



OPEN

SUBJECT AREAS:

INFORMATION
TECHNOLOGY

APPLIED MATHEMATICS

Received
24 July 2014Accepted
10 October 2014Published
4 November 2014Correspondence and
requests for materials
should be addressed to
X.R.G. (gengxr@sina.
com.cn)

A High-Order Statistical Tensor Based Algorithm for Anomaly Detection in Hyperspectral Imagery

Xiurui Geng¹, Kang Sun¹, Luyan Ji² & Yongchao Zhao¹

¹Key Laboratory of Technology in Geo-spatial Information Processing and Application System, Institute of Electronics, Chinese Academy of Sciences, Beijing, China, ²Ministry of Education Key Laboratory for Earth System Modelling, Centre for Earth System Science, Tsinghua University, Beijing, China.

Recently, high-order statistics have received more and more interest in the field of hyperspectral anomaly detection. However, most of the existing high-order statistics based anomaly detection methods require stepwise iterations since they are the direct applications of blind source separation. Moreover, these methods usually produce multiple detection maps rather than a single anomaly distribution image. In this study, we exploit the concept of *coskewness tensor* and propose a new anomaly detection method, which is called COSD (coskewness detector). COSD does not need iteration and can produce single detection map. The experiments based on both simulated and real hyperspectral data sets verify the effectiveness of our algorithm.

The research on the anomaly detection of hyperspectral data has drawn much attention recently in many fields^{1–3}. The so-called anomaly detection is basically to find out “abnormal” pixels from an image where the targets and their associated background are both unknown. Many anomaly detection methods have been proposed, among which RX Detector (RXD)^{4–6} is the most typical one. It has been applied to both multi and hyper-spectral successfully in terms of anomaly detection. In fact, the expression of RXD is equivalent to the Mahalanobis distance. There are many anomaly detection operators derived from RXD, such as modified RX (MRX), normalized RX (NRX), weighted RX, Causal RX⁷ and adaptive causal anomaly detector algorithm (ACAD)⁸. The low probability detector (LPD)⁹ is another anomaly detector used frequently. LPD determines whether a pixel is abnormal or not according to the relationship between any pixel of the image and the unity vector multiplied by the inversion of the sample auto-correlation matrix. The uniform target detector (UTD) is an evolved version of LPD which has a translational shift of the origin of the image to the mean vector. Kwon¹⁰ proposed a new anomaly detection method, dual window-based eigen separation transform anomaly detector (DWEST). DWEST model involves two local windows, namely inner and outer windows, which are designed to maximize the separation between anomalies and background. The inner window is used to detect the anomalies presented in it, while the outer window is used to model the background of the anomalies. By moving these two local windows in an image, we can calculate the local mean and covariance matrix for each window and their differences. Consequently, anomalies can be extracted by projecting the differential mean between two windows onto the eigenvector associated with the largest positive eigenvalue of the differential covariance matrix. Similar to DWEST, nested spatial window-based target detector (NSWTD) is presented in¹¹. NSWTD model involves three nested local windows, namely, inner, middle and outer windows. The first two windows are used to extract the smallest and largest anomalies respectively, while the outer window is used to model the local background. Moreover, the other key difference of this model from the DWEST and RX-based algorithms is to use the orthogonal projection divergence (OPD) instead of eigenvector projection or sample covariance matrix as a measurement. Based on a nonparametric model, the combined F-Test anomaly detector (CFT) is presented by Rosario¹². The main assumption of this method has an asymptotic behavior of Fisher’s F distribution for data sets which are examined by a common statistical test. Some other anomaly detection methods can be seen at^{13–15}.

The anomaly detectors mentioned above are all basically conducted from the statistical perspective. The statistics used include first-order statistics (e.g., mean vector) and second-order statistics (e.g., covariance matrix). It is not sufficient to use the first-order and the second-order statistics since the distributions of scatter points in feature space for most real images are not normal distributions.



In this study, the third-order statistical tensor is introduced to extract anomalies. In fact, there are many approaches in existing literature for anomaly detection using high-order statistics^{16–17}. However, all these approaches are the direct applications of the blind signal separation (BSS) methods (e.g., FastICA), which generally involve step-by-step iterations to reach the optimal solution. As a result, they are apt to be trapped in local minima. In order to address the convergence issue, Geng¹⁸ introduced the concept of coskewness tensor to hyperspectral data analysis and proposed a target detection method based on higher order singular value detection (HOSVD). Nevertheless, both the BSS-based techniques and HOSVD are in the domain of feature extraction, aiming at the extraction of not only the anomalies but also the other independent components in the image.

In this paper, combining the concept of third-order statistical tensor and the idea of RXD, we present a new anomaly detection method termed coskewness tensor detector (COSD). The proposed method can directly get the distribution of the anomaly of a hyperspectral image without any iteration, which can therefore avoid problems of the BSS-based methods.

