Hyperspectral Anomaly Detection Using Dual Window Density

Bing Tu, *Member, IEEE*, Xianchang Yang, *Student Member, IEEE*,
Chengle Zhou, *Graduate Student Member, IEEE*, Danbing He, *Student Member, IEEE*,
and Antonio Plaza, *Fellow, IEEE*

Abstract— Hyperspectral anomaly detection is one of the most active topics in hyperspectral image (HSI) analysis. The fine spectral information of HSIs allows us to uncover anomalies with very high accuracy. Recently, an intrinsic image decomposition (IID) model has been introduced for low-rank IID (LRIID) in multispectral images. Inspired by the LRIID, which is able to effectively recover the reflectance and shading components of the multispectral image, this article adapts the LRIID for obtaining the reflectance component of HSIs (which is the key feature for the discrimination of different objects). In order to exploit the reflectance component, we also propose a new dual window density (DWD)-based detector for anomaly detection, which is based on the idea that anomalies are usually rare pixels and, thus, exhibit low density in the image. The density analysis of DWD is intended not only to circumvent the Gaussian assumption regarding the distribution of HSI data, but also to mitigate the contamination of background statistics caused by anomalies. The dual window operation of our DWD is specifically designed to adaptively calculate the density of each pixel under test, so as to identify anomalies with nonspecific sizes. Our experimental results, obtained on a database of real HSIs including Airport, Beach, and Urban scenes, demonstrate the superiority of the proposed method in terms of detection performance when compared to other widely used anomaly detection methods.

Index Terms— Anomaly detection, density, dual window, hyperspectral image (HSI), intrinsic image decomposition (IID).

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) provide rich spectral information about the materials on the ground surface, reflecting diagnostic features that can be used to distinguish between different substances. This enables the fact that ground objects, that are undetectable in wideband remote sensing images, can be uncovered in HSIs [1]. Owing to this advantage, HSI anomaly detection has become a very active research topic, aiming at identifying anomalous objects (which exhibit distinct spectral characteristics from background and with low probability of occurrence in the images) [2]–[6]. Since the spectral information of the target is not known *a priori* in HSI anomaly detection, it is generally very challenging to detect such anomalies.

The Reed–Xiaoli (RX) detector [7] is one of the most typical HSI anomaly detection methods. It estimates the mean and covariance matrix of the global image, and then uses the Mahalanobis distance from the mean as the detector output. However, there are often some problems involved with the direct use of RX: 1) using all the pixels to compute the background statistics (the mean and the covariance matrix) often leads to contamination, since the anomalous pixels (which are greatly different from background) are also used in the calculation and, consequently, the background statistics are not completely accurate; 2) the RX assumes that the background follows a Gaussian distribution, which may not be realistic in HSI data [8]; and 3) the detection of anomalies may be complicated [9]. As a result, several works have attempted to improve the original RX version. Representative techniques include the kernel-RX [10], [11], iterative-RX [12], segment-RX [13], subspace-RX [14], regularized-RX [15], topology based-RX [16], weighted-RX [17], and linear filter based-RX [17], among others.

In addition, several local window-based methods have been developed to address the aforementioned problems. In fact, the local RX (LRX) is the most typical window-based approach, in which the RX is conducted in a local window around the pixel under test. The detection performance of LRX is usually better than that of RX, since the background pixels in local windows tend to be more compliant with the Gaussian distribution. However, Since the LRX uses only a single window, an unsatisfactory local distribution may be obtained when the target size is large, and the contamination of the background statistics may become even larger with the

increase of the number of anomalies in the local background. Several methods have been proposed to mitigate this issue. For example, Kwon *et al.* [18] developed a dual window-based eigenseparation transform (DWEST), which implemented the RX using inner and outer windows. Liu and Chang [19] developed a nested spatial window-based target detection (NSWTD) method, in which different spatial windows are nested to detect anomalies. In [9], a multiple-window anomaly detection (MWAD) is designed by using a variety of windows to detect anomalies with different sizes. Bo *et al.* [20] proposed a local summation anomaly detection (LSAD) method, which adopts a local summation strategy to combine multiple local distributions from neighboring local windows. Moreover, the collaborative representation-based detector (CRD) [21] and the background joint sparse representation (BJSR)-based detector [22] can also be considered as window-based anomaly detectors. The CRD assumes that the background pixels can be represented by their neighboring pixels in a local window (whereas the anomalies cannot), and the BJSR can adaptively select the most representative local background bases. More recently, [23] uses an adaptive neighborhood set for the RX detector (RX-ASSN), in which only spectrally similar neighborhood pixels are considered.

Opposite to the aforementioned methods, other approaches are based on adaptively segmenting the original image. For instance, the cluster-based anomaly detector (CBAD) [24] detects anomalies in different spectrally homogeneous clusters rather than in fixed square windows, and the binary partition tree (BPT)-based detector can achieve more homogeneous background regions [25]. In addition, many low-rank and sparse-based hyperspectral anomaly detection methods have been proposed and show their excellent performance [26]–[29]. In fact, whether the image is processed using segmentation or locally selecting a neighborhood set, highly discriminative features of different objects are required. However, these methods ignore the fact that the spectral signatures are often mixed due to the complex image acquisition conditions and the geometrical structure of natural scenes, leading to inaccurate segmentation results or unsatisfactory neighborhood sets.

In order to address the aforementioned issues, in this article we introduce (for the first time in the literature) the intrinsic image decomposition (IID) into HSI anomaly detection. IID is a well-known problem in computer vision [30]–[34]. Its purpose is to extract the reflectance and shading components from a single image. Recently, an optimization-based IID method [35] has been employed for feature extraction-based HSI classification, obtaining good results [36]. This indicates that HSIs comprise two main components: 1) the spectral reflectance component, which is determined by the material and 2) the shading component, which is not only caused by the illumination conditions and the scene geometry, but also by the reflected solar radiation and thermal radiation of ground objects. More recently, a low-rank IID (LRIID) model [37] has been proposed, which extends the Retinex theory [38] to the multispectral IID domain. Inspired by this approach, and assuming that the Retinex theory also holds on the HSI domain, in this article we adapt the LRIID model to be suitable

for recovering the reflectance and shading components from HSIs. In our context, the reflectance component is used for HSI anomaly detection and the shading component is removed.

In order to exploit the reflectance component, we propose a new dual window density (DWD)-based anomaly detector that is based on the density peak clustering (DPC) algorithm [39]. Recently, many DPC (density measurement)-based hyperspectral processing algorithms have been proposed [40]–[44]. In [44], the DPC algorithm is applied to detect noisy labels for hyperspectral classification. The density measurement-based hyperspectral anomaly detection methods have also been developed. Tu *et al.* [45] provided a DPC-based detector using single window and [46] detects anomalies by the spatial density background purification. In this article, the use of the reflectance component (with pure and highly discriminative spectral features) brings significant benefits to HSI anomaly detection (in general) and to our newly developed DWD (in particular). Specifically, our DWD performs adaptive density integration by using inner and outer windows to detect anomalies. Besides that, the advantages of the DWD are mainly reflected in the following aspects.

1) Since anomalous objects are rarely present and, consequently, can be assumed to exhibit low densities, detecting anomalies using the concept of pixel density is feasible. This concept does not require the calculation of the mean and covariance matrices of the background statistics, avoiding the potential contamination caused by anomalies and allowing for real-time implementations.

2) As opposed to many detection methods (that assume that the background follows a Gaussian distribution), density analysis does not require data distribution assumptions but only analyze the density attribute of the data itself. That is, density analysis can extract the data's structure with non-Gaussian distribution [47], which may be more suitable for HSIs.

3) By adaptively integrating the density in the inner and outer windows, the dual window operation of our newly developed DWD is capable of recognizing the size of anomalies, and detecting anomalies with different sizes.

The remainder of this article is organized as follows. Section II describes some related works. Section III describes the proposed HSI anomaly detection method in detail. Section IV provides the experimental setup and presents our extensive experimental results. Section V concludes this article with some remarks and hints at plausible future research lines.

II. RELATED WORK

In this section, we briefly revisit some related works, including the RX anomaly detector, the LRIID model, and the DPC algorithm.

A. RX Detector

The RX is one of the most typical hyperspectral anomaly detection methods which is defined as

$$RX(\mathbf{x}) = (\mathbf{x} - \mu)^T \Gamma^{-1} (\mathbf{x} - \mu) \tag{1}$$

where $\mathbf{x}$ is a test pixel, μ and Γ are the mean and covariance matrix of the global image, respectively. $RX(\mathbf{x})$ is the output of RX, which is the Mahalanobis distance, equivalent to the standard deviation away from the mean.

B. LRIID

Let a HSI be denoted as $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_I] \in \mathbb{R}^{I \times D}$, where I and D are, respectively, the number of pixels and the number of spectral bands. In the IID problem, each pixel $\mathbf{x}_i$ can be modeled as the product of a reflectance component $\mathbf{r}_i$ and a shading component $\mathbf{s}_i$ as follows:

$$\mathbf{x}_i = \mathbf{r}_i. * \mathbf{s}_i. \tag{2}$$

The LRIID extends the Retinex theory to the multispectral IID domain as follows.

1) When the reflectance difference between two pixels is significant, the Retinex assumes that the difference between such pixels is mainly caused by the *reflectance* component (the shading tends to be consistent). This leads to $\mathbf{s}_i = \mathbf{s}_j$, which can also be written as $\mathbf{x}_i. * \mathbf{r}_j = \mathbf{x}_j. * \mathbf{r}_i$. We can define an energy function E_{sc} as follows:

$$E_{\text{sc}} = \sum_{i,j \in \mathbf{N}} ||w_{i,j}(\mathbf{x}_i. * \mathbf{r}_j - \mathbf{x}_j. * \mathbf{r}_i)||^2 \tag{3}$$

where $\mathbf{N}$ is a neighborhood set (in which only horizontally adjacent and vertically adjacent pixels are considered) and $w_{i,j}$ contains the weights that reflect the similarity between the two pixels.

