

Nonparametric Spectral-Spatial Anomaly Detection

M. Imani*

Department of Electrical Engineering, Tarbiat Modares University, Tehran.

Received 13 January 2018; Revised 31 August 2018; Accepted 02 March 2019

*Corresponding author: maryam.imani@modares.ac.ir (M. Imani).

Abstract

Due to the abundant spectral information contained in the hyperspectral images, they are suitable data for anomalous targets detection. The use of spatial features in addition to spectral ones can improve the anomaly detection performance. An anomaly detector, called nonparametric spectral-spatial detector (NSSD), is proposed in this work, which utilizes the benefits of spatial features and local structures extracted by the morphological filters. The spectral-spatial hypercube has obtained high dimensionality. Thus, accurate estimates of the background statistics in small local windows may not be obtained. Applying conventional detectors such as local Reed Xiaoli (RX) to the high dimensional data is not possible. To deal with this difficulty, a non-parametric distance, without any need to estimate the data statistics, is used instead of the Mahalanobis distance. According to the obtained experimental results, the detection accuracy improvement of the proposed NSSD method compared to Global RX, Local RX, weighted RX, linear filtering based RX (LF-RX), background joint sparse representation detection (BJSRD), Kernel RX, sub-space RX (SSRX), and RX and uniform target detector (RX-UTD), on average, is 47.68%, 27.86%, 13.23%, 29.26%, 3.33%, 17.07%, 15.88%, and 44.25%, respectively.

Keywords: *Spectral-spatial Information, Anomaly Detection, Hyperspectral Image, Morphological Profile.*

1. Introduction

Many narrow and contiguous spectral bands are acquired by the hyperspectral imaging sensor from different portions of the spectrum (visible, near-infrared and mid-infrared) [1]. The spectral signature of each material in the ground surface operates as a fingerprint for it, which simplifies the discrimination between that material and others with high details [2]. This characteristic of the hyperspectral images makes them appropriate data for classification and target detection [3]-[8]. Target detection is actually a two-class classification problem. Providing an accurate target detection in the hyperspectral images has attracted lots of attention because of its importance in many civilian and military applications. Target detection can be done supervised when the spectral signature of the

interested target is determined or unsupervised when there is no knowledge about the target. The second one is called anomaly detection, where the detector tries to find pixels with different spectral signatures with respect to the associated background [9]-[10]. Generally, in the anomaly detection problems, materials with a significantly different spectrum with respect to the surrounding pixels are known as anomalies. The background pixels generally compose the dominant portion of the image, while anomalies compose the minority of it.

One of the well-known anomaly detection methods is the Reed-Xiaoli (RX) anomaly detector. RX, which uses the Mahalanobis distance, has been successfully used in many hyperspectral anomaly detection problems [11].

The RX algorithm works based on the estimate of background statistics in a global or local manner. The mean vector and covariance matrix are globally estimated in the global RX detector while the mean vector is locally estimated in the local RX. The local RX detector assumes a multivariate normal distribution for data that is not valid in many complicated background data. A non-linear version of RX, called Kernel RX, has been proposed to reduce the high-order correlation between spectral channels and improves the anomaly detection performance in non-linear data [12].

In a RX detector, the anomalies and noisy signals may contaminate the background pixels, and so, cause a non-accurate estimate of the mean vector and covariance matrix. In order to deal with this difficulty, the weighted RX and the linear filtering based RX (LF-RX) have been introduced in [13]. The weighted RX method assigns larger weights to the pixels that belong to the background with higher probability. Thus, it estimates the background statistics with a higher accuracy. The LF-RX method does a similar duty with the filtering operator.

Other different versions of RX have also been proposed. RX and uniform target detector (RX-UTD) tries to remove noise and background for improvement of the RX detector performance [14]. The sub-space RX (SSRX) detector, by assuming that the background and target pixels can be represented in different sub-spaces, uses the eigenvectors associated with the largest eigenvalues of background covariance matrix for background feature extraction [15]. The sparse representation has also been proposed to model the background clutter in the background joint sparse representation detection (BJSRD) [16].

Most anomaly detection methods just use the spectral information of the hyperspectral image and ignore the valuable spatial information contained in the local neighborhood. With the contribution of neighboring pixels and incorporation of contextual information during anomaly detection, the performance of detection can be significantly improved. Although the spectral-spatial features have been used in many classification problems, they have been used

seldom in the target and anomaly detection applications.

