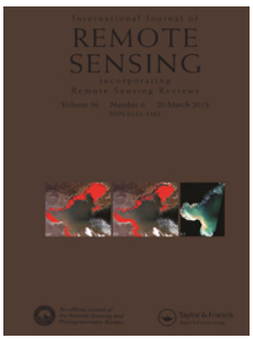


International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: http://www.tandfonline.com/loi/tres20

Comparison of spectral and spatial windows for local anomaly detection in hyperspectral imagery

Zhiyong Li, Jonathan Li, Shilin Zhou & Saied Pirasteh

To cite this article: Zhiyong Li, Jonathan Li, Shilin Zhou & Saied Pirasteh (2015) Comparison of spectral and spatial windows for local anomaly detection in hyperspectral imagery, International Journal of Remote Sensing, 36:6, 1570-1583, DOI: 10.1080/01431161.2015.1017666

To link to this article: http://dx.doi.org/10.1080/01431161.2015.1017666

Published online: 16 Mar 2015.

International Journal of Remote Sensing, 2015
Vol. 36, No. 6, 1570–1583, http://dx.doi.org/10.1080/01431161.2015.1017666

Taylor & Francis
Taylor & Francis Group

Comparison of spectral and spatial windows for local anomaly detection in hyperspectral imagery

Zhiyong Li[a,b]*, Jonathan Li[b,c], Shilin Zhou[a], and Saied Pirasteh[b]

[a]*School of Electronic Science and Engineering, National University of Defence Technology, Changsha, Hunan 410073, China;* [b]*Department of Geography & Environmental Management, Faculty of Environment, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1;* [c]*School of Information Science and Engineering, Xiamen University, Xiamen, Fujian 361005, China*

(*Received 13 August 2014; accepted 6 January 2015*)

A new way of implementing two local anomaly detectors in a hyperspectral image is presented in this study. Generally, most local anomaly detector implementations are carried out on the spatial windows of images, because the local area of the image scene is more suitable for a single statistical model than for global data. These detectors are applied by using linear projections. However, these detectors are quite improper if the hyperspectral dataset is adopted as the nonlinear manifolds in spectral space. As multivariate data, the hyperspectral image datasets can be considered to be low-dimensional manifolds embedded in the high-dimensional spectral space. In real environments, the nonlinear spectral mixture occurs more frequently, and these manifolds could be nonlinear. In this case, traditional local anomaly detectors are based on linear projections and cannot distinguish weak anomalies from background data. In this article, local linear manifold learning concepts have been adopted, and anomaly detection algorithms have used spectral space windows with respect to the linear projection. Output performance is determined by comparison between the proposed detectors and the classic spatial local detectors accompanied by the hyperspectral remote-sensing images. The result demonstrates that the effectiveness of the proposed algorithms is promising to improve detection of weak anomalies and to decrease false alarms.

1. Background and introduction

Anomaly detection is a primary procedure of auto target curing in hyperspectral image processing. Anomaly detectors can handle original hyperspectral data to identify pixels with distinct spectra from the background (Stein et al. 2002; Yver and Marion 2007; Huck and Guillaume 2010). The anomaly detectors do not require prior spectral information, reflectance spectrum retrieval, and atmospheric compensation. Nowadays, it is a very active research topic in the field of hyperspectral image processing.

Ahlberg, Renhorn, and Forskningsinstitut (2004) stressed that algorithms for target and anomaly detection in hyperspectral imagery can be divided into two categories as (a) detectors that are based on unstructured models, and (b) detectors that are based on structured models. The unstructured models refer to no specific geometric structure on the data. They are also called probabilistic, statistical, and/or data-driven models. Anomaly detection methods based on unstructured models are traditional and familiar. They encompass the well-known Reed-Xiaoli (RX) detector proposed by Reed and

*Corresponding author. Email: lzylmz75@gmail.com

Yu (1990) and its improvement, e.g. the subspace RX (SSRX) detector and RX after orthogonal subspace projection (Schaum 2007; Borghys et al. 2012). Detectors using structured models were based on linear subspace or the linear mixture model (LMM), such as orthogonal subspaces projection (OSP) developed by Harsanyi, Farrand, and Chang (1994), and low probability detection (LPD) (Harsanyi, Farrand, and Chang 1994; Chang 2005). It was interesting that both structured and unstructured detectors performed the same functional form of matched filter or linear projection (Chang and Chiang 2002).

In general, the anomaly detection algorithm can be implemented either in the global or the local area of the image scene. There were two ways to implement local detectors (Bachega, Theiler, and Bouman 2011), which are sliding spatial windows and segmentation of the image scene. The fundamental concept of both methods involves local data of the image scene, and it is more suitable for a single statistical model. The RX algorithm was derived by hypothesis tests that are based on multivariate Gaussian distribution. It is always considered to be a local method, such as local RX and quasi local RX (Caefer et al. 2008; Borghys et al. 2012).

Multivariate Gaussian distribution is frequently used in anomaly detection. It usually implies the assumption of statistical independence between the dimensions. However, the description of the geometric dispersion of hyperspectral image datasets is not accommodated, even for the local spatial data. Because the data dimensions (or bands) are dependent on the hyperspectral images, the geometric dispersion would be some lower-dimensional manifold. Furthermore, a nonlinear spectral mixture occurs more often in real environments; the nonlinear manifolds would be more common. In this case, anomaly detectors based on linear projections cannot deal with these data efficiently. Therefore, there were many kernel-based detectors, such as kernel RX and kernel OSP (Nasrabadi 2014). They could handle this problem by using the kernel functions to implement a linear projection in high-dimensional feature space. Another two kinds of detectors were developed to treat this problem. One kind was to use a local topological structure to calculate the weighted vertex volume (WVV) and extract the ‘outliers’ in the local spatial image (Messinger and Chester 2011). Outlier detection aims at discovering anomalous or inconsistent patterns from a dataset. This was one of the major tasks in data mining when working on a large dataset (Zhao and Saligrama 2009; Ramaswamy, Rastogi, and Shim 2000). Similar studies have been shown in the literature of Du and Zhang (2014a, 2014b). The patches with a local linear structure have been used to fit the manifold structure of the entire hyperspectral image. The other kind of detectors focused on manifold embedding methods. Agovic et al. (2007) have studied anomaly detection in transportation corridors using manifold embedding. They investigated the usefulness of manifold embedding methods for feature representation in anomaly detection problems. The study focused on both the linear methods, such as multidimensional scaling (MDS), and the nonlinear methods including locally linear embedding (LLE) and isometric feature mapping (ISOMAP).