2) When the reflectance difference between two pixels is small, the Retinex assumes that the difference between the two pixels is mainly caused by the *shading* component (the reflectance tends to be consistent). This leads to $\mathbf{r}_i = \mathbf{r}_j$ and the energy function E_{rc} can be defined as follows:

$$E_{\text{rc}} = \sum_{i,j \in \mathbf{N}} ||v_{i,j}(\mathbf{r}_i - \mathbf{r}_j)||^2 \tag{4}$$

where $v_{i,j}$ contains the weights that reflect the similarity between the two pixels. The LRIID introduces a low-rank basis for the reflectance, denoted as matrix $\mathbf{B}_r$, and the reflectance vectors can be linearly represented as $\mathbf{r}_i = \mathbf{B}_r \tilde{\mathbf{r}}_i$. The energy functions in (3) and (4) can be now separately written as

$$E_{\text{sc}} = \sum_{i,j \in \mathbf{N}} ||w_{i,j}(\mathbf{L}_i \mathbf{B}_r \tilde{\mathbf{r}}_j - \mathbf{L}_j \mathbf{B}_r \tilde{\mathbf{r}}_i)||^2$$
$$E_{\text{rc}} = \sum_{i,j \in \mathbf{N}} ||v_{i,j}(\mathbf{B}_r \tilde{\mathbf{r}}_i - \mathbf{B}_r \tilde{\mathbf{r}}_j)||^2 \tag{5}$$

where $\mathbf{L}_i$ is a diagonal matrix composed of spectral elements of $\mathbf{x}_i$. In order to avoid the ambiguity of the IID problem, LRIID also provides another solution with a generic constraint on the coefficient sum as $\mathbf{M}\tilde{\mathbf{R}} = \mathbf{C}$, where $\tilde{\mathbf{R}}$ consists of $\tilde{\mathbf{r}}$ for all pixels. The energy functions can now be combined (in matrix form) as follows:

$$E_r = ||\mathbf{W}_{\mathbf{L},\mathbf{B}_r}\tilde{\mathbf{R}}||^2 + \lambda_1||\mathbf{V}_{\mathbf{B}_r}\tilde{\mathbf{R}}||^2 + \lambda_2||\mathbf{M}_r\tilde{\mathbf{R}} - \mathbf{C}||^2 \tag{6}$$

where $\mathbf{W}_{\mathbf{L},\mathbf{B}_r}$ are jointly determined by the neighborhood set $\mathbf{N}$, the weight $w_{i,j}$, the reflectance basis $\mathbf{B}_r$, and the spectral
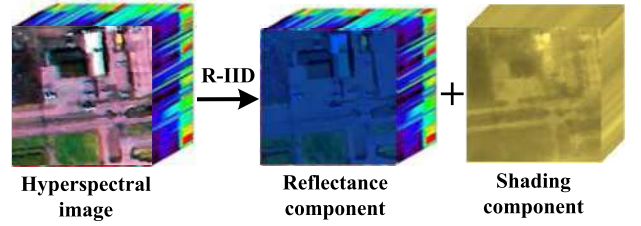

Fig. 1. Graphical illustration of the RIID model.

vector $\mathbf{x}_i$. Likewise, $\mathbf{V}_{\mathbf{B}_r}$, λ_1 and λ_2 are the weights that allow combining the three functions. When the shading base matrix $\mathbf{B}_s$ is known, the energy function can be formulated as follows (for a more detailed description of the LRIID model, we refer readers to the work [37]):

$$E_s = ||\mathbf{W}_{\mathbf{B}_s}\tilde{\mathbf{S}}||^2 + \lambda_1||\mathbf{V}_{\mathbf{L},\mathbf{B}_s}\tilde{\mathbf{S}}||^2 + \lambda_2||\mathbf{M}_s\tilde{\mathbf{S}} - \mathbf{C}||^2. \tag{7}$$

C. DPC

In this section, we briefly describe the DPC algorithm (a detailed description is available in [39]). It relies on the Euclidean distance between two pixels $\mathbf{x}_i$ and $\mathbf{x}_j$, computed as follows:

$$d_{ij} = ||\mathbf{x}_i - \mathbf{x}_j||. \tag{8}$$

The local density, ρ_i, of pixel $\mathbf{x}_i$ can then be defined [48] as

$$\rho_i = \sum_j \exp\left[\left(-\frac{d_{ij}}{d_c}\right)^2\right] \tag{9}$$

where d_c is the specified cutoff distance that indicates the radius of the search region, and is specified to control the weight degradation rate.

III. PROPOSED APPROACH

In this section, we introduce a new HSI anomaly detection method which is based on two main ideas. First, a Retinex-based IID (RIID) model is used to extract the *reflectance* component and remove the *shading* component, in order to obtain pure and discriminative features. Fig. 1 graphically illustrates the IID operation. Second, in order to avoid the Gaussian distribution assumption (and the contamination of the background statistics), and further detect anomalies with different sizes, a new DWD anomaly detector is designed. Fig. 2 provides a schematic of our DWD. Hereinafter, the newly developed HSI anomaly detection method is referred to as R-DWD.

A. RIID

We apply the RIID (an IID model inspired by the LRIID) to obtain the pure reflectance component in HSIs, with the aim of capturing material reflectance properties. Our assumption is that Retinex theory still plays a role in HSI data. Given the amount of evidence already available about the low-rank nature of the reflectance space [37], a low-dimensional subspace of $\mathbf{X} \in \mathbb{R}^{I \times D}$ (designated by $\mathbf{X}' \in \mathbb{R}^{I \times D'}$) is first

extracted for subsequent processing. In order to keep the low-rank nature, we use isometric sampling for band selection. Next, we introduce the assumptions of Retinex theory in HSI data. For two neighboring pixels $\mathbf{x}_i$ and $\mathbf{x}_j$ in the scene, we can obtain the same relations given in Section II-B. By adopting the weights w_{ij} and v_{ij}, we combine these two relations as follows:

$$w_{ij}\mathbf{x}_i.*\mathbf{r}_j + v_{ij}\mathbf{r}_i = w_{ij}\mathbf{x}_j.*\mathbf{r}_i + v_{ij}\mathbf{r}_j \tag{10}$$

where w_{ij} and v_{ij} are determined by a normalized cosine distance formulated as

$$d_{\mathbf{x}_i,\mathbf{x}_j} = 1 - \frac{\mathbf{x}_i \cdot \mathbf{x}_j}{|\mathbf{x}_i||\mathbf{x}_j|}. \tag{11}$$

As introduced in [37]

$$w_{ij} = \frac{1}{1+e^{\alpha(d_{\mathbf{x}_i,\mathbf{x}_j}-\beta)}}, \quad v_{ij} = 1 - w_{ij} \tag{12}$$

where α and β are the parameters of the sigmoid function. In our implementation, we assume that $\mathbf{x}_j$ denotes the pixels in the local window $\mathbf{W}_i$ of $W_I \times W_I$, centered on pixel $\mathbf{x}_i$. The relation in (10) can now be written as follows:

$$\mathbf{r}_i = \sum_{j \in \mathbf{W}_i} \frac{v_{ij}\mathbf{r}_j - w_{ij}\mathbf{x}_i.*\mathbf{r}_j}{v_{ij} - w_{ij}\mathbf{x}_j}. \tag{13}$$

Based on (13) and the inherent data constraint of (2), the resulting shading and reflectance components can be obtained by optimizing the following energy function:

$$E(\mathbf{r}_i, \mathbf{s}_i) = \sum_{i \in \mathbf{X}'} \left(\mathbf{r}_i - \sum_{j \in \mathbf{W}_i} \frac{v_{ij}\mathbf{r}_j - w_{ij}\mathbf{x}_i.*\mathbf{r}_j}{v_{ij} - w_{ij}\mathbf{x}_j} \right)^2 + \sum_{i \in \mathbf{X}'} (\mathbf{x}_i./\mathbf{s}_i - \mathbf{r}_i)^2. \tag{14}$$

Let $\mathbf{R}$ and $\mathbf{S}$ be the matrices of the reflectance component $\mathbf{r}_i$ and the shading component $\mathbf{s}_i$, respectively. The shading component is not taken into account in this article, since it does not directly reflect the material features.

B. DWD Anomaly Detector

Considering that the distribution of the reflectance component (decomposed from HSI data with high dimensionality) is still complex and non-Gaussian, and assuming that density analysis can extract the non-Gaussian structure of the data [47], we developed a DPC-based algorithm to obtain the densities of all pixels. Because of the low probability of existence of anomalies in the image, objects in low-density areas are considered to be anomalies. Taking advantage of this fact, we can detect anomalies by directly using the concept of pixel density (without the need to compute the mean and the covariance matrix of the background).

Specifically, we use the DPC to compute the density in local windows. The resulting IID reflectance feature matrix $\mathbf{R} = [\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_I] \in \mathbb{R}^{I \times D'}$ is divided into N local windows, with T pixels represented as $\mathbf{Y}^n = [\mathbf{r}_1^n, \ldots, \mathbf{r}_t^n, \ldots, \mathbf{r}_T^n]$. For two pixels under test, that is, $\mathbf{r}_i^n$ and $\mathbf{r}_j^n$, belonging to the nth window, the Euclidean distance can be calculated as follows:

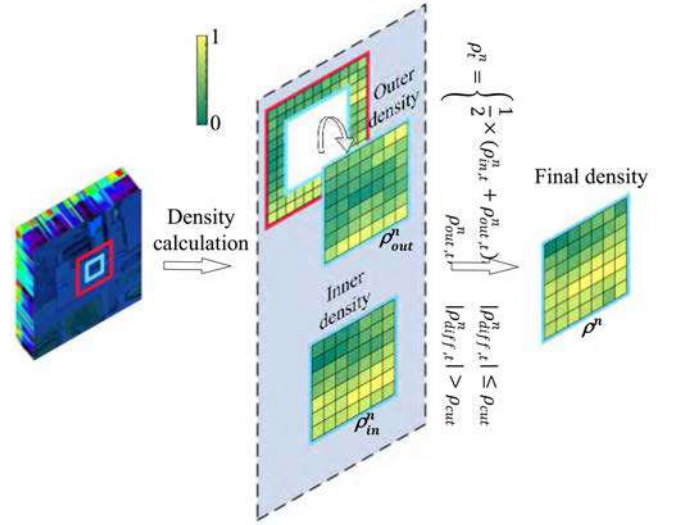$$d_{ij}^n = ||\mathbf{r}_i^n - \mathbf{r}_j^n||. \tag{15}$$

Fig. 2. Schematic of the proposed DWD detector.

In this way, we can obtain a distance matrix $\mathbf{D}^n = [\mathbf{d}_1^n, \ldots, \mathbf{d}_t^n, \ldots \mathbf{d}_T^n]$, where the column $\mathbf{d}_t^n = \{d_{t1}^n, \ldots, d_{tt}^n, \ldots, d_{tT}^n\}$ denotes the distance between the tth pixel, $\mathbf{r}_t^n$, and all pixels under test within the nth window. Then, the upper triangular elements in matrix $\mathbf{D}^n$ are extracted and sorted as vector $\mathbf{d}$ in an ascending order. The cutoff distance d_c can be estimated as follows:

$$d_c = \mathbf{d}([T(T-1) \cdot P\%]) \tag{16}$$

where $[\cdot]$ represents the integer operation, P is a free-control parameter, and $P\%$ is the percentage which determines the position of d_c in $\mathbf{d}$. Consequently, the local density of the pixels under test in the nth window can be computed with (9). For convenience of the subsequent analysis, (9) is rewritten into an average form as follows:

$$\varrho\mathbf{P}^n = \frac{1}{T} \sum_{t=1}^{T} \exp\left[\left(-\frac{\mathbf{D}^n}{d_c} \right)^2 \right] \tag{17}$$

where $\varrho\mathbf{P}$ is the capital form of ρ and $\varrho\mathbf{P}^n$ is the density matrix of the nth window.

