LOW-RANK TENSOR DECOMPOSITION BASED ANOMALY DETECTION FOR HYPERSPECTRAL IMAGERY

*Shuangjiang Li, Wei Wang, Hairong Qi** *Bulent Ayhan, Chiman Kwan* *Steven Vance*

University of Tennessee, EECS Dept.
Knoxville, TN, USA

Signal Processing Inc.
Rockville, MD, USA

Jet Propulsion Laboratory
Pasadena, CA, USA

ABSTRACT

Anomaly detection becomes increasingly important in hyperspectral image analysis, since it can now uncover many material substances which were previously unresolved by multispectral sensors. In this paper, we propose a Low-rank Tensor Decomposition based anomaly Detection (LTDD) algorithm for Hyperspectral Imagery. The HSI data cube is first modeled as a dense low-rank tensor plus a sparse tensor. Based on the obtained low-rank tensor, LTDD further decomposes the low-rank tensor using Tucker decomposition to extract the core tensor which is treated as the "support" of the anomaly spectral signatures. LTDD then adopts an unmixing approach to the reconstructed core tensor for anomaly detection. The experiments based on both simulated and real hyperspectral data sets verify the effectiveness of our algorithm.

***Index Terms*—** Hyperspectral imaging, anomaly detection, low-rank approximation, tensor decomposition

1. INTRODUCTION

Hyperspectral images (HSI) attract more and more interests in recent years as a suitable tool for target detection and recognition in many applications including search-and-rescue operations, mine detection and military usages. Among all these usages, Anomaly Detection (AD) has received a lot of attention for various applications. The aim of anomaly detection is to detect pixels in the hyperspectral data cube whose spectra differ significantly from the background spectra with no *a-priori* knowledge [1].

Many real world problems involve signals that are multidimensional or tensor data. HSI can also be modeled as a three-dimensional tensor, with the first two dimensions indicating the spatial domain and the third dimension indicating the spectral domain [2]. The classical matrix-based anomaly detection methods require to rearrange the tensor into a two-dimensional matrix, then process in the column space, and finally rebuild the tensor. To avoid the rearranging and rebuilding steps where the spatial correlation would be lost, the tensor-based anomaly detection methods can be used directly to process the HSI by employing multilinear data analysis.

*This work was supported in part by NASA NNX12CB05C.

In this paper, we propose a Low-rank Tensor Decomposition based anomaly Detection (LTDD) algorithm. The HSI data cube is then modeled as a dense low-rank tensor plus a sparse tensor. The dense low-rank tensor captures the spectral background as well as the anomaly signatures while the sparse tensor represents the noises and arbitrary errors within the data. Based on the obtained low-rank tensor, LTDD further decomposes the low-rank tensor using Tucker decomposition [3] to extract the core tensor from the low-rank tensor. The core tensor is treated as the "support" of the anomaly spectral signatures, where we apply an unmixing approach to the reconstructed core tensor for anomaly detection.

The rest of this paper is organized as follows. In Section 2, we discuss some related works. In Section 3, we introduce some background on tensor decomposition. In Section 4, we discuss the proposed LTDD algorithm. Performance evaluation on using LTDD and existing algorithms is shown in Section 5. Section 6 concludes the paper.

2. RELATED WORKS

Most common statistical ADs are outlined in the tutorial overview by Matteoli et al. [4]. In the literature, the most popular approach is the Reed-Xiaoli (RX) detector [5], which is derived from the generalized likelihood ratio test. The RX requires that the covariance matrix be estimated from the neighborhood pixels of the target pixel, i.e., the local background. The RX by definition is prone to high false alarms because the local Gaussian assumption is largely inaccurate. Local RX [6] is then proposed to force stability upon the pixels that are locally defined with respect to a window size to reduce the bias and error when estimating the mean and covariance, respectively.

Another anomaly detection approach is called Projection Pursuit (PP) [7, 8]. It searches for an optimal projection that the anomaly component will become the most obvious in a subspace. However, searching the optimal projection is usually very computationally intensive, which hinders the real-time data processing.

