NOVEL HYPERSPECTRAL ANOMALY DETECTION METHODS BASED ON UNSUPERVISED NEAREST REGULARIZED SUBSPACE

Zengfu Hou [1], Yu Chen [1], Kun Tan [1, *], Peijun Du [2]

[1] Key Laboratory for Land Environment and Disaster Monitoring of NASG, China University of Mining and Technology, Xuzhou 221116, China - tankuncu@126.com

[2] School of Geographic and Oceanographic Sciences, Nanjing University

KEY WORDS: Anomaly detection, Hyperspectral imagery, Unsupervised nearest regularized subspace, Outlier Removal, Local summation, Dual-window sliding

ABSTRACT:

Anomaly detection has been of great interest in hyperspectral imagery analysis. Most conventional anomaly detectors merely take advantage of spectral and spatial information within neighboring pixels. In this paper, two methods of Unsupervised Nearest Regularized Subspace-based with Outlier Removal Anomaly Detector (UNRSORAD) and Local Summation UNRSORAD (LSUNRSORAD) are proposed, which are based on the concept that each pixel in background can be approximately represented by its spatial neighborhoods, while anomalies cannot. Using a dual window, an approximation of each testing pixel is a representation of surrounding data via a linear combination. The existence of outliers in the dual window will affect detection accuracy. Proposed detectors remove outlier pixels that are significantly different from majority of pixels. In order to make full use of various local spatial distributions information with the neighboring pixels of the pixels under test, we take the local summation dual-window sliding strategy. The residual image is constituted by subtracting the predicted background from the original hyperspectral imagery, and anomalies can be detected in the residual image. Experimental results show that the proposed methods have greatly improved the detection accuracy compared with other traditional detection method.

1. INTRODUCTION

Hyperspectral imagery contains a wealth of spectral information and is widely used in the field of target detection because of the unique advantages. Since it is difficult for researchers to obtain enough prior knowledge to characterize the statistical information of target categories, the detection without a priori spectral information of target, which is called anomaly detection, has been of significant interest(Niu and Wang 2016). Anomaly detection models the background and using the differences between pixels and the background to detect anomalous pixels which is a small quantity of pixels in the hyperspectral image whose spectral characteristics differ significantly from those of a large proportion of pixels in the hyperspectral data cube(Vafadar and Ghassemian 2017). Anomaly detection is an important application in the field of hyperspectral remote sensing, which can be widely applied to detect location of crop stress in precision farming, to find scarce minerals in geology, to analyse oil and environmental pollution, and to detect landmines for public safety (Li, Zhang et al. 2015, Taghipour, Ghassemian et al. 2016, Zhao, Du et al. 2016).

Many anomaly detection algorithms have been proposed. The classical detection algorithm called the Reed-Xiaoli (RX)(Reed and Yu 1990) detector, is developed by Reed and Xiaoli in 1990, which has been considered as the benchmark for performance evaluation of hyperspectral anomaly detectors. It is a second-order matched filtering algorithm that calculates Mahalanobis distance of testing pixel and the background to complete the anomaly detection. When the entire image is considered for background modeling, it is also called Global RX (GRX). If the RX detector estimates the background using local statistics, it is also called Local RX (LRX). In real hyperspectral imagery, the background information is very complicated and cannot be described merely using multivariate normal distribution, and it may be difficult to use the estimated covariance and mean vector to represent the background statistics because of the existence of noise and anomalies. Under the circumstances, some improved algorithms such as Weighted-RX(W-RXD)(Guo, Zhang et al. 2014) and Linear Filter-Based RXD(LF-RXD)(Guo, Zhang et al. 2014) aim at improving the probability of anomaly detection by improving the estimation of the background information. Some kernel-based detection algorithms, such as the classical nonlinear kernel RX(Kwon and Nasrabadi 2005) detection algorithm, have also achieved better detection performance than the conventional algorithms in anomaly detection.