Anomalies in the image usually have varying sizes. If we only use the density obtained in a single window for detection purposes, this may lead to false-positive detections. In order to illustrate this fact, we provide a simple graphical example in Fig. 3. In the figure, we place a target with size of 3×3 pixels in a background window with size of 12×12 pixels (in this case, the target is supposed to have a relatively low density and can be identified as an anomaly). However, if we place the same target in a background window with size of 4×4 pixels, the target will have a higher density, since it contains more pixels than the background, leading to a wrong identification. To address this issue, a dual window operation is used for the density computation, which can effectively identify the relative size of the target with respect to the background and better characterize the density of the pixel under test.

Specifically, the dual window refers to an inner window with a size of $W_{\text{in}} \times W_{\text{in}}$ and an outer window with size
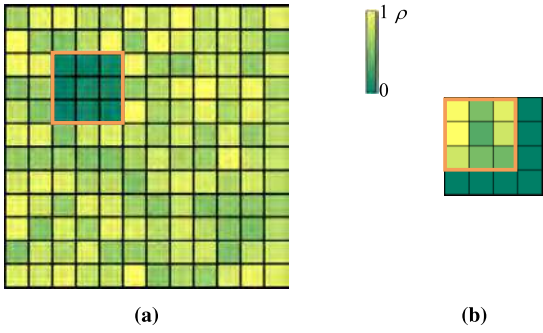
Fig. 3. Density of the same target in different windows. (a) 3×3 target in a 12×12 window. (b) 3×3 target in a 4×4 window.

☐ Inner window ☐ Outer window

Fig. 4. Illustration of four possible scenarios in anomaly detection.

of $W_{\text{out}} \times W_{out}$, in which the outer window is centered on the inner window. The densities obtained in the nth inner and outer windows are denoted as the matrices $\varrho \mathbf{P}_{\text{in}}^{n}$ and $\varrho \mathbf{P}_{\text{out}}^{n}$, respectively. We note that $\varrho \mathbf{P}_{\text{out}}^{n}$ only consists of the densities of the center $W_{\text{in}} \times W_{\text{in}}$ pixels, which are extracted from the entire outer window. Since we focus on the pixels under test within the inner windows, the outer windows are viewed as the background and need to be pure. That is, we assumed that anomaly pixels appear only in the inner window. However, this is hardly the case in practice, when the dual windows move across the whole image, there are many instances that the anomalies are outside the inner window or across the boundaries of the inner window. As a result, we provided a strategy to address this issue by eliminating the interference of the anomalies that are outside the inner window or across the boundaries of the inner window, which is as follows.

With (15), we can compute the distance matrix $\mathbf{D}$ of the pixels within an outer window. $\mathbf{D}$ is a symmetric matrix in which each row or column represents the distance between a pixel and all other pixels. By summing $\mathbf{D}$ by row or column, we obtain the vector $\mathbf{d}$, which contains the sum-distance between each pixel and all other pixels within the outer window. Generally, the larger the sum-distance, the higher the degree to which a pixel is out of the whole background. According to the value in $\mathbf{d}$, we rank the pixels in the outer window in the ascending order, and think that the first 10% of the pixels are out of the background and the last 90% of the pixels belong to the background. The last 90% of the pixels are then used to compute the average pixel vector $\mathbf{u}$ for the replacement of the first 10% of the pixels. Therefore, a new and purer pixel set is obtained, which will be used to calculate the density in the following. This operation can be considered as a coarse prerecognition of anomalies with the purpose of eliminating the interference of the anomalies that are outside the inner window or across the boundaries of the inner window.

Then for a certain pixel under test $\mathbf{r}_{t}^{n}$, $\rho_{\text{in},t}^{n}$ and $\rho_{\text{out},t}^{n}$ are (separately) its density in the inner window and its density in the outer window, so that one of the following scenarios may appear.

1) If $\rho_{\text{in},t}^{n}$ and $\rho_{\text{out},t}^{n}$ are both relatively low, there should be an anomalous target with small size in the inner window,

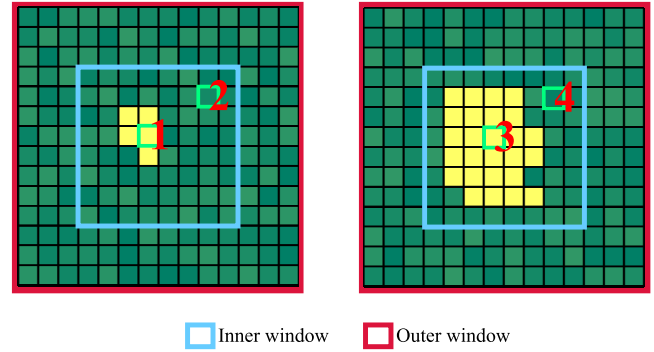and $\mathbf{r}_{t}^{n}$ is supposed to be an anomalous pixel (e.g., pixel 1 in Fig. 4). Here, $\rho_{\text{in},t}^{n}$ and $\rho_{\text{out},t}^{n}$ will be combined for detection.

2) If $\rho_{\text{in},t}^{n}$ and $\rho_{\text{out},t}^{n}$ are both relatively high, there should be an anomalous target with small size or no anomalies in the inner window, and $\mathbf{r}_{t}^{n}$ is supposed to be a normal background pixel (e.g., pixel 2 in Fig. 4). Here, $\rho_{\text{in},t}^{n}$ and $\rho_{\text{out},t}^{n}$ will be combined for detection.

3) If $\rho_{\text{in},t}^{n}$ is relatively high and $\rho_{\text{out},t}^{n}$ is relatively low, there should be an anomalous target with large size in the inner window and $\mathbf{r}_{t}^{n}$ is supposed to be an anomalous pixel (e.g., pixel 3 in Fig. 4). Here, $\rho_{\text{out},t}^{n}$ will be used for detection.

4) If $\rho_{\text{in},t}^{n}$ is relatively low and $\rho_{\text{out},t}^{n}$ is relatively high, there should be an anomalous target with large size in the inner window and $\mathbf{r}_{t}^{n}$ is supposed to be a normal background pixel (e.g., pixel 4 in Fig. 4). Here, $\rho_{\text{out},t}^{n}$ will be used for detection.

Let us define $\varrho \mathbf{P}_{\text{diff}}^{n} = \varrho \mathbf{P}_{\text{out}}^{n} - \varrho \mathbf{P}_{\text{in}}^{n}$ as the difference density matrix between $\varrho \mathbf{P}_{\text{in}}^{n}$ and $\varrho \mathbf{P}_{\text{out}}^{n}$, and $\rho_{\text{diff},t}^{n}$ is the element of $\varrho \mathbf{P}_{\text{diff}}^{n}$, namely the difference value between the density of $\mathbf{r}_{t}^{n}$ in the outer window and in the inner window. When $\rho_{\text{in},t}^{n}$ and $\rho_{\text{out},t}^{n}$ are both relatively low or high, the absolute value of $\rho_{\text{diff},t}^{n}$ tends to be small. When one is relatively high and the other is relatively low, the absolute value of $\rho_{\text{diff},t}^{n}$ tends to be large. As a result, we can decide which density to be used for each pixel under test in the inner window as follows:

$$\rho_{t}^{n} = \begin{cases} \frac{1}{2}\left(\rho_{\text{in},t}^{n} + \rho_{\text{out},t}^{n}\right), & \left|\rho_{\text{diff},t}^{n}\right| \leq \rho_{\text{cut}} \\ \rho_{\text{out},t}^{n}, & \left|\rho_{\text{diff},t}^{n}\right| > \rho_{\text{cut}} \end{cases} \quad (18)$$

where ρ_{cut} is the cutoff density, set as 3×10^{-3} in this article. Then, in order to obtain a separable density distribution, we propose the following method:

$$\rho_{t}^{n} = \rho_{t}^{n} \cdot \exp\left(\frac{d_{c}^{n}}{d_{tk}^{n}} - 1\right) \quad (19)$$

where d_{c}^{n} is the cut-off distance of the nth inner window, d_{tk}^{n} is the distance between pixel $\mathbf{r}_{t}^{n}$ and $\mathbf{r}_{k}^{n}$, and $\mathbf{r}_{k}^{n}$ is the pixel with the highest density in the nth inner window. Finally, the density is normalized by the tanh function as follows:

$$\rho_{t}^{n} = \frac{\exp(\rho_{t}^{n}) - \exp(\rho_{t}^{n})}{\exp(\rho_{t}^{n}) + \exp(\rho_{t}^{n})}. \quad (20)$$

$\varrho \mathbf{P}^n$ is the final density of the nth inner window. A density map can be constructed for the whole image by combining the local densities of all of the N windows. Anomalies can be determined using the final density map. Due to the fact that anomalies are usually rare and exhibit low density in the image, if the density of a pixel under test is smaller than a given threshold, the pixel under test will be determined as an anomaly.

IV. EXPERIMENTAL RESULTS

A. Data Sets

In our experiments, the Airport-Beach-Urban (ABU) data set is used to evaluate the proposed method, which includes an Airport scene, a Beach scene, and an Urban scene. This data set is available online from[1] The color composite images and reference maps are separately shown in the first and second rows of Figs. 14–16. The images are portions with the size of 100×100 pixels extracted from larger images available on the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) web site[2] The ground truth of anomalies is manually labeled with the help of the Environment for Visualizing Images (ENVI) software. The Airport-(1)–(3) images were captured in Los Angeles area, with 7.1-m per pixel spatial resolution, and the Airport-(4) image was captured in Gulfport, with 3.4-m per-pixel spatial resolution. The Beach-(1)–(4) images were captured in Cat Island, San Diego, Bay Champagne, and Pavia, and their spatial resolution is 17.2, 7.5, 4.4, and 1.3 m per pixel, respectively. The Urban-(1) and (2) images were captured in Texas Coast and have 17.2-m per pixel spatial resolution. Urban-(3) was captured in Gainesville, with 3.5-m per pixel spatial resolution. Urban-(4) and (5) were captured in Los Angeles, with a spatial resolution of 7.1 m per pixel. All images were obtained by the AVIRIS sensor except the Beach-(4) image, which was captured by the Reflective Optics System Imaging Spectrometer (ROSIS-03) sensor.

B. Verification of Hypothesis

We first focus on verifying the hypothesis about the performance of the DWD discussed in Section III-B. Fig. 5 shows a small portion extracted from the Urban-(4) image (namely block 1 shown in Fig. 16 [Urban-(4)]), in which the white pixels represent anomalies and the black pixels represent the background. The anomalies are located in a large (outer) window, with a size of 16×16 pixels in Fig. 5(a) and in a small (inner) window in Fig. 5(b), with a size of 8×8 pixels. Fig. 6(a) and (b) shows the densities computed for each pixel with the outer window and the inner window, respectively. As can be seen, the densities of anomalies are low in the outer window whereas the densities are relatively high in the inner window, which is consistent with our previous statements in Section III-B.