In this article, an improved method for the local detectors has been proposed. Borrowed from the local linear concepts in manifold learning methods (Seung and Lee 2000), the sliding windows would be established in spectral space, and the local dataset could be considered to be a linear manifold. The detector based on a linear projection has been implemented in these local datasets. These detectors can be called spectral space window anomaly detection (SSW-AD). An improved method has encompassed two major detectors: RX based on the spectral space window (SSW-RX) and low

probability anomaly detection (LPAD) (Li, Wang and Zheng 2014) based on the spectral space window (SSW-LPAD).

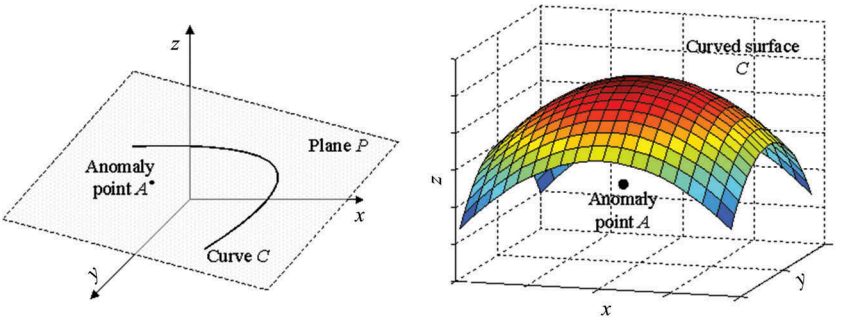
2. Nonlinear manifold and anomalies

As a kind of multivariate dataset, the distribution of the hyperspectral dataset could be considered to be a lower-dimensional manifold in spectral space. Usually, the dimension of this manifold is called the intrinsic dimension (Verveer and Duin 1995). If the spectral data fit the description of the LMM, the manifold is linear and can be regarded as a hyperplane or a linear hypersimplex (Koppen 2000). In this case, the whole dataset can be linearly projected into a subspace whose dimension is an intrinsic dimension. The dispersion of the original data can be represented accurately in this linear subspace.

On the other hand, the lower-dimensional manifold would be nonlinear if spectral data fit the description of the nonlinear mixed model, such as a hyper-curved surface or a nonlinear hypersimplex (Koppen 2000). Different from the case of a linear manifold, the whole dataset cannot be described accurately by using a linear transform (e.g. translation, rotation, and scaling) in intrinsic dimensional subspace. It needs a higher-dimensional linear subspace to contain this manifold.

In anomaly detection, nonlinear properties of the data manifold have considerable effects on the detectors based on the linear projection. An example in 3-dimensional space (3-D) can illustrate this problem, see Figure 1. Figure 1(*a*) shows a 2-D curve *C* and one anomaly point *A* in the 3-D space, and Figure 1(*b*) shows a 3-D curved-surface *C* and one anomaly point *A* in the 3-D space. In Figure 1(*a*), the curve *C* can be considered to be a nonlinear manifold and its intrinsic dimension is 1. The points of the curve *C* and anomaly point *A* are all located on the plane *P*. In order to suppress the energy of points in curve *C* effectively, the detectors on linear orthogonal projections will project all points to the orthogonal space of plane *P*. Obviously, the energy of anomaly point *A* will be submerged into the points of curve *C* in the projected subspace. It will be hard to detect *A* from this projected image. The case of Figure 1(*b*) has the same problem.

Next, the detectors based on a statistical model are considered, such as the RX algorithm. Multivariate Gaussian distribution is usually used as a statistical model, the probability density function is:



(*a*) (*b*)

Figure 1. Nonlinear manifold and anomaly point. (*a*) 2-D curve and anomaly point; (*b*) 3-D curve surface and anomaly point.

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{l/2}|\mathbf{M}|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \mathbf{M}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right), \tag{1}$$

where $\boldsymbol{x}$ is the multivariate random vector, $\mathbf{M}$ is the covariance matrix, $\boldsymbol{\mu}$ is the mean vector, and superscript l is the dimension. This model implies that the data of dimensions are statistically independent of each other. However, if the data among the dimensions are nonlinearly correlated, the whole dataset will be distributed in a nonlinear manifold, such as the curved-surface C in Figure 1(b). However, the geometric structure of the multivariate Gaussian distributed model will be an ellipsoid. The anomaly point A is included in this ellipsoid and cannot be separated from this ellipsoid in the projected data.

Nevertheless, if detectors have handled the nonlinear manifold data based on linear projections, the weak anomalies (close to the background manifold) would be ignored, even for the local data of the image scene. Therefore, if a small area is segmented from this manifold, it could be considered to be a linear structure. At the same time, the detector implementations based on linear projections could be carried out efficiently. This is the motivation of the proposed method.

3. Method

It is important to compare the performance between the proposed algorithm and algorithms based on the spatial window. Here, the algorithms based on the spatial windows will be introduced first; they were the local RX (LRX) and local OSP (LOSP) algorithms (Chao, Huijie, and Wei 2009). Next, two detectors based on the spectral space windows will be presented: SSW-RX and SSW-LPAD.