In the pattern recognition applications domain, the parametric phrase is predicated to methods that do not use the data statistics such as mean vector (the first statistics) and covariance matrix (the second statistics). For instance, the local RX anomaly detector calculates the Mahalanobis distance containing the mean vector and covariance matrix. Therefore, local RX is called a parametric anomaly detector. In contrast to parametric methods, there are non-parametric ones, where there is no need to estimate the data statistics. In this paper, a non-parametric spectral-spatial detector (NSSD) is proposed, which uses the benefits of the morphological profile (MP) [17]-[20]. To utilize the contextual information, the extracted morphological features are added to the original spectral bands. Thus, the discriminability between anomaly and background pixels is significantly increased. The local RX detector and its different versions use the mean vector and covariance matrix in the local windows. Actually, they utilize the representatives of each local region, and so miss some useful information. In contrast, the proposed NSSD method uses all the pixels contained in a local window instead of using the mean vector and covariance matrix. Thus, it utilizes the information of all pixels contained in the neighborhood region. Moreover, NSSD uses both the spectral and spatial features to separate the anomaly signals from the background clutter. Note that because of two reasons, the use of spectral-spatial features in a local RX detector is not efficient and reasonable: 1- because of high dimensionality of the spectral-spatial hypercube; the mean vectors and covariance matrices may not be accurately estimated, and the singularity problem may occur. 2- Estimation of covariance matrices with increased dimensionality in each local window around each testing pixel significantly increases the computation time.

The experiments done on a hyperspectral image acquired by Air-borne Visible/Infrared Imaging Spectrometer (AVIRS) shows the superior performance of NSSD compared with other anomaly detectors with a reasonable computation time.

2. NSSD

Most anomaly detection methods just use the spectral signatures of data samples, and ignore the spatial information contained at the neighborhood locations. Spatial information consisting of geometrical features and local structures in a scene has a great importance for analyzing images acquired by remote sensors. The spatial information can be exploited for modeling, extraction, and detection of objects in the scene. Thus in an anomaly detection problem, the use of spatial information in addition to spectral one can significantly increase the discriminability between the anomaly and background pixels. Different methods have been proposed for extraction of spatial features in the hyperspectral images. The morphological transformations provide a multi-scale decomposition of an image. They are known as one of the most efficient methods for spatial feature extraction in the hyperspectral images. Thus, in this paper, we add the spatial features extracted by the morphological filters to the original spectral bands.

To provide the morphological profile (MP) from the hyperspectral images (HS), the opening and closing filters by reconstructions are applied to the first principal components (PCs) of data. The principal components of data are extracted by the principal component analysis (PCA) method and an MP is acquired from each PC. To this end, the covariance matrix of HS, Σ , is estimated. For extraction of n PCs from HS, n eigenvectors of Σ corresponding to the largest eigenvalues compose the PCA projection matrix. From each PC image, an MP is achieved. To this end, m opening and closing operators by reconstruction are applied to the obtained PC image to provide an MP with $p = 2m + 1$ spatial features, as follows:

$$MP_m(PC_i) = \{\varphi_1^*(PC_i), \dots, \varphi_m^*(PC_i), PC_i, \gamma_1^*(PC_i), \dots, \gamma_m^*(PC_i)\} \quad ; i = 1, \dots, n \quad (1)$$

where, γ_λ^* and $\varphi_\lambda^*(x)$ are the morphological opening and closing operators by reconstruction, respectively, and λ is the structure element. By

stacking the $MP_m(PC_i); i = 1, \dots, n$ on the HS, the spectral-spatial cube of the hyperspectral image, denoted by MP+HS, is obtained. The dimensionality of the stacked MP+HS is high, and so, because of two reasons, applying local RX to it is not reasonable. First, because of an increased dimensionality, the accurate estimate of mean vectors and covariance matrix in the local windows containing a limited number of samples may not be obtained, and the singularity problem may occur. Secondly, because of the increased dimensionality of mean vectors and covariance matrices, the computation time increases significantly. To deal with these problems, the proposed NSSD method uses a non-parametric distance instead of the parametric Mahalanobis distance. NSSD uses all pixels of the local window instead of the background statistics (mean vector and covariance matrix), as follows:

$$d_i = \sum_{j=1}^s (\mathbf{x}_i - \mathbf{x}_{ij})^T (\mathbf{x}_i - \mathbf{x}_{ij}) \quad (2)$$

where, s is the number of pixels in the local window and \mathbf{x}_{ij} is the j th neighbor of sample under test \mathbf{x}_i . The samples \mathbf{x}_i and \mathbf{x}_{ij} are the spectral-spatial feature vectors of MP+HS associated with the i th central pixel and j th neighbor of the i th central pixel, respectively. In this way, the information of all pixels of data, instead of just their representatives, i.e., mean vector and covariance matrix, are used to model the local background. Thus, the performance of anomaly detection is improved. The benefits of the proposed NSSD method can be represented as follows:

1-NSSD uses both the spectral and spatial information in the hyperspectral images for anomaly detection.

2-Estimation of mean vector and covariance matrix, in a small local window may not be accurate. Thus, NSSD uses all pixels of the local window instead of statistics estimate. It benefits from the information of all samples in the local region instead of just their representatives (mean vector and covariance matrix).