As shown in Fig. 7, two portions of the Urban-(4) image (namely, blocks 1 and 2, shown in Fig. 16 [Urban-(4)]) are extracted to simulate four scenarios. In Fig. 7(a) and (b), there

[1]http://jiong.tea.ac.cn/people/xudongkang/Codes.html
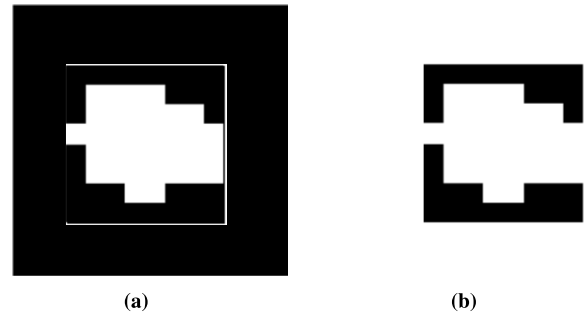[2]http://aviris.jpl.nasa.gov/

Fig. 5. Portion of the Urban-(4) image. (a) Outer window. (b) Inner window. White pixels: anomalies. Black pixels: background.
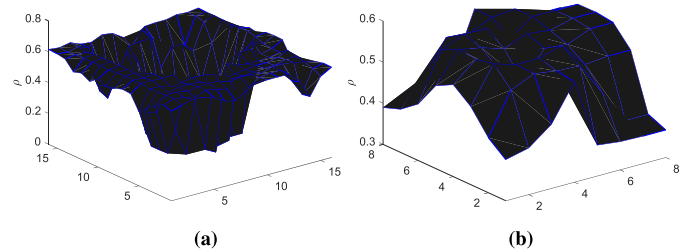

Fig. 6. Densities of pixels. (a) Densities calculated by the outer window. (b) Densities calculated by the inner window.

is a small anomaly and large target, respectively. Twelve pixels are selected as test samples, including six anomalous pixels and six background pixels, numbered from 1 to 12.

Fig. 8 presents the densities of each test pixel in both the inner window (ρ_{in}) and the outer window (ρ_{out}). As shown in Fig. 8(a), which corresponds to Fig. 7(a), pixels 1–3 (which are anomalous pixels) have low densities, and pixels 4–6 (which are background pixels) have high densities in both the inner and outer windows, which confirms assumptions 1 and 2 in Section III-B. In Fig. 8(b), we can observe that pixels 7–9 (which are anomalous pixels) exhibit high ρ_{in} but low ρ_{out}, while pixels 10–12 (which are background pixels) exhibit lower ρ_{in} but higher ρ_{out}, which confirms assumptions 3 and 4 in Section III-B.

Moreover, Fig. 9 shows the density after the process indicated in (19) and (20). As can be intuitively seen in the figure, the densities of anomalous pixels 1–3 and 7–9 (which are expected to be low) become smaller, whereas the densities of background pixels 4–6 and 10–12 (which are expected to high) become higher, resulting in a more separable density distribution. This means that we can achieve more discriminative identification results for anomaly detection after applying the proposed DWD method.

C. Discussion on Segmented Adaptive Local Regions

In this section, we discuss why fixed windows are used in this article instead of adaptive local regions based on image segmentation. Fig. 10 shows different superpixel segmentation results for the Airport-(2) image obtained by the SLIC superpixel algorithm [49]. Superpixel segmentation groups neighboring pixels with similar spectral features into the same
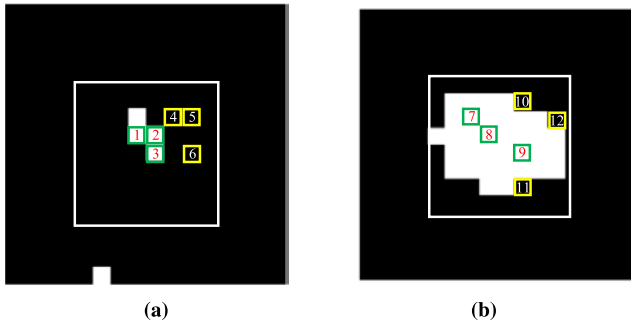
(a)   **(b)**

Fig. 7. Four scenarios of anomalous targets in a real image. (a) Small anomalous target in the inner and outer windows. (b) Large anomalous target in the inner and outer windows. White pixels: anomalies. Black pixels: background.
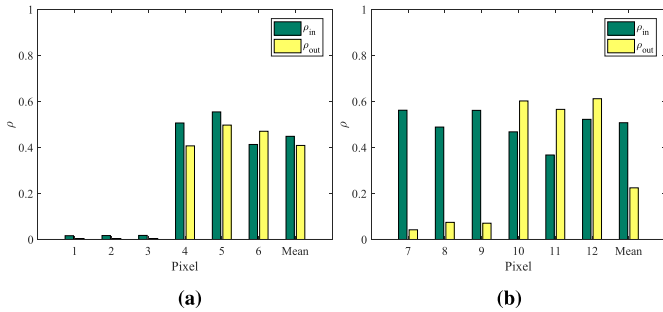

(a)   **(b)**

Fig. 8. Density of some selected test pixels in both the inner and outer windows. Test pixels (a) 1 to 6 and (b) 7 to 12.
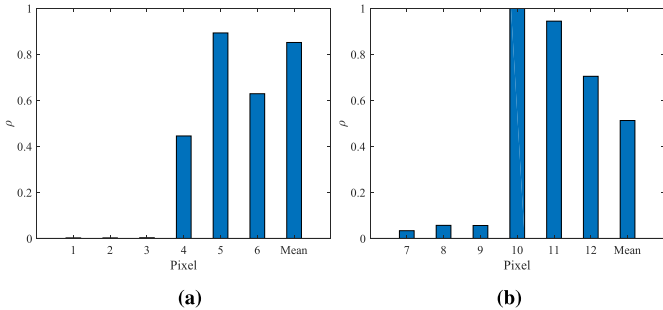

(a)   **(b)**

Fig. 9. Density of some selected test pixels after applying the DWD operation. Test pixels (a) 1 to 6 and (b) 7 to 12.

local region (superpixel). As can be seen from Fig. 10, with the decrease in the *number* of superpixels (that is, the increase in the *size* of superpixels), the anomalous objects (i.e., the airplanes) are always grouped into independent small regions which contain a very small amount of background pixels. According to Section IV-B, in such a superpixel regions the anomalies tend to exhibit high density, whereas background pixels tends to exhibit low densities. Consequently, if we detect anomalies in these superpixel regions, there is a risk that we do not identify targets of interest but normal background pixels.

D. Experimental Setup

1) Evaluation Metrics: In order to conduct an exhaustive quantitative evaluation, we use receiver operating characteristic (ROC) curves [50] and the area under the curve (AUC)

(a)   **(b)**
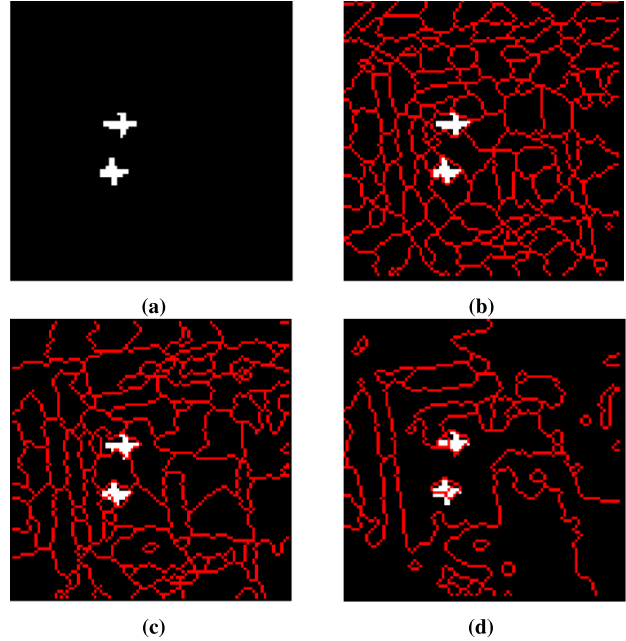

(c)   **(d)**

Fig. 10. Different superpixel segmentation results for the Airport-(2) image obtained by the SLIC superpixel algorithm. (a) Reference map with the targets of interest. (b) SLIC segmentation with 125 superpixels. (c) SLIC segmentation with 95 superpixels. (d) SLIC segmentation with 45 superpixels.

TABLE I
PARAMETER SETTINGS FOR THE PROPOSED R-DWD METHOD IN
EACH ANALYZED IMAGE

Images	W_{in}	W_{out}	P	α	β
Airport-(1)	11	35	1	13000	10^{-3}
Airport-(2)	20	34	1	3000	10^{-5}
Airport-(3)	7	35	1	9000	5×10^{-4}
Airport-(4)	15	23	5	11000	5×10^{-4}
Beach-(1)	14	34	4	13000	5×10^{-6}
Beach-(2)	6	12	21	1000	10^{-4}
Beach-(3)	19	35	5	7000	10^{-3}
Beach-(4)	16	34	5	1000	5×10^{-3}
Urban-(1)	5	13	20	1000	5×10^{-5}
Urban-(2)	21	35	30	1000	5×10^{-4}
Urban-(3)	14	28	19	1000	5×10^{-3}
Urban-(4)	13	35	30	1000	5×10^{-3}
Urban-(5)	21	35	17	11000	5×10^{-4}

as our metrics for evaluating the detection performance of different methods.

