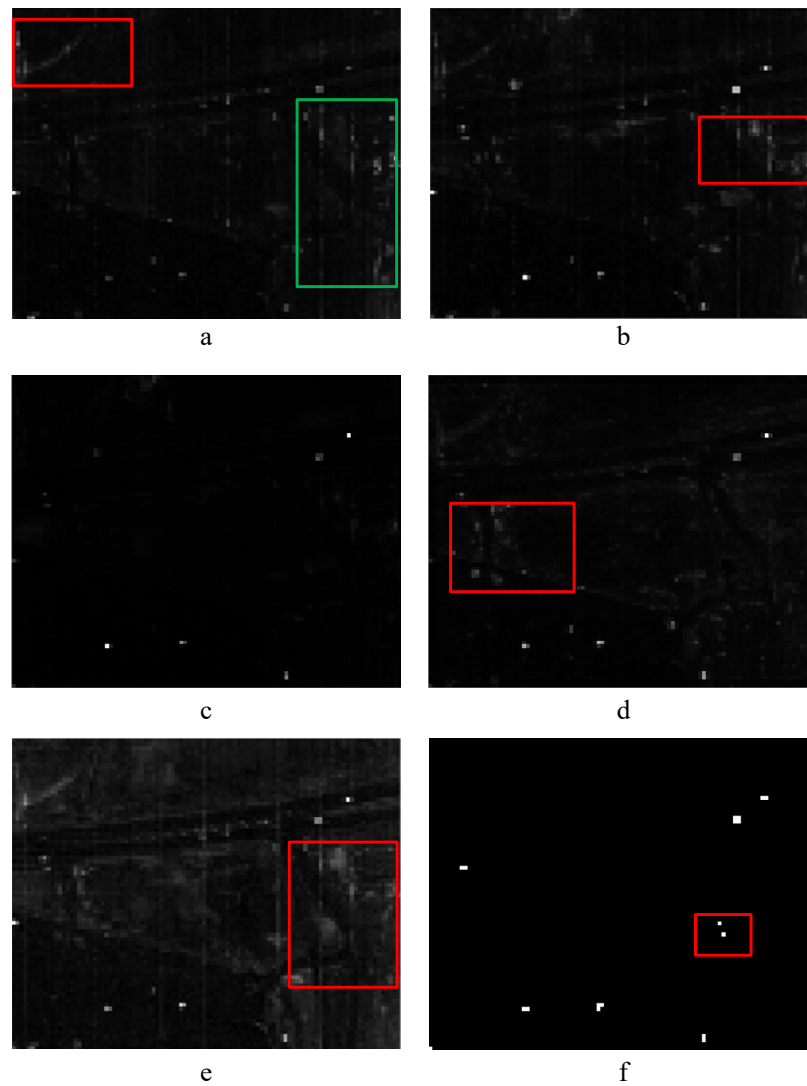
Figure 6. Detection results of (**a**) GRX, (**b**) FEE, (**c**) FEBPAD, (**d**) LSAD-CR-IDW, (**e**) LRSRD, and (**f**) Triple-S for HYDICE Urban data set.
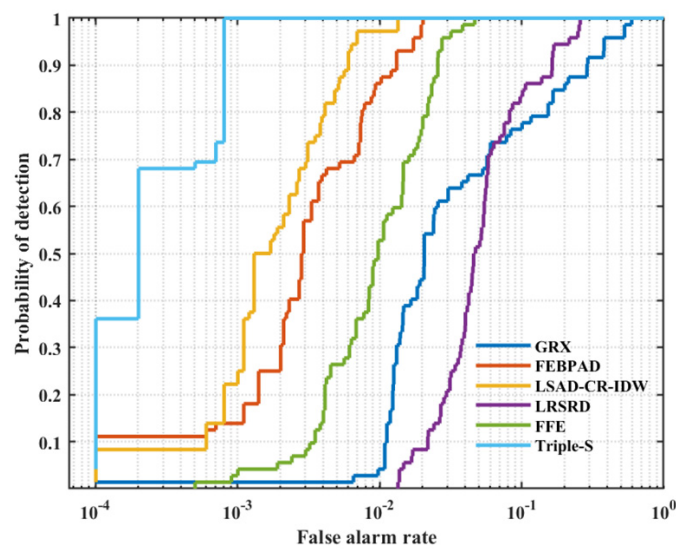
Figure 7. Receiver operating characteristic (ROC) curves for AVIRIS San Diego airport dataset.
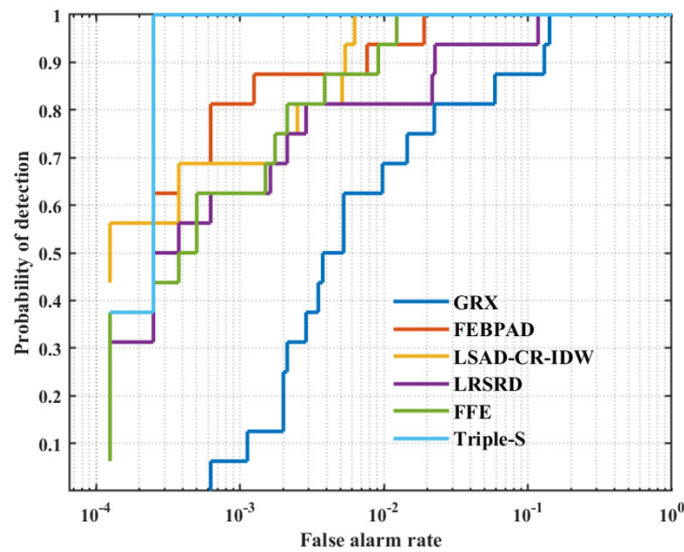
Figure 8. Receiver operating characteristic (ROC) curves for HYDICE Urban dataset.

