BACKGROUND JOINT SPARSE REPRESENTATION FOR HYPERSPECTRAL IMAGE SUB-PIXEL ANOMALY DETECTION

*Jiayi Li, Hongyan Zhang, and Liangpei Zhang**

The State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, P.R. China
Corresponding author: zlp62@whu.edu.cn

ABSTRACT

A novel sparsity-based sub-pixel anomaly detection framework is proposed for hyperspectral imagery. The proposed approach consists of the following steps. First, a joint sparsity model is utilized to simultaneously represent the surrounding local background pixels and to automatically prune the rough overcomplete dictionary as a reliable, compact base for the following center test pixel representation. An unconstrained linear unmixing approach based on the compact dictionary is then utilized to decompose the abundance of the center test pixel. The unmixing result is finally compared to the former background joint sparse representation step, and the energy disparity is utilized to reflect the anomaly test result. The experimental results confirm that the proposed algorithm outperforms the classical RX-based anomaly detector and the orthogonal subspace projection based detector, and gives a desirable and stable performance.

Index Terms—joint sparse representation (JSR), sub-pixel, anomaly detection (AD), hyperspectral imagery

1. INTRODUCTION

Hyperspectral images (HSIs), spanning the spectrum in many contiguous and very narrow spectral bands, can be potentially used to discriminate different objects on the basis of their unique spectral signatures. With the high spectral resolution band-rich information, a subtle signal which cannot be known as a prior can now be captured. In view of this, hyperspectral imaging has become an emerging technique in many applications in land-cover data exploitation.

Anomaly detection (AD), in which no a priori information about the spectra of the targets of interest is provided, has been widely used in many applications, such as detecting crop stress locations in precision agriculture, rare minerals in geology, oil pollution in environmental research, landmines in the public safety and defense domain, and man-made objects in reconnaissance and surveillance applications [1].

Considering the complicated remote sensing background scene, the large number of possible objects of interest, and the uncertainty, AD still faces some obstacles. Firstly, the way of defining a target of interest is ambiguous, as some pixels detected as a target in a local region belong to the background in a global view. Secondly, the burdensome construction of an accurate model of the complicated background exacerbates the difficulty of the extraction of anomalies, as the background model cannot be precisely obtained by single or multiple-Gaussian models. Moreover, with the limitation of the spatial resolution, sub-pixel objects are a fundamental challenge for AD in HSI.

In order to deal with the above problems, we integrate the linear mixing model (LMM) for the center test pixel and the recently emerging sparse representation theory for the local background of the test pixel to a uniform framework for the sub-pixel AD task, which is named as background joint representation detection (BJSRD). With the highly redundant background information in the hyperspectral scene, the background pixels in a fixed neighborhood are assumed to belong to several subspaces of an overcomplete dictionary. Assuming that the center test pixel shows similar spectral characteristics to its surrounding neighboring pixels, the active dictionary atoms in the previous joint sparse representation (JSR) [2] act as active spectral components of the center test pixel, and are utilized to calculate the fractional abundances and the corresponding tolerance residual by the unconstrained LMM. Comparing the reconstruction results of the background JSR and the unmixing procedure of the center test pixel, we check the difference and judge whether the center test pixel is an anomaly target or not.

2. PROPOSED FRAMEWORK

2.1. Spatial Window Setting

Unlike the well-used dual spatial window in most AD methods, we utilize a different spatial window, as shown in Fig. 1. For the test pixel, we first open a center area sized c^2 (usually only the center test pixel itself for the sub-pixel detection scenario), which is colored in blue in Fig. 1, and we then adopt a guard window sized $g^2 - c^2$ that is effective in preventing some potential target pixels being mixed with the outer background pixels. The guard window is marked by a purple dashed line in Fig. 1. On the outside of the guard window, we finally set a local background window (colored

in red) with $n = b^2 - g^2$ pixels, which is assumed to be dissimilar to the center anomaly. The main goal of the proposed method is to explore the differentiation between the center test window and the outer local background window in the semantic feature space. The semantic feature space is constructed by the pixels in the same hyperspectral scene, without estimating the statistical distribution of the complicated local background window, as is usually done in the other hyperspectral anomaly detectors.