In recent years, the method based on signal sparse representation has been applied to hyperspectral image target detection. This method aims to use background and target signals as few as possible to concisely represent the entire image information in an overcomplete dictionary composed of background information and target information (Liu, Lin et al. 2010, Chen and Chang 2013, Xu, Wu et al. 2016). However, this method merely takes advantage of spectral information and rarely gives the consideration to spatial information. It's difficult to obtain satisfactory performance when applied to the anomaly detection. Collaborative-Representation-based Detector (CRD)(Li and Du 2015) for hyperspectral anomaly detection is directly based on the concept that each background pixel can be approximately represented by its spatial neighborhoods, while anomalies cannot, and it has achieved satisfactory detection. Compared with the CRD algorithm, the proposed Collaborative Representation-Based with Outlier Removal Anomaly Detector (CRBORAD)(Vafadar and Ghassemian 2017) method for hyperspectral imagery anomaly detection has further improved performance, which removes outlier pixels that are significantly different from majority of pixels. The Local Summation Anomaly Detection (LSAD)(Du, Zhao et al. 2016) method proposed by Du et al. makes full use of spatial information of neighboring pixels of testing pixel by using the local window summation strategy, which greatly improves the accuracy of the anomaly detection. However, the statistics of the background information are easily contaminated by the abnormal target under the circumstances using a single local window for anomaly detection, which leads to a higher false alarm rate.

In this work, we propose Unsupervised Nearest Regularized Subspace-based with Outlier Removal Anomaly Detector (UNRSORAD) and Local Summation UNRSORAD (LSUNRSORAD) methods, which make full use of various local spatial distributions information with the neighboring pixels of the testing pixel through local summation dual-window sliding strategy. The residual image is constituted by subtracting the predicted background from the original imagery, and anomalies can be detected in the residual image. The detection results are assessed using Receiver Operating Characteristic (ROC)(Hanley and Mcneil 1982, Crichton 2002) curves and Area Under Curve (AUC)(Hanley and Mcneil 1982) values. By comparison with five popular and classical methods, the proposed UNRSORAD and LSUNRSORAD method provide higher detection accuracies. The rest of the paper is organized as follows: In Section 2, the proposed methods and main concepts will be presented. Section 3 shows the experimental and illustrates the superiority of the proposed methods. Finally, Section 4 draws our conclusions.

2. PROPOSED METHODS

In this section we introduce the proposed UNRSORAD and LSUNRSORAD methods. Before explaining the proposed methods, we provide a short review of Unsupervised Nearest Regularized Subspace (UNRS)(Li and Du 2014) algorithm. It is an important technique used in our proposed approach. Then we explain the outliers removal strategy, then the dual-window sliding strategy.

2.1 Unsupervised Nearest Regularized Subspace(UNRS) Algorithm

Let a given hyperspectral image dataset $X \in R^d$ be expressed as:

$$X = \{x_i\}_{i=1}^{N} \tag{1}$$

where N = the total number of the image pixels
x_i = each pixel in the image

For each testing pixel $y \in R^d$(of size $d \times 1$), we assume an appriximation $y' \in R^d$ (of size $d \times 1$) calculated via the linear combination of the surrounding selected data can be expressed as:

$$y' = X_s \alpha \tag{2}$$

where X_s = the surrounding background pixels
α = weight vector

$X_s = \{x_i\}_{i=1}^{s}$ represents the surrounding background pixels collected inside the outer windows while outside the inner window(Fig. 1), in which s is the number of chosen surrounding background pixels between dual-window, which can be calculated by:

$$s = win_{out} \times win_{out} - win_{in} \times win_{in} \tag{3}$$

where win_{out} = the size of outer windows
win_{in} = the size of inner windows

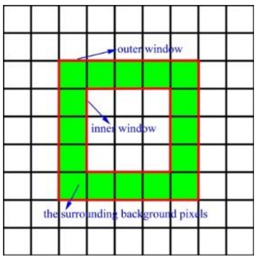

Fig. 1 dual-window

In eqn. (2), $\alpha \in R^{s\times 1}$ is weight vector. If we want to calculate the y', the weight vector α must be known. Therefore, it is converted to find the sum-to-one constraint weight vector α which minimizes the objective function as:

$$f(\alpha) = \|y - X_s\alpha\|_2^2 = \|y - \sum_{i=1}^{s} \alpha_i x_i\|_2^2 \tag{4}$$