The corresponding AUC values are summarized in Table 1. As an important aspect to evaluate the algorithm efficiency, the running times of each algorithm are listed in Table 2. Based on these detection results and the performance of the algorithms for comparison, several related assessments were carried out both qualitatively and quantitatively to evaluate the effectiveness of our Triple-S anomaly detection algorithm, which are described in the discussion section in detail.

Table 1. AUC scores of different methods.

Dataset	GRX	FFE	FEBPAD	LSAD-CR-IDW	LRSRD	Triple-S
AVIRIS San Diego Airport	0.9139	0.9873	0.9951	0.9974	0.9873	**0.9997**
HYDICE Urban	0.9745	0.9979	0.9981	0.9986	0.9883	**0.9998**

Table 2. Computing time of different methods.

| Dataset	GRX	FFE	FEBPAD	LSAD-CR-IDW	LRSRD	Triple-S	Unit
AVIRIS San Diego Airport	0.11	30.98	3.671	2236.68	73.19	15.91	second
HYDICE Urban	0.09	12.61	2.696	15.6138	65.27	13.92	second

4. Discussion

In this section, we first analyze the parameter setting of our proposed Triple-S algorithm. In order to achieve a fair comparison, the related parameter setting of the algorithms for comparison are also introduced according to the corresponding references. Next, we specifically evaluate the experimental results mentioned above qualitatively based on the comparison of the results images shown in Figures 5a–f and 6a–f against the ground-truth images in Figures 2b and 3b, respectively. Quantitative analysis is also applied to validate the accuracy of the Triple-S anomaly detector, including the ROC curve shown in Figures 7 and 8, the corresponding AUC values shown in Table 1 and running time shown in Table 2.

4.1. Parameter Setting of Triple-S

We investigate and demonstrate the influence of parameter settings on the results of initial oversegmentation and region merging in the proposed spectra-based selective search

method. The proposed spectra-based selective searching algorithm has two important parameters, namely K and T.