3.1. Local RX (LRX)

The RX algorithm is a constant false alarm rate (CFAR) adaptive anomaly detector that is derived from the generalized likelihood ratio test (GLRT). It has been considered to be the benchmark anomaly detection algorithm for hyperspectral data (Yver and Marion 2007; Messinger and Chester 2011). The RX algorithm is based on the assumption that the background can be modelled as a multivariate Gaussian distribution, the detector was:

$$D_{\mathrm{RXD}}(\boldsymbol{x}_{\mathrm{i}}) = (\boldsymbol{x}_i - \boldsymbol{\mu})^T \mathbf{M}^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}) \begin{cases} \geq \eta, & \text{anomaly present} \\ < \eta, & \text{anomaly absent} \end{cases} \tag{2}$$

where $\boldsymbol{x}_i$ is the pixel spectra of the image data, and subscript i refers to the ith pixel of the datasets. $\boldsymbol{\mu} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i$ is the mean spectrum, and $\mathbf{M} = \frac{1}{N}\sum_{i=1}^{N} (X_i - \boldsymbol{\mu})(X_i - \boldsymbol{\mu})^T$ is the covariance matrix.

Usually, the RX algorithm is regarded as a local detector, but it can also handle global data. In the implementation of LRX, the covariance matrix and the mean spectrum of the background are estimated locally in a window around the pixel under test (PUT). A double sliding window is used: a guard window and the outer window are defined, and the background statistics are determined using the pixels between the two windows (see Figure 2).

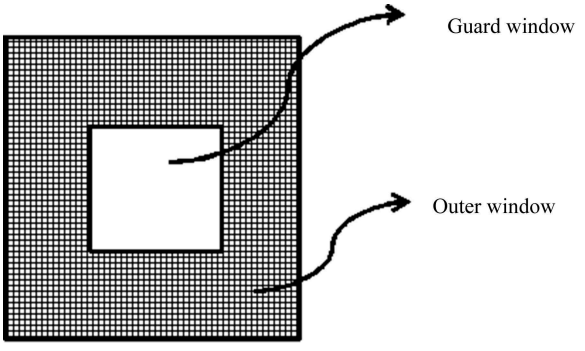


Figure 2. Double spatial window of LRX.

3.2. *Local orthogonal subspace projection (LOSP)*

The local orthogonal subspace projection (LOSP) algorithm (Chao, Huijie, and Wei 2009) was based on the assumption that the variation of terrain was small in the local area of the image scene. Hence, the mean spectrum of the windows could represent the background and be used for constructing the orthogonally projected operator. If the spectrum of the PUT is $\boldsymbol{d}$, the mean spectrum of the sliding window is $\bar{\boldsymbol{d}}$ and then the detector of the LOSP is:

$$P_{\mathrm{LOSP}} = \kappa \boldsymbol{d}^{\mathrm{T}} \left[\mathbf{I} - \bar{\boldsymbol{d}} (\bar{\boldsymbol{d}}^{\mathrm{T}} \bar{\boldsymbol{d}})^{-1} \bar{\boldsymbol{d}}^{\mathrm{T}} \right], \tag{3}$$

where $\bar{\boldsymbol{d}} = \sum_{i}^{m} \boldsymbol{y}_i$, $\boldsymbol{y}_i$ is the pixels' spectra in the window except PUT. κ is a scalar for normalization.

3.3. *Spectral space window RX (SSW-RX)*

In SSW-RX, the RX algorithm was implemented on sliding windows in spectral space. By calculating the Euclidean distance between any two pixel spectra, the list of nearest neighbours can be established for each PUT. Then, double sliding windows were defined as the guard window and the outer window. The size of the outer window defined the k-nearest neighbours, and the size of the guard window was defined as the G-nearest neighbours. The pixel spectra between the two windows were used to calculate statistical parameters, such as the mean vector and the covariance matrix. G represents the number of anomalous pixels with similar spectra in the whole image scene.

It is worth noting that the radii of these windows are variable because the density is different throughout the entire manifold. Theoretically, taking into account the requirements of the minimum number of pixels for the detection algorithm, the entire number of pixels between the two windows should form a linear manifold. Thus, k and G should be mutative if the density of the manifold is different everywhere. To estimate an appropriate k is very complex, while the efficiency of computing k is low. This is in order to determine whether or not the various neighbourhood areas of every pixel are linear manifolds. Moreover, the constant values have been chosen to make a convenient calculation and comparison (Parzen 1962; Cremers, Timo, and Christoph 2002).

3.4. Spectral space window LPAD (SSW-LPAD)

In the SSW-LPAD, the LPAD algorithm is implemented on sliding windows in spectral space. The LPAD algorithm was derived from the assumption that the geometric structure of dataset is a linear manifold (Li, Wang and Zheng 2014; Bachega, Theiler, and Bouman 2011; Borghys et al. 2012). If an image dataset is considered to be a linear manifold, anomalies will not be included in this manifold because of spectral independence. The distances along the special direction could be used to evaluate whether the points are anomalies or background data. The purpose of the LPAD was to find these particular directions.

First, a mathematical description of the linear manifold in high-dimensional space was considered. It could be developed from the general plane equation in a 3-dimensional space. Suppose there was a dataset Ω which is distributed in a linear manifold in l-dimensional space, all of the vectors $\boldsymbol{\alpha} = (a_1, a_2, \ldots a_l)$ in dataset Ω should meet the following formula:

$$x_1 a_1 + x_2 a_2 + \cdots + x_l a_l + x_{l+1} = 0, \tag{4}$$

where the nonzero vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_{l+1})$ is the normal vector of this linear manifold. With respect to the linear algebra, Equation (3) means all the vectors $\boldsymbol{\alpha}$ of Ω could linearly express each other. After removing the mean vector of the dataset, the matrix form of Equation (4) could be rewritten as:

$$\mathbf{A}x = 0, \tag{5}$$

where $\mathbf{A}$ is an $N \times l$ matrix. The row vectors in $\mathbf{A}$ represent l-dimensional spectra of the dataset. N is the number of samples in the dataset. $\boldsymbol{x}$ is the normal vector.