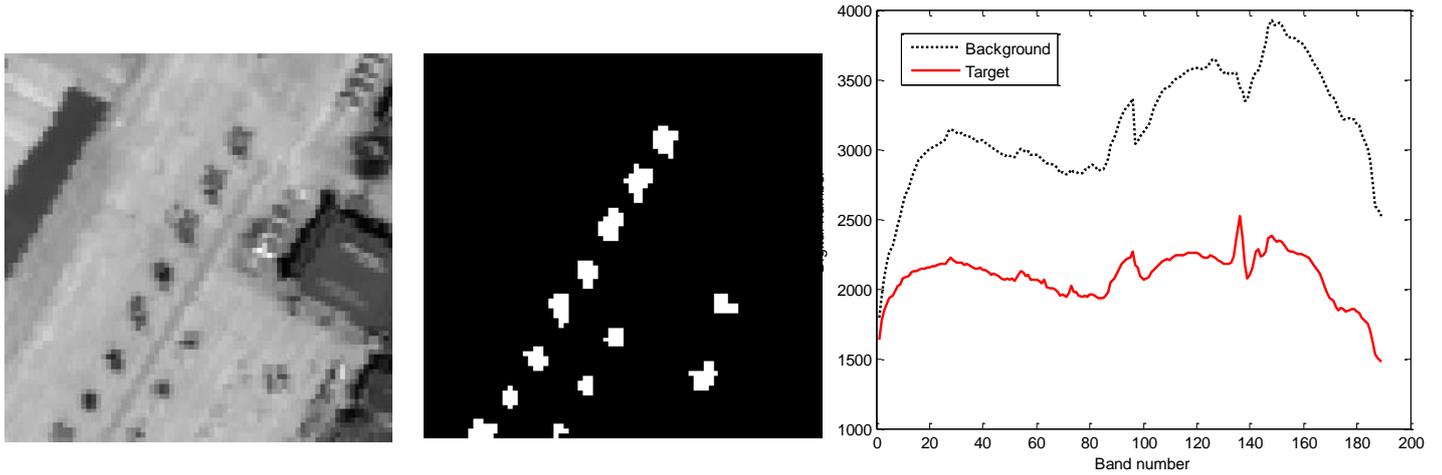


Figure 2. The grey level image, target map, and spectral signatures of target and background in real AVIRIS image.

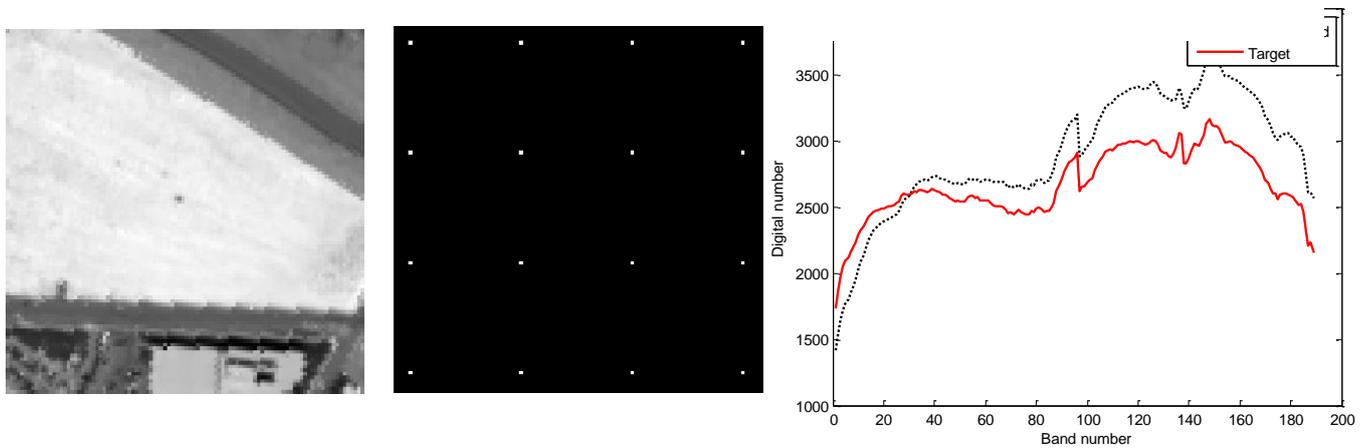


Figure 1. The grey level image, target map, and the spectral signatures of target and background in synthetic image.

3-NSSD does not need to estimate the mean vectors and covariance matrices for each local window around each pixel under test. Thus, it is much faster than many other detectors such as local RX.

Generally, the spectral and spatial features contained in MP+HS allows more discrimination between background and anomaly pixels than the spectral information contained in HS.