First, we randomly selected K to be 0.02, 0.2, 2, 20, 200, and 2000 for the two hyperspectral datasets, and the corresponding results of oversegmentation are shown in Figure 9a–l. As shown in Figure 9a,g, if the value of K is too small, such as $K = 0.02$, the oversegmentation results are too trivial to identify the category information of the features in the image. When K is set in an appropriate range, the hyperspectral image can be divided into different regions with a data-driven shape and size according to the shape and category of the surface features. When K gradually increases and is greater than the appropriate range, such as $K = 20$, most of the pixels in the hyperspectral image are classified as background. However, the background pixels around the anomaly regions are regarded as anomalies, and some anomaly pixels will be regarded as background. As K becomes larger, most of the anomaly pixels are submerged into the background. It is difficult to distinguish abnormal targets from the background in this case. For a better initial oversegmentation effect, the constant parameter K was set to 2 for the AVIRIS San Diego airport dataset and 1 for the HYDICE urban dataset in our experiment.
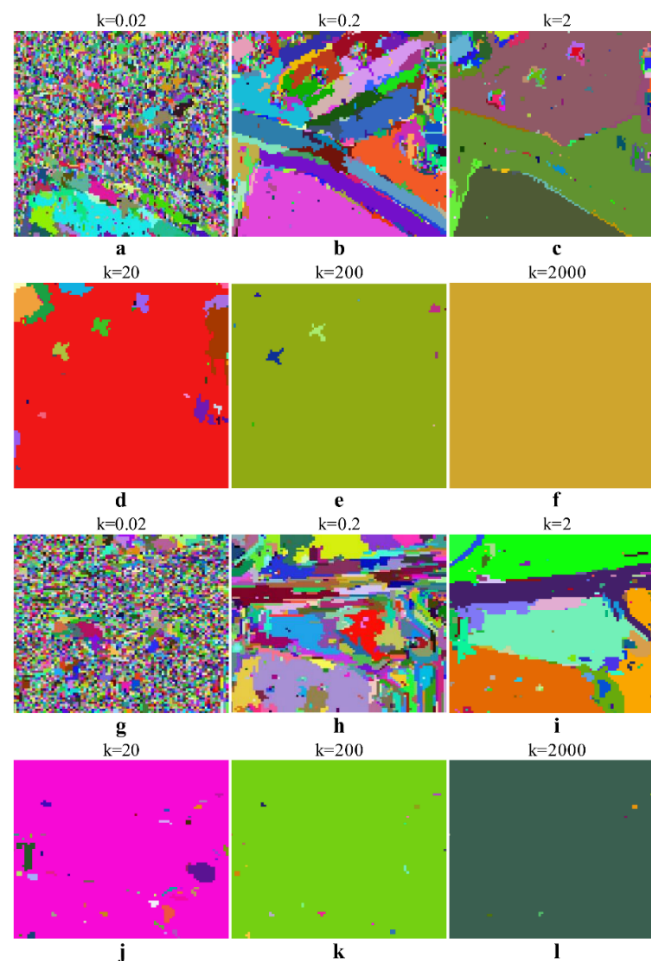
Figure 9. Different oversegmentation results for the two real-world HSIs with K = 0.02, 0.2, 2, 20, 200, and 2000; (**a–f**) correspond to the AVIRIS San Diego airport data and (**g–l**) correspond to the HYDICE urban data.

Next, we get threshold T of region merging through the spectral similarity of a series of adjacent regions obtained by the step of oversegmentation. The value range of the parameter T is dependent on the similarity matrix of each pair of adjacent regions. The minimum value of the similarity matrix is the lower limit of the threshold T, while the maximum value of the similarity matrix is the upper limit of the threshold T. We traversed

the value of T within this range in steps of 0.05. Choosing 0.05 as the step not only ensures that a reasonable threshold T can be obtained, but also reduces the complexity of the traversal process. The optimal thresholds T are determined by the best detection result and the highest AUC value. The best detection performance of the proposed Triple-S anomaly detector could be achieved when $T = 1.15$ for the AVIRIS San Diego airport dataset and $T = 1.4$ for the HYDICE urban dataset.