2) Parameter Settings: This section analyzes the influence of different parameters on the detection performance of the proposed method in terms of AUC. The considered parameters are: the window sizes: W_{in} and W_{out}, the parameters of the weight function: α and β, and the parameter controlling the cut-off distance: P. Figs. 11 and 12 show the impact of parameters, W_{in}, W_{out}, α and β on the obtained AUCs. When analyzing one of the parameters, the other parameters remain
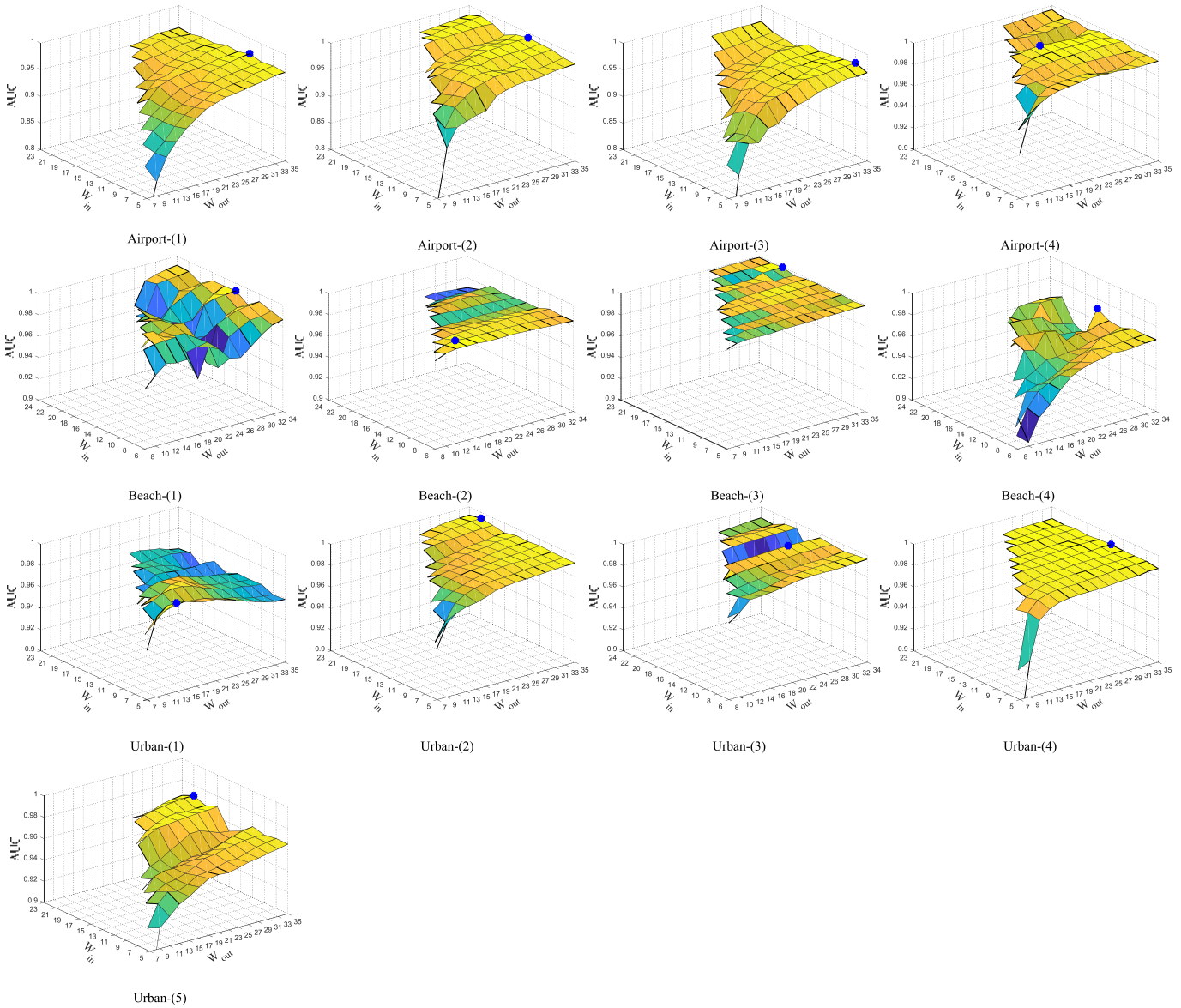
Fig. 11.   Influence of parameters W_{in} and W_{out} in different analysis scenarios.

fixed. As shown in Fig. 11, when W_{in} is fixed the AUC increases with the change of W_{out} in most cases. Although the AUC may still improve when W_{out} is larger than 35, considering the computational complexity, we analyze values of W_{out} within the range [7]–[35]. We can also see that the detection performance is more sensitive to W_{out} than to W_{in}. As shown in Fig. 12, When the value of α is small and the value of β is large, the proposed method can achieve better detection performance. Regarding parameter P, Fig. 13 illustrates the influence of this parameter on the AUCs obtained for different images. Due to the different spatial and spectral characteristics of the considered images, we report different parameter settings for the proposed R-DWD method when applied to the different images considered in our experiment (see Table I). In addition, the local window size W_I described in Section III-A is set to 3 pixels for most

images, except Airport-(2) and Urban-(3), in which this parameter is respectively set to 7 and 11 pixels.

E. Detection Performance

In this section, the proposed R-DWD method is compared to five widely used anomaly detection methods: the RX detector [7], the LRX detector [7], the CRD detector [21], the local kernel RX (L-KRX) detector [10], the low-rank and sparse matrix decomposition (LRaSMD)-based Mahalanobis distance detector (LSMAD) [51], and the density measurement-based SDBP detector [46]. Considering that another IID approach called optimization-based IID has been successfully applied to HSI classification [36], the O-DWD (DWD combined with the optimization-based IID) is also included in our comparison. In order to illustrate the benefits of using the RIID, the results obtained by the DWD without the RIID operation are also
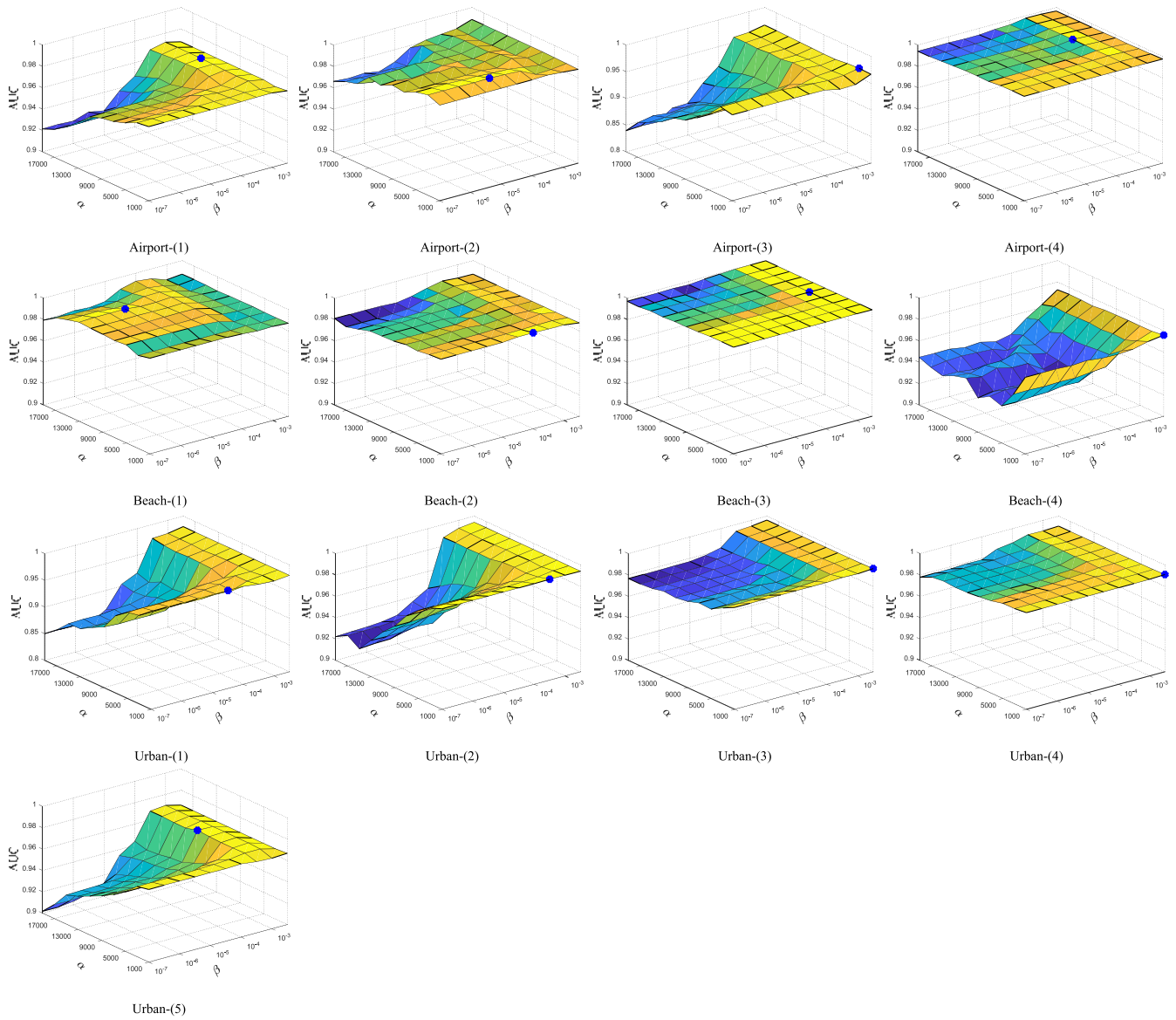
Fig. 12.   Influence of parameters α and β in different analysis scenarios.
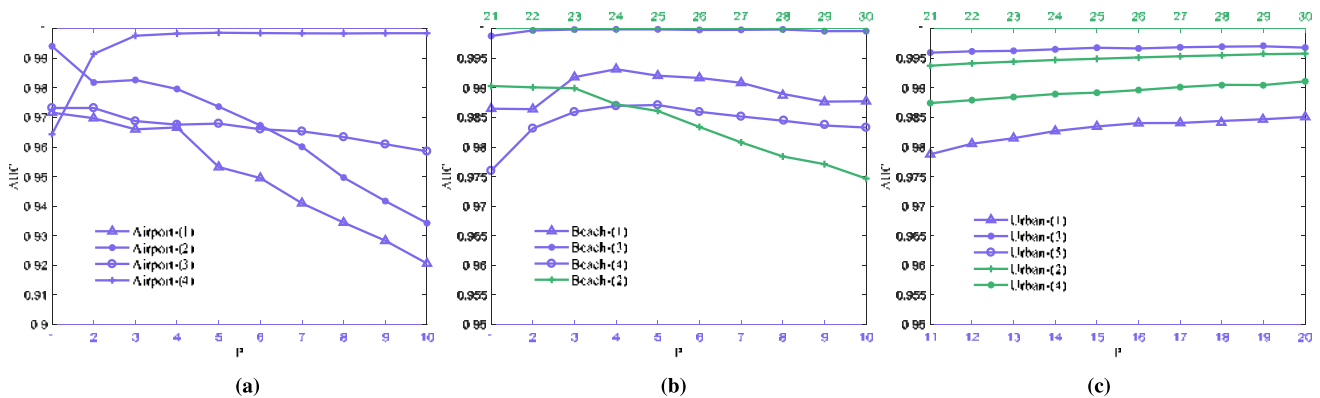
Fig. 13.   Influence of parameter P in different analysis scenarios. (a) Airport scene. (b) Beach scene. (c) Urban scene.

reported. To optimally select the relevant parameters (window size W_{in} and W_{out}) of the LRX, CRD and the L-KRX, we sample values of W_{in} in the range [3]–[21] and W_{out} in the range [5]–[25], and select those values that perform best in terms of the AUC metric. For the proposed R-DWD, the default parameter settings for each image are shown

TABLE II

AUC Scores Obtained on the ABU Data Set by the R-DWD, O-DWD, DWD, RX, LRX, CRD, L-KRX, LSMAD and SDBP Anomaly Detection Methods

Airport scene

Images	R-DWD	O-DWD	DWD	RX	LRX	CRD	L-KRX	LSMAD	SDBP
Airport-(1)	0.9716	0.9518	**0.9723**	0.8281	0.9665	0.9665	0.9005	0.8720	0.9615
Airport-(2)	**0.9940**	0.9856	0.9876	0.8404	0.9874	0.9744	0.8590	0.9018	0.9842
Airport-(3)	0.9732	0.9568	**0.9804**	0.9288	0.9559	0.9564	0.8680	0.9225	0.9662
Airport-(4)	**0.9986**	0.9940	0.9955	0.9526	0.8740	0.9757	0.9592	0.9588	0.9930
Average	**0.9844**	0.9721	0.9839	0.8874	0.9459	0.9682	0.8966	0.9132	0.9762