3. Experiments

The performance of the NSSD method was evaluated compared to global RX, local RX, weighted RX, LF-RX, BJSRD, Kernel RX, SSRX, and RX-UTD. MATLAB R2014a installed on a Laptop with windows 7 containing 64 bit operation system with Intel Core™ i5 and 4 GB Memory was used for doing the experiments. The receiver operating characteristic (ROC) curve and area under the curve (AUC) were used to

quantitatively assess the performance of anomaly detection [21]. To qualitatively assess anomaly detectors, the detection maps were also illustrated. To assess the performance of the anomaly detectors, the detection maps were compared with the target maps. By counting the correct and false detected samples, the false alarm rate (FAR) and the probability of detection (PD) were calculated as follows:

$$FAR = \frac{N_{fd}}{N}, PD = \frac{N_{cd}}{N_t} \quad (3)$$

where, N is the number of all hyperspectral pixels, N_t indicates the total number of anomaly samples, N_{fd} and N_{cd} are the number of falsely detected anomalous pixels and the number of correctly detected ones, respectively.

Table 1. Performance of NSSD and other anomaly detectors in terms of AUC and computation time in real AVIRIS image. AUC is area under ROC curve, where a higher value for AUC means a higher detection accuracy.

Method	Global RX	Local RX	Weighted RX	LF-RX	BJSRD	Kernel RX	SSRX	RX-UTD	NSSD
AUC	46.33	68.71	81.66	65.84	87.44	86.02	80.17	48.24	93.53
Computation time	0.21	80.16	163.90	45.77	1848.21	237.70	0.24	0.24	4.68

Table 2. Performance of NSSD and other anomaly detectors in terms of AUC and computation time in synthetic image. AUC is area under ROC curve, where a higher value for AUC means a higher detection accuracy.

Method	Global RX	Local RX	Weighted RX	LF-RX	BJSRD	Kernel RX	SSRX	RX-UTD	NSSD
AUC	47.21	64.47	80.77	64.54	94.80	68.74	76.97	52.16	95.37
Computation time	0.28	136.01	276.58	75.49	4261.08	1058.18	0.35	0.30	6.00

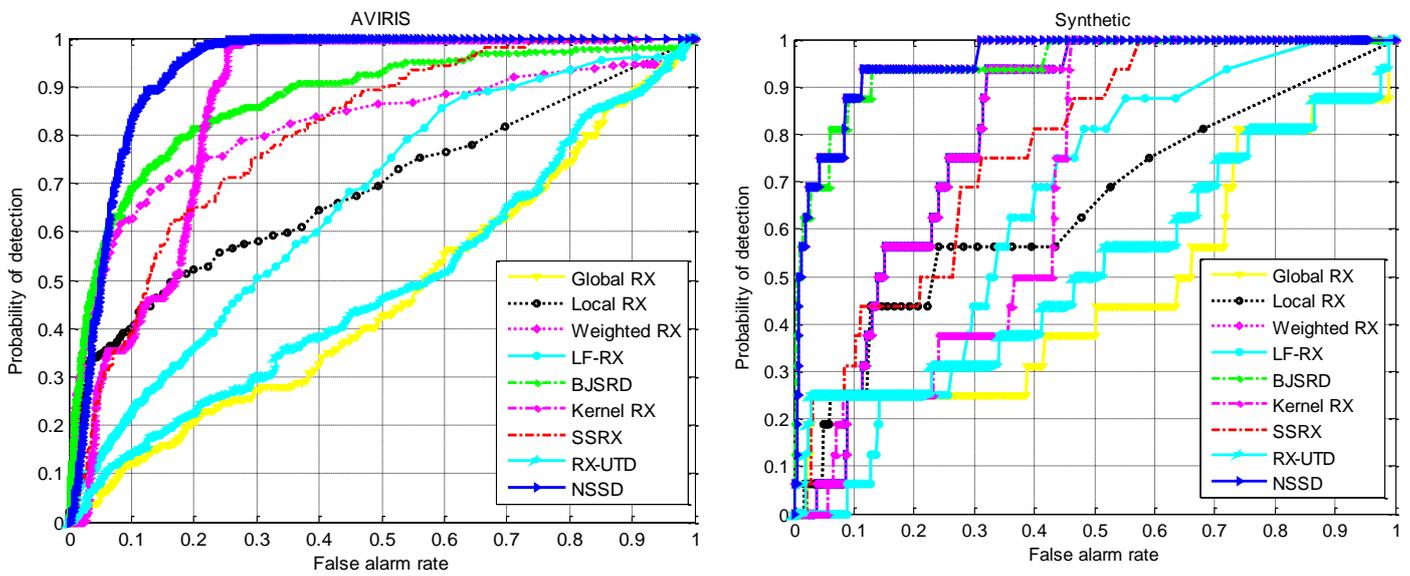


Figure 3. ROC curves for NSSD and other anomaly detectors for real AVIRIS image and the synthetic dataset. The ROC curve of each method that is nearer to the top left corner illustrates a more detection accuracy.