Considering the hyperspectral image dataset, the main parts of the geometrical structure would be composed of background data, and it would be a linear manifold. The anomalies would be located outside the manifold. Therefore, the anomalies could be distinguished from the manifold efficiently if we could find a decision boundary along the normal direction of the manifold properly. In fact, Equation (5) is a homogeneous linear equation. In this case, the extraction of normal vectors is converted to solve homogeneous linear equations. Matrix $\mathbf{A}$ is a singular matrix because the row vectors could be denoted as a linear combination of the others. Therefore, the solution of Equation (5) is not unique and has a solution set. The Moore–Penrose pseudoinverse (Barata and Hussein 2012) was used to calculate that. If some rows of $\mathbf{A}$ are linearly independent, the solution set (in terms of least squares) of Equation (5) can be derived from the Moore–Penrose pseudoinverse:

$$x = (\mathbf{E} - \mathbf{U}^{+}\mathbf{U})\xi, \tag{6}$$

where $\mathbf{U}^{+}$ denotes the Moore–Penrose pseudoinverse, where $\mathbf{U}^{+} = \mathbf{U}^{\mathrm{T}}(\mathbf{U}\mathbf{U}^{\mathrm{T}})^{-1}$. $\mathbf{U}$ is a $k \times l$ matrix, it should be constructed by k independent rows of $\mathbf{A}$. $\mathbf{E}$ is the identity matrix and $\boldsymbol{\xi}$ is an arbitrary l-dimensional vector.

Because $\boldsymbol{\xi}$ is an arbitrary l-dimensional vector, the spectrum of PUT could be used to replace it. Thus, the vertical distance (along the normal direction of the manifold) between the PUT and background linear manifold would be calculated as:

$$D(\alpha) = \alpha^{\mathrm{T}}(\mathbf{E} - \mathbf{U}^{+}\mathbf{U})\alpha = \alpha^{\mathrm{T}}(\mathbf{E} - \mathbf{U}^{\mathrm{T}}\left(\mathbf{U}\mathbf{U}^{\mathrm{T}}\right)^{-1}\mathbf{U})\alpha, \tag{7}$$

where D represents the function of distance and $\boldsymbol{\alpha}$ represents the spectrum of PUT. Usually, matrix $\mathbf{U}$ could be composed of the endmembers of the background. However, for convenient calculation, the larger eigenvectors were always used to replace the background endmembers. Thus, the equation became:

$$D(\alpha) = \alpha^{\mathrm{T}}(\mathbf{E} - \mathbf{V}^{+}\mathbf{V})\alpha = \alpha^{\mathrm{T}}(\mathbf{E} - \mathbf{V}^{\mathrm{T}}\left(\mathbf{V}\mathbf{V}^{\mathrm{T}}\right)^{-1}\mathbf{V})\alpha, \tag{8}$$

where $\mathbf{V}$ is composed of several significant eigenvectors of the data covariance matrix. This detector is called low probability anomaly detection (LPAD) because it is similar to the LPD detector. The significant difference is the unit vector in LPD, and it was replaced by the spectrum of PUT in LPAD.

In this article, we implemented the LPAD in the sliding windows. For every PUT, spectra between the double windows were involved in the calculation of the covariance matrix. The number of significant eigenvectors in SSW-LPAD was the same as the intrinsic dimension of the global data.

4. Data collection and description

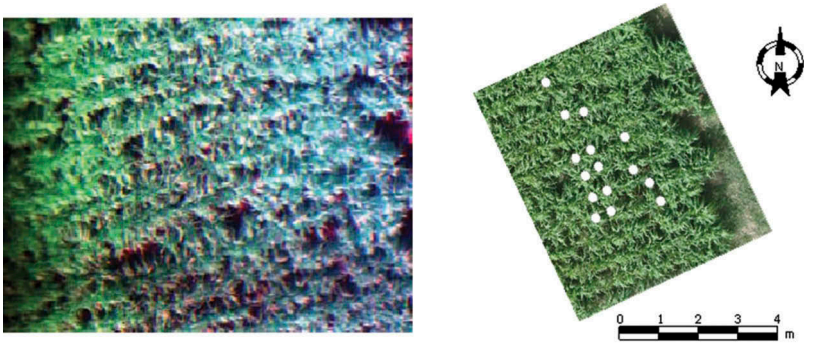
To assess the performance and compare the effectiveness of different algorithms, two groups of hyperspectral images were used in these experiments: (a) the field imaging spectrometer system (FISS) and (b) the operative modular imaging spectrometer (OMIS) that was produced by the Shanghai Institute of Technical Physics (SITP) of the Chinese Academy of Sciences. The basic parameters of the two image datasets are depicted in Table 1.

The data of FISS were collected from the vehicle platform, and it was placed 30 m from the ground. The image was 230 × 300 pixels and 86 bands were selected. The anomalies in FISS data were small planks depicted in a green colour and put in the corn field (see Figure 3). Figure 3(*a*) shows the hyperspectral image cube and Figure 3(*b*) shows the digital pictures. There are in total 15 planks, and their sizes are about 6–10 cm. In Figure 3(*b*), the white dots indicate the actual positions of the planks.

The data of OMIS were achieved by the actual flight of the airborne platform. The image size is 250 × 250 pixels. Because of the low signal noise ratio, the bands of vapour absorption have been removed; only 80 bands have been processed in this research. The original image and ground truth data are shown in Figure 4. The anomalies have two parts: (a) the vehicles in the centre of the image (i.e. labelled from T1 to T7 in Figure 4(*b*)), and (b) the eight white tiles in the top of the image.