Results

Although there are many anomaly detection algorithms based on the 2nd-order statistic, they are all generally derived from RXD. Similarly, we can also derive corresponding algorithms based on COSD. Therefore, the experiments in this study only focus on comparing the performances between COSD, FastICA and RXD. In order to facilitate the comparison between COSD and FastICA, we used skewness as a measure of non-Gaussianity in FastICA. In addition, since FastICA can produce a lot of independent components, we just select the one with the greatest skewness as the detection result.

Evaluation with simulated data. The simulated image of two bands with 50×50 pixels was first used in this experiment. Simulation data consists of two parts, abnormal targets (8×8) and their background. The background pixels fit a Gaussian distribution. The target was located in the upper left corner of the image and randomly scattered outside the background in the feature space of the image (see Fig. 1).

The comparisons between RXD, FastICA and COSD are given in Fig. 2. By a visual comparative analysis, the performance of COSD is superior to that of RXD. Fig. 3 shows the receiver operating characteristic (ROC) curve of detection rate versus false alarm rate for the three algorithms (see reference [15] for their definitions of detection rate and false alarm rate). Clearly, the detection performance of COSD is comparable to that of FastICA, and both are better than that of RXD.

Evaluation with real hyperspectral data. The hyperspectral data of 100×100 pixels from OMIS-II (Operational Modular Imaging Spectrometer) is used to test these methods. The hyperspectral imaging system was developed by Shanghai Institute of Technical Physic, Chinese Academy of Sciences (SITP). The data, which was acquired by the Aerial Photogrammetry and Remote Sensing Bureau in Xi'an, China in 2003, includes 64 bands from visible to thermal infrared with 3.6 m spatial resolution and 10 nm spectral resolution in the visible and near infrared region (60 bands). There were small man-made targets simulated as common objects within the scene, which were distributed at two locations around the top right corner of the image (marked by the rectangles in Fig. 4 a), consisting of tens of pixels. From the true color (approximately) composition image (see Fig. 4 a), it is hard to find any information of the targets in the rectangles.

From Fig. 4b, we can see two small bright blocks at the top right corner, which are the man-made targets. It indicates that RXD can distinguish them as the abnormal pixels from the background. Fig. 4c is the result of COSD, where the two man-made targets are highlighted significantly as the abnormal pixels while the rest is greatly suppressed as the background. The result of NCOSD in Fig. 4d indicates that it is good enough to extract anomalies by only using the skewness information. Fig. 5 shows the ROC (receiver operating characteristic) curve of detection rate versus the false alarm rate for both RXD and COSD. It illustrates that the detection capability of COSD has a significant advantage over that of RXD.

Now we turn to the comparison between FastICA and COSD. Usually, FastICA gets all the independent components through iterations, and the one with the maximum skewness is chosen as the anomaly detection result. However, due to the local optimum, the first independent component of FastICA is not always corresponding to the global maximum skewness, thus different surface objects may be detected (see Fig. 6, which shows the inconsistency of anomaly detections. Besides, the skewness-based FastICA does not have a global convergence. Nevertheless, our COSD does not have these problems.

Discussion

In this study we proposed a new method of using high-order statistic tensor to detect the anomaly of a hyperspectral image and analyzed in detail the application of the skewness tensor in the anomaly detection. Compared to the traditional methods based on the second-order statistics, COSD has a better capacity to extract the abnormal objects. Moreover, COSD can directly get the distribution of the

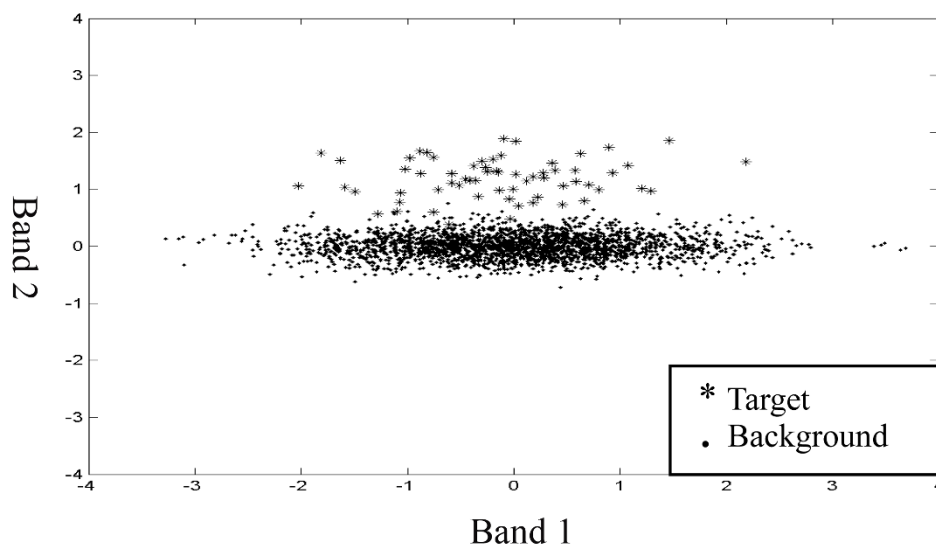


Figure 1 | The scatter plot of the simulated data.