4.2. Parameter Setting of the Comparison Algorithms

The experiments of other comparison methods were carried out by the corresponding source code on GitHub and default parameters given by the authors. See the supplementary materials for the details. Among the five algorithms for comparison, LSAD-CR-IDW and LRSRD needed to preset parameters based on the data sets. Therefore, we chose the optimal parameters of the comparison algorithms LSAD-CR-IDW and LRSRD for the two test images used in the experiment in order to ensure fair comparison. Firstly, for the local algorithms, such as LSAD-CR-IDW, the sizes of the dual-window depended on the size of the anomaly in the image of each dataset. According to the principle of CRD involved in LSAD-CR-IDW, the inner window should contain all abnormal pixels near the under-tested pixel, and the size of the outer window should be larger than the size of the inner window. At the same time, it should ensure that there was no anomaly in the remaining pixel set after the outer window minus the inner window. Improper selection of dual-window sizes would cause the background set to be polluted by abnormal pixels in the calculation process, thus decreasing the detection performance. The anomalies were three aircraft which were composed of 23–25 pixels for the AVIRIS San Diego airport hyperspectral dataset, and the anomaly for the HYDICE urban hyperspectral dataset were the vehicles embedded in different backgrounds which contained 1, 2, 2, 2, 2, 3 and 4 pixels. Therefore, according to the size of the anomaly in the above two datasets and the selection principle of the inner and outer window sizes, we determined the inner and outer window sizes of the two datasets in a reasonable range. For the San Diego dataset, the size of the outer window ranged from 7 to 19, and the inner window size ranged from 5 to 13. For the urban dataset, the outer window size ranged from 5 to 17, and the inner window size ranged from 3 to 11. In our comparison experiment, the detection result graph and AUC value were used to judge anomaly detection performance. In order to compare the detection performance of the algorithms fairly, the optimal window sizes for the local anomaly detector LSAD-CR-IDW corresponded to the best detection result and the highest AUC value. Furthermore, regularization parameter λ of the LSAD-CR-IDW detector was set as 100 in our comparison experiments, as suggested in [40]. For LRSRD, the number of clusters K and the number of pixels P picked up for background dictionary construction in each cluster were two important parameters that affected the performance of anomaly detection. According to the original literature [38], K and P were fixed to 15 and 20, respectively.

4.3. Qualitative Analysis

For the AVIRIS San Diego airport hyperspectral dataset, the detection results of the proposed Triple-S algorithm as well as the methods for comparison are shown in Figure 5a–f. From the detection result image Figure 5a, the detection performance of the GRX algorithm was not satisfactory. The position and shape of the aircraft as the anomaly of the dataset were difficult to identify compared with the ground-truth image as shown in Figure 2b. At the same time, several background pixels were misjudged as anomaly pixels in the red, green and yellow boxes as shown in Figure 5a. In addition, the overall background of the image was not well suppressed. The detection result image of the FFE algorithm Figure 5b looked much better than the result image of the GRX algorithm. The locations and shapes of the three aircraft were been detected but it was still not clear enough. Similar to the GRX algorithm, a number of background pixels in the red, green and yellow boxes were misjudged as anomalies. The overall background of the image was also not well suppressed. In the detection result image of the FEBPAD algorithm, the center

position of the three aircraft could be clearly detected, but the edge pixels of the aircraft could not be detected. Meanwhile, a few background pixels in the position of the red region were misjudged as anomalies according to the ground-truth image. But as the name suggests, background purification, this method suppressed background pixels very well among the five algorithms for comparison. In the anomaly detection result of the LSAD-CR-IDW algorithm, not only the locations but also the shapes of three aircraft were detected clearly. Most of the abnormal pixels were correctly detected. However, the disadvantage was that the algorithm detected the spatial distribution of the entire image at the same time. A large number of background pixels had a higher probability of being abnormal pixels, and the false alarm rate was very high. The detection performance of LRSRD was also not satisfactory, because the locations and shapes of the three aircraft were also hardly detected and at the same time many background pixels in the red, green and yellow boxes were misjudged as anomalies. This could be attributed to the inaccurate and incomplete background dictionary. Compared with the other five algorithms, the detection result image of the proposed Triple-S algorithm looked better, in which not only the position and shape of the three aircraft were clearly detected, but also the background pixels were completely suppressed. Owing to the inherent peculiarity of the Triple-S method, it can make an unambiguous distinction between anomalous targets and background. Therefore, the three abnormal aircraft pixels were markedly distinct from the background pixels of the result image, which was similar to the ground-truth map.

