Anomaly Detection in Hyperspectral imagery based on Low-Rank and Sparse Decomposition

Xiaoguang Cui, Yuan Tian, Lubin Weng and Yiping Yang.
Institute of Automation, Chinese academy of Sciences.

ABSTRACT

This paper presents a novel low-rank and sparse decomposition (LSD) based model for anomaly detection in hyperspectral images. In our model, a local image region is represented as a low-rank matrix plus spares noises in the spectral space, where the background can be explained by the low-rank matrix, and the anomalies are indicated by the sparse noises. The detection of anomalies in local image regions is formulated as a constrained LSD problem, which can be solved efficiently and robustly with a modified "Go Decomposition" (GoDec) method. To enhance the validity of this model, we adapts a "simple linear iterative clustering" (SLIC) superpixel algorithm to efficiently generate homogeneous local image regions i.e. superpixels in hyperspectral imagery, thus ensures that the background in local image regions satisfies the condition of low-rank. Experimental results on real hyperspectral data demonstrate that, compared with several known local detectors including RX detector, kernel RX detector, and SVDD detector, the proposed model can comfortably achieves better performance in satisfactory computation time.

Keywords: Anomaly detection, hyperspectral imagery, low-rank and sparse, superpixels.

1. INTRODUCTION

In recent years, hyperspectral images are being increasingly used in target classification and detection due to their capability to supply a very dense spectral sampling of spectral signatures for different materials. In this paper, we focus on anomaly detection, which can be referred as a particular case of target detection. Anomaly detection aims at distinguishing unusual pixels of which spectral signatures are significantly different from their surrounding background. Due to its ability to find valuable targets without requiring prior spectral information about the interest targets, anomaly detection plays an important role in hyperspectral imagery analysis [1,2].

Anomaly detection methods in hyperspectral imagery mainly estimate the distance between candidate targets and the background in a certain measure space to detect anomalies. Common methods usual select the background in a local manner using a sliding window around the pixel under test and assume that the background is composed of one single class. The well known RX detector [3], which is based on the generalized likelihood ratio test and an assumed Gaussian distribution of the background, was first used in multispectral and later was introduced in hyperspectral images, presenting acceptable performance for anomaly detection. The nonlinear version of the RX algorithm was proposed by [4], referred to as kernel RX detector; it acquires improved performance over the conventional RX detector by using kernel method to model non-Gaussian distributions of the background. Another nonlinear local method is based on the support vector data description (SVDD) [5]. The SVDD operator estimates an enclosing hypersphere around the background in a high-dimensional feature space and treats pixels that lie outside the hypersphere as outliers. Although these methods can reasonably be adopted in homogeneous areas, they are often unstable when the local background contains sparse and large corruptions or multiple classes because the statistical characterization of the background tends to be disturbed in this situation. In addition, these methods are computationally intensive due to their point-by-point calculation manner.

To overcome the shortcomings in the conventional methods, in this paper, we propose a novel low-rank and sparse decomposition based model for anomaly detection in hyperspectral images. The framework of the proposed method is shown in Fig.1. We represent a local image region as a low-rank matrix plus spares noises in the spectral space, where the background can be explained by the low-rank matrix, and the anomalies are indicated by the sparse noises. We treat anomaly detection in a local image region as a constrained *Low-Rank and Sparse Decomposition* (LSD) problem, and modify a *Go Decomposition* (GoDec) [6] algorithm to solve it efficiently and robustly. Furthermore, we adapt a *simple linear iterative clustering* (SLIC) superpixel algorithm [7] to efficiently generate homogeneous local image regions so as to ensure that the background in local image regions is low-rank.