The anomaly detectors mentioned above are all basically conducted from the statistical perspective. Recently, there

have been many approaches using high-order statistics [9,10]. Specifically, Geng et al.'s work published in Nature [11] proposed to use high-order statistical tensor based algorithm for AD. In their work, a third-order statistical tensor (i.e., coskewness tensor) was introduced and proposed a detection method based on higher order singular value detection to extract anomalies. The idea of sparse and low-rank modeling has been applied in many areas, e.g., background modeling in surveillance video processing [12], band selection in HSI [13], nonlinear hyperspectral image unmixing [14] and pedestrian tracking and recognition [15]. Low-rank and tensor based methods have also been applied in HSI, for example, Zhang et al. [16] proposed to use low-rank matrix recovery for HSI restoration. Similarly, the tensor based HSI denoising was introduced in [17].

3. BACKGROUND ON TENSOR DECOMPOSITION

Tensor decompositions allow us to approximate tensor data sets by models depending on few parameters, i.e., less parameters than the total number of entries of the tensor. This reduction of degree of freedom allows us to capture the essential structures in multidimensional data sets. Generally, tensor decomposition takes two forms: the Canonical Polyadic (CP) decomposition and Tucker decomposition [18]. The goal of tensor decomposition in data reduction applications is to find a good multilinear rank $(r_1, r_2, \cdots, r_d)$ approximation to the order-d tensor $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$, i.e., $\min_{\mathcal{A}} \|\mathcal{T} - \mathcal{A}\|_F^2$ with $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ restricted to be of rank $(r_1, r_2, \cdots, r_d)$ and $r_i \leq n_i$. This approximation problem is well-posed, but does not have a known closed-form solution [19]. For example, CP approximates a tensor with a sum of d rank-one tensors. It often can be unique up to permutation of the rank-1 terms and up to scaling/counterscaling of the factors in the same term. Tucker approximation is useful for dimensionality reduction of large tensor datasets and is also important when one wishes to estimate signal subspaces from tensor data [17]. A reduced-rank approximation is simply obtained by restricting the factor matrices of $\mathcal{T}, \mathbf{U}_i \in \mathbb{R}^{n_i \times n_i}$, where $\mathcal{T} = \mathcal{C} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \cdots \times_d \mathbf{U}_d$ and $\mathcal{C} \in \mathbb{R}^{n_1 \times n_2 \times \cdots n_d}$ is the core tensor, to the first $r_1, r_2, \cdots, r_d$ columns.

4. LTDD ALGORITHM

In this section, we present the LTDD algorithm by first discussing how to model the HSI tensor as a dense low-rank tensor plus a sparse tensor as well as how to solve for the dense low-rank tensor and the core tensor, following which, we introduce the unmixing based approach for the anomaly detection.

4.1. Robust Low-rank Tensor Decomposition

In HSI, the data volume often displays a low-rank structure due to significant correlations in the spectra of neighboring pixels [16]. In practice, the underlying tensor data is often low-rank, even though the actual observed data may not be due to noises and arbitrary errors. Therefore, the major part of the variation in the data is often governed by a relatively small number of latent factors. Based on this assumption, we model the HSI data as the three-dimensional tensor $\mathcal{G}^{I_1 \times I_2 \times I_2} = \mathcal{S} + \mathcal{L}$, where $\mathcal{S}$ is a sparse tensor and $\mathcal{L}$ is a dense low-rank tensor. Moreover, $\mathcal{L}$ lives in a low-dimensional manifold that changes slowly across different spectral bands, whereas $\mathcal{S}$ corresponds to more transient and spontaneous objects. In the context of AD for HSI, $\mathcal{L}$ is the superposition of spectral background and the anomaly signatures, while $\mathcal{S}$ is the noises and arbitrary errors.

In order to separate the low-rank tensor from the noisy tensor data subject to noise and arbitrary errors, we employ the idea of robust PCA (rPCA) [20] to high-order tensor data [21] as follows:

$$\min_{\mathcal{S},\mathcal{L}} \text{Trank}(\mathcal{L}) + \lambda \|\mathcal{S}\|_0, \; s.t. \; \mathcal{G} = \mathcal{S} + \mathcal{L}, \tag{1}$$

where Trank$(\cdot)$ denotes the Tucker decomposition. The algorithm to solve Eq. (1) is known as HoRPCA (Higher Order RPCA) proposed in [21]. Once we get optimal $\mathcal{L}$ result, the Tucker decomposition of the core tensor $\mathcal{C}$ can be reconstructed as follows:

$$\mathcal{C} = \mathcal{L} \times_1 (\mathbf{U}^{(1)})^T \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T \tag{2}$$

where $\mathbf{U}^{(i)}$ is the left factor matrix from the SVD of $\mathbf{L}^{(i)}$. Thus, we can recover the Tucker decomposition, i.e., the subspaces along each mode, of the low-rank tensor $\mathcal{L}$ from its corrupted version without the need to specify the target Tucker-rank.