The ROC curve shows the relationship between FAR and PD. If an ROC curve is closer to the top left corner, it means that the detector has a more accuracy and a higher AUC value. Two hyperspectral datasets (a real data and a synthetic data) were used for evaluation of the anomaly detectors. The real hyperspectral image was acquired by AVIRS over San Diego, CA, USA which contained 224 spectral bands. The spectral channels with low SNR and also the water absorption bands were removed and 186 spectral bands were retained for doing the experiments. The wavelengths ranging within this hyperspectral image with spatial resolution of 3.5 m per pixel is from 0.37 to 2.51 μm . The whole scene of image was 400×400 and a sub-image of it containing 80×80 pixels was chosen for the experiments. The grey level illustration of band 19 of AVIRS

hyperspectral image, the target map, and the spectral signatures of anomalous target and background are shown in figure 1. The synthetic dataset was generated based on a part of real hyperspectral image acquired by the AVIRS over San Diego with 186 spectral bands. A region with 100×100 pixels was chosen to generate the synthetic data. For simulation of anomalous pixels, the target implantation method was used, where a linear mixing model was utilized [22]. To this end, a simulated anomaly pixel having spectrum z was generated by implanting a desired anomalous spectrum in a background pixel [22]:

$$z = f \cdot t + (1 - f) \cdot b \tag{4}$$

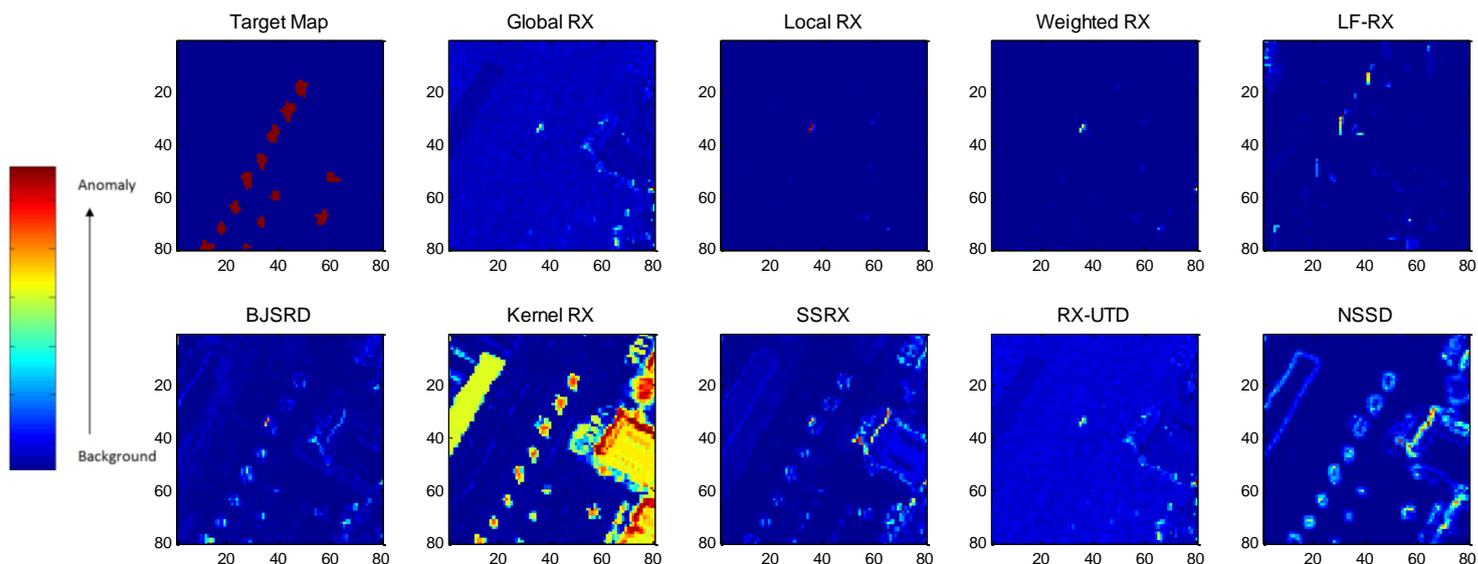


Figure 4. Detection maps obtained by NSSD and other anomaly detectors for real AVIRIS image.