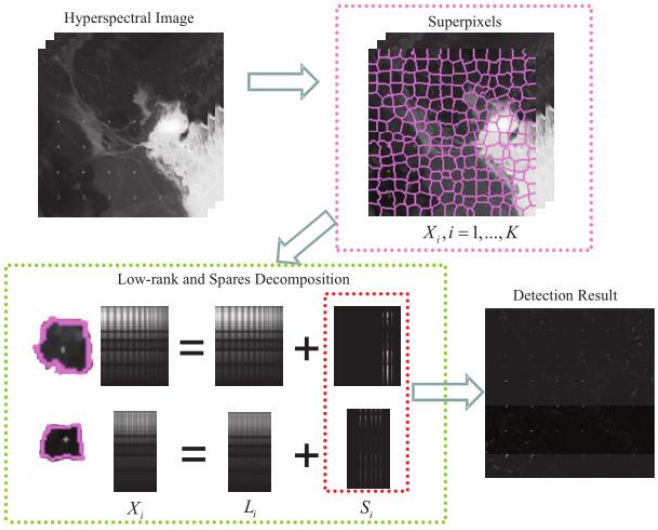


Figure. 1 The framework of the proposed model. For a given hyperspectral image, firstly, we partition it into K non-overlapping homogeneous superpixels $X_i, i = 1,...K$. Secondly, we decompose each X_i into a low-rank matrix L_i and a sparse matrix S_i which represents the background and anomalies respectively. Finally, we generate the detection result by calculating the ℓ_2-norm of each column in S_i.

2. LOW-RANK AND SPARSE DECOMPOSITION FOR ANOMALY DETECTION

2.1 Problem formulation

For a given hyperspectral image, we partition it into non-overlapping local regions. Let $X_i = [x_1, x_2, ..., x_n]$ be a local region matrix of size n pixels, each column of which is a spectral vector x_j corresponding to a pixel j. We represent the local region as a low-rank matrix plus spares noises in the spectral space, where the background can be explained by the low-rank matrix and the anomalies are indicated by the sparse noises, as the following form:

$$X_i = L_i + S_i + G_i \tag{1}$$

where L_i is an unknown low-rank matrix representing the background, S_i is an unknown sparse matrix representing anomalies, and G_i denotes the residual noise. Though representing local image regions as the above formulation, we cast the anomaly detection as a LSD problem, which will be solved via a modified GoDec method in the following section. After obtaining S_i, we calculate a saliency score to each pixel using the ℓ_2-norm and run a threshold-based test with a given threshold η to decide whether a pixel j is anomaly, as follows:

$$score(j) = \left\| S_i(:, j) \right\|_2 \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{2}$$

where $\left\| S_i(:, j) \right\|_2$ denotes the ℓ_2-norm of the j-th column of S_i, H_0 and H_1 is the anomaly-absent and the anomaly-present hypothesis, respectively.

2.2 Modify GoDec for LSD

In our model, a meaningful decomposition must meet both of the following requirements: 1) the rank range of L_i ($rank(L_i)$) needs to reflect the number of end-members in local regions. 2) the cardinality range of S_i ($card(S_i)$) should approximately equal to the size of desired anomalies. Consequently, we want problem (1) can be solved with pre-defined $rank(L_i)$ and $card(S_i)$, i.e.

$$X_i = L_i + S_i + G_i \quad s.t.\ rank(L_i) = r, card(S_i) = k \tag{3}$$

To this end, we fist introduce a recently developed algorithm called GoDec to deal with LSD, then we give a modification of GoDec to make it more suitable for our model.

2.2.1 GoDec

GoDec formulates the approximated LSD problem in (3) as:

$$\min_{L_i, S_i} \left\| X_i - L_i - S_i \right\|_F^2 \quad s.t.\ rank(L_i) \le r, \left\| S_i \right\|_0 \le k \tag{4}$$

where $card(S_i)$ is measured by the ℓ_0-norm. Then the optimization problem of (4) can be solved by alternatively assigning the low-rank approximation of $X_i - S_i$ to L_i and the sparse approximation of $X_i - L_i$ to S_i until convergence:

$$\begin{cases} L_i^t = \sum_{j=1}^{r} \lambda_i U_i V_i^T,\ SVD(X_i - S_i^{t-1}) = U \Lambda V^T \\ S_i^t = P_\Omega(X_i - L_i^t) \end{cases} \tag{5}$$

where Ω is the nonzero subset of the first k largest entries of $\left| X_i - L_i^t \right|$. The SVD in the updating L_i^t is a time-consuming operation that often impractical when X_i is of large size. In order to significantly accelerate the operational speed, GoDec replaces SVD with bilateral random projections (BRP) to give a fast low-rank approximation. Though BRP, GoDec significantly reduces the time cost meanwhile keeps a near optimal approximation. In addition, it was theoretically proved that the decomposition error of GoDec monotonically decreases and converges to a local minimum [6].