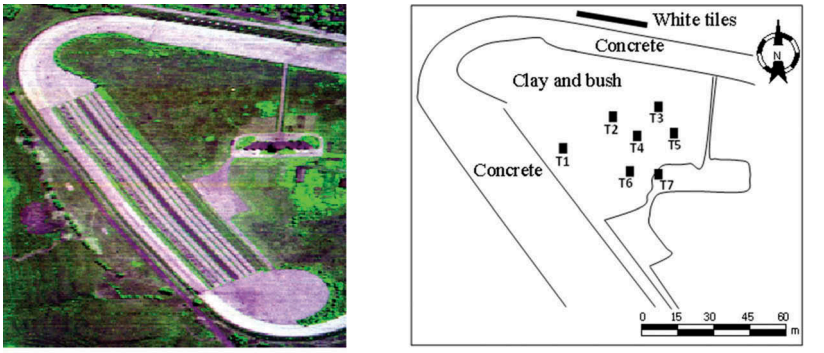
Table 1. Main parameters of the two datasets.

	FISS dataset	OMIS dataset
Spectral coverage (μm)	0.4–0.9	0.4–12.5
Spectral resolution (nm)	1.4	10–500
Number of bands	344	128
Instantaneous field of view (mrad)	1	3–5
Sensor height	30	1000



(a) (b)

Figure 3. FISS data and target positions, southwest, Beijing, China (116° 11′ E, 39° 52′ N). (a) False-colour image (R, band 167; G, band 243; B, band 340); (b) digital photo.



(a) (b)

Figure 4. OMIS data and anomaly positions, southwest, Beijing, China (116°10′ E, 39°51′N). (a) False-colour image (R, band 125; G, band 45; B, band 15); (b) anomaly positions.

5. Results and comparative analysis

Three groups of algorithms were compared by using two hyperspectral images. First were global detectors, including the global RX and the global LPAD; second were detectors based on spatial windows, including LRX and LOSP; third were detectors based on spectral space windows, including SSW-RX and SSW-LPAD. For LRX, LOSP, SSW-RX, and SSW-LPAD, the size of the outer window was set as 15 × 15 pixels; the guard window was 5 × 5 pixels. The number of superlative eigenvectors was nine for FISS data and six for OMIS data. Figure 5 shows the projected image of six algorithms for FISS data. Figure 6 shows the projected image of six algorithms for OMIS data.

The thresholds were calculated by exploiting a constant false alarm rate (CFAR) algorithm in projected images. The experimental receiver operating characteristic (ROC) curves (Matteoli, Diani, and Corsini 2010) were calculated to assess the performance. They have been derived by counting the number of targets accurately detected, and the corresponding number of false alarms (see Figure 7). CFAR used the lognormal distribution. Figure 7 shows the six ex-ROC curves for FISS and OMIS data.

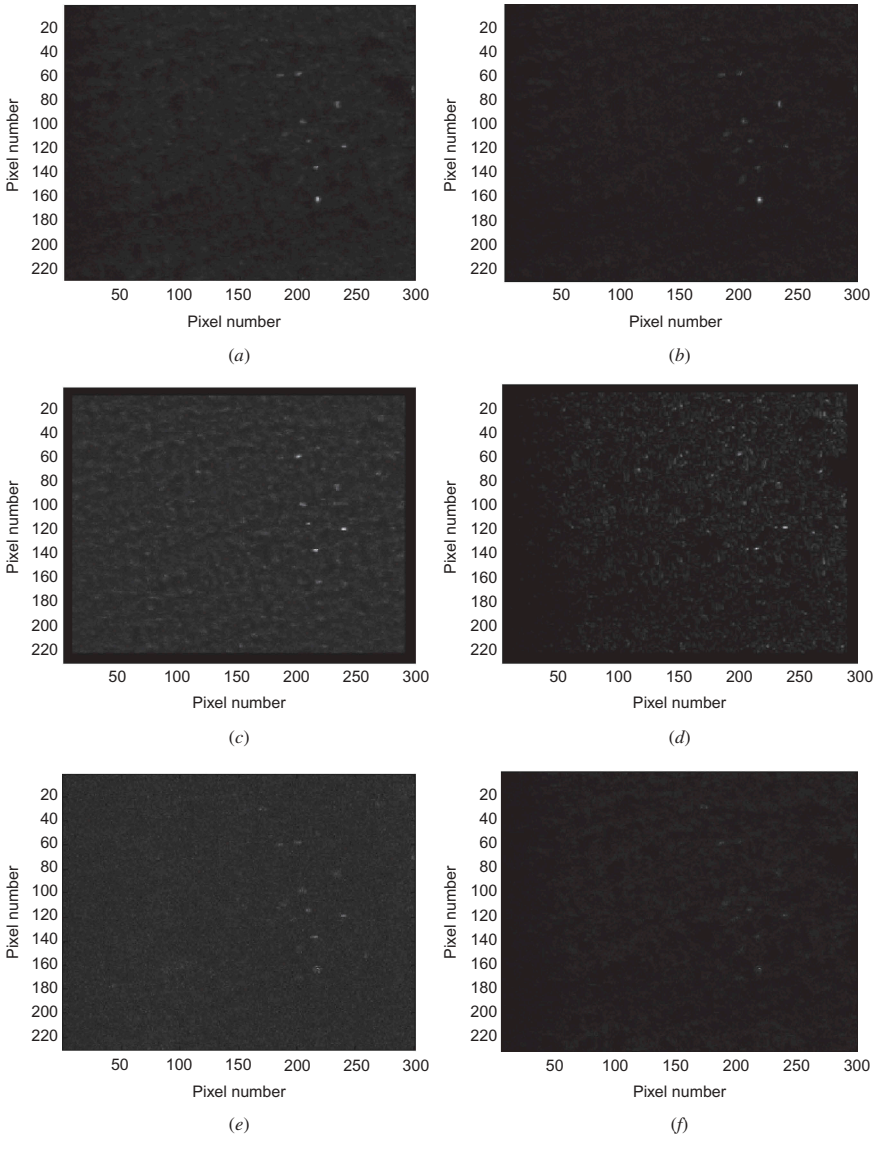


Figure 5. FISS data projected images of six algorithms. (*a*) Global RX; (*b*) global LPAD; (*c*) LRX; (*d*) LOSP; (*e*) SSW-RX; (*f*) SSW-LPAD.