Then, the original space is 'shifted' via centering on the focal point y , and then it can be calculated:

$$f(\alpha) = \|\sum_{i=1}^{s}(x_i - y)\|_2^2 = \|\sum_{i=1}^{s} \alpha_i x_i\|_2^2 = \alpha^T Z Z^T \alpha \tag{5}$$

where $z_i = x_i - y \in R^{d\times 1}$
ZZ^T = symmetric matrix called

$Z \in R^{s\times d}$, and $ZZ^T \in R^{s\times s}$ is a symmetric matrix called Gram matrix, denoted by G. In order to estimate the weight vector α, we consider using a Lagrange multiplier to solve the function under the sum-to-one constraint condition(Li and Du 2014). Finally, the value of α can be expressed as:

$$\alpha = \frac{\sum_{j=1}^{s} G^{-1}}{\sum_{i=1}^{s}\sum_{j=1}^{s} G^{-1}} \tag{6}$$

where i, j = the rows and columns of the matrix index

Note that matrix $G^{-1} \in R^{s\times s}$. We apply this technique by adding the regularized term to the objective function. Then, it can be modified as:

$$f(\alpha, \lambda) = \alpha^T G \alpha + \lambda \alpha^T \alpha \tag{7}$$

where λ = a constant

Thus, we convert the above problem to an equivalent problem solved by using Lagrange multiplier method. Finally, we can obtain the value of α which minimizes the new cost function:

$$\alpha = \frac{\sum_{j=1}^{s} (G + \lambda \boldsymbol{I})^{-1}}{\sum_{i=1}^{s}\sum_{j=1}^{s} (G + \lambda \boldsymbol{I})^{-1}} \tag{8}$$

where $\boldsymbol{I}$ = a identify matrix

2.2 Outliers Removal Strategy

The dual-window makes use of two windows, called inner window and outer window to capture the characteristics of targets and background respectively(Liu and Chang 2008). In order to make full use of spatial information, many anomaly detection algorithms use the dual-window, such as LRX, CRD, UNRS and so on, while the dual-window can still not rule out the influence of anomalous pixels between outer window and inner window. In order to improve the detection accuracy, we adopt outliers removal strategy(Vafadar and Ghassemian 2017) to remove the outlier pixels in the double window that are obviously different from others as shown in Fig. 2. The light green square is anomalous pixel that is significantly different from other dual-window within pixels, and the green squares are background pixels. The UNRS algorithm aims to represent testing pixels(the blue square in Fig. 2) linearly with the background pixels(the green squares in Fig. 2), while the existence of anomalous pixels will affect the accuracy of the linear representation of the test pixels. The approximation of testing pixels sample(the light blue square in Fig. 2) is generated via a linear combination from the background pixels after removing the anomalous pixels, so what is the outlier on earth? It requires to find a suitable maximum and minimum threshold based on statistical theorem. Pixels with intensity values greater than maximum threshold or smaller than minimum threshold are considered the outliers. Similar to reference (Vafadar and Ghassemian 2017), we calculate mean and standard deviation of dual-window within pixels intensities values and construct threshold values simply by:

$$\tau_{\max} = \mu + 2 \times \sigma \quad (9)$$

$$\tau_{\min} = \mu - 2 \times \sigma \quad (10)$$

where μ = the mean of the background pixels X_s
σ = the standard deviation of the background pixels X_s

$\tau_{\max}$ and $\tau_{\min}$ represent the maximum and minimum of the background pixels intensities, respectively. Pixels with intensity value greater than $\tau_{\max}$ or less than $\tau_{\min}$ are removed. Therefore, X_s can be replaced by X_s' with the size of $d \times s'$ in which s' is the number of background pixels after outliers removal, and X_s' is used to predict the testing pixel y'. Once the representation process is finished, we can obtain the residual image by:

$$r_1 = \|y - y'\|_2 = \|y - X_s' \, \alpha'\|_2 \quad (11)$$

where α' = the new weight vector after removing the outliers
r_1 = a distance value

If the distance r_1 is larger than a threshold, then y is declared as an anomalous pixel. The proposed UNRSORAD algorithm is different from the UNRS algorithm. We adopt the outliers removal strategy to rule out the influence of anomalous pixels based on the UNRS method. The overall description of the UNRSORAD algorithm is given as **Algorithm 1**

