Anomaly detection in hyperspectral images through spectral unmixing and low rank decomposition

Conference Paper · July 2016
DOI: 10.1109/IGARSS.2016.7729476

CITATIONS
0

READS
158

7 authors, including:

Bulent Ayhan
Signal Processing Inc.
61 PUBLICATIONS 658 CITATIONS
SEE PROFILE

Chiman Kwan
Signal Processing, Inc.
254 PUBLICATIONS 2,740 CITATIONS
SEE PROFILE

Steven Douglas Vance
California Institute of Technology
91 PUBLICATIONS 344 CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

- Europa and icy world seismology View project
- A High Performance Approach to Local Active Noise Reduction in Noisy Cabins View project

All content following this page was uploaded by Chiman Kwan on 24 July 2016.
The user has requested enhancement of the downloaded file. All in-text references underlined in blue are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.
ANOMALY DETECTION IN HYPERSPECTRAL IMAGES THROUGH SPECTRAL UNMIXING AND LOW RANK DECOMPOSITION

Ying Qu, Rui Guo, Wei Wang and Hairong Qi

EECS Department
The University of Tennessee, Knoxville, TN 37996
{yqu3, rguo1, wwang34, hqi}@utk.edu

Bulent Ayhan, Chiman Kwan
Signal Processing, Inc.
Rockville, MD

Steven Vance
Jet Propulsion Lab
Pasadena, CA

ABSTRACT
Anomaly detection has been known to be a challenging, ill-posed problem due to the uncertainty of anomaly and the interference of noise. In this paper, we propose a novel low rank anomaly detection algorithm in hyperspectral images (HSI), where three components are involved. First, due to the highly mixed nature of pixels in HSI, instead of using the raw pixel directly for anomaly detection, the proposed algorithm applies spectral unmixing algorithms to obtain the abundance vectors and uses these vectors for anomaly detection. Second, for better classification, a dictionary is built based on the mean-shift clustering of the abundance vectors to better represent the highly-correlated background and the sparse anomaly. Finally, a low-rank matrix decomposition is proposed to encourage the sparse coefficients of the dictionary to be low-rank, and the residual matrix to be sparse. Anomalies can then be extracted by summing up the columns of the residual matrix. The proposed algorithm is evaluated on both synthetic and real datasets. Experimental results show that the proposed approach constantly achieves high detection rate while maintaining low false alarm rate regardless of the type of images tested.

Index Terms— Hyperspectral image, anomaly detection, mean-shift clustering, low-rank, sparsity.

1. INTRODUCTION
Anomaly detection is to identify salient pixels inside an image. It can be modeled as an unsupervised binary classification problem between the background class and the anomaly class. The challenge of anomaly detection is that we do not have any prior knowledge of the anomaly objects or the background. However, there are two principal properties that could distinguish anomaly objects from their background, i.e., compared to the other objects, the anomaly objects are sparse and possess distinctive features as compared to their surrounding background. Hyperspectral images (HSI) can provide spectral characteristics of different materials on the ground surface, which potentially increases the probability of separating salient components from background [1].

Therefore, HSI-based anomaly detection has been intensively studied recently.

Among all the anomaly detection algorithms, the GRX detector proposed by Reed and Yu [2] is one of the most popular statistical approaches. GRX assumes that the probability density function of the background can be modeled as Gaussian normal distribution. Then the saliency can be identified by measuring the Mahalanobis distance of a pixel vector to its background. However, normally the noise and background components contained in real HSI is nonuniform, which cannot be simply modeled using Gaussian distribution. To overcome these limitations, several algorithms have been proposed to improve the basic GRX, including the subspace RX [3] and local RX [4]. Local approaches are able to improve the detection rate, but the drawback is that they may struggle with isolated noise pixels, which will increase the false alarm rate dramatically. Also the size of the local window needs to be defined according to the size of anomaly which is usually unknown.

In addition to the statistical approaches, low-rank based methods have also drawn much attention recently. This group of techniques assumes that the matrix of anomaly is sparse in the image. Therefore after removing the anomaly, the background matrix should have low rank [5–7]. [6] and [7] are based on robust principal component analysis (RPCA) [8, 9], which have been performing well on detecting anomalies. However, as shown in [10], RPCA is based on the assumption that the data has a single low-rank subspace, therefore, the sparse noise can easily be identified as anomalies. [5] is based on GoDec [11] which also considers the noise during the low-rank and sparse estimation. Unfortunately, in practice, anomalies are prone to being identified as noise.