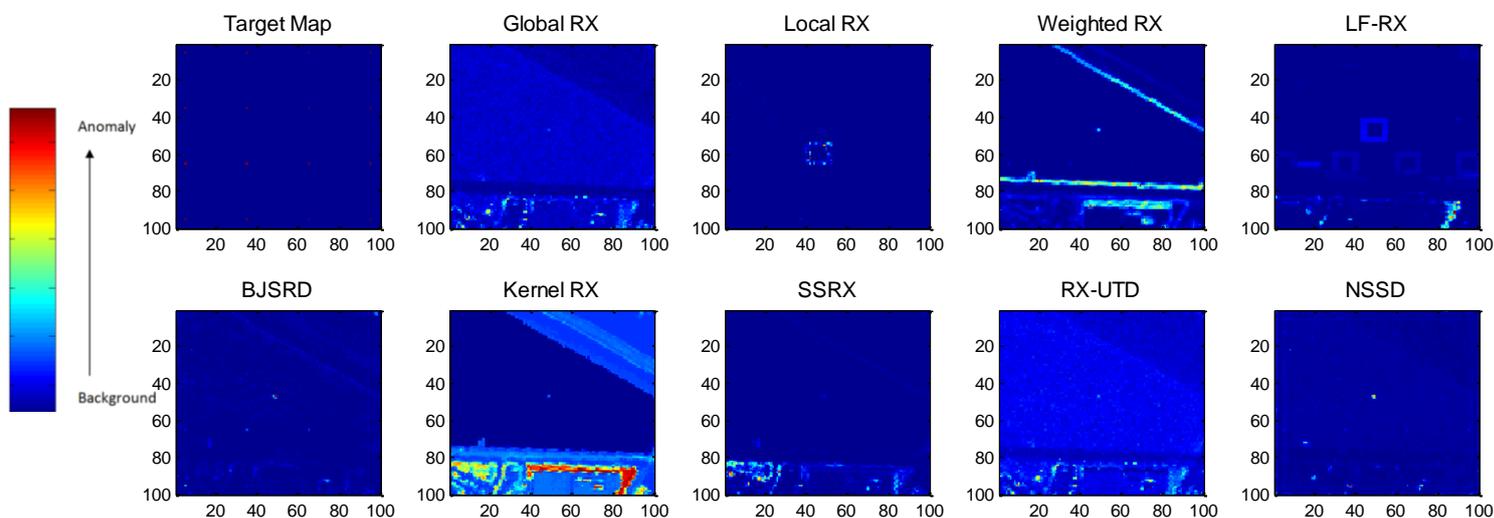


Figure 5. Detection maps obtained by NSSD and other anomaly detectors for synthetic image.

where, f determines the specified abundance fraction, and the spectrum of desired anomaly and background pixels are denoted by t and b , respectively. In this work, the synthetic data was composed of implanting 16 anomalous pixels distributed in four rows and columns by the f values of 0.05, 0.1, 0.2, and 0.4 for different rows, respectively where the anomalies were unchanged in the same row. The spectrum of the desired anomaly was chosen from a point of the San Diego hyperspectral image outside the selected area, corresponding to a plane. The image of band 19 of generated synthetic data, its target map, and the spectral signatures of target and background are shown in figure 2.

The AUC values and the computation time of NSSD and the other detectors are reported in tables 1 and 2 for the real AVIRIS hyperspectral image and the synthetic data, respectively. The ROC curves are shown in figure 3. The detection maps are also shown in figures 4 and 5. The output of each anomaly detector is a measure of belonging the pixel to the anomalies. In each detection map, the output values are shown with colors in the range of blue to red. A color nearer to red indicates that the sample belongs to anomalous, while a color nearer to blue shows that the pixel belongs to background. According to figures 4 and 5, the anomalous targets available in the target maps were detected with a higher accuracy in the detection maps obtained by the

proposed NSSD method. As we can see in the obtained results, the following conclusions can be found:

- 1) NSSD provides the highest AUC value. It detects anomalies with the lowest number of false detection points. The superior performance of NSSD is due to two reasons. First, it uses both the spectral and spatial information of the hyperspectral image instead of just using the spectral one. Second, instead of just using the representatives of each local window, i.e., the local mean vector and covariance matrix, it uses all pixels contained in the spatial neighborhood region.
- 2) After NSSD, BJSRD, which utilizes the benefits of sparse representation, obtains the highest AUC value in both datasets. Moreover, Kernel RX, which increases the separation between non-linear background and anomaly signals, obtain a high AUC value in real AVIRIS data. Also the weighted RX, which removes the undesired effect of contamination of background with anomalous pixels through a weighting manner, provides a high AUC value in the synthetic dataset.
- 3) The lowest AUC is acquired by global RX and RX-UTD because they miss the local background information for anomaly detection.
- 4) Although BJSRD provides a good detection performance because of optimizing the sparse representation objective function, it needs a high computation time. The computations of kernel space also increase the running time of Kernel RX.
- 5) NSSD not only provides the highest AUC but also is a fast method compared to other detectors such as BJSRD, Kernel RX, weighted RX, LF-RX, and local RX.

Thus, generally, NSSD is an efficient anomaly detector, which is simple and fast compared to many other popular and state-of-the-art anomaly detectors. Among the competitor anomaly detectors, BJSRD that uses the sparse representation uses the local information of other pixels. As shown in the experimental results, after the NSSD proposed method, BJSRD provides the best detection performance.