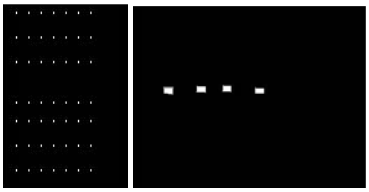
The idea is that by first modeling $\mathcal{G} = \mathcal{L} + \mathcal{S}$, we will be able to reduce the noise and errors and then condense the raw HSI into a compact format where only the anomaly reflectance lie in (e.g., anomaly detection). The basic assumption here is that the low-rank core tensor $\mathcal{C}$, which denotes pixels that have distinct spectral differences and a small number of quantity, is "supported" by these anomaly spectral signatures.

4.2. Anomaly Detection through Unmixing

After obtaining the core tensor $\mathcal{C}$ from the low-rank tensor $\mathcal{L}$ of the original HSI, we introduce a final anomaly extraction through unmixing process on the reconstructed core tensor data cube. Note that this step is different from the common unmixing process which is to estimate the material signatures and abundances. The purpose for the unmixing step

here is to extract the anomaly signature purely as the core tensor data cube is only supported by these anomaly signatures. Therefore, the number of endmembers for the unmixing process can then be pre-determined to the size of the core tensor data cube, which in practice is very small and makes the unmixing process very efficient in computation. For the proposed LTDD algorithm, we adopt the popular nFINDR algorithm [22] with fixed endmember parameter q. After obtaining the endmember signatures, we use the non-negative constrained least squares (NNLS) [23] to find the "anomaly" abundance map(s).

(a) Synthetic (b) AF image (c) Mastcam Solday 183



(d) GT for (a) (e) GT for (b)

Fig. 1. First row: visual images of the three HSI datasets, second row: ground truth image of the synthetic and AF image respectively.

5. EXPERIMENTAL RESULTS

In this section, we will evaluate the performance of the proposed LTDD algorithm on one synthetic and two real HSI datasets and compare with existing approaches. We first discuss some experimental settings, followed by some visual and quantitative performance results.

5.1. Experimental Settings

We use three HSI datasets for evaluation purpose. 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 a testing image from 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 Mars Rover Mast camera solday 183 dataset [1] from both the right and left cameras after calibration which has a dimension of $598 \times 670 \times 12$. For both the synthetic and AF images, there are ground truth (GT) data available for performance evaluation purpose. However, for the Mastcam Solday 183 image, we do not have GT. The objective is to find the hydration materials which mostly appears in the drilled hole and the cracks of the soil surface as provided by NASA. The visual images of the three data sets as well as the available GT are shown in Fig. 1.

We evaluate all the algorithms based on quantitative metrics. With a gray image of the anomaly response, we use different threshold values to convert the detection results into different binary images, from which we can identify the anomaly pixels. When compared to the ground truth of the anomaly pixels, we can calculate the detection ratio (dr) and the false alarm ratio (fr). By arranging the x-axis as the fr while the y-axis as the dr, the closer the curve approaching the top-left corner of the graph, the more effective and robust the detection performance, meaning that the approach provides high detection ratio with low false alarm ratio.

We also compare LTDD with some existing approaches. The first one is the benchmark RX algorithm [5]. RX has been applied to both multi and hyper-spectral images successfully in terms of anomaly detection. In fact, the expression of RX is equivalent to the Mahalanobis distance. The second one is a variant of RX, LRX, which estimates the background using local statistics (i.e., mean and covariance matrix within pixels of a given window size). Lastly, we also compare with the projection pursuit (PP) algorithm [7]. PP is a technique that uses one or more linear combinations of the original features to maximize some measure of "interestingness". Then, the optimal projection can be searched by either some evolution algorithm or performing the exhaustive search of the pixel-spectrum. In the experiments, we use Matlab `Tensorlab` toolbox [24] for Tucker decomposition. For LRX, we set the local window size as 11×11.