For the HYDICE urban hyperspectral dataset, the detection results of the proposed Triple-S algorithm as well as the methods for comparison are shown in Figure 6a–f. In the detection result image of the GRX algorithm shown in Figure 6a, most of the locations of anomaly could be detected but not clearly enough compared with the ground-truth image as shown in Figure 3b. Meanwhile, a number of background pixels were misjudged as anomalies in the red and green boxes. Therefore, the suppression of overall background in this image was not satisfactory. From Figure 6b, the detection result of FFE looked better than the result of the GRX algorithm. Most of the abnormal pixels could be clearly detected. However, similar to the GRX algorithm, there were still a large number of background pixels with high anomaly scores in the red box according to the ground-truth image. The overall background of the image was also not well suppressed. Different from the above two algorithms, FEBPAD had a better effect of background suppression. At the same time, most of the abnormal pixels were successfully detected and clearly displayed, leaving a few pixels that should have been judged to be abnormal in the background. In the anomaly detection result of the LSAD-CR-IDW algorithm, we see that most of the anomalies could be clearly detected. However, a mass of background pixels was misjudged as anomalies, such as those pixels that were framed in red box. Although the LRSRD algorithm could detect the locations of anomalies, many background pixels were misjudged as anomalies at the same time. Moreover, the background pixels that were misjudged as abnormal had comparable scores to the real abnormal pixels, making the extraction of abnormal pixels difficult. From the overall detection result in Figure 6e, the false detection rate of this algorithm was also the highest. Compared with the other five algorithms, the detection result image of the proposed Triple-S algorithm looked best. Almost all abnormal pixels had been successfully detected, and only two background pixels were misjudged as abnormal pixels as shown in the red box of Figure 6f. It could be said that the anomaly pixels and background pixels were completely separated with maximum distinction.

4.4. Quantitative Analysis

In order to distinguish the ROC curves of these six algorithms for easier analysis, when drawing the ROC curve, we performed a logarithmic operation on the horizontal axis (FPR). Since the horizontal axis of the ROC curve was FPR and the vertical axis was TPR, it could be seen that when the ROC curve got a larger TPR under a smaller FPR, that is,

when it approached the upper left corner of the coordinate space earlier, the algorithm had better anomaly detection performance.

For the AVIRIS San Diego airport hyperspectral dataset, Figure 7 shows the ROC curves of different anomaly detectors for the quantitative comparison of detection performance. As we can see from Figure 7, the proposed Triple-S method was superior to the other detection methods for this dataset. When the false-alarm rate changed from 0.0001 to 1, the detection probability of the Triple-S method was always higher than those of other methods. The reason was that the proposed Triple-S algorithm had a purer background and the separation between the background and anomaly was maximal, which improved detection accuracy, compared with those algorithms for comparison. When the false alarm rate was low, the probability of detection was much higher than other methods, which corresponded to the relatively sharp rise of the detection probability curve of the Triple-S method. The ROC curve of the proposed Triple-S detection method was closest to the upper-left corner than those of the other algorithms. Moreover, the corresponding AUC values of these methods were calculated and reported in the first row of Table 1. Except for GRX, the other five algorithms had an AUC value close to 1. In the GRX, all of the pixels in the entire image were used as the background to calculate the background statistics, so the background statistics were seriously polluted by noise and abnormalities, resulting in poorer detection performance. The Triple-S algorithm we proposed had the highest AUC value. The AUC score of the Triple-S method was 0.9997, due to the full use of spatial and spectral information making the distinction between background and anomalies more accurate and complete.

Figure 8 shows the ROC curve as a quantitative comparison of different anomaly detectors for the HYDICE Urban hyperspectral dataset. Due to the number of anomaly pixels in the Urban dataset being smaller than that of the San Diego airport dataset, the steps in the shape of the ROC curve were more obvious. All of the six ROC reached the maximum value faster. Among them, the ROC curve of the proposed Triple-S detection method was closest to the top-left corner, indicating superior detection performance. All of this proves the effectiveness and superiority of our proposed algorithm in the detection of hyperspectral anomalies. Moreover, the corresponding AUC values of these methods are calculated and reported in the second row of Table 1. The AUC score of the Triple-S method was 0.9998, which was larger than those of the other detectors. The reason was that the Triple-S algorithm not only effectively detected abnormal pixels, but also successfully suppressed background pixels. This can be attributed to the regions composed of adjacent pixels with the same spectral information, which effectively combined the advantages of hyperspectral information with image spatial information.

The runtime performances of all the aforementioned methods were also recorded and compared, as presented in Table 2. The running time of the GRX was shortest due to the simple algorithm process. The calculation time of FEBPAD was shorter than that of FFE, and the reason was that the calculation order of FrFT was directly given in FEBPAD while there was a process of finding the best calculation order in FFE. The running time of LSAD-CR-IDW algorithm was the longest since the calculation process was the most complicated involving LSAD and CRD. The computing time of our proposed Triple-S algorithm depended on the size of the test image. The larger the size of the test image, the longer the corresponding calculation time.