2.2.2 Modify GoDec

It is noteworthy that the ℓ_0-norm constraint in (4), used as a metric of sparsity of S_i, is not suitable for our model. Although ℓ_0-norm is a universal form for representing a random sparse matrix, it is unreliable for measuring structured sparse matrixes. S_i is a typical matrix with structured sparsity, whose columns with most non-zero entries indicate the position of anomalies. According to the sparsity structure of S_i, we quantify the cardinality of S_i by the $\ell_{2,0}$-norm to obtain a reliable metric, as follows:

$$card(S_i) = \left\| \left\| S_i(:, j) \right\|_2 \right\|_0 \tag{6}$$

The ℓ_0-norm constraint often leads S_i to contain numerous unrelated large noise entries, in contrast, the $\ell_{2,0}$-norm constraint encourages S_i to generate more non-zero entries in columns which are corresponding to anomalies, leading to a more accurate LSD solution, as shown in Fig.2.

3. SUPERPIXEL ALGORITHM FOR HYPERSPECTRAL IMAGERY

The selection of the rank r takes a decisive role in GoDec algorithm. r should be chosen to be exactly the same as the actual number of background classes in local image regions. If r is underestimated, S_i tends to be disturbed by undesired background pixels; conversely, if r is overestimated, anomalies could be assigned to the background matrix L_i, hence the anomalies cannot be detected. To avoid the impact of the incorrect selection of r, we want GoDec to perform on a homogeneous local image region composed of one single class, so that GoDec can yield a reliable result with setting r to 1.

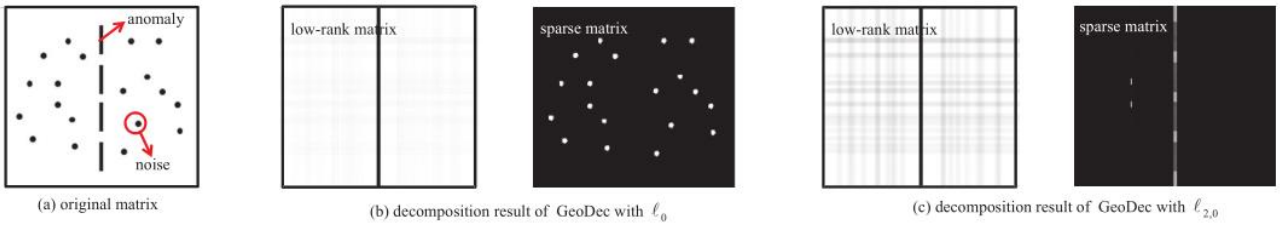


Figure. 2 An example for illustrating the difference between GoDec with ℓ_0-norm and GoDec with $\ell_{2,0}$-norm.

For this purpose, we adapt the SLIC algorithm to efficiently generate homogeneous local image regions in hyperspectral imagery for GoDec. SLIC is a very practical, memory efficient superpixel algorithm. Despite its simplicity, it has the ability to produces superpixels at a lower computational cost while achieving state-of-the-art boundary adherence. SLIC is an adaptation of k-means for superpixel generation with using a distance measure in the labxy color-image plane space. Since SLIC was initially designed for color images, we next adapt it to hyperspectral imagery via a distance measure that combines both the spectral distance and the spatial distance.

During the clustering process of SLIC, a distance measure D is required to determine the nearest cluster center for each pixel. It is necessary to consider both the spectral distance and the spatial distance for designing D. The spectral distance is used to ensure that pixels in a same cluster have strong spectral correlation, while the spatial distance is critical to control the spatial compactness. To combine the two distances into a single measure, we define D as:

$$\begin{aligned} d_{spe}(a,b) &= \arccos\left[\sum_{i=1}^{n_b} a_i b_i \Big/ \left(\sum_{i=1}^{n_b} a_i^2\right)^{1/2} \left(\sum_{i=1}^{n_b} b_i^2\right)^{1/2}\right] \\ d_{spa}(a,b) &= \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2} \\ D(a,b) &= \frac{d_{spe}}{N_{spe}} + \lambda \frac{d_{spa}}{N_{spa}} \end{aligned} \tag{7}$$

where d_{spe} denotes the spectral angle between a pixel a at position (x_a, y_a) and a pixel b at position (x_b, y_b), n_b is the number of bands; d_{spa} is the Euclidean distance between a and b; it is necessary to normalize d_{spe} and d_{spa} by their respective maximum distances within a cluster, N_{spe} and N_{spa}, then D is given by the sum of the normalized spectral distance and the normalized spatial distance weighted by λ. λ is a positive parameter that trades off the two distances. A large λ leads to produce compact superpixels, and vice versa. In practice, λ should be properly chose to offer a good balance between spectral similarity and spatial proximity.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The data employed in this paper is an AVIRIS hyperspectral imagery with size of 200×200 which contains 120 bands. As shown in Fig.3, there is a single anomaly containing two pixels lies on the central section, 36 targets with various sizes and shapes were implanted in the image, each of them has the same spectrum as the original anomaly.