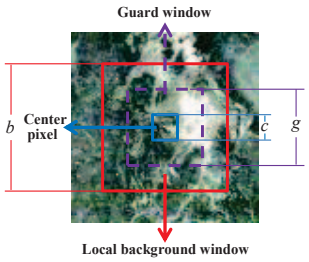


Fig. 1. The spatial windows used in the proposed method.

2.2. Proposed AD Method for the Center Test Pixel

For a test pixel, we first construct the spatial window, and then set the center pixel as $\boldsymbol{s}_c \in \mathbb{R}^B$ and the local background pixels as $\mathbf{S} = [\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n] \in \mathbb{R}^{B \times n}$

2.2.1. Background Joint Sparse Representation and Active Spectral Component Selection

We construct the original dictionary from the hyperspectral scene for the background JSR. We stack all the pixels in the scene as a matrix $\mathbf{H} \in \mathbb{R}^{B \times H}$, in which H is the number of pixels, and B refers to the number of spectral bands. Assuming that there are C_{bn} categories of background materials and C_{tn} categories of anomaly targets, columns of the matrix $\mathbf{H}$ are reordered by category, without loss of generality. In view of this, the matrix can be represented as $\mathbf{H} = [\mathbf{B}_{C_{b1}} \ldots \mathbf{B}_{C_{bn}} \mathbf{T}_{C_{t1}} \ldots \mathbf{T}_{C_{tn}}]$, where $\mathbf{B}_{C_{bi}} \in \mathbb{R}^{B \times N_{bi}}$ belongs to background category C_{bi}, and $\mathbf{T}_{C_{ti}} \in \mathbb{R}^{B \times N_{ti}}$ belongs to target category C_{ti}.

Considering that one pixel in a fixed patch $\mathbf{S}$ named $\boldsymbol{s}_j$ belongs to class C_{bi}, it can be compactly represented by a linear combination of the pixels of $\mathbf{B}_{C_{bi}}$, as shown as $\boldsymbol{s}_j = \mathbf{B}_{C_{bi}} \boldsymbol{\beta}_{C_{bi},j} + \boldsymbol{\xi}_j$, where $\boldsymbol{\xi}_j$ is the random noise and $\boldsymbol{\beta}_{C_{bi},j}$ is an unknown N_{bi}-dimensional coefficient vector. For the pixel $\boldsymbol{s}_j$ without judging its class label, we construct the dictionary $\mathbf{A}$ by removing the pixels in the guard window and $\boldsymbol{s}_j$ from the original dictionary $\mathbf{H}$. That is to say, the remaining $N = H - b^2$ pixels of the $\mathbf{H}$ is stacked as the overcomplete dictionary $\mathbf{A}$ to represent the pixel $\boldsymbol{s}_j$. With the widely accepted knowledge that the background is dominant in the whole scene, it is considered that the property of each background sub-dictionary $\hat{\mathbf{B}}_{C_{bi}}$ of $\mathbf{A}$ should reasonably approach that of $\mathbf{H}$, and the anomaly sub-dictionary $\hat{\mathbf{T}}_{C_{ti}}$ is assumed to be weaker than that of $\mathbf{H}$, or even lost. In view of this, $\boldsymbol{s}_j$ can also be sparsely represented by:

$$\boldsymbol{s}_j = \underbrace{\left[\hat{\mathbf{B}}_{C_{b1}} \ldots \hat{\mathbf{B}}_{C_{bn}} \hat{\mathbf{T}}_{C_{t1}} \ldots \hat{\mathbf{T}}_{C_{tn}}\right]}_{\mathbf{A}} \underbrace{\left[\boldsymbol{\alpha}'_{C_{b1},j} \ldots \boldsymbol{\alpha}'_{C_{bn},j} \boldsymbol{\alpha}'_{C_{t1},j} \ldots \boldsymbol{\alpha}'_{C_{tn},j}\right]'}_{\boldsymbol{\alpha}_j} + \boldsymbol{\xi}_j \quad (1)$$

where $\boldsymbol{\alpha}_j$ is the sparse coefficient vector, and can be obtained by solving this optimization problem:

$$\boldsymbol{\alpha}_j = \arg\min \left\| \mathbf{A}\boldsymbol{\alpha}_j - \boldsymbol{s}_j \right\|_2 \quad s.t. \ \left\| \boldsymbol{\alpha}_j \right\|_0 \le N_{bi} \qquad (2)$$