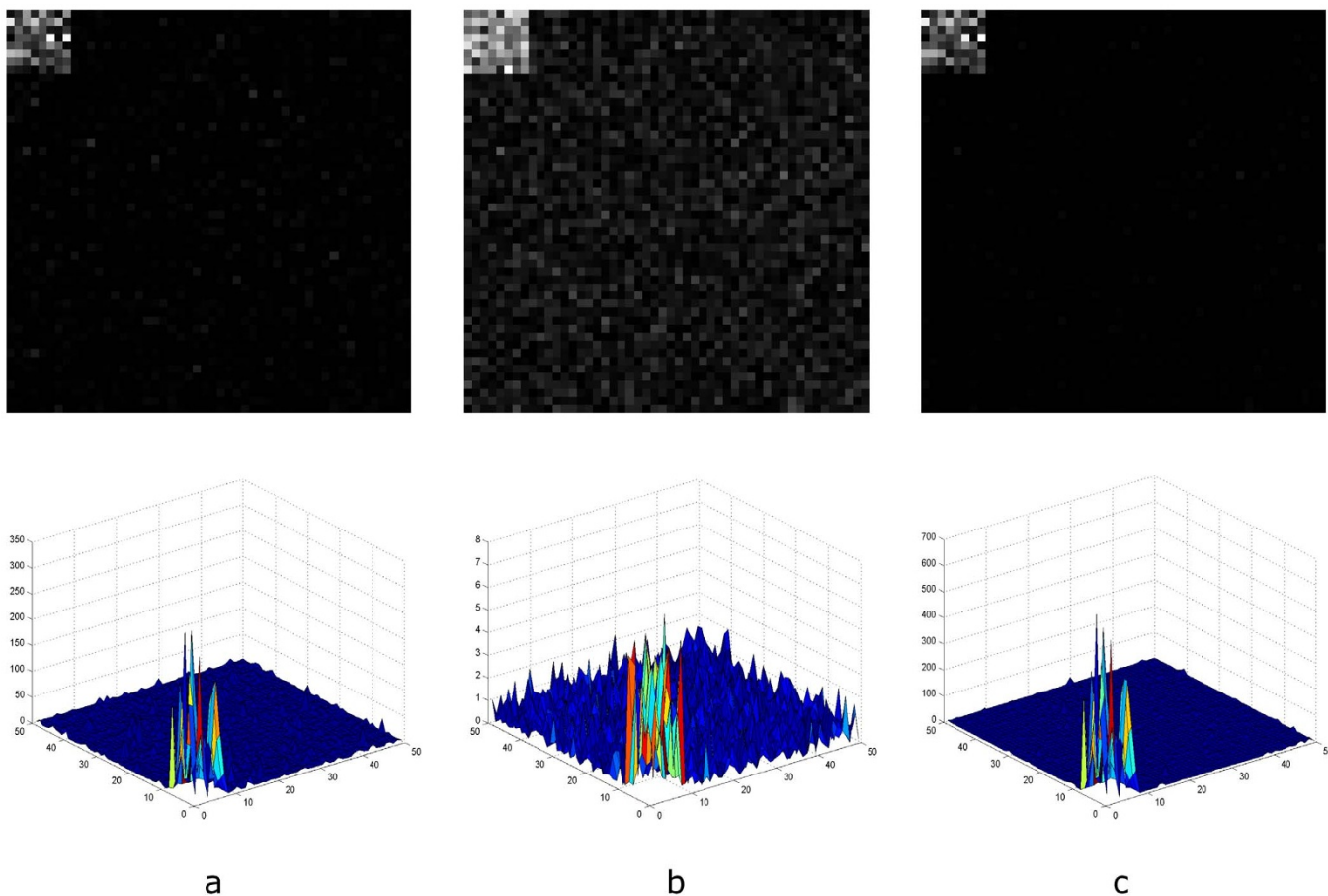


Figure 2 | Anomaly detection result. (a) RXD; (b). FastICA; (c). COSD.

abnormal objects by using a higher-order statistic tensor, compared to the traditional methods based on blind signal separation methods. Since COSD does not need iteration, it can avoid the shortcomings of the blind signal separation methods. By the experiment with simulated data, it shows that the detection performance of the COSD is

better than that of RXD. In the experiment using real hyperspectral data, it is illustrated that the COSD can highlight the man-made targets as the anomalies out of the image successfully. It is noteworthy that, the obtained abnormal pixels might not be the ones of interest due to the uncertainty of the abnormal pixels in an image.