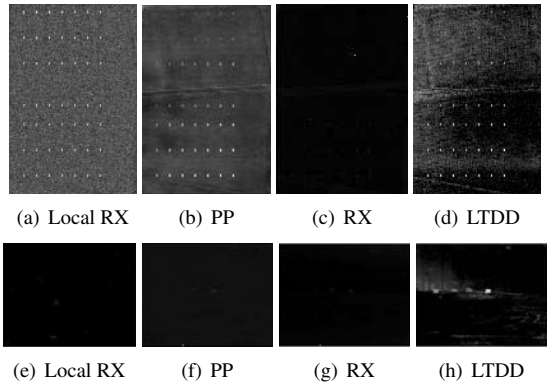
5.2. Results Analysis

We first conduct comparisons between different approaches based on two HSI datasets (synthetic and AF image) with ground truth.

Fig. 2 visually demonstrates the detection results on both the synthetic and AF images. From the detection results of the synthetic HSI data, we can find that the local RX performs the best. This is mainly because all the pixels of anomaly are small points with uniform size, the LRX is the most suitable approach to detect this kind of anomalies. The proposed LTDD algorithm also performs well and is comparable to PP with a false alarm rate around 0.04 as shown in Fig 3(a).

For AF image, we see that LTDD performs best among all the approaches. This is largely due to the homogeneous background of the AF image, which is perfect to be modeled as a low-rank tensor, LTDD is then able to find the "support" anomalies by way of unmixing.

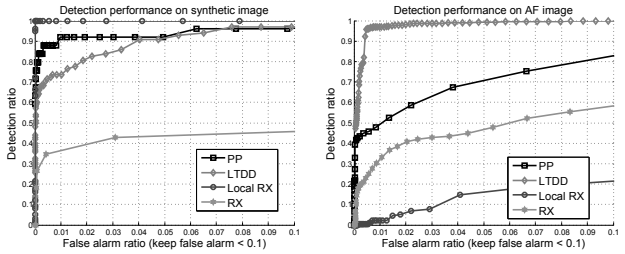
[1] `http://pds-imaging.jpl.nasa.gov/volumes/msl.html`



(a) Local RX (b) PP (c) RX (d) LTDD

(e) Local RX (f) PP (g) RX (h) LTDD

Fig. 2. Detection results of the synthetic image (first row) and AF image (second row) with different approaches.

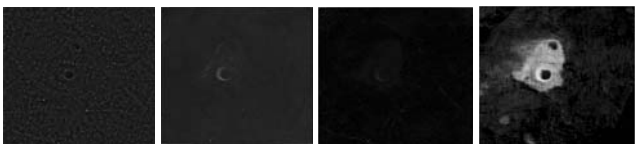


(a) Synthetic (b) AF image

Fig. 3. Quantitative evaluation of the detection results on synthetic image data (left) and AF image (right).

The reason for LRX and RX to perform worse might be because that RX is prone to be dominated by the noticeable outliers in the lower part of the AF image. Same case goes for the PP algorithm. While for LRX it is difficult to predefine a good window size to model the background and the anomaly target in the window patch. Also the PP algorithm cannot easily find the projection for the subspace of anomaly.

For the Mastcam Solday 183 data, since we do not have exact GT, in order to detect the hydrated "anomaly" materials which mostly appears in the drilled hole and the cracks of the soil surface as provided by NASA. We plot the visual detection result of various approach. Based on visual inspection, LTDD gives the best detection result visually.



(a) LRX (b) PP (c) RX (d) LTDD

Fig. 4. Visual detection results of the Mastcam sol183 data with various approaches.

We also compare the computational time between different approaches as listed in Table 1. All the experiments are carried out on a laptop with 2 GHz Intel Core i7 CPU and 8 GB memory running Matlab R2014b. We notice that RX is the most efficient algorithm. LTDD requires much less computational cost as compared to LRX and PP, while keeping acceptable or comparable anomaly detection performance. Both LRX and PP require much processing time since LRX basically runs RX in a patch by patch fashion and PP uses an exhaustive search to find the optimal projection if any.

	Synthetic	**AF image**	**Mastcam Solday 183**
LTDD	39	135	480
RX	4	15	10
Local RX	900	3400	Na
PP	644	7614	1930

Table 1. Comparison of processing time between different approaches (in seconds).