5. Conclusions

In this study, an innovative hyperspectral anomaly detection method, called the spectra-based selective searching (Triple-S) algorithm, utilizing image oversegmentation and merging, is proposed, fully combining and utilizing the spatial and spectral information of hyperspectral images. The adaptive regions composed of adjacent pixels with the same spectral characteristics are referred to as the processing unit rather than a single pixel. In addition to the difference between the anomaly and background, the spectral similarities between the anomaly pixels and between the background pixels are both exploited to

achieve efficient and high-performance anomaly detection. Graph-based image segmentation and the greedy algorithm are adapted to hyperspectral images to selectively search for abnormal targets in the graph without exhaustive searching by making full use of spatial and spectral information. The Triple-S algorithm can effectively and quickly detect anomaly targets of freewill shape and size as well as completely suppress the background, without making assumptions about the background distribution or strenuously searching for a suitable background dictionary. The experimental results reported in the manuscript show that the Triple-S method outperforms other state-of-the-art anomaly detection methods no matter whether from qualitative or quantitative comparison. From the detection results, the Triple-S algorithm can accurately extract the anomaly from the background, and the anomaly and the background have a satisfactory degree of separation. From the ROC curve and AUC value, the Triple-S algorithm showed the highest detection accuracy compared with other algorithms. Therefore, it is highly efficient and competitive in terms of detection accuracy as well as separation between background and anomaly. However, it should be noted that the need to artificially set the initial oversegmentation degree parameter K and the regional merging threshold T parameter in advance are the main limitations of the proposed spectra-based selective searching algorithm for practical applications. Improper values of K and T parameters result in the failure of the proposed algorithm. The focus of our future work is to improve the parameter setting method so that it can automatically select the best parameters according to the HSI data itself. Furthermore, the detection performance of the proposed algorithm is not satisfactory when the hyperspectral image dataset to be detected has a distinct hierarchy in the spatial distribution. The initial improvement plan for the Triple-S algorithm is to add the overall judgment of the image spatial structure and perform corresponding processing according to different spatial characteristics, making it applicable to the datasets with distinct hierarchical spatial information.

Abbreviations

HSI	Hyperspectral image
RX	Reed–Xiaoli
GRX	Global Reed–Xiaoli
LRX	Local Reed–Xiaoli
KRX	Kernel Reed–Xiaoli
CBAD	Cluster-based anomaly detection
FrFT	Fractional Fourier transform
FFE	Fractional Fourier entropy
LRaSMD	low rank and sparse matrix decomposition
FEBPAD	feature extraction and Background Purification anomaly detector
LRASR	low-rank and sparse representation
CRD	Collaborative representation-based detector
LSAD-CR-IDW	local summation anomaly detection based on collaborative representation and inverse distance weight
Triple-S	Spectra-based selective searching
ED	Euclidean distance
LD	Lance distance
SAM	Spectral angle measure
r	Correlation coefficient
Ent	Information entropy of the spectral difference curve
IDC	Internal difference of a component
ROC	Receiver operating characteristic curve
AUC	The area under the ROC curve value
TPR	True positive rate
FPR	False positive rate
TP	True positive
TN	True negative
FP	False positive
FN	False negative
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
HYDICE	Hyperspectral Digital Imagery Collection Experiment