For more analysis of the separation between anomalous targets and background, a separability

diagram is used to illustrate the range of output anomaly detectors for both anomaly and background signals. Since different detectors result in a wide range of output detection values, the logarithm value of detectors outputs are plotted to do a better comparison between different separability diagrams. Individual boxes are used to show the output range of each detector for anomalous targets and background. These boxes enclose the main part of samples, where the smallest 10% and the biggest 10% are excluded. The lines at the top and bottom of each box are the extreme values. The middle line in each box is the mean of pixels, and the box encloses the middle 80% of main pixels. The position of boxes with respect together reflects the compactness and tendency of pixels distributions in each detector. In other words, they reflect the separability of anomaly and background pixels. The separability diagrams for AVIRIS and synthetic datasets are shown in figure 6. As we can see, the boxes of anomaly target and background in the proposed NSSD method are well-separated from each other. This means the good ability of NSSD in discrimination between target and background.

4. Conclusion

A nonparametric spectral-spatial anomaly detector is proposed in this paper. The proposed NSSD method extracts the local structures and spatial features of the hyperspectral images and adds them to the original spectral ones. The spectral and spatial features are given to a nonparametric distance measure to obtain the detection output. The proposed NSSD does not compute the mean vector and covariance matrix of data in each local region. Thus, it has two main advantages. First, it does not encounter the singularity problem. Second, it runs relatively fast. The better performance of NSSD compared to several anomaly detectors is shown on both the real and synthetic hyperspectral images in terms of ROC curve, AUC value, computation time, and detection maps. The fused spectral and spatial features in addition to using a non-parametric distance for calculating distances in high dimensional MP+HS hypercube provide the superior performance of NSSD.