The influence of the two parameters (*K* and *G*, the size of the guard window and outer window) on detection performances are depicted in Figure 8. The SSW-LPAD algorithm and OMIS dataset were used here. Usually, guard window *G* could be determined with the targets of spatial size and image resolution.

First, on the basis of the fixed *K*, the experimental ROC curves on different *G* were calculated, such as in Figure 8(*a*). Obviously, *G* should not be too large. At the same time, it should not be too small. Since the spectra are mixed, there would be many pixels that have a similar spectrum as the targets in the image. If *G* is small, these pixels will involve in the orthogonal projection, and the separation between targets and background will decrease.

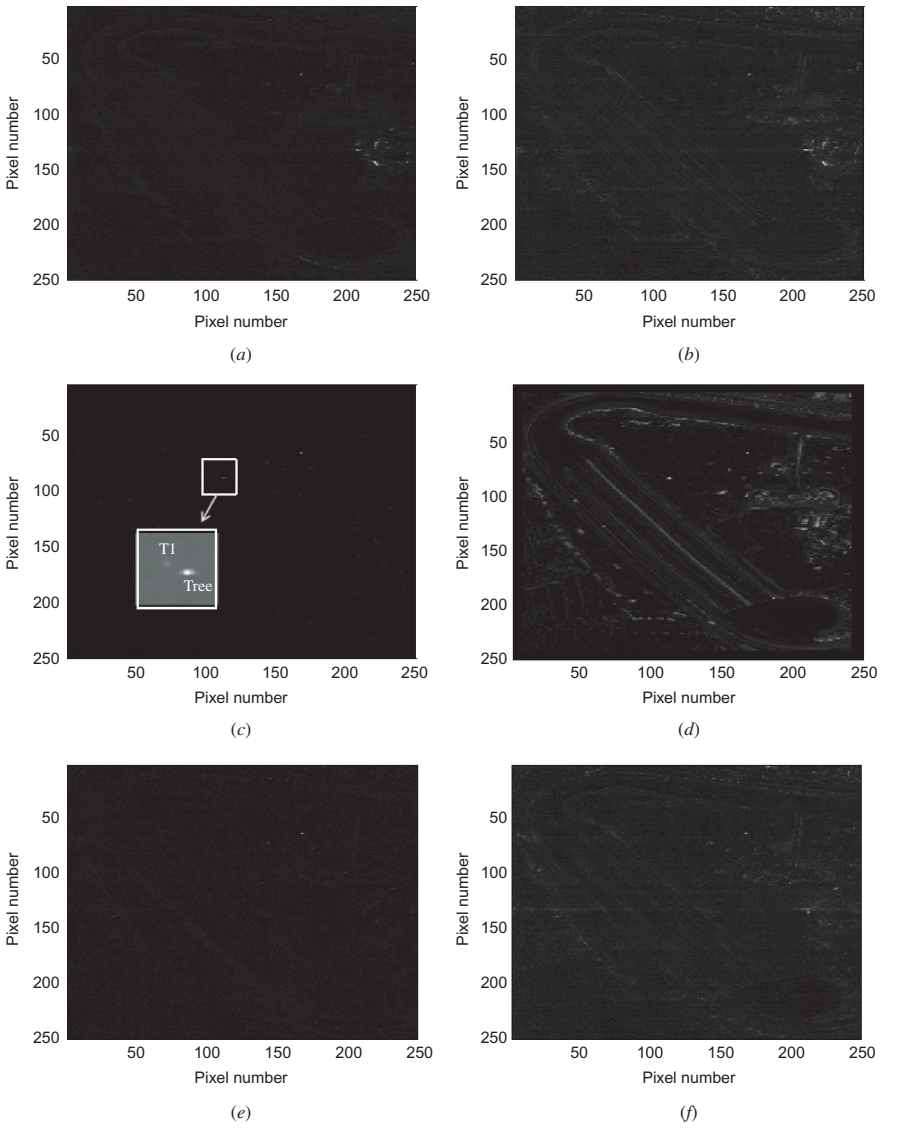


Figure 6. OMIS data projected images of six algorithms. (*a*) Global RX; (*b*) global LPAD; (*c*) LRX; (*d*) LOSP; (*e*) SSW-RX; (*f*) SSW-LPAD.

On the other hand, the experimental ROC curves on different K were calculated based on the fixed G, such as in Figure 8(*b*). Different K were tested, and only six curves are shown here. For the first few easily detected anomalies, six curves were similar. However, for the last few weak anomalies, the performances of the detector with K greater than 200 are more promising to have a better output. It meant that outer windows should be big enough to form the local linear manifolds (i.e. at least 5–6 times more than the G). A small K would lead to the decrease in detection performance. However, the larger value of K was not necessary. In Figure 8(*b*), the detection performances of detectors with K greater than 200 are similar.

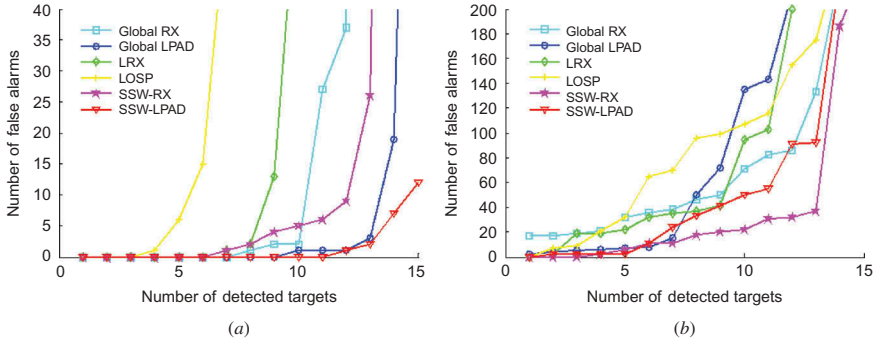


Figure 7. Experimental ROC curves of six algorithms of FISS and OMIS data. (*a*) Experimental ROC curves of FISS data; (*b*) experimental ROC curves of OMIS data.

