

IDENTIFY ANOMALY COMPONENT BY SPARSITY AND LOW RANK¹

Wei Wang, Shuangjiang Li, Hairong Qi
University of Tennessee, EECS Dept.
Knoxville, TN

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

Steven Vance
Jet Propulsion Lab
Pasadena, CA

ABSTRACT

Traditional anomaly detection methods either model the global background or the local neighborhood, that bring some apparent drawbacks, such as the unreasonable assumption of uni-modular background in global detectors, or the high false alarms by sliding windows in local detectors. In this paper, a source component-based anomaly detection approach is proposed. It first extracts the source components in the spectral image data cube by using the blind source component separation and then identifies the components that are anomaly (or salient) to other components. We interpret the anomaly detection as a matrix decomposition problem with the minimum volume constraint for the multi-modular background and sparsity constraint for the anomaly pixels. Experimental results show that the approach is promising for anomaly detection in spectral data cube.

1. INTRODUCTION

Anomalies in the remote sensing domain refer to pixels with two properties, i.e., distinct spectral differences and a small number of quantity with regard to the surrounding background pixels in the spectral feature space. For example, an airplane on the runway of an airport presents a spectrum that is obviously different from that of the background runway. Searching for those spectrally distinct and rarely appeared objects is the main task of anomaly detection. Generally speaking, unlike target detection, which needs the targets spectra in advance, anomaly detection methods require no prior spectral information about targets, but they do require a sufficient spectral difference between the targets of interest and their backgrounds. Hyperspectral imagery (HSI), a 3-D “image cube”, provides a wealth of spectral information to uniquely identify various materials by their spectrum, which makes it possible to distinguish different objects of interest based on their spectral signatures.

In the literature of anomaly detection (AD), the most popular approach is the Reed–Xiaoli (RX) detector, which was introduced in [1]. It is built on the concept that a hypothesis testing can be formulated for a pixel vector and the conditional probability density functions (pdfs) under the two hypotheses (without and with anomaly) are assumed

to be Gaussian. The solution to the resulting generalized likelihood ratio test turns out to be the Mahalanobis distance between the pixel under test and the background. Two typical variations of the RX have been studied: global RX, which estimates the background statistics (i.e., mean and covariance matrix) of the entire image, and local RX, which estimates the background using local statistics. There have been quite some other detectors proposed that belong to these two categories, e.g., the global detectors [2,3] and the local detectors [4,5]. The global detectors assume that all of the non-target pixels come from a homogenous background. However, this assumption might not be accurate in practice, since, in general, the background may not come from a single uni-modular space and thus it is difficult to estimate the statistics. In contrast, the local detectors use sliding windows to obtain nearby pixels for background statistics, such that the Gaussian assumption is more reasonable in a local region. However, the lack of global information in these local detectors may cause false alarms, e.g., the isolated pixels embedded in other homogeneous background may be detected as anomalies due to their small number in just its local neighborhood. In addition, the local detectors usually need to define the size of the moving dual window, but the sizes of targets are not easy to be estimated in advance, especially if the targets have different sizes.

To overcome the drawbacks of the aforementioned two categories of approaches, this work proposes a nonparametric anomaly detection approach that estimates the background without assuming a pdf or estimating its covariance matrix, neither does it need local detection with a sliding window. The proposed detection approach is based on the idea that we find out the source components of the whole image data cube and identify which components are the most salient as compared to others. In the first step of source component retrieval, we can make use of the well-developed unmixing techniques, such as VCA [6], MVC-NMF [7], *et al.* With the source components, we explore the two properties of AD that are sparsity and saliency. That is, instead of dealing with the original spectral data, we focus on the source components. Explicitly, we expect to select several source components that are sparsely distributed spatially, and also have the largest accumulated distance to all the other source components.

The uniqueness of the proposed approach includes: 1. It does not need to estimate the background statistics to model

¹ This work was supported in part by NASA NNX 12CB05C.

anomaly, instead, it uses saliency to define anomaly. 2. It does not need to use sliding window for local region modeling. We describe all the materials in image by source components. Therefore, our approach considers the whole statistics in the image without background estimation.

2. SOURCE COMPONENT EXTRACTION

In highly mixed hyperspectral data, each pixel is a mixture of responses from multiple materials. The basic formation model is expressed as:

$$X = AS + E \quad (1)$$

where the columns of $X \in \mathbb{R}^{l \times n}$ denote the observation vectors of n pixels measured at l spectral bands. $A \in \mathbb{R}^{l \times c}$ is the material signature matrix whose columns correspond to the spectral signatures of c components or endmembers. The abundance vectors are represented by the columns of $S \in \mathbb{R}^{c \times n}$, which satisfy two physical constraints: first, each element of S is non-negative; second, the sum of column elements equals 1. E denotes the possible noise and errors.