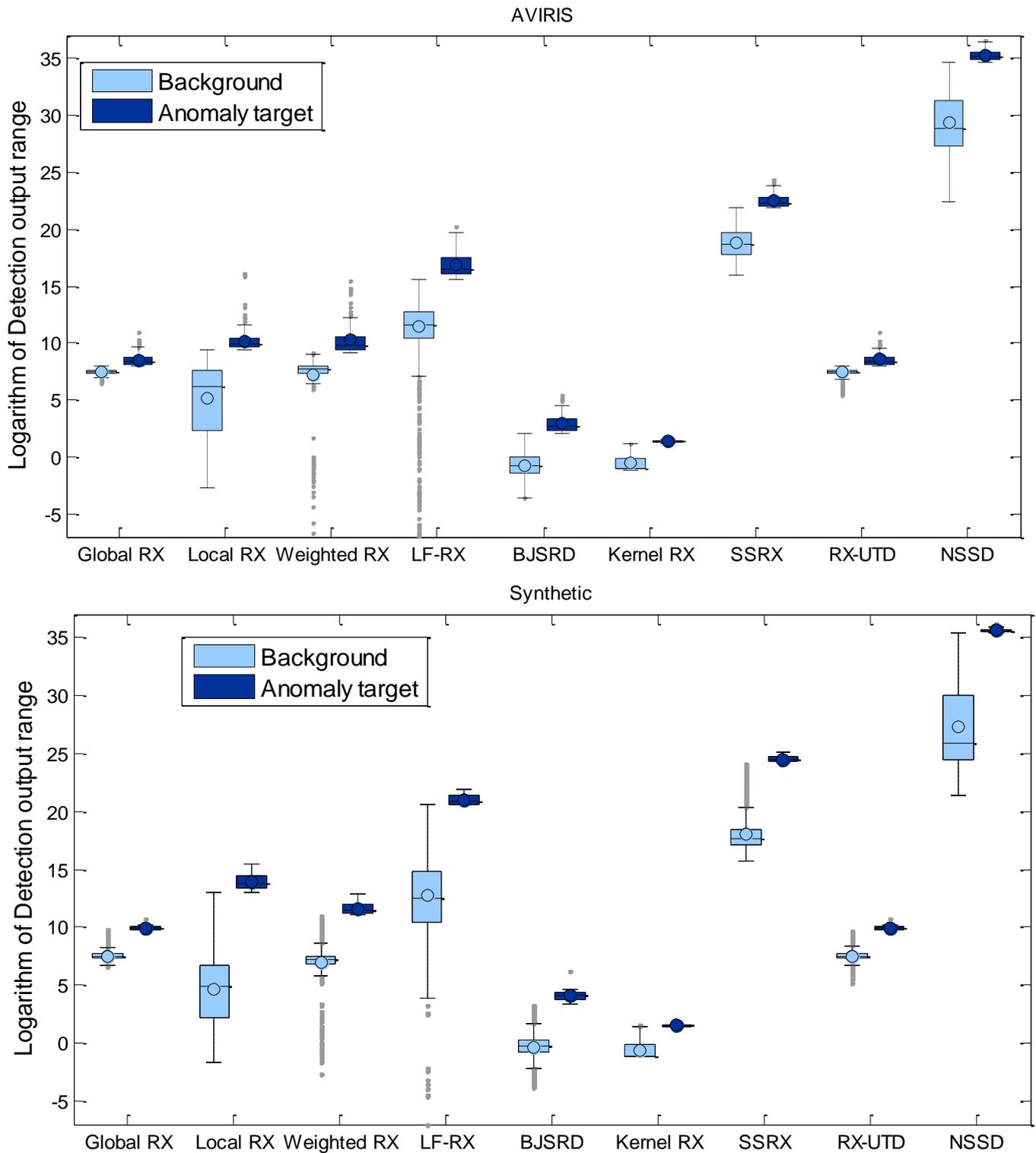


Figure 6. Separability diagrams for AVIRIS and synthetic datasets (Individual boxes show the output range of each detector for anomaly targets and background).

References

[1] Makki, I., Younes, R., Francis, C., Bianchi, T. & Zucchetti, M. (2017). A survey of landmine detection using hyperspectral imaging, ISPRS Journal of Photogrammetry and Remote Sensing, vol. 124, pp. 40-53.

[2] Imani, M. & Ghassemian, H. (2015). Feature reduction of hyperspectral images: discriminant analysis and the first principal component, Journal of AI and Data Mining, vol. 3, no. 1, pp.1-9.

[3] Heras, D. B., Argüello, F. & Quesada-Barriuso, P. (2014). Exploring ELM-based spatial-spectral

classification of hyperspectral images, *International Journal of Remote Sensing*, vol. 35, no. 2, pp. 401-423.

[4] Wang, Q., Lin, Q., Li, M. & Tian, Q. 2009. A new target detection algorithm: spectra sort encoding, *International Journal of Remote Sensing*, vol. 30, no.9, pp. 2297-2307.

[5] Imani, M. & Ghassemian, H. (2017). Local Histogram and Discriminative Learning-Based Hyperspectral Data Classification, *Remote Sensing Letters*, vol. 8, no.1, pp. 86–95.

[6] Zheng, X., Yuan, Y. & Lu, X. (2016). A target detection method for hyperspectral image based on mixture noise model, *Neurocomputing*, vol. 216, pp. 331-341.

[7] Chen, J. & Jiao, L. (2015). Hyperspectral imagery classification using local collaborative representation, *International Journal of Remote Sensing*, vol. 36, no.3, pp. 734-748.

[8] Imani, M. & Ghassemian, H. (2017). Feature extraction of hyperspectral images using boundary semi-labeled samples and hybrid criterion, *Journal of AI and Data Mining*, vol. 5, no. 1, pp. 39-53.

[9] Du, B., Zhao, R., Zhang, L. & Zhang, L. (2016). A spectral-spatial based local summation anomaly detection method for hyperspectral images, *Signal Processing*, vol. 124, pp. 115-131.

[10] Guo, Q., Pu, R., Gao, L. & Zhang, B. (2016). A novel anomaly detection method incorporating target information derived from hyperspectral imagery, *Remote Sensing Letters*, vol. 7, no.1, pp. 11-20.

[11] Reed, I. S. & Yu, X. (1990). 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.

[12] Kwon, H. & Nasrabadi, N. M. (2005). Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery, *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 388–397.

[13] Guo, Q., Zhang, B., Ran, Q., Gao, L., Li, J. & Plaza, A. (2014). 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.

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

[15] Schaum, A. P. (2007). Hyperspectral anomaly detection beyond RX, *Proc. SPIE*, 6565(656502).

[16] Li, J., Zhang, H., Zhang, L. & Ma, L. (2015). 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.

[17] Wen, J., Zhao, J. & Wang, C. (2015). Improved morphological component analysis for interference hyperspectral image decomposition, *Computers & Electrical Engineering*, vol. 46, pp. 394-402.

[18] Imani, M. & Ghassemian, H. (2017). Edge Patch Image-Based Morphological Profiles for Classification of Multispectral and Hyperspectral Data, *IET Image Processing*, vol. 11, no. 3, pp. 164-172.

[19] Li, Z., Li, L., Zhang, R. & Ma, J. (2011). An improved classification method for hyperspectral data based on spectral and morphological information, *International Journal of Remote Sensing*, vol. 32, no. 10, pp. 2919-2929.

[20] Han, M. & Zhang, C. (2017). Spectral-spatial classification of hyperspectral image based on discriminant sparsity preserving embedding, *Neurocomputing*, vol. 243, pp. 133-141.

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

[22] Xu, Y., Wu, Z., Li, J., Plaza, A. & Wei, Z. (2016). Anomaly detection in hyperspectral images based on low-rank and sparse representation, *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 1990–2000.