Similar to the pixel $\boldsymbol{s}_j$, all the pixels in the local background window can also be simultaneously sparsely represented by the overcomplete dictionary $\mathbf{A}$, as shown as $\mathbf{S} = [\boldsymbol{s}_1 \ \boldsymbol{s}_2 \ \ldots \ \boldsymbol{s}_n] = \mathbf{A}[\boldsymbol{\alpha}_1 \ \boldsymbol{\alpha}_2 \ldots \boldsymbol{\alpha}_n] + [\boldsymbol{\xi}_1 \ \boldsymbol{\xi}_2 \ \ldots \ \boldsymbol{\xi}_n]$. Since the local background window contains only one or a few kinds of background categories, it is assumed that all the pixels in the local background window can be linearly represented in the same low-dimensional feature subspace with different compact coefficients, where this subspace can be associated with a background sub-dictionary $\tilde{\mathbf{B}}_{C_{bi}}$ or a combination of a few background sub-dictionaries. In this way, the represented vectors of the local background matrix $\mathbf{S}$ can be better solved as follows:

$$\boldsymbol{\Psi} = \arg\min \left\| \mathbf{A}\boldsymbol{\Psi} - \mathbf{S} \right\|_F \quad s.t. \ \left\| \boldsymbol{\Psi} \right\|_{row,0} \le L_0 \qquad (3)$$

where $\boldsymbol{\Psi}$ is the set of all the sparse coefficient vectors $\tilde{\boldsymbol{\alpha}}_j$ ($j = 1, \ldots, n$), and L_0 denotes the upper bound of the number of non-zero rows of $\boldsymbol{\Psi}$. In this paper, we utilize simultaneous orthogonal matching pursuit (SOMP) [2] to gain the row-sparse coefficient matrix $\boldsymbol{\Psi}$, as well as selecting the active atom subset $\tilde{\mathbf{A}}$ of $\mathbf{A}$. It is natural that the atoms of $\tilde{\mathbf{A}}$ effectively reflect the characteristic of the feature space associated with the principal background components of $\mathbf{S}$, which is to say that the subset $\tilde{\mathbf{A}}$ filters out the bases associated with the underlying anomaly subspaces from the overcomplete dictionary $\mathbf{A}$.

2.2.2. Center Pixel Unmixing

Induced from the orthogonality utilized in each iteration of SOMP, each atom of $\tilde{\mathbf{A}}$ can be considered as the most representative signature of the local background pixel set. Suppose $\boldsymbol{s}_c$ is a mixed pixel, then an unmixing approach can achieve a desirable performance when the component of $\boldsymbol{s}_c$ belongs to one or several background categories associated with $\tilde{\mathbf{A}}$. For simplicity, we utilize an unconstrained linear unmixing model $\boldsymbol{s}_c = \tilde{\mathbf{A}}\tilde{\boldsymbol{\alpha}} + \boldsymbol{\xi}_c$ in this work, where $\tilde{\boldsymbol{\alpha}}$ denotes the fractional abundance of the "most representative spectral component" and $\boldsymbol{\xi}_c$ denotes the tolerance error.

The solution of this model can be easily and analytically derived as:

$$\tilde{\boldsymbol{\alpha}} = \left(\tilde{\mathbf{A}}^T \tilde{\mathbf{A}}\right)^{-1} \tilde{\mathbf{A}}^T \boldsymbol{s}_c \tag{4}$$

It is believed that $\tilde{\mathbf{A}}\tilde{\boldsymbol{\alpha}}$ can approximate the center pixel $\boldsymbol{s}_c$ well with weak errors, if and only if the pixel $\boldsymbol{s}_c$ is similar to the principal component of the local background set $\mathbf{S}$. In contrast, the linear unmixing will work in vain with a large residual when the pixel $\boldsymbol{s}_c$ is significantly different and is considered as an anomaly.