To overcome the drawbacks of the aforementioned approaches, we propose a low rank anomaly detection algorithm which contains three components, spectral unmixing, dictionary construction using mean shift clustering, and low-rank decomposition. Instead of using the raw pixels as features to detect anomalies, we apply the spectral unmixing algorithm first to obtain the abundance vectors, as features for anomaly detection. We then construct a dictionary to de-
scribe the highly-correlated background and sparse anomalies where the dictionary atoms are determined using the mean-shift clustering algorithm. In this way, isolated noise can be represented using the same coefficients as the background. Finally, a low-rank decomposition algorithm is proposed to extract the sparse anomalies.

The rest of the paper is organized as follows. Sec. 2 describes the proposed algorithm with the three components elaborated. Sec. 3 evaluates the algorithm using both synthetic and real data sets. Conclusion is drawn in Sec. 4.

2. PROPOSED ALGORITHM

There are three featured components in the proposed anomaly detection algorithm design that contribute to the performance gain, namely, using the abundance vectors derived from spectral unmixing instead of the raw data for anomaly detection, using mean-shift clustering on abundance vectors to construct a dictionary describing both the background and the anomaly, and using low-rank decomposition to extract anomaly objects. The design flow is illustrated in Fig. 1.

![Flow chart of the proposed algorithm](image)

**Fig. 1.** The flow chart of the proposed algorithm

### 2.1. Abundance as Feature Vector

It is important to find suitable features which can describe the background and anomaly accurately. Since each pixel in HSI is a mixture of a few pure materials called endmembers, we conduct the unmixing operation first to find the endmembers and the corresponding mixing coefficients (i.e., abundance) according to the model:

\[
X = AS + N
\]

where \(X\) is the observed mixture, \(N\) represents the noise, and \(A\) and \(S\) denote endmembers and their abundance vectors, respectively. To extract the features, some state-of-the-art algorithms such as MVC-NMF [12–14] are employed to perform the unmixing procedure. The proposed approach then takes \(S\) as features of HSI, and builds a dictionary to represent both background and anomalies.

### 2.2. Mean-Shift based Dictionary Construction

Dictionary plays an important role in the proposed algorithm, where a nonparametric mean-shift clustering algorithm [15] is used to estimate the clusters by given a small bandwidth. Taking the columns of \(S\) as samples, \(k\) clusters are estimated based on the input data. To better represent the data, we choose both the center and the edge of each cluster as dictionary atoms in \(D\) for \(k\) clusters. Compared to traditional clustering approaches like k-means, the mean-shift method formulates the clustering problem using statistical distribution of samples, and finds the modes of the data which indicate the center of clusters. So it does not require prior knowledge of the number of clusters and can work on arbitrarily shaped clusters. The advantage of this step is that, isolated noise with similar pixels will be grouped with similar coefficients. Therefore, the noise is no longer sparse anymore, which lowers the probability of being confused with the sparse anomaly.

### 2.3. Low-Rank based Anomaly Detection

Upon finding \(S\) (Sec. 2.1) and constructing \(D\) (Sec. 2.2), the next step is to find which features in \(S\) belongs to anomaly. Since background is highly correlated, we can use the predefined dictionary \(D\) to represent it. The matrix \(S\) can be further decomposed into two parts:

\[
S = DZ + E
\]

where \(DZ\) is the dictionary represented background and \(E\) is the anomaly component. To solve this ill-posed problem in Eq. (2), we encourage \(Z\) to be low rank and \(E\) to be sparse. Note that the features of anomalies can be different from each other, some anomalies could be separated by one feature, while some anomalies could be separated by another feature. Thus it is reasonable that the columns of matrix \(E\) has a few nonzero values. Based on the analysis above, we use \(l_1\) norm to encourage the matrix \(E\) to be sparse. It has been proved in [16] that minimizing the nuclear norm of a matrix is a good surrogate for minimizing the rank of a matrix, as long as matrix \(E\) is sufficiently sparse. Therefore, the proposed objective function can be written as

\[
\min_{Z,E} \|Z\|_n + \lambda \|E\|_1
\]

s.t. \(S = DZ + E\).

where \(\|Z\|_n\) denotes the nuclear norm, i.e., the sum of singular values of \(Z\). Parameter \(\lambda > 0\) is used to balance the tradeoff between low-rank and sparsity. Then \(E\) and \(Z\) in Eq. (3) can be solved by soft threshold method and lemma 3.2 in [10], respectively. After low-rank decomposition, the anomaly image can be obtained by summing up each column of \(E\).