Algorithm 1 The UNRSORAD algorithm

Input: Hyperspectral Data $X = \{x_i\}_{i=1}^{N}$, window size (win_{out}, win_{in}) and the parameter λ

for all pixels **do**

1) For each testing pixel y, a 2D matrix constitutes based on the dual-window within pixels by $X_s = \{x_i\}_{i=1}^{s}$;
2) Removing outlier pixels in the matrix X_s;
3) Calculating the weight vector α' by eqn. (8);
4)Calculating the distance measurement by eqn. (11) and obtaining the detection result;

end for

Output: Anomaly detection map

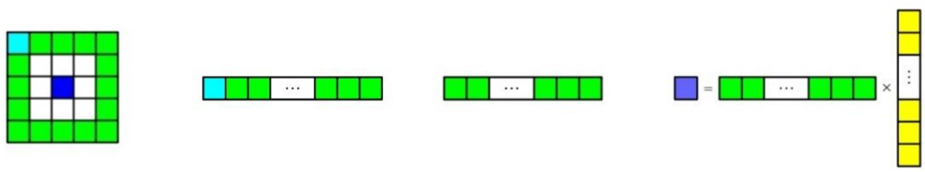

Fig. 2 outliers removal and liner representation process

2.3 Dual-Window Sliding Strategy

For a testing pixel, the traditional detection algorithm exploits only one local window to estimate the local background statistics. While the local background distribution of every local window for the testing pixel is so penurious and unitary that the most ideally representative local distribution for test pixel cannot be achieved. Du et al. put forward a single-window sliding local summation strategy(Du, Zhao et al. 2016), which can take full advantage of the local background statistics of the testing pixel. Therefore, in order to make full use of the local background distribution of every local window for the testing pixel, in this work, we proposed the dual-window sliding strategy shown in Fig. 3, which is different from the single-window sliding. The inner window is merely used to restrict the testing pixel and the size of the inner window is slightly larger than the radius of the anomalous objects. The green squares represent the background pixels and the blue represents the test pixel. When the inner windows size is three and the outer window size is five, we need to calculate nine ($win_{in} \times win_{in}$) times after window sliding completed. And then we need to calculate the result.:

$$r = \sum_{i=1}^{\mathrm{win_{in}} \times \mathrm{win_{in}}} r_i \qquad (12)$$

The surrounding pixels collected inside the outer window while outside the inner window are used for estimating local background statistics information of test pixel, while the inner window is only used to constrain the anomalous targets. Meanwhile, sliding strategy is used to accurately represent the local background statistics of the testing pixel by local dual-window summation. With respect to the size of the internal window, it's usually set to the maximum radius of the anomalous target in the case of prior knowledge, so that the anomalous target can be included in the inner window as much as possible to exclude its influence on the surrounding pixels.

Fig. 3 dual-window sliding strategy

It is notable that the LSUNRSORAD algorithm is different from the UNRSORAD algorithm. For each test pixel, the UNRSORAD algorithm merely uses one dual-window, while the LSUNRSORAD algorithm uses $win_{in} \times win_{in}$ dual-windows. The result of the LSUNRSORAD is $win_{in} \times win_{in}$ times differ accumulation of the result of the UNRSORAD. The overall description of the LSUNRSORAD algorithm is given as **Algorithm 2**.

Algorithm 2 The LSUNRSORAD algorithm

Input: Hyperspectral Data $X = \{x_i\}_{i=1}^{N}$, window size (win_{out}, win_{in}) and the parameter λ

for all pixels **do**

  for each window **do**

    1) For each testing pixel y, a 2D matrix constitutes based on the dual-window within pixels by $X_s = \{x_i\}_{i=1}^{s}$;

    2) Removing outlier pixels in the matrix X_s;

    3) Calculating the weight vector α' by eqn. (8);

    4)Calculating the distance measurement by eqn. (11);

  end for

  5) Calculating the final detection result by eqn. (12);

end for

Output: Anomaly detection map

3. EXPERIMENTS

In order to evaluate the effectiveness and superiority of the proposed methods, we compare the results derived by the proposed methods with those derived by five methods which are GRX, LRX, UNRS, CRD and CRBORAD. The hyperspectral image used in the experiments is a part of Sand Diego airport in USA. This data is acquired by the AVIRIS sensor with the spatial resolution of 3.5 meters. In our experiment, 189 bands of the raw data were remained after removing the corresponding water vapor bands, low signal-to-noise ratio bands. The whole image size is 400 rows ×

400 columns × 189 bands, and an 80 × 80 × 189 region (Samples: 94 to 173, Lines: 321 to 400) of that was cut out for experiments. The whole image and the target map of anomalies in the test region are illustrated in Fig.4, which have been used in the reference (Vafadar and Ghassemian 2017).