Beach scene

Images	R-DWD	O-DWD	DWD	RX	LRX	CRD	L-KRX	LSMAD	SDBP
Beach-(1)	0.9931	0.9903	0.9895	0.9807	0.9956	0.9916	0.9753	0.9737	**0.9959**
Beach-(2)	**0.9903**	0.9736	**0.9903**	0.9106	0.9777	0.9416	0.9012	0.9133	0.9876
Beach-(3)	**0.9999**	0.9998	0.9998	**0.9999**	0.9997	0.9996	0.9997	**0.9999**	**0.9999**
Beach-(4)	**0.9871**	0.9609	0.9866	0.9538	0.9391	0.9561	0.9771	0.9528	0.9763
Average	**0.9926**	0.9812	0.9915	0.9612	0.9780	0.9722	0.9633	0.9599	0.9899

Urban scene

Images	R-DWD	O-DWD	DWD	RX	LRX	CRD	L-KRX	LSMAD	SDBP
Urban-(1)	0.9851	0.9675	0.9867	0.9907	0.9968	0.9948	0.9885	0.9904	**0.9989**
Urban-(2)	0.9958	0.9737	0.9956	0.9946	0.9231	0.9410	0.9959	0.9951	**0.9984**
Urban-(3)	**0.9970**	0.9812	0.9962	0.9513	0.9800	0.9634	0.9632	0.9657	0.9946
Urban-(4)	0.9911	0.9531	0.9903	0.9887	0.9696	0.9816	0.9928	0.9864	**0.9972**
Urban-(5)	0.9711	0.9582	0.9744	0.9692	0.9538	0.9521	0.9638	0.9668	**0.9877**
Average	0.9880	0.9667	0.9885	0.9789	0.9646	0.9665	0.9808	0.9809	**0.9954**

in Table I. The O-DWD and DWD also use the optimal parameters for experiments.

Table II shows the detection performance in terms of the AUC score for each compared method. The best score for each row is highlighted in bold. As can be seen from Table II, the proposed R-DWD method obtains the highest AUCs compared to other methods in most of the considered images. The average AUCs obtained by our R-DWD method are 0.9844, 0.9926, and 0.9880 for the Airport, Beach, and Urban scenes, respectively. Compared to the RX, LRX, CRD, L-KRX, and LSMAD, our R-DWD increases the average AUC of these methods by 0.0621, 0.0257, and 0.0137 for the Airport, Beach, and Urban scenes, respectively. When comparing to the SDBP (also a density measurement-based method), the R-DWD performs much better for the Airport and Beach scenes, whereas the average AUC of the SDBP is higher for the Urban scene. Compared to the DWD, we can observe detection performance improvements that come with the RIID operation in some images. However, in some images, the DWD performs better

than the R-DWD. This shows that reflectance component can effectively enhance spectral features of images, but not a few images. The reason may be that some useful information is removed with the removal of shading component. Therefore, at present, we can directly use the DWD for hyperspectral anomaly detection in some cases. Our future works will study how to combine reflection component and shading component to improve the generalization performance of our method. Compared to the O-DWD, the AUCs obtained by our R-DWD are significantly higher. If we take the Urban scene as an example, the average AUC for the O-DWD is 0.9667, whereas the average AUC for our R-DWD is 0.9880. This indicates that the RIID is more suitable for HSI anomaly detection than the optimization-based IID. In addition, the volatility of the AUC for the R-DWD method is also smaller than for the other tested methods, which reveals a better stability of our newly developed R-DWD.

Figs. 14–16 show the detection maps obtained by the different tested methods on the Airport, Beach, and Urban

Airport-(1)

Airport-(2)

Airport-(3)

Airport-(4)

(a)   (b)   (c)   (d)   (e)   (f)   (g)   (h)   (i)   (j)   (k)
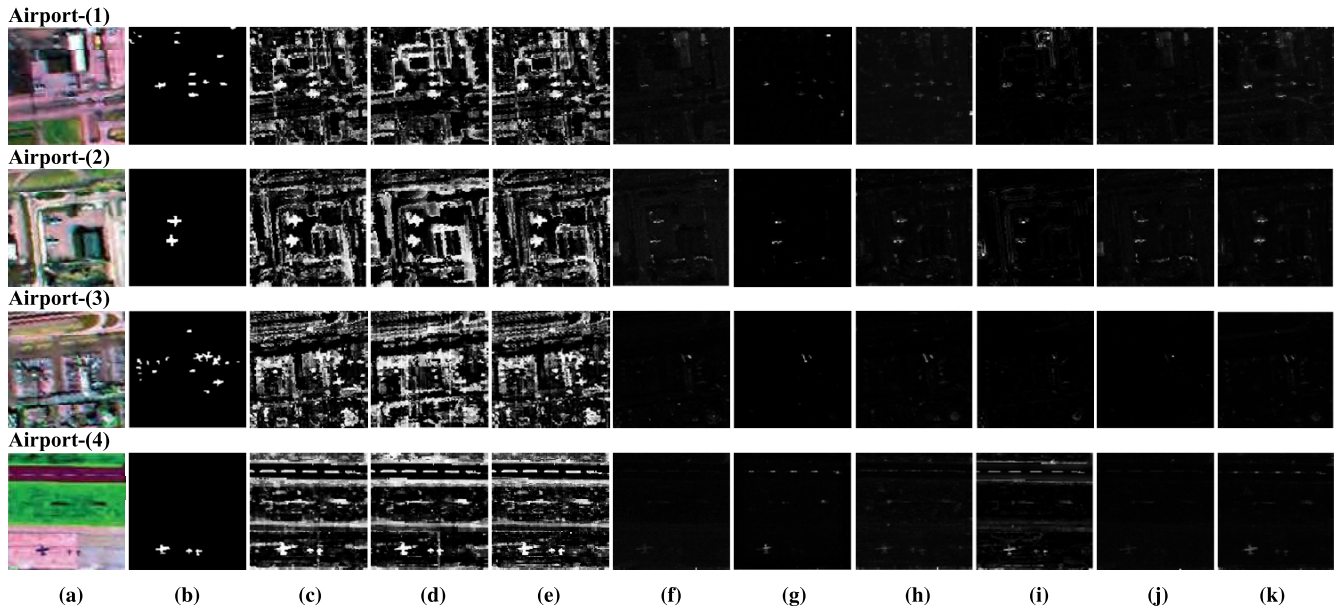
Fig. 14. Detection maps obtained by the different tested methods on the Airport scene. (a) Color composite of the HSI. (b) Reference map showing the targets to be detected. (c) Result of the R-DWD method. (d) Result of the O-DWD method. (e) Result of the DWD method. (f) Result of the RX method. (g) Result of the LRX method. (h) Result of the CRD method. (i) Result of the L-KRX method. (j) Result of the LSMAD method. (k) Result of the SDBP method.

Beach-(1)

Beach-(2)

Beach-(3)

Beach-(4)

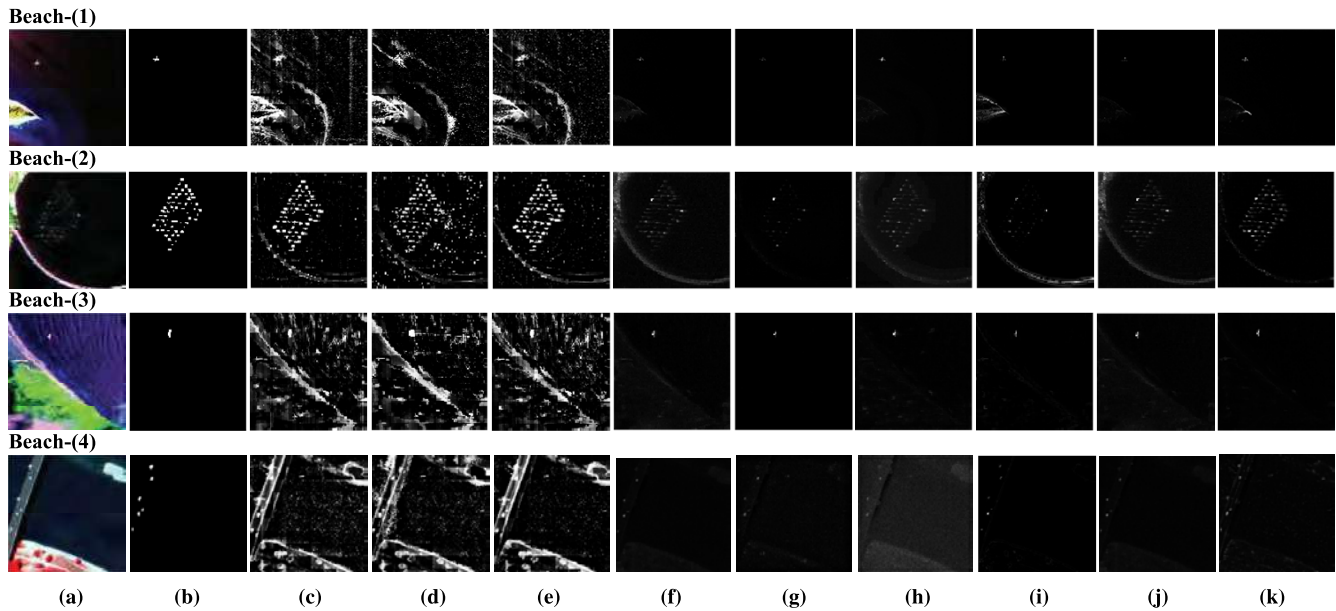(a)   (b)   (c)   (d)   (e)   (f)   (g)   (h)   (i)   (j)   (k)

Fig. 15. Detection maps obtained by the different tested methods on the Beach scene. (a) Color composite of the HSI. (b) Reference map showing the targets to be detected. (c) Result of the R-DWD method. (d) Result of the O-DWD method. (e) Result of the DWD method. (f) Result of the RX method. (g) Result of the LRX method. (h) Result of the CRD method. (i) Result of the L-KRX method. (j) Result of the LSMAD method. (k) Result of the SDBP method.