3. EXPERIMENTAL RESULTS

In this section, the proposed algorithm is evaluated with three hyperspectral image datasets including both synthetic and real data. Although there are numerous detection algorithms, we
focus on comparing the proposed approach to some representative approaches including Global RX (GRX) [2], Local RX (LRX) and RPCA based approach IACRPCA [7]. For quantitative comparison, detection results are converted to binary images according to different thresholds. Based on these binary images, ROC curves are generated by calculating the detection rate versus false alarm rate.

3.1. Dataset Description

First, we evaluate the algorithms on hyperspectral images added with synthetic targets. The real data have 79 bands of resolution $150 \times 103$. Forty-nine anomalous dots are implanted [17] based on the linear mixing model in Eq. (4). Target $t$ is fractionally implanted with background spectrum $b$ by varying $f$ from 0.05 to 1 as follows:

$$z = f \cdot t + (1 - f) \cdot b \quad (4)$$

In order to test how the algorithm works with anomaly targets of different reflections, we also perform algorithms on a group of images provided by Air Force (AF) with resolution $267 \times 342 \times 124$. The anomalies in AF images are four aluminum panels with different colors (Black, Green, Tan, and Silver).

The last HSI dataset has anomaly targets with irregular shapes, acquired from the Mars Rover Mast camera solday. This data has a dimension of $598 \times 670 \times 12$. The visual images are shown in Fig. 2 (a), Fig. 4 (a) and Fig. 6 (a), respectively.

There are three free parameters in the proposed algorithm, including the number of endmembers, $c$, used in the unmixing step, the bandwidth, $bw$, used in the clustering step, and the $\lambda$ used in the low-rank approximation step. In addition, in LRX, we set the window size as $11 \times 11$. Table 1 shows all the parameters used in the following experiments.

<table>
<thead>
<tr>
<th>Name of HSI</th>
<th>$c$</th>
<th>$bw$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic target</td>
<td>25</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>Air Force images</td>
<td>20</td>
<td>0.2</td>
<td>0.02</td>
</tr>
<tr>
<td>MastCam solday</td>
<td>9</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

3.2. Detection Performance

Fig. 2 shows the comparison of different algorithms on the synthetic targets. By visual comparison, the LRX algorithm keeps the best contrast between anomalies and background. That’s because LRX is appropriate for uniformly distributed targets. However the noise in the background is also mistakenly selects by LRX detector. IACRPCA performs better than GRX, but it also chooses some sparse noise. This is because some noise is isolated and easily being confused as anomalies. Compared to the other algorithms, the proposed algorithm can find all the targets with lowest false alarm rate 0.075. As illustrated in the ROC curve given by Fig. 3, it keeps the false alarm rate below 0.26.

Fig. 2. Detection results of synthetic target (a) Synthetic targets (b) GRX (c) LRX (d) IACRPCA (e) Proposed (f) Ground truth target

Fig. 3. ROC curve of detection performance with different algorithms on synthetic targets

We also evaluate the performance on real HSI dataset. For the air force images, there are four targets made with different colors, and the reflection of the targets are similar to that of the background in several bands. As shown in Fig. 4, both GRX and LRX fail to detect the target. The detection capability of the proposed algorithm is comparable to that of IACRPCA. The ROC curve given by Fig. 5 demonstrates that the proposed approach again has a higher detection rate over that of IACRPCA, GRX and LRX. Meanwhile it keeps the false alarm rate below 0.014.

The last group of HSI images is from Mars Rover Mast camera provided by NASA without ground truth. The proposed algorithm successfully detects the hydrated ‘anomaly’ which mostly appears in the drilled hole and crackles of the soil surface. As shown in Fig. 6, the detected target from the proposed algorithm has less background portion as compared to that of IACRPCA, and both approaches are superior to that of GRX and LRX.

4. CONCLUSION

In this paper, we analyzed the drawbacks of some existing algorithms for anomaly detection in hyperspectral images. Based on the experiments, we found that current anomaly detection approaches tend to confuse isolated noise with anomaly. We proposed a unique method based on spectral
unmixing and low-rank decomposition. Experimental results showed that the proposed algorithm is able to detect the anomaly successfully while suppressing the noise in the background. For future work, we will focus on improving the computational efficiency of the proposed algorithm through parallel computation using multiple CPUs or GPUs.

5. REFERENCES