2.2.3. Saliency Difference Comparison

Considering the joint sparse background representation as the benchmark, we calculate the noise matrix $\boldsymbol{\Sigma}$ of the local background pixels: $\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Psi} - \mathbf{S}$. It is assumed that there is no significant difference between the tolerance error $\boldsymbol{\xi}_c$ and the elements of $\boldsymbol{\Sigma}$ if $\boldsymbol{s}_c$ is not an abnormal pixel. We compare the energy difference of the residuals of the former two steps, and we consider the ratio of the energies as the discrepancy and the detection value:

$$t = n\left\|\boldsymbol{\xi}_c\right\| / \left\|\boldsymbol{\Sigma}\right\|_F \tag{5}$$

where n is the number of local background pixels.

2.3. Computational Burden

For each pixel to be tested, the computational burden for the proposed AD algorithm is $O\left(nL_0BH + \left(L_0B + n^2\right)H^2\right)$, as the local background representation costs $O\left(nL_0BH\right)$ by SOMP, the second least squares calculation costs $O\left(L_0BH^2\right)$, and the final comparison step costs $O\left(n^2H^2\right)$. Overall, the computational burden for the whole scene is $O\left(nL_0BH^2 + \left(L_0B + n^2\right)H^3\right)$

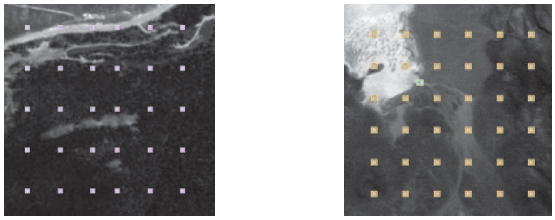
3. SYNTHETIC EXPERIMENTS

In this section, we investigate the effectiveness of the proposed algorithm with two hyperspectral images. The detectors of global RX (GRX) [3], low possibility detection (LPD) based on OSP [4], and a recent low rank subspace decomposition method for global RX detection (referred to as LRD-GRX) [5] were used as benchmarks in this paper. To obtain a fair comparison, each detection map was linearly normalized by its maximum value. Based on the LMM, a synthetic sub-pixel anomaly with spectrum $\boldsymbol{x}$ was generated by fractionally implanting a desired pixel target with a $\boldsymbol{t}$ reflectance spectrum in a given pixel in the background (i.e. host pixel) with a $\boldsymbol{b}_i$ reflectance spectrum, as follows:

$$\boldsymbol{x}_i = \mathbf{exp}\left(-c\rho_i^2\right) f\boldsymbol{t} + \left[1 - \mathbf{exp}\left(-c\rho_i^2\right) f\right]\boldsymbol{b}_i \tag{6}$$

where $f \in (0,1]$ denotes the implant fraction, and $\boldsymbol{b}_i$ and $\boldsymbol{x}_i$ are the reflectance and implanted spectrum of the ith neighbor of the host pixel, respectively. ρ_i is the spatial Euclidean distance between the ith adjacent pixel and the host pixel, and c is a constant, which is set as 1.7 in this paper to control the degree of diffusion. All the experiments were carried out using MATLAB on a PC with one 3.50 GHz processer and 16.0 Gb of RAM. For the computational purpose of the background JSR we restrict the search of the active background bases, i.e., dictionary construction, in a larger "search window" of 9×9 pixels in the image.

3.1. Data Sets



(a) (b)

Fig. 2. Synthetic experiment settings. (a) The ground truth of the 30 synthetic targets and the grayscale map of the HyMap scene. The 15 left targets with panels are associated with the pure spectrum of F1, while each of the 15 right targets with panels were generated by mixing the host pixel with F2. (b) The ground truth of all the 36 synthetic targets and the grayscale map of the AVIRIS scene. The 18 upper targets with panels are associated with the pixel located at (75,73), while each of the 18 lower targets with panels were generated by mixing the host pixel with the pixel located at (75,74)

A HyMap image which was captured at the location of the small town of Cook City, Montana, USA, on July 4, 2006, was used in this experiment. This image is a part of the "Target Detection Test" [6] published by the Rochester Institute of Technology (RIT). From the self-test reflectance data, we cut a sub-region sized 260×260 with 120 high signal-to-noise ratio bands. In this experiment, we took the spectrum of pure pixel targets F1 and F2 available at the full scene as two different targets, and each of the spectra were implanted in 15 host pixels, with consideration of their adjacent effects, in a constant window size of 5×5.