The initial choices of different parameters are important for many algorithms. In order to show the superiority of the proposed methods in comparison with others, we set the same parameters for different methods in the following real data set: $win_{out} = 9$, $win_{in} = 7$, and $\lambda = 0.1$. We use a desktop computer with Microsoft 7 OS, Inter Core(TM)2 Duo CPU E7400 @ 2.80GHz and 4 GB RAM. The codes are run by MATLAB® 2016b. It is notable that these parameters are merely empirical parameters not the optimal parameters. We can also find the optimal parameters through the trial-and-error testing by a considerable amount of experiments. Detection results of different methods are shown in Fig. 5. It can be found that the detection result of LSUNRSORAD is more capable of highlighting anomaly target. Fig. 6 illustrates ROC curves of different methods, and Table 1 collects ROC's Area Under Curve (AUC) values of the proposed UNRSORAD and LSUNRSORAD methods and others.

As can be seen from Fig. 6 and Table 1, the proposed methods show better performance in comparison with others. The classical GRX detection method has inacceptable performance, while the AUC values of CRBORAD and LSUNRSORAD are superior in anomalies detection than other detection algorithms in our experiment.

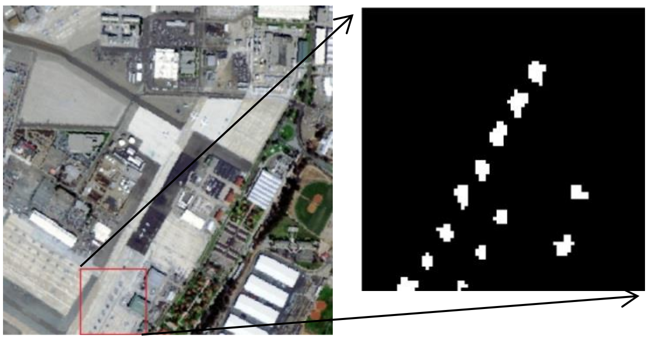

Fig. 4 The San Diego airport hyperspectral data with its test region and the target map