References

1. Borengasser, M.; Hungate, W.S.; Watkins, R. *Hyperspectral Remote Sensing—Principles and Applications*; CRC Press: Boca Raton, FL, USA, 2008.
2. Plazaa, A.; Benediktsson, J.A.; Boardman, J.W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* **2009**, *113*, S110–S122. [CrossRef]
3. Bioucas-Dias, J.M.; Plaza, A.; Camps-Valls, G.; Scheunders, P.; Nasrabadi, N.M.; Chanussot, J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geosci. Remote Sens. Mag.* **2013**, *1*, 6–36. [CrossRef]
4. Rasti, B.; Scheunders, P.; Ghamisi, P.; Licciardi, G.; Chanussot, J. Noise reduction in hyperspectral imagery: Overview and application. *Remote Sens.* **2018**, *10*, 482. [CrossRef]
5. Nasrabadi, N.M. Hyperspectral target detection: An overview of current and future challenges. *IEEE Signal Process. Mag.* **2014**, *31*, 34–44. [CrossRef]
6. Bitar, A.W.; Cheong, L.F.; Ovarlez, J.P. Sparse and low-rank matrix decomposition for automatic target detection in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5239–5251. [CrossRef]
7. Zhang, Y.; Du, B.; Zhang, L.; Liu, T. Joint sparse representation and multitask learning for hyperspectral target detection. *IEEE Trans. Geosci. Remote Sens.* **2016**, *55*, 894–906. [CrossRef]
8. Manolakis, D.; Shaw, G. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Process. Mag.* **2002**, *19*, 29–43. [CrossRef]
9. Lorenz, S.; Salehi, S.; Kirsch, M.; Zimmermann, R.; Unger, G.; Vest Sørensen, E.; Gloaguen, R. Radiometric correction and 3D integration of long-range ground-based hyperspectral imagery for mineral exploration of vertical outcrops. *Remote Sens.* **2018**, *10*, 176. [CrossRef]
10. Zhang, L.; Wei, W.; Tian, C.; Li, F.; Zhang, Y. Exploring structured sparsity by a reweighted Laplace prior for hyperspectral compressive sensing. *IEEE Trans. Image Process.* **2016**, *25*, 4974–4988. [CrossRef]
11. Huyan, N.; Zhang, X.; Zhou, H.; Jiao, L. Hyperspectral anomaly detection via background and potential anomaly dictionaries construction. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2263–2276. [CrossRef]