6. CONCLUSION

In this paper, we presented LTDD based on low-rank and sparse tensor representation of the HSI data. The low-rank tensor was further decomposed into a core tensor using Tucker decomposition. An unmixing approach was then adopted for anomaly detection from the reconstructed core tensor. Experiments were conducted on both synthetic and real HSI datasets. Compared to the existing statistical anomaly detection approaches, LTDD performs well through both visual inspection and quantitative analysis while requiring much less processing time.

7. REFERENCES

[1] Chein-I Chang and Shao-Shan Chiang, "Anomaly detection and classification for hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 40, no. 6, pp. 1314–1325, 2002.

[2] Tao Lin and Salah Bourennane, "Survey of hyperspectral image denoising methods based on tensor decompositions," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, pp. 1–11, 2013.

[3] Ledyard R Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

[4] Stefania Matteoli, Marco Diani, and Giovanni Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 25, no. 7, pp. 5–28, 2010.

[5] Irving S Reed and Xiaoli Yu, "Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 38, no. 10, pp. 1760–1770, 1990.

[6] Yuri P Taitano, Brian A Geier, and Kenneth W Bauer, "A locally adaptable iterative rx detector," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, pp. 11, 2010.

[7] José A Malpica, Juan G Rejas, and María C Alonso, "A projection pursuit algorithm for anomaly detection in hyperspectral imagery," *Pattern recognition*, vol. 41, no. 11, pp. 3313–3327, 2008.

[8] Wei Wang, Huijie Zhao, and Chao Dong, "Anomalies detection approach using projection pursuit in hyperspectral image based on parallel genetic algorithm," in *7th International Symposium on Instrumentation and Control Technology*, 2008.

[9] Hsuan Ren and Yang-Lang Chang, "A parallel approach for initialization of high-order statistics anomaly detection in hyperspectral imagery," in *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*. IEEE, 2008, vol. 2, pp. II–1017.

[10] Xiurui Geng, Luyan Ji, Yongchao Zhao, and Fuxiang Wang, "A small target detection method for the hyperspectral image based on higher order singular value decomposition (HOSVD)," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 6, pp. 1305–1308, 2013.

[11] Xiurui Geng, Kang Sun, Luyan Ji, and Yongchao Zhao, "A high-order statistical tensor based algorithm for anomaly detection in hyperspectral imagery," *Scientific reports*, vol. 4, 2014.

[12] Shuangjiang Li and Hairong Qi, "Recursive low-rank and sparse recovery of surveillance video using compressed sensing," in *Proceedings of the International Conference on Distributed Smart Cameras*. 2014, ICDSC '14, pp. 1:1–1:6, ACM.

[13] Shuangjiang Li and Hairong Qi, "Sparse representation based band selection for hyperspectral images," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*, Sept 2011, pp. 2693–2696.

[14] Wei Wang and Hairong Qi, "Unsupervised nonlinear unmixing of hyperspectral image using sparsity constrained probabilistic latent semantic analysis," in *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (Whispers)*, 2013.

[15] Xin Li, Rui Guo, and Chao Chen, "Robust pedestrian tracking and recognition from flir video: A unified approach via sparse coding," *Sensors*, vol. 14, no. 6, pp. 11245–11259, 2014.

[16] Hongyan Zhang, Wei He, Liangpei Zhang, Huanfeng Shen, and Qiangqiang Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, pp. 4729–4743, 2014.

[17] Lieven De Lathauwer, "A survey of tensor methods," in *Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on*. IEEE, 2009, pp. 2773–2776.

[18] Tamara G Kolda and Brett W Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[19] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle, "On the best rank-1 and rank-$(r_1, r_2, \cdots, r_n)$ approximation of higher-order tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2000.

[20] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright, "Robust principal component analysis?," *Journal of the ACM (JACM)*, vol. 58, no. 3, pp. 11, 2011.

[21] Donald Goldfarb and Zhiwei Qin, "Robust low-rank tensor recovery: Models and algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 35, no. 1, pp. 225–253, 2014.

[22] Michael E Winter, "N-findr: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*. International Society for Optics and Photonics, 1999, pp. 266–275.

[23] Rasmus Bro and Sijmen De Jong, "A fast non-negativity-constrained least squares algorithm," *Journal of chemometrics*, vol. 11, no. 5, pp. 393–401, 1997.

[24] L Sorber, M Van Barel, and L Tensorlab De Lathauwer, "v2.0," *Available online, URL: http://www. tensorlab. net*, 2014.