Another standard hyperspectral AD data set is the AVIRIS image of the Lunar Crater Volcanic Field (LCVF) in Northern Nye Country, Nevada, USA. We cut a sub-image sized 200×200 of the full scene. We removed some noisy spectral bands, leaving 195 bands for the subsequent processing. A single anomaly containing two pixels (located at (75, 73) and (75, 74)) in the sub-scene lie on the border of the dry lakebed in the scene. In order to simulate a series of sub-pixels to make a quantitative analysis possible, we also took these two real anomaly pixels as the target, and each of the spectra were implanted in 18 host pixels, with consideration of their adjacent effects, in a constant window size of 5×5.

For both data sets, each implant was accomplished at five fractions: 5%, 10%, 20%, 30%, and 40%. The synthetic image and ground truth of these two data sets are shown in Fig. 2.

3.2. Detection Performance

The first line of each fraction item in Tables I and II reports the average false alarm rate (AFAR) [7], which is considered as a suitable summary metric to evaluate the detection task obtained by all the AD methods at each implant fraction for the two hyperspectral data sets. It is demonstrated that the proposed BJSRD algorithm shows a significantly superior detection result, especially in dealing with the extremely difficult case of the 5% fraction, where the rest of the AD methods mostly fail. To further analyze the comparison, a statistical separability box map [8] and the ROC curves for the HyMap data set at the 20% fraction, and those for the AVIRIS data set at the 10% fraction, are shown in Fig. 3 and Fig. 4. Both the ROC curves and the gap between the target and background of the proposed method indicate a desirable detection performance.

TABLE I
AFAR VALUES/RUNNING TIMES OF THE ALGORITHMS OBTAINED WITH THE HYMAP SYNTHETIC IMAGE AT THE FIVE IMPLANT FRACTIONS

AD	GRX	LPD	LRD-GRX	Proposed
5%	0.5415 3.593±0.0506	0.4678 2.944±0.0416	0.5109 33.140±0.1210	**0.2781** 64.848±1.6478
10%	0.4209 3.598±0.0320	0.2584 2.913±0.0330	0.1370 33.148±0.1416	**0.0742** 65.289±1.2869
20%	0.1684 3.573±0.0448	0.0240 2.969±0.0410	0.0921 31.823±0.0928	**0.0074** 50.499±0.1979
30%	0.0396 3.565±0.0358	0.0027 2.936±0.0578	0.0115 29.356±0.0980	**0.0011** 50.013±0.3909
40%	0.0079 3.570±0.0265	7.19e-4 2.904±0.0338	0.0060 31.884±0.1207	**3.01e-4** 49.981±0.2785

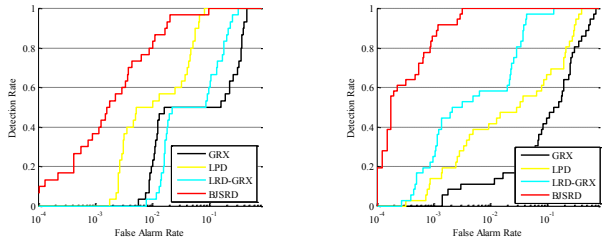
TABLE II
AFAR VALUES/RUNNING TIMES OF THE ALGORITHMS OBTAINED WITH THE AVIRIS SYNTHETIC IMAGE AT THE FIVE IMPLANT FRACTIONS