scenes. As can be intuitively seen, the R-DWD provides the best detection maps. Actually, the R-DWD, O-DWD and DWD can effectively detect most of the anomalies and preserve the shapes of such anomalies. In the detection maps obtained by the RX, LRX, CRD, L-KRX, LSMAD, and SDBP, it is difficult to visually identify anomalies or the shapes of the anomalous objects. Both LRaSMD and RIID can be regarded as matrix decomposition-based methods that can effectively extract the useful component for anomaly detection, but our RIID performs much better than the LSMAD. Compared to

the O-DWD, the RDWD generates fewer false alarms. This reveals the advantages of detecting anomalies in the reflectance component obtained by RIID, which extracts the material spectral information and enhances spectral discrimination, thus leading to the detection of anomalies with different sizes by means of the DWD. If we take the Urban scene as an example, the resulting map obtained by the R-DWD appears much cleaner than those obtained by the O-DWD. For the comparison of R-DWD and DWD, the results are similar in vision. However, if we take Figs. 14 and 15 as an example,
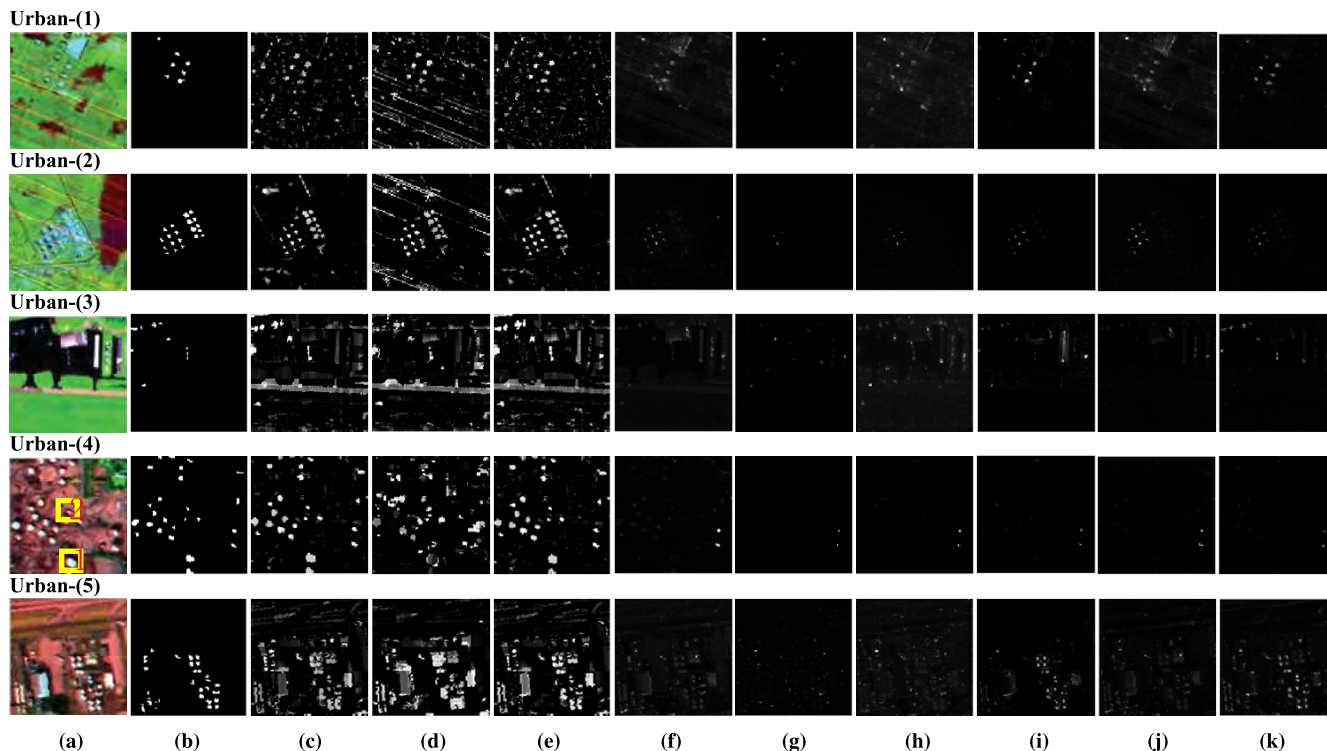
Fig. 16. Detection maps obtained by the different tested methods on the Urban scene. (a) Color composite of the HSI. (b) Reference map showing the targets to be detected. (c) Result of the R-DWD method. (d) Result of the O-DWD method. (e) Result of the DWD method. (f) Result of the RX method. (g) Result of the LRX method. (h) Result of the CRD method. (i) Result of the L-KRX method. (j) Result of the LSMAD method. (k) Result of the SDBP method.
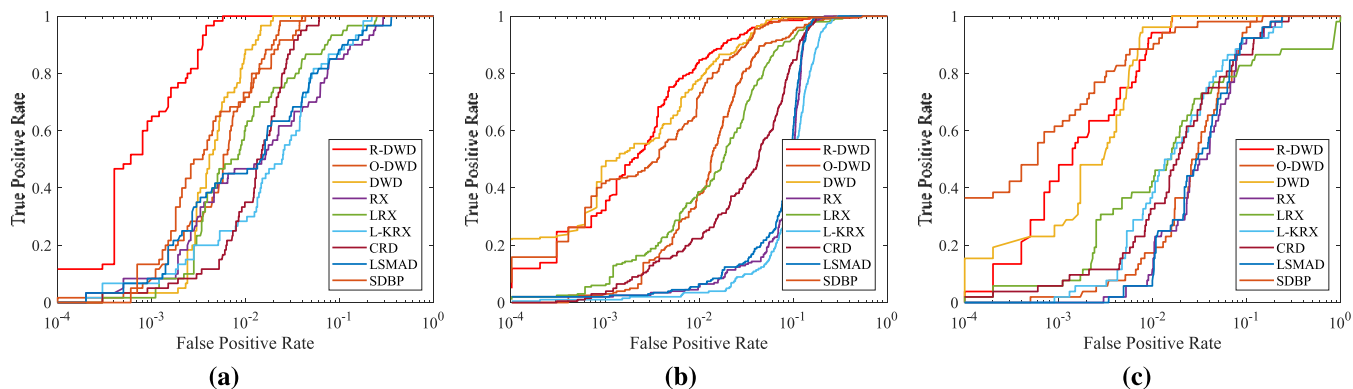

Fig. 17. ROC curves obtained for the different tested methods on (a) Airport-(4) image, (b) Beach-(2) image, and (c) Urban-(3) image.

the R-DWD and DWD still produces a certain amount of false alarms in the overall detection result.

For further evaluation, Airport-(4), Beach-(2), and Urban-(3) have been selected for comparing the ROC curves obtained by the different tested methods. As shown in Fig. 17, the proposed R-DWD, O-DWD and DWD outperform other detection methods in most cases. In other words, the true-positive rates obtained by the proposed methods are always higher than those obtained by the other methods when the false alarm rate varies from 10^{-4} to 1. As can be seen in Fig. 17(a), the ROC curve of the R-DWD rise relatively sharply. When the false alarm rate is very small, the true-positive rate for the R-DWD is superior.

In addition, Fig. 18 shows the final binary detection results obtained by the proposed method through threshold segmentation. As shown in Fig. 18, the detection results are much more clean than the detection maps in Figs. 14–16 and most of the anomalies are detected.

F. Comparison of Computational Time

In this section, we compare the computational time of some comparison detectors: R-DWD, O-DWD, DWD, RX, LRX, and CRD, which is summarized in Table III. The experiment is conducted on the Airport-(4), Beach-(4), and Urban-(5) images on a computer with a 2.7-GHz CPU and 8-GB memory and implemented in MATLAB. As can be seen from Table III, the RX consumes the least time compared with other algorithms. Also as a sliding window-based algorithm, the R-DWD is much faster than LRX and CRD. The speed of the R-DWD is also faster than that of the O-DWD since the RIID takes less
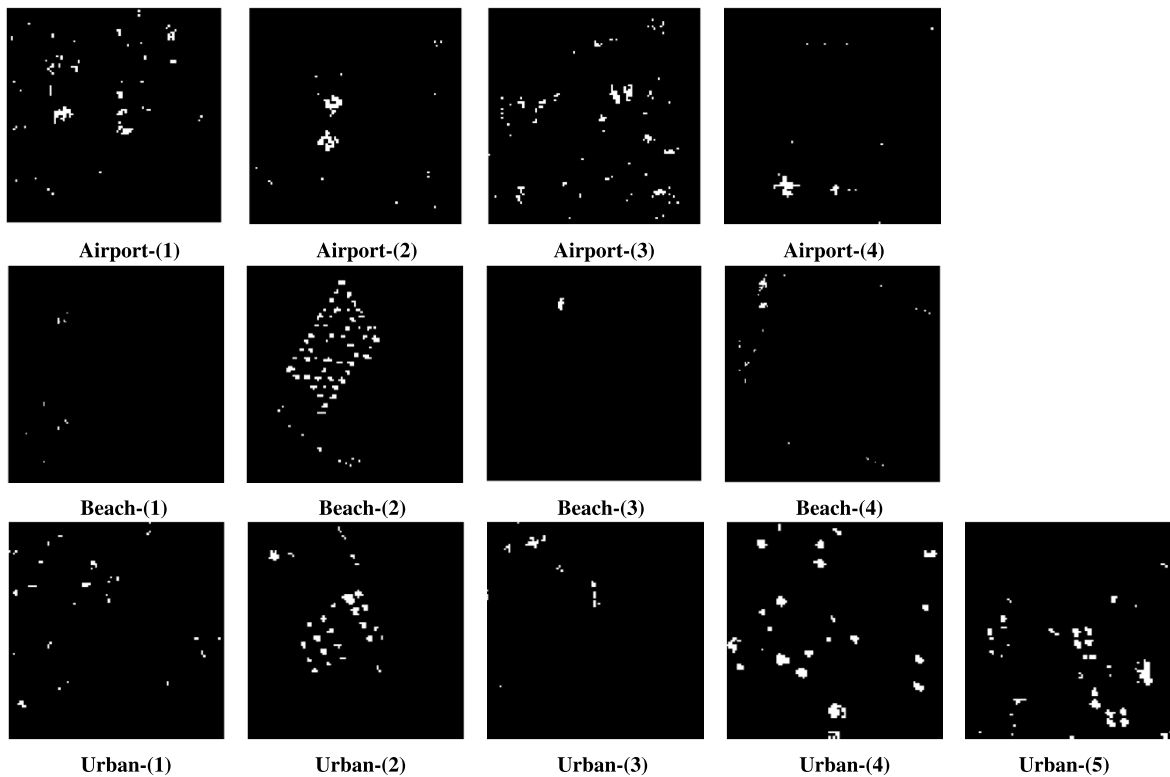
Fig. 18.   Binary detection results obtained by the proposed method through threshold segmentation.

TABLE III
COMPUTATION TIME (IN SECONDS) OF DIFFERENT
COMPARISON DETECTORS

	R-DWD	O-DWD	DWD	RX	LRX	CRD
Airport-(4)	4.95s	37.68s	5.80s	0.27s	58.52s	172.49s
Beach-(4)	19.64s	69.60s	22.11s	0.55s	37.45s	47.34s
Urban-(5)	25.26s	65.16s	19.30s	0.44s	65.50s	63.77s

time than the optimized-based IID. For the speed comparison between the R-DWD and DWD, it depends on the determined window size. When the selected window size is the same, the R-DWD is faster than DWD since that the spectral bands have been reduced after the RIID operation.