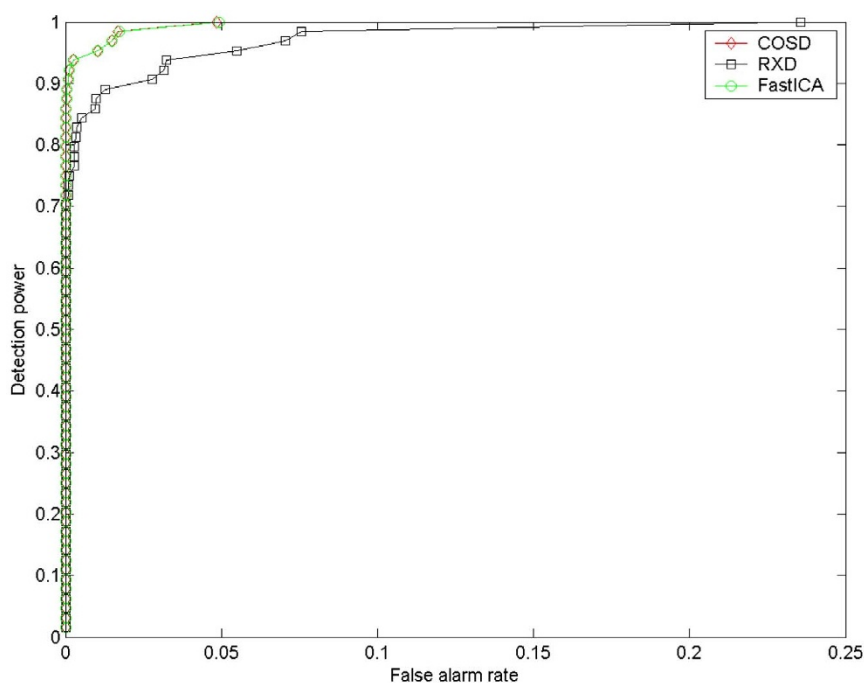


Figure 3 | ROC curves.

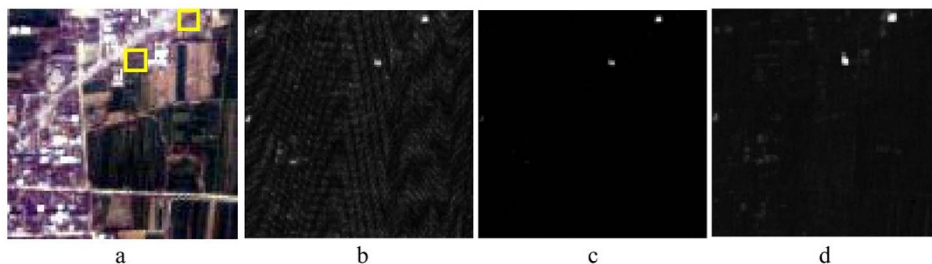


Figure 4 | Anomaly detection results for the real hyperspectral image: (a). True true color composition image, (b). RXD, (c). COSD, (d). NCOSD.

However, the introduction of a higher-order statistic tensor will benefit a lot in the anomaly detection for hyperspectral images.

Although the introduction of coskewness tensor benefits anomaly detection a lot, COSD may suffer from larger computational complexity. Figure 7 shows the computational complexity (measured by the required float operations, flops) for RXD, and COSD. Assume the size of the hyperspectral data is N pixels and L bands. The flops required for RXD is $\left(\frac{3}{2}L^2 + L\right)N$ while that for COSD is $\left(\frac{7}{6}L^3 + L^2 + L\right)N$. We can see that from Fig. 7, COSD is more sensitive to the number of bands. When L is relatively large (for instance, >50), the computational complexity of COSD is about $\frac{7}{9}L$ times of that of RXD.

It is noticeable that since the COSD method can be considered as the extension of the RXD method in formula expression from the 2nd-order statistics (covariance matrix) to the 3rd-order statistics (coskewness tensor), all the other 2nd-order statistics based anomaly detection methods (such as modified RX, weighted RX, causal RX, DWEST, NSWTD) can be simply extended to those 3rd-order statistics based ones or even higher-order statistics based ones. The advantage of our COSD algorithm in detecting anomaly of hyperspectral image ensures a rationality of this extension.

In conclusion, the anomalies generally show strong features in the high-order statistics. Thus, this paper presents a new anomaly

detection method COSD based on third-order statistical tensor. Formally, the COSD is