In order to evaluate the detection performance, a comparison is preformed between the proposed method and three common local anomaly detectors, i.e., RX, kernel-RX and SVDD. Since the covariance matrix used in RX is often unstable to estimate in high-dimensional data samples, we select 20 bands via a maximum-variance principle component

analysis (MVPCA) based band selection method [8] for both RX and kernel-RX, so as to increase their detection efficiency. The modified GoDec is performed on local image regions with setting $r=1$, $k=5$ (the possible largest number of pixels containing in a target). Here, we provide the comparative analysis mainly on the separability between anomalies and the background, which refers to the ability of a detector to separate anomalies from the background. To make fair comparisons, all the detectors are preformed on two different kinds of local regions, i.e. the local regions generated by a hollow sliding window and the local regions generated by SLIC. The hollow sliding window is an usual approach to collect local samples, we set the size of outer window and guard window to 17×17 and 9×9, according to the size of desired targets. As to SLIC, we generate 200 superpixels (each one contains about 200 pixels) in the image to produce non-overlap homogeneous local regions for local detectors. As shown in Fig.4, if λ is too large, spatial distances outweigh spectral distances, resulting in compact superpixels that missing the spectral proximity; conversely, if λ is too small, superpixels are generated without respect to spatial compactness, which will yield lots of disjoint pixels that would cause errors in the process of enforcing connectivity. In addition, it is advisable to generate superpixels with small size, so as to ensure pixels in a superpixel is homogeneous. Here, we set λ=1.5 in the experiment.

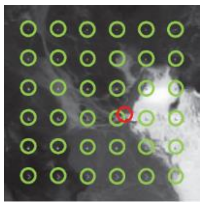


Figure. 3 A single band of the AVIRIS hyperspectral subimage with showing the position of the anomalies. The original anomaly is marked by a red circular, and implanted targets are marked by green circulares.

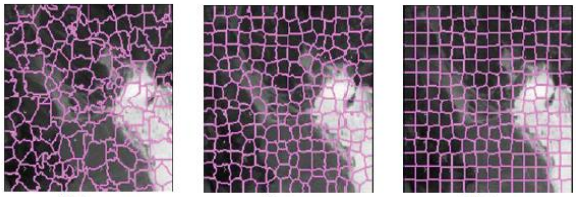
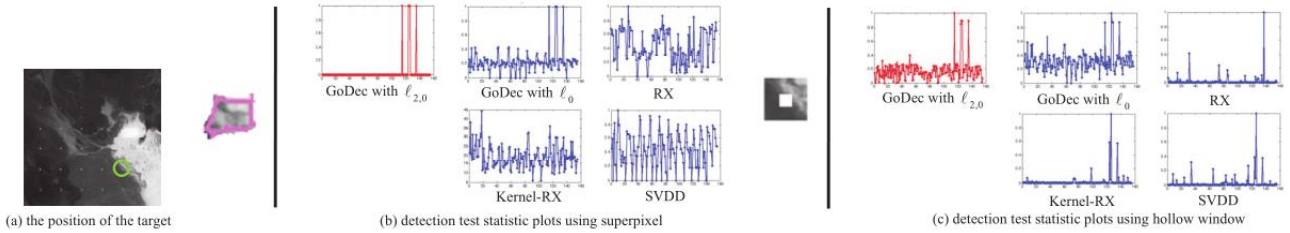


Figure. 4 Image segmented using Algorithm 1 into superpixels of (approximate) size 200 with λ=0.5, λ=1.5 and λ=10.