AD	GRX	LPD	LRD-GRX	Proposed
5%	0.4497 4.404±0.0197	0.4676 3.767±0.0203	0.2282 29.117±0.3388	**0.0254** 41.171±0.6330
10%	0.1992 4.431±0.0283	0.1012 3.770±0.0190	0.0164 29.261±0.1770	**5.58e-4** 45.167±0.5083
20%	0.0233 4.415±0.0179	0.0040 3.774±0.0266	0.0019 29.332±0.1301	**8.54e-5** 35.398±0.1017
30%	4.26e-4 4.448±0.0362	2.63e-4 3.768±0.0187	4.26e-4 29.350±0.1071	**4.23e-5** 35.144±0.1061
40%	1.63e-4 4.436±0.0206	1.13e-4 3.766±0.0178	1.62e-4 29.538±0.2242	**2.50e-5** 31.959±0.1489

The running times for each algorithm are shown in the second line of each fraction item in Tables I and II. The running times are averaged over 10 runs, and the second line of each item in Tables I and II shows the standard deviation. Although the proposed method takes more time than the rest of the AD methods, it is still acceptable, considering the excellent detection performance.

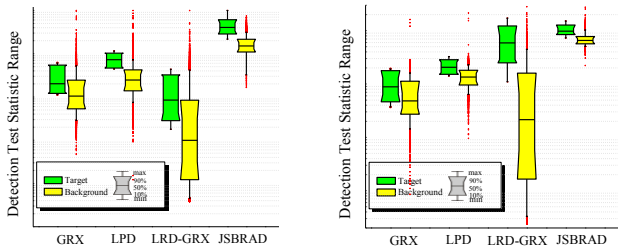
4. CONCLUSION

This paper proposes a sparse representation based framework for hyperspectral image sub-pixel anomaly detection. The proposed algorithm utilizes the redundant background information in the hyperspectral scene, and automatically deals with the complicated multiple background classes without estimating the statistical information of the background. A linear unmixing approach is utilized to determine the anomaly abundance in the test pixel. Finally, whether the center test pixel is an anomaly or not is judged by the differential reconstruction results of the two former steps. The proposed BJSRD method was tested on AVIRIS and HyMap hyperspectral images, and the experimental results confirm the effectiveness of the proposed hyperspectral anomaly detector.



(a) (b)

Fig 3. The ROC curves of the AD algorithms for the synthetic data sets. (a) The HyMap data set at the 20% fraction; and (b) the AVIRIS data set at the 10% fraction .



(a) (b)

Fig 4. Statistical separability analysis of the AD algorithms. (a) The HyMap data set at the 20% fraction; and (b) the AVIRIS data set at the 10% fraction.

5. REFERENCES

[1] S. Khazai, A. Safari, B. Mojaradi, and S. Homayouni, "An Approach for Subpixel Anomaly Detection in Hyperspectral Images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.,* vol. 6, pp. 769–778, 2013.

[2] H. Zhang, J. Li, Y. Huang, and L. Zhang, "A Nonlocal Weighted Joint Sparse Representation Classification Method for Hyperspectral Imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10.1109/JSTARS.2013.2264720 2013.

[3] 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, pp. 1760–1770, 1990.

[4] J. C. Harsanyi, *Detection and Classification of Subpixel Spectral Signatures in Hyperspectral Image Sequences*. University of Maryland, Baltimore County, 1993.

[5] S.-Y. Chen, S. Yang, K. Kalpakis, and C.-I. Chang, "Low-rank decomposition-based anomaly detection," in *SPIE Defense, Security, and Sensing*, 2013, pp. 87430N–87430N-7.

[6] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, and S. Hager, "Development of a web-based application to evaluate target finding algorithms," in *Proc. IEEE Int. Geoscience and Remote Sensing Symp. 2008, IGARSS 2008,* pp. II-915–II-918

[7] P. Bajorski, E. J. Ientilucci, and J. R. Schott, "Comparison of basis-vector selection methods for target and background subspaces as applied to subpixel target detection," in *Defense and Security*, 2004, pp. 97–108.

[8] Zhang, L., Zhang, L., Tao, D., Huang, X., 2014, "Sparse Transfer Manifold Embedding for Hyperspectral Target Detection", *IEEE Trans. Geosci Remote Sens.*, vol. 52, no. 2, pp. 894–906.