We use the minimum volume constrained non-negative matrix factorization (MNV-NMF) for source endmember extraction. MVC-NMF explores two important facts: first, the spectral data are non-negative; second, the constituent materials occupy the vertices of a simplex, and the volume of the simplex determined by the actual materials is the minimum among all possible simplexes that circumscribe the data scatter space. Combining the goal of minimum approximation error with the volume constraint, we arrive at the following constrained optimization problem:

$$\text{minimize } f(A, S) = \frac{1}{2} \|X - AS\|_F^2 + \lambda J(A) \quad (2)$$

$$\text{subject to } A \geq 0, S \geq 0, 1_c^T S = 1_n^T$$

where $J(A)$ is the penalty function, calculating the simplex volume determined by the estimated endmembers. The regularization parameter λ is used to control the tradeoff between the accurate reconstruction and the volume constraint.

The simplex volume is calculated based on the connection between the volume and the determinant, which leads to:

$$J(A) = \frac{1}{2(c-1)!} \det^2 \left(\begin{bmatrix} 1_c^T \\ \tilde{A} \end{bmatrix} \right) \quad (3)$$

where \tilde{A} is a low-dimensional transform of A given by PCA.

$$\tilde{A} = U^T (A - \mu 1_c^T) \quad (4)$$

The detailed optimization is referred to [7]. After this step, each sample in X is converted to be represented by S , with each element in s_i , $i=1:n$, denoting the percentage of the corresponding endmember or source component.

3. SALIENT COMPONENT EXTRACTION

Byers et al [8] found the accumulated distance can be used for clutter removal and saliency discovery. The saliency object is defined as the one having the largest accumulated distance to the other objects. Therefore, in our application, after we obtain the c source components, we need to find out which component is the most salient one, then we can obtain the saliency mapping in the spatial domain of the image. In this way, we may even have the potential for subpixel anomaly detection.

After the first step, we convert the spectral data from X into S , whose column represents the abundance of the source components in each spectrum. With S , we expect to identify some components that rarely appear in the coefficient matrix S but with the largest accumulated distance to the other source components. Therefore, the matrix S can be further decomposed into two parts:

$$S = R + D \quad (5)$$

where R is the spectral data with salient components being removed, and D is the detected anomaly components. Since the source components in D have the uniqueness that their accumulated distances to the other remaining components in R are all relatively large, the components in R will form a compact shape in the hyperspace (small area size if in a 2-D plane), thus the volume of the simplex that circumscribes the data space that with only components in R will be small. Then, the objective function becomes:

$$\begin{aligned} &\text{minimize } V(\text{AR}) + \|D\|_1 \\ &\text{subject to } S = R + D \end{aligned} \quad (6)$$

where AR represents the original data with the source components in D being removed, or the orthogonal data subspace to the subspaces in D . Fortunately, the volume of (AR) can be approximated by the rank of (AR) , also

$$\text{rank}(\text{AR}) \leq \text{rank}(A)\text{rank}(R) \quad (7)$$

$\text{rank}(A)$ is fixed, so minimize $\text{rank}(\text{AR})$ equals to minimizing $\text{rank}(R)$, then (6) becomes:

$$\begin{aligned} &\text{minimize: } \text{rank}(R) + \|D\|_1 \\ &\text{subject to: } S = R + D \end{aligned} \quad (8)$$

Obviously, Eq. (8) is the popular form of matrix low rank decomposition, we can simply apply the RPCA algorithm [9] to solve it. Finally, we sum up all the elements in each column of D to produce the anomaly mapping image for detection with a threshold.

Compared to the unmixing approaches, the advantage of the proposed approach is that we can discriminate which components is the anomaly, and still hold the advantage of the unmixing technique, i.e., preserve the percentage of anomaly in each pixel.

4. EXPERIMENTAL RESULTS

4.1. Data Sets

To evaluate the performance, the proposed approach is applied onto three datasets. The first one is a synthetic hyperspectral data, where 49 dots are manually added to the

94-band image data of spatial size 150×103 . The 49 dots are with different anomaly percentage varying from 5% to 100%. The second dataset is the air force (AF) image with 4 aluminum panels (Black, Green, Tan, and Silver), representing the anomaly in the scene. The dimension of the AF image is $267 \times 342 \times 124$. The third dataset is the Mastcam sol183 dataset from both the right and left camera after calibration which has a dimension of $598 \times 670 \times 12$. The images of the three data sets are shown in Figure 1.

In experiments, we apply the MVCNMF first, with the predefined number of source components c being a large value, so that the source components can be well separated. Then, we apply the second step for anomaly components extraction.