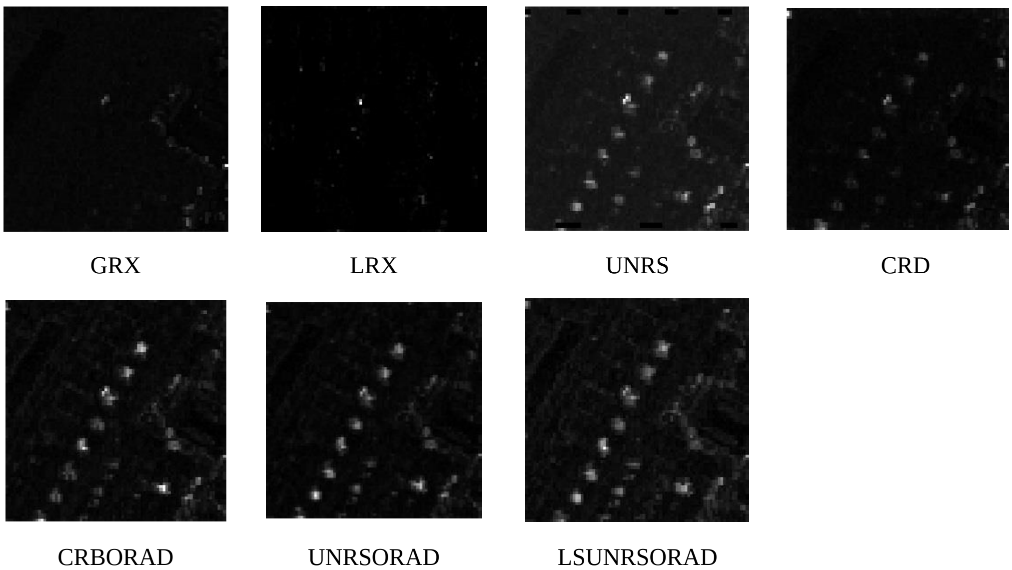

Fig. 5 Detection results of the different methods

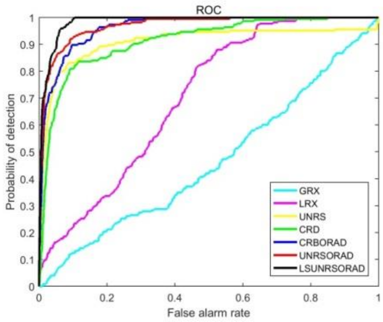

Fig. 6 ROC curves of different methods

Table.1 AUC values of all the methods

Method	GRX	LRX	UNRS	CRD	CRBORAD	UNRSORAD	LSUNRSORAD
AUC	0.46432	0.70167	0.90905	0.91313	0.96272	0.97005	0.98438

4. CONCLUSIONS

We proposed UNRSORAD and LSUNRSORAD methods for anomaly detection in hyperspectral imagery, which take full advantage of the spectral and spatial information of hyperspectral image. We adopt unsupervised nearest regularized subspace method to obtain the weight vector, and in order to obviate interference of outliers we use the outlier removal strategy. The LSUNRSORAD method exploits a dual-window local summation strategy which implements a series of the local dual-window that contains the local background information used to predict the test pixel. The experiment has proven that the UNRSORAD and LSUNRSORAD outperform the traditional anomaly detection methods, such as GRX, LRX, UNRS, CRD and CRBORAD algorithms.

ACKNOWLEDGMENTS

This research is supported in part by Natural Science Foundation of China(No.41471356), the Xuzhou Scientific Funds(KC16SS092), and Priority Academic Program Development of Jiangsu Higher Education Institutions.

REFERENCES

Chen, S. Y. and C. Chang (2013). Low-rank decomposition-based anomaly detection. SPIE Defense, Security, and Sensing.

Crichton, N. (2002). "Receiver operating characteristic ({ROC}) curves."

Du, B., R. Zhao, L. Zhang and L. Zhang (2016). "A spectral-spatial based local summation anomaly detection method for hyperspectral images ." Signal Processing 124(C): 115-131.

Guo, Q., B. Zhang, Q. Ran, L. Gao, J. Li and A. Plaza (2014). "Weighted-RXD and Linear Filter-Based RXD: Improving Background Statistics Estimation for Anomaly Detection in Hyperspectral Imagery." IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing 7(6): 2351-2366.

Hanley, J. A. and B. J. Mcneil (1982). "The meaning and use of the area under a receiver operating characteristic (ROC) curve." Radiology 143(1): 29.

Kwon, H. and N. M. Nasrabadi (2005). "Kernel RX-algorithm: a nonlinear anomaly detector for hyperspectral imagery." IEEE Transactions on Geoscience & Remote Sensing 43(2): 388-397.

Li, J., H. Zhang, L. Zhang and L. Ma (2015). "Hyperspectral Anomaly Detection by the Use of Background Joint Sparse Representation." IEEE Journal of Selected Topics in Applied

Earth Observations & Remote Sensing 8(6): 2523-2533.

Li, W. and Q. Du (2014). Unsupervised nearest regularized subspace for anomaly detection in hyperspectral imagery. Geoscience and Remote Sensing Symposium.

Li, W. and Q. Du (2015). "Collaborative Representation for Hyperspectral Anomaly Detection." IEEE Transactions on Geoscience & Remote Sensing 53(3): 1463-1474.

Liu, G., Z. Lin, S. Yan, J. Sun, Y. Yu and Y. Ma (2010). "Robust recovery of subspace structures by low-rank representation." IEEE Transactions on Pattern Analysis & Machine Intelligence 35(1): 171-184.

Liu, W. M. and C. I. Chang (2008). Multiple-Window Anomaly Detection for Hyperspectral Imagery. Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International.

Niu, Y. and B. Wang (2016). "Hyperspectral Anomaly Detection Based on Low-Rank Representation and Learned Dictionary." Remote Sensing 8(4): 289.

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 & Signal Processing 38(10): 1760-1770.

Taghipour, A., H. Ghassemian and F. Mirzapour (2016). Anomaly detection of hyperspectral imagery using differential morphological profile. Electrical Engineering.

Vafadar, M. and H. Ghassemian (2017).Hyperspectral anomaly detection using outlier removal from collaborative representation. International Conference on Pattern Recognition and Image Analysis.

Xu, Y., Z. Wu, J. Li, A. Plaza and Z. Wei (2016). "Anomaly Detection in Hyperspectral Images Based on Low-Rank and Sparse Representation." IEEE Transactions on Geoscience & Remote Sensing 54(4): 1990-2000.

Zhao, R., B. Du, L. Zhang and L. Zhang (2016). "Beyond Background Feature Extraction: An Anomaly Detection Algorithm Inspired by Slowly Varying Signal Analysis." IEEE Transactions on Geoscience & Remote Sensing 54(3): 1757-1774.