V. CONCLUSION

This article introduces a new HSI anomaly detection method called R-DWD. Our method is composed of main ingredients: the RIID and the DWD-based anomaly detector. The RIID is adapted from the low rank multispectral IID model, called LRIID, which can effectively extract the reflectance component of HSIs for anomaly detection purposes. In order to exploit this reflectance component for anomaly detection purposes, a DWD is proposed that circumvents the Gaussian assumption of the distribution of the background in HSIs, improving the calculation of background statistics. The dual window operation of DWD exhibits the ability to detect anomalies with varying sizes and accurately recognize their shape.

Our experimental results, conducted on a large database of HSIs including airport, beach, and urban scenes, demonstrate the superiority of the proposed method with regards to state-of-the-art competitors. As with any new approach, there are still some unresolved issues that may present challenges over time. Specifically, the proposed method still produces a certain number of false alarms in the detection result, and the parameter-setting of our algorithm is laborious. These aspects deserve further research and improvements in future work. We are also planning on developing an efficient implementation of the proposed approach by resorting to high performance computing architectures such as graphics processing units (GPUs).

REFERENCES

[1] Q. Tong, B. Zhang, and L. Zheng, *Hyperspectral Remote Sensing: The Principle, Technology and Application*. Beijing, China: Higher Education Press, 2006.
[2] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.
[3] 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.
[4] 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.

[5] S. M. Schweizer and J. M. F. Moura, "Hyperspectral imagery: Clutter adaptation in anomaly detection," *IEEE Trans. Inf. Theory*, vol. 46, no. 5, pp. 1855–1871, Aug. 2000.

[6] P. Bajorski, "Target detection under misspecified models in hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 470–477, Apr. 2012.

[7] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.

[8] 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.

[9] W. Liu and C.-I. Chang, "Multiple-window anomaly detection for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 644–658, Apr. 2013.

[10] 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.

[11] 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.

[12] Y. P. Taitano, B. A. Geier, and K. W. Bauer, "A locally adaptable iterative RX detector," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, pp. 1–10, Dec. 2010.

[13] N. Gorelnik, H. Yehudai, and S. R. Rotman, "Anomaly detection in non-stationary backgrounds," in *Proc. 2nd Workshop Hyperspectral Image Signal Process.*, Jun. 2010, pp. 1–4.

[14] R. Zhao, B. Du, and L. Zhang, "A robust nonlinear hyperspectral anomaly detection approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1227–1234, Apr. 2014.

[15] N. M. Nasrabadi, "Regularization for spectral matched filter and RX anomaly detector," *Proc. SPIE*, vol. 6966, Apr. 2008, Art. no. 696604.

[16] B. D. Bartlett, A. Schlamm, C. Salvaggio, and D. W. Messinger, "Anomaly detection of man-made objects using spectropolarimetric imagery," *Proc. SPIE*, vol. 8048, no. 1, pp. 289–293, 2011.

[17] Q. Guo, B. Zhang, Q. Ran, L. Gao, J. Li, and A. Plaza, "Weighted-RXD and linear filter-based RXD: Improving background statistics estimation for anomaly detection in hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2351–2366, Jun. 2014.

[18] H. Kwon, "Adaptive anomaly detection using subspace separation for hyperspectral imagery," *Opt. Eng.*, vol. 42, no. 11, pp. 3342–3351, Nov. 2003.

[19] C.-I.-C. Weimin Liu, "A nested spatial window-based approach to target detection for hyperspectral imagery," in *Proc. IEEE Int. IEEE Int. IEEE Int. Geosci. Remote Sens. Symp., (IGARSS)*, Sep. 2004, p. 268.

[20] B. Du, R. Zhao, L. Zhang, and L. Zhang, "A spectral-spatial based local summation anomaly detection method for hyperspectral images," *Signal Process.*, vol. 124, pp. 115–131, Jul. 2016.

[21] W. Li and Q. Du, "Collaborative representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1463–1474, Mar. 2015.

[22] J. Li, H. Zhang, L. Zhang, and L. Ma, "Hyperspectral anomaly detection by the use of background joint sparse representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2523–2533, Jun. 2015.

[23] M. Ben Salem, K. Saheb Ettabaa, and M. S. Bouhlel, "An adaptive spatial and spectral neighborhood for the RX anomaly detector," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2018, pp. 8484–8487.

[24] 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.

[25] M. A. Veganzones, J. Frontera-Pons, F. Pascal, J.-P. Ovarlez, and J. Chanussot, "Binary partition trees-based robust adaptive hyperspectral RX anomaly detection," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 5077–5081.

[26] D. Ma, Y. Yuan, and Q. Wang, "Hyperspectral anomaly detection via discriminative feature learning with multiple-dictionary sparse representation," *Remote Sens.*, vol. 10, no. 5, p. 745, 2018.

[27] Y. Qu *et al.*, "Anomaly detection in hyperspectral images through spectral unmixing and low rank decomposition," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 1855–1858.

[28] Y. Qu *et al.*, "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.

[29] W. Wang, S. Li, H. Qi, B. Ayhan, C. Kwan, and S. Vance, "Identify anomaly component by sparsity and low rank," in *Proc. IEEE 7th Workshop Hyperspectral Image Signal Process. Evol. Remote Sens. (Whispers)*, Mar. 2015, pp. 1–4.

[30] M. F. Tappen, W. T. Freeman, and E. H. Adelson, "Recovering intrinsic images from a single image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1459–1472, Sep. 2005.

[31] A. Bousseau, S. Paris, and F. Durand, "User-assisted intrinsic images," *ACM Trans. Graph.*, vol. 28, no. 5, pp. 1–10, Dec. 2009.

[32] R. Grosse, M. K. Johnson, E. H. Adelson, and W. T. Freeman, "Ground truth dataset and baseline evaluations for intrinsic image algorithms," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 2335–2342.

[33] L. Shen, C. Yeo, and B.-S. Hua, "Intrinsic image decomposition using a sparse representation of reflectance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2904–2915, Dec. 2013.

[34] Q. Zhao, P. Tan, Q. Dai, L. Shen, E. Wu, and S. Lin, "A closed-form solution to retinex with nonlocal texture constraints," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1437–1444, Jul. 2012.

[35] J. Shen, X. Yang, X. Li, and Y. Jia, "Intrinsic image decomposition using optimization and user scribbles," *IEEE Trans. Cybern.*, vol. 43, no. 2, pp. 425–436, Apr. 2013.

[36] X. Kang, S. Li, L. Fang, and J. A. Benediktsson, "Intrinsic image decomposition for feature extraction of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2241–2253, Apr. 2015.

[37] Q. Huang *et al.*, "Multispectral image intrinsic decomposition via subspace constraint," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018.

[38] E. H. Land and J. J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Amer.*, vol. 61, no. 1, p. 1, Jan. 1971.

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

[40] X. Luo, R. Xue, and J. Yin, "Information-assisted density peak index for hyperspectral band selection," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1870–1874, Oct. 2017.

[41] G. Tang, S. Jia, and J. Li, "An enhanced density peak-based clustering approach for hyperspectral band selection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 1116–1119.

[42] Y. Chen, S. Ma, X. Chen, and P. Ghamisi, "Hyperspectral data clustering based on density analysis ensemble," *Remote Sens. Lett.*, vol. 8, no. 2, pp. 194–203, Feb. 2017.

[43] B. Tu, C. Zhou, D. He, S. Huang, and A. Plaza, "Hyperspectral classification with noisy label detection via superpixel-to-pixel weighting distance," *IEEE Trans. Geosci. Remote Sens.*, early access, Jan. 13, 2020, doi: 10.1109/TGRS.2019.2961141.

[44] B. Tu, X. Zhang, X. Kang, J. Wang, and J. A. Benediktsson, "Spatial density peak clustering for hyperspectral image classification with noisy labels," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5085–5097, Jul. 2019.

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

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

[47] X. Yang, L. He, and H. Lu, "A clustering algorithm for datasets with different densities," in *Proc. Int. Conf. Comput. Technol. Develop.*, 2009, pp. 504–507.

[48] K. Sun, X. Geng, and L. Ji, "Exemplar component analysis: A fast band selection method for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 998–1002, May 2015.

[49] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[50] J. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251–255, Apr. 2008.

[51] 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.

Bing Tu (Member, IEEE) received the M.S. degree from the Guilin University of Technology, Guilin, China, in 2009, and the Ph.D. degree from the Beijing University of Technology, Beijing, China, in 2013.

From 2015 to 2016, he was a Visiting Researcher with the Department of Computer Science and Engineering, University of Nevada, Reno, NV, USA, which was supported by the China Scholarship Council. Since 2018, he has been an Associate Professor with the School of Information Science and Engineering, Hunan Institute of Science and Technology, Yueyang, China. His research interests include sparse representation, pattern recognition, and analysis in remote sensing.

Xianchang Yang (Student Member, IEEE) received the B.S. degree from the Hunan Institute of Science and Technology, Yueyang, China, in 2020, where he is pursuing the M.S. degree.

His research interests include image processing, pattern recognition, hyperspectral image classification, and anomaly detection.

Chengle Zhou (Graduate Student Member, IEEE) received the B.S. degree from the Hunan Institute of Science and Technology, Yueyang, China, in 2019, where he is pursuing the M.S. degree.

His research interests include image processing, pattern recognition, hyperspectral image classification, anomaly detection, and noisy label detection.

Danbing He (Student Member, IEEE) is pursuing the B.S. degree with the School of Information Science and Technology, Hunan Institute of Science and Technology, Yueyang, China.

Her research interest includes hyperspectral image analysis.

Antonio Plaza (Fellow, IEEE) received the M.Sc. and Ph.D. degrees in computer engineering from the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain, in 1999 and 2002, respectively.

He is the Head of the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura. He has authored more than 600 publications, including over 200 JCR journal articles (over 160 in IEEE journals), 23 book chapters, and around 300 peer-reviewed conference proceeding papers. His research interests include hyperspectral data processing and parallel computing of remote sensing data.

Dr. Plaza was a member of the Editorial Board of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER from 2011 to 2012 and the *IEEE Geoscience and Remote Sensing Magazine* in 2013. He was also a member of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS). He is also a fellow of the IEEE for contributions to hyperspectral data processing and parallel computing of Earth observation data. He was a recipient of the recognition of Best Reviewers of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2009 and the Best Reviewer of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2010, for which he has served as an Associate Editor from 2007 to 2012. He was also a recipient of the Best Column Award of the *IEEE Signal Processing Magazine* in 2015, the 2013 Best Paper Award of the JSTARS journal, and the Most Highly Cited Paper (2005–2010) in the *Journal of Parallel and Distributed Computing*. He received best paper awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. He has served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) from 2011 to 2012 and the President for the Spanish Chapter of the IEEE GRSS from 2012 to 2016. He has reviewed more than 500 manuscripts for over 50 different journals. He has served as the Editor-in-Chief for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2013 to 2017. He has guest edited ten special issues on hyperspectral remote sensing for different journals. He is also an Associate Editor of the IEEE ACCESS (receiving the recognition as an Outstanding Associate Editor of the journal in 2017). Additional information: http://www.umbc.edu/rssipl/people/aplaza