12. Matteoli, S.; Diani, M.; Corsini, G. A tutorial overview of anomaly detection in hyperspectral images. *IEEE Aerosp. Electron. Syst. Mag.* **2010**, *25*, 5–28. [CrossRef]
13. Liu, J.; Feng, Y.; Liu, W.; Orlando, D.; Li, H. Training data assisted anomaly detection of multi-pixel targets in hyperspectral imagery. *IEEE Trans. Signal Process.* **2020**, *68*, 3022–3032. [CrossRef]
14. Qu, Y.; Wang, W.; Guo, R.; Ayhan, B.; Kwan, C.; Vance, S.; Qi, H. Hyperspectral anomaly detection through spectral unmixing and dictionary-based low-rank decomposition. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4391–4405. [CrossRef]
15. Xie, W.; Jiang, T.; Li, Y.; Jia, X.; Lei, J. Structure tensor and guided filtering-based algorithm for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 4218–4230. [CrossRef]
16. Yuan, Y.; Ma, D.; Wang, Q. Hyperspectral anomaly detection by graph pixel selection. *IEEE Trans. Cybern.* **2016**, *46*, 3123–3134. [CrossRef]
17. Taghipour, A.; Ghassemian, H. Hyperspectral anomaly detection using attribute profiles. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1136–1140. [CrossRef]
18. Reed, I.S.; Yu, X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1760–1770. [CrossRef]
19. Molero, J.M.; Garzón, E.M.; García, I.; Plaza, A. Analysis and optimizations of global and local versions of the RX algorithm for anomaly detection in hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2013**, *6*, 801–814. [CrossRef]
20. Liu, W.; Chang, C.I. Multiple-window anomaly detection for hyperspectral imagery. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2013**, *6*, 644–658. [CrossRef]
21. Zhao, L.; Lin, W.; Wang, Y.; Li, X. Recursive local summation of RX detection for hyperspectral image using sliding windows. *Remote Sens.* **2018**, *10*, 103. [CrossRef]
22. Matteoli, S.; Veracini, T.; Diani, M.; Corini, G. A locally adaptive background density estimator: An evolution for RX-based anomaly detector. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 323–327. [CrossRef]
23. Guo, Q.; Zhang, B.; Ran, Q.; Gao, L.; Li, J.; Plaza, A. Weighted-RXD and linear filter-based RXD: Improving background statistics estimation for anomaly detection in hyperspectral imagery. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 2351–2366. [CrossRef]
24. Matteoli, S.; Diani, M.; Corsini, G. Improved estimation of local background covariance matrix for anomaly detection in hyperspectral images. *Opt. Eng.* **2010**, *49*, 1–16. [CrossRef]
25. Nasrabadi, N.M. Regularization for spectral matched filter and RX anomaly detector. *Proc. SPIE* **2008**, *6966*, 696604-1–696604-12. [CrossRef]
26. Kwon, H.; Nasrabadi, N.M. Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 388–397. [CrossRef]
27. Carlotto, M.J. A cluster-based approach for detecting man-made objects and changes in imagery. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 374–387. [CrossRef]
28. Wang, Q.; Guo, Q.; Zhou, J.; Lin, Q. Nonlinear joint fractional Fourier transform correlation for target detection in hyperspectral image. *Opt. Laser Technol.* **2012**, *44*, 1897–1904. [CrossRef]
29. Tao, R.; Zhao, X.; Li, W.; Li, H.; Du, Q. Hyperspectral anomaly detection by fractional Fourier entropy. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2019**, *12*, 4920–4929. [CrossRef]
30. Ma, Y.; Fan, G.; Jin, Q.; Huang, J.; Mei, X.; Ma, J. Hyperspectral Anomaly Detection via Integration of Feature Extraction and Background Purification. *IEEE Geosci. Remote Sens. Lett.* **2020**, 1–5. [CrossRef]
31. Du, B.; Zhang, L. A discriminative metric learning based anomaly detection method. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6844–6857. [CrossRef]
32. Du, B.; Zhang, L. Random-selection-based anomaly detector for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1578–1589. [CrossRef]
33. Zhao, R.; Du, B.; Zhang, L. A robust nonlinear hyperspectral anomaly detection approach. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 1227–1234. [CrossRef]
34. Ranney, K.I.; Soumekh, M. Hyperspectral anomaly detection within the signal subspace. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 312–316. [CrossRef]
35. Vafadar, M.; Ghassemian, H. Anomaly detection of hyperspectral imagery using modified collaborative representation. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 577–581. [CrossRef]
36. Chen, Y.; Nasrabadi, N.M.; Tran, T.D. Sparse representation for target detection in hyperspectral imagery. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 629–640. [CrossRef]
37. Li, J.; Zhang, H.; Zhang, L.; Ma, L. Hyperspectral anomaly detection by the use of background joint sparse representation. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2015**, *8*, 2523–2533. [CrossRef]
38. Xu, Y.; Wu, Z.; Li, J.; Plaza, A.; Wei, Z. Anomaly detection in hyperspectral images based on low-rank and sparse representation. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1990–2000. [CrossRef]
39. Wei, L.; Qian, D. Collaborative representation for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1463–1474. [CrossRef]
40. Tan, K.; Hou, Z.; Wu, F.; Du, Q.; Chen, Y. Anomaly detection for hyperspectral imagery based on the regularized subspace method and collaborative representation. *Remote Sens.* **2019**, *11*, 1318. [CrossRef]

41. Felzenszwalb, P.; Huttenlocher, D. Efficient Graph-Based Image Segmentation. *Int. J. Comput. Vis.* **2004**, *59*, 167–181. [CrossRef]

42. Lance, G.N.; Williams, W.T. Computer programs for hierarchical polythetic classification ("similarity analyses"). *Comput. J.* **1966**, *9*, 60–64. [CrossRef]

43. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]

44. Sweet, J.N. The spectral similarity scale and its application to the classification of hyperspectral remote sensing data. In Proceedings of the IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, Greenbelt, MD, USA, 27–28 October 2003; pp. 92–99. [CrossRef]

45. Van de Sande, K.E.A.; Uijlings, J.R.R.; Gevers, T.; Smeulders, A.W. Segmentation as selective search for object recognition. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 1879–1886. [CrossRef]

46. Uijlings, J.R.; Van De Sande, K.E.; Gevers, T.; Smeulders, A.W. Selective search for object recognition. *Int. J. Comput. Vis.* **2013**, *104*, 154–171. [CrossRef]

47. Kerekes, J. Receiver operating characteristic curve confidence intervals and regions. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 251–255. [CrossRef]

48. Ferri, C.; Hernández-Orallo, J.; Flach, P. A coherent interpretation of AUC as a measure of aggregated classification performance. In Proceedings of the 28th International Conference on Machine Learning, ICML 2011, Bellevue, WA, USA, 28 June–2 July 2011; pp. 657–664.