Fig.5 shows the detection test statistic plots around a target using GeoDec and the other three detectors with the two kinds of local regions. The target (containing 5 pixels) is difficult to detect because its spectrum is partly similar to the background. From Fig.5, we can obtain the following analysis: 1) In the superpixel region, the modified GeoDec (GeoDec-$\ell_{2,0}$) suppresses the background to a low steady range, while the target pixels have a much higher statistical value and can be easily separated. The original GeoDec (GeoDec-ℓ_0) leads some background pixels to present high statistical values, this limits its performance. The other three detectors output the background pixels with greater fluctuations that would cause false alarms, this is mainly because the background statistic characteristics are disturbed by large corruptions brought by the target pixels. 2) In the hollow window region, the separability of GoDec goes bad because the hollow window cannot provide homogeneous regions so that the rank of the background tends to be underestimated when setting $r=1$, but it still produces comparable results. The other three detectors achieve much better performance in the hollow window region, the reason is that the effect of undesired target pixels can be avoided by using a guard window. However, they still cannot provide a satisfied result because the background statistical estimation is unstable when the background contains multiple classes.

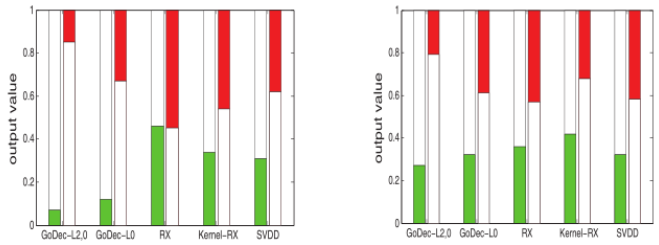


(a) the position of the target (b) detection test statistic plots using superpixel (c) detection test statistic plots using hollow window

Figure. 5 The detection test statistic plots around a target using GeoDec and the other three detectors with the two kinds of local regions..

To further investigate the separability between all the anomalies and the background, we plot the range of output test statistic value as shown in Fig.6. Each method is evaluated using a group of box comprised of a green box and a red box. The green box and the red box represents the output value range of all the background pixels and that of all the anomalies, respectively. The gap between the green box and the red box in each group reflects the ability of a method to separate anomalies from the background. From Fig.6, we can draw a consistent conclusion with analysis in Fig.5, that the modified GoDec is more robust to corruptions compared with other detectors, moreover, it's separability ability can be considerably enhanced with the help of superpixels.

The two critical factors that impact the speed of the detectors are the computational complexity in a single local region and the selection manner of local regions. Although GoDec significantly reduces the time cost compared with directly using SVD, it is slower than the other detectors because of its time-consuming iterative calculation. However, the total processing time of a detector is mainly decided by the other factor, i.e. the selection manner of local regions. Select regions using the hollow window leads to a pixel-by-pixel calculation, which is $O(N)$ complex. In contrast, the complexity can be reduced to $O(K)$ by using superpixels as local regions. K is usually two order of magnitudes smaller than N, thus saving considerable computation times.



(a) (b)

Figure. 6 The range of output test statistic value of different detectors performed on (a)superpixel regions and (b)hollow window regions. 1% pixels with largest value in the background and 10% pixels with smallest value of the anomalies are removed in the statistical process.

5. CONCLUSION

In this paper, a novel LSD based model is proposed for anomaly detection in hyperspectral imagery. This model successfully transforms anomaly detection into a LSD problem, and efficiently solves this problem by modified the GeDec algorithm. The experimental results show that the modified GeDec has stronger ability to separate anomalies from the background compared with other common local detectors. Moreover, a hyperspectral imagery version of SLIC is proposed to provide homogeneous local regions for local anomaly detectors, by which local detectors can achieve better background statistical estimation at a much lower computational cost.

REFERENCES

[1] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine*,

vol. 19, no. 1, pp. 29–43, (2002).
[2] S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE A&E Systems Magazine*, vol. 25, no. 7, pp. 5–28, (2010).
[3] I. S. Reed and X. Yu, "Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Acoustics, Speech, & Signal Processing*, vol. 38, no. 10, pp. 1760–1770, (1990).
[4] H. Kwon and N. Nasrabadi, "Kernel rx-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 388–397, (2005).
[5] A. Banerjee, P. Burlina, and C. Diehl, "A support vector method for anomaly detection in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 8, pp. 2282–2291, (2006).
[6] T. Zhou and D. C. Tao, "Godec: Randomized low-rank & sparse matrix decomposition in noisy case," *in Proceedings of the* 28*th International Conference on Machine Learning*, (2011).
[7] R. Achanta, A. Shaji, and K. S. etc., "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, (2012).
[8] C. I. Chang, Q. Du, T. L. Sun, and M. L. G. Althouse, "A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification," IEEE Transactions on Geoscience and Remote Sensing, vol. 37, no. 6, pp. 2631–2641, (1999).