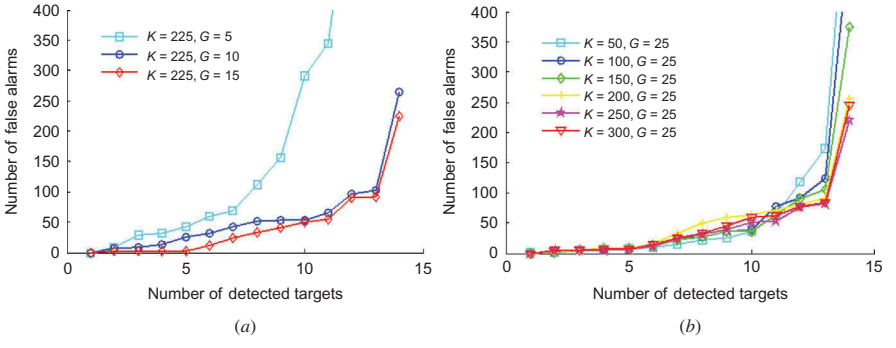


Figure 8. Experimental ROC curves of SSW-LPAD algorithms and OMIS dataset based on different K (the size of the guard window) and G (the size of the outer window). (*a*) Experimental ROC curves of different G; (*b*) experimental ROC curves of different K.

The experiments and analyses in this study have revealed that

(1) The detectors on spectral space windows (SSW-RX and SSW-LPAD) outperformed the global and spatial local detectors. They were better in background suppression, primarily for the SSW-RX detector (see Figure 5(*e*) and Figure 6(*e*)). The texture of the background almost disappeared in the projected images. Optimal detection results of FISS data were SSW-LPAD and SSW-RX for OMIS data. In addition, (2) among the detectors on the unstructured model, SSW-RX outperformed global RX and LRX, especially for weak anomaly detection. However, we should be cautious in handling the trade-offs between the accuracy of statistical parameter estimation in local data and rationality of the statistical model in global data. We should make a proper choice between the global and local algorithms, for example, in Figure 7(*a*), the global RX had a better performance than LRX for the first few targets. This is because the terrain of the FISS data was almost unique and the estimation of statistical parameters in global RX was better than LRX. This is the well-known case of the 'curse of dimensionality' (Koppen 2000). Moreover, (3) among the detectors on a structured model, SSW-LPAD outperformed the global LPAD and LOSP, especially

for FISS data. All targets could be detected in the case of few false alarms (see Figure 7(*a*)). In these algorithms, LOSP showed unsatisfactory results (see Figures 5(*d*) and 6(*d*)). The main reason was that using a mean spectrum was not good enough to suppress background energy in the windows. According to the properties of the manifolds, local spatial data may not necessarily coincide with the one-dimensional linear manifold because of the complexity of the terrain. Therefore, LOSP would obtain better results if more endmembers of the background could be used in the projections. Finally, (4) although performances of SSW-AD were better than spatial local detectors, it was more outstanding in the projected image than others for the particular target, such as target T1 in Figure 6(*c*). This case indicated that the LRX could distinguish the targets whose spectra are similar to the spectral neighbours but different from the spatial neighbours. That also meant, for the SSW-LPAD algorithm, an inappropriate neighbourhood of the local manifold may lead to failure of detection.

In summary, there were better performances for the local detectors based on spectral space windows. This case showed that it was reasonable and adequate to handle the nonlinear manifold by using the detectors based on spectral space windows. These detectors could improve the detection of weak anomalies hidden in the nonlinear structures and decrease the false alarms. However, there was a critical limitation we would encounter. The computational efficiency of SSW-AD was low. The calculation of the nearest neighbour's list was very time-consuming. The time complexity was $O((M \times N)^2)$, where O was the function of the time complexity and M and N were the image sizes. Thus, this paper recommends motivating further research to find a proper way to speed up the aforementioned limitation.

6. Conclusion

Algorithms of SSW-AD have been derived from geometric structure features of low-dimensional manifolds. By applying the local linear concepts in nonlinear manifolds, we establish the sliding window for neighbours in spectral space and implemented detectors based on linear projection in these local areas. Thus, the problem of global nonlinear data processing was converted into a local linear way; it improved the performance of detection. In this article, two local anomaly detection algorithms were presented: the SSW-RX and the SSW-LPAD. The experimental results and comparative analysis based on real hyperspectral image data demonstrated the effectiveness of the proposed algorithms. At the same time, it also validated the improvement of local linear concepts when dealing with the nonlinear manifold data in anomaly detection.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Natural Science Foundation of China [grant number 40901216].

Reference

Agovic, A., A. Banerjee, A. Ganguly, and V. Protopopescu. 2007. "Anomaly Detection in Transportation Corridors using Manifold Embedding." Proceedings of the First International Workshop on Knowledge Discovery from Sensor Data, ACM KDD Conference, San Jose, CA. http://www-users.cs.umn.edu/~aagovic/kdd-anomaly2.pdf.

Ahlberg, J., I. Renhorn, and T. Forskningsinstitut. 2004. *Multi-and Hyperspectral Target and Anomaly Detection*. Sweden: Linköping: FOI.

Bachega, L. R., J. Theiler, and C. A. Bouman. 2011. "Evaluating and Improving Local Hyperspectral Anomaly Detectors." Applied Imagery Pattern Recognition Workshop (AIPR), 1–8, doi:10.1109/AIPR.2011.6176369 IEEE.