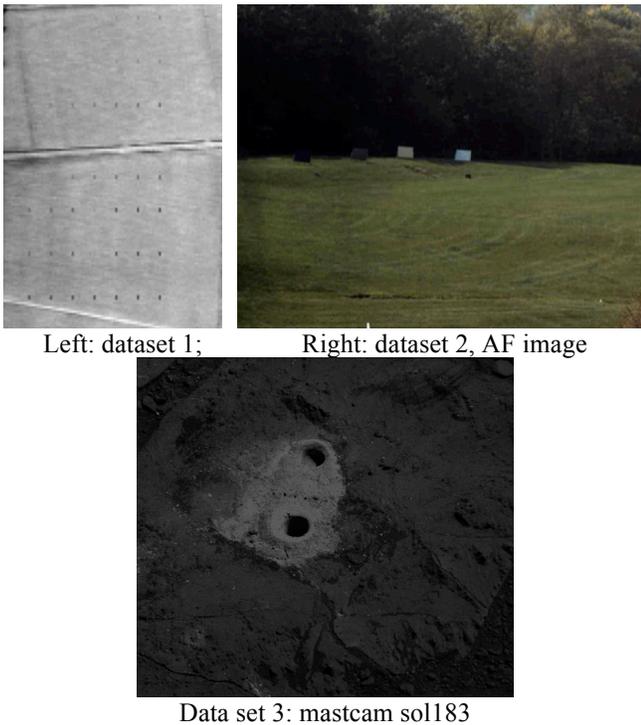


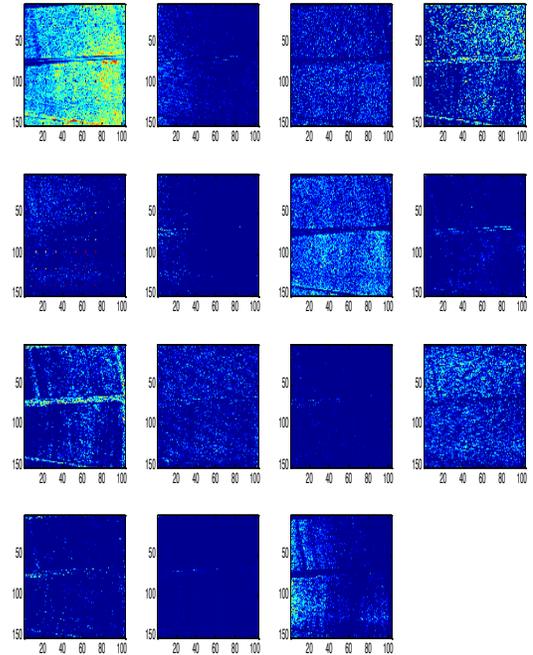
Figure 1. Sample spectral images of the three data sets.

4.2. Results Analysis

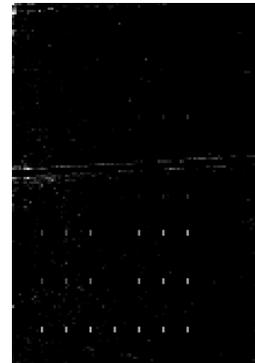
The anomaly detection results on dataset 1 are shown in Figure 2. 15 source components are extracted in the first unmixing step, however, it is difficult to discriminate which component is anomaly to others, which forms multi-modular of the background. From the results of the second step, we can find the components of the dot targets and the thin line in the middle of the image that are both detected. Although the dot targets are the synthetic anomaly, the thin lines also have the two properties of anomaly in this image, therefore both two components are selected as anomaly components in the results.

The detection results on dataset 2 are shown in Figure 3. The number of source components is also defined as 15. From the 15 components, it is difficult to select which

components are anomaly. Especially when the 4 targets have different characteristics, and they appear in different source components. However, in the detection results, we again successfully identified the anomaly components, which perfectly align with the four targets from different sources.



(a) extracted components in step 1.



(b) identified anomaly components

Figure 2. Results on data set 1.

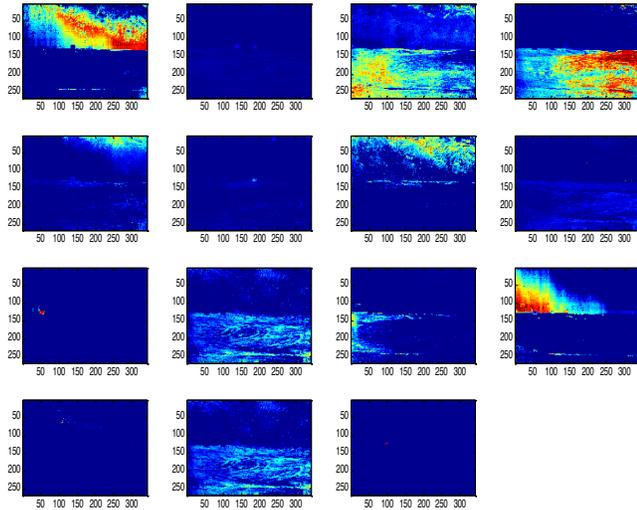
The detection results on data set 3 are shown in Figure 4. 9 source components are extracted in the first step, and the identified anomaly components seem to correctly reflect the characteristic of the hydration materials, which mostly appears in the drilled hole and the cracks of the soil surface. A reference image is also shown in Figure 4(c), provided by NASA, although it is from another data set.

5. CONCLUSION

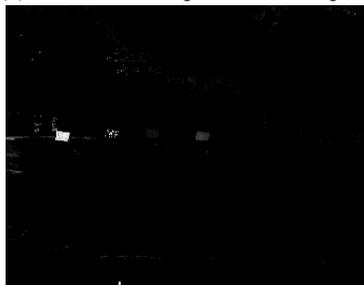
In this work, we presented an approach that combines source component extraction by unmixing and anomaly components identification by volume and sparsity constraint.

The anomaly has been interpreted as saliency and converted to be a low-rank matrix decomposition problem. The experimental results showed that we are potentially able to identify anomaly components with good performance.

The future work will be to compare our approach to other existing techniques with quantitative evaluation, such as the local/global RX algorithm [2,3], sparse representation based algorithm [10], and subspace based algorithms [11].



(a) extracted components in step 1.



(b) identified anomaly components

Figure 3. Results on data set 2.

REFERENCES

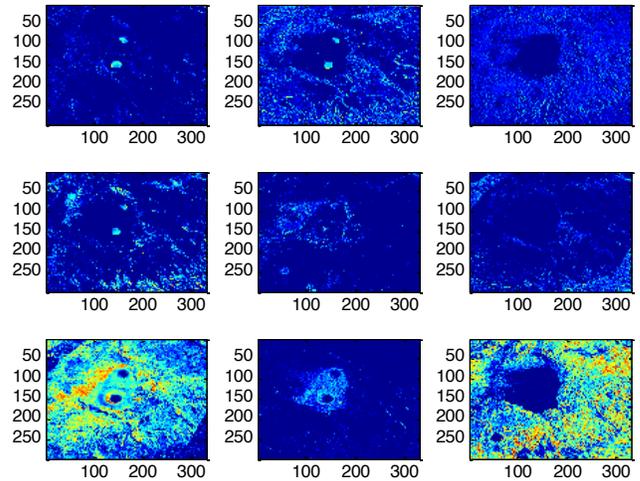
[1] I. S.Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, 1990.

[2] C. I. Chang and S. S. Chiang, "Anomaly detection and classification for hyperspectral imagery," *IEEE Geosci. Remote Sens.*, vol. 40, no. 6, pp. 1314–1325, Jun. 2002.

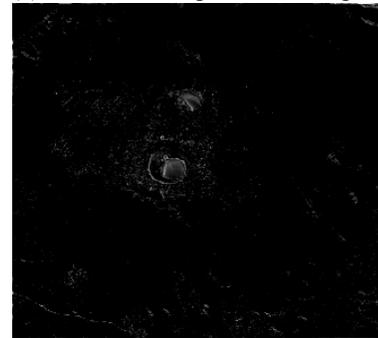
[3] L. He, Q. Pan, *et al*, "Anomaly detection in hyperspectral imagery by maximum entropy and nonparametric estimation," *Pattern Recog. Lett.*, vol. 29, no. 9, pp. 1392–1403, Jul. 2008.

[4] N. M. Nasrabadi, "Regularization for spectral matched filter and RX anomaly detector," in *Proc. SPIE*, 2008, vol. 6966, pp. 1–12.

[5] W. Liu and C.-I. Chang, "Multiple-window anomaly detection for hyperspectral imagery," in *Proc. IGASS*, 2008.



(a) extracted components in step 1.



(b) identified anomaly components

(c) hydration mapping in another image data

Figure 4. Results on data set 3.

[6] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[7] L. Miao, H. Qi, "A Constrained Non-Negative Matrix Factorization Approach to Unmix Mixed Hyperspectral Data", *ICIP* 2007.

[8] S. Byers and A. Raftery. Nearest-neighbor clutter removal for estimating features in spatial point processes. *Journal of the American Statistical Association*, 1998.

[9] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, p. 11, May 2011.

[10] W. Li, Q. Du, "Collaborative Representation for Hyperspectral Anomaly Detection," *IEEE Trans on GRS*, 2015.

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