Barata, J. C. A., and M. S. Hussein. 2012. "The Moore–Penrose Pseudoinverse: A Tutorial Review of the Theory." *Brazilian Journal of Physics* 42: 146–165. doi:10.1007/s13538-011-0052-z.

Borghys, D., I. Kasen, V. Achard, and C. Perneel. 2012. "Comparative Evaluation of Hyperspectral Anomaly Detectors in Different Types of Background." Proceedings SPIE 8390, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII, Baltimore, MD, April 23.

Caefer, C. E., J. Silverman, O. Orthal, D. Antonelli, Y. Sharoni, and S. R. Rotman. 2008. "Improved Covariance Matrices for Point Target Detection in Hyperspectral Data." *Optical Engineering* 47 (7). doi:10.1117/1.2965814.

Chang, C.-I.. 2005. "Orthogonal Subspace Projection (OSP) Revisited: A Comprehensive Study and Analysis." *IEEE Transactions on Geoscience and Remote Sensing* 43 (3): 502–518. doi:10.1109/TGRS.2004.839543.

Chang, C.-I., and -S.-S. Chiang. 2002. "Anomaly Detection and Classification for Hyperspectral Imagery." *IEEE Transactions on Geoscience and Remote Sensing* 40 (6): 1314–1325. doi:10.1109/TGRS.2002.800280.

Chao, D., Z. Huijie, and W. Wei. 2009. "Hyperpectral Image Anomaly Detection Based on Local Orthogonal Subspace Projection." *Optics and Precision Engineering* 17 (8): 2004–2010.

Cremers, D., K. Timo, and S. Christoph. 2002. "Nonlinear Shape Statistics in Mumford–Shah Based Segmentation." *Computer Vision – ECCV, LNCS* 2351: 93–108.

Du, B., and L. Zhang. 2014a. "A Discriminative Metric Learning Based Anomaly Detection Method." *IEEE Transactions on Geoscience and Remote Sensing* 52 (11): 6844–6857. doi:10.1109/TGRS.2014.2303895.

Du, B., and L. Zhang. 2014b. "Target Detection Based on a Dynamic Subspace." *Pattern Recognition* 47 (1): 344–358. doi:10.1016/j.patcog.2013.07.005.

Harsanyi, J. C., W. Farrand, and C. I. Chang. 1994. "Detection of Subpixel Signatures in Hyperspectral Image Sequences." Proceeding of American Society for Photogrammetry and Remote Sensing, Reno, NV, 236–247.

Huck, A., and M. Guillaume. 2010. "Asymptotically CFAR-Unsupervised Target Detection and Discrimination in Hyperspectral Images With Anomalous-Component Pursuit." *IEEE Transactions on Geoscience and Remote Sensing* 48 (11): 3980–3991.

Koppen, M. 2000. "The Curse of Dimensionality." 5th Online World Conference on Soft Computing in Industrial Applications (WSC5)." http://yaroslavvb.com/papers/koppen-curse.pdf.

Li, Z. Y., L. L. Wang, and S. Y. Zheng. 2014. "Applied Low Dimension Linear Manifold in Hyperspectral Imagery Anomaly Detection." *Proceeding of SPIE* 9142: 91421–91429.

Matteoli, S., M. Diani, and G. Corsini. 2010. "A Tutorial Overview of Anomaly Detection in Hyperspectral Images." *IEEE Aerospace & Electronic Systems Magazine* 25 (7): 5–28. doi:10.1109/MAES.2010.5546306.

Messinger, D. W., and F. Chester. 2011. "A Graph Theoretic Approach to Anomaly Detection in Hyperspectral Imagery." 3rd Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lisbon, 1–4.

Nasrabadi, N. M.. 2014. "Hyperspectral Target Detection: An Overview of Current and Future Challenges." *IEEE Signal Processing Magazine* 31 (1): 34–44. doi:10.1109/MSP.2013.2278992.

Parzen, E. 1962. "On Estimation of a Probability Density Function and Mode." *The Annals of Mathematical Statistics* 33: 1065–1076. doi:10.1214/aoms/1177704472.

Ramaswamy, S., R. Rastogi, and K. Shim. 2000. "Efficient Algorithms For Mining Outliers From Large Data Sets." In *SIGMOD '00 Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, 427–438. New York: ACM. doi:10.1145/342009.335437.

Reed, I. S., and X. Yu. 1990. "Adaptive Multiple-band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution." *IEEE Transactions on Acoustics Speech and Signal Processing* 38 (10): 1760–1770. doi:10.1109/29.60107.

Schaum, A. 2007. "Hyperspectral Anomaly Detection: Beyond RX." Proceeding of SPIE6565: Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery XII, Orlando, FL, April 9. doi:10.1117/12.718789.

Seung, H., and D. Lee. 2000. "The Manifold Ways of Perception." *Science* 290 (5500): 2268–2269. doi:10.1126/science.290.5500.2268.

Stein, D. W. J., S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker. 2002. "Anomaly Detection from Hyperspectral Imagery." *IEEE Signal Processing Magazine* 19 (1): 58–69. doi:10.1109/79.974730.

Verveer, P. J., and R. P. W. Duin. 1995. "An Evaluation of Intrinsic Dimensionality Estimators." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (1): 81–86. doi:10.1109/34.368147.

Yver, B., and R. Marion. 2007. "A Theoretical Framework for Hyperspectral Anomaly Detection Using Spectral and Spatial A Priori Information." *IEEE Geoscience and Remote Sensing Letters* 4 (3): 436–440. doi:10.1109/LGRS.2007.898080.

Zhao, M., and V. Saligrama. 2009. "Anomaly Detection with Score Functions Based on Nearest Neighbor Graphs." In *Proceedings of Advances in Neural Information Processing Systems*, Vancouver, BC, December 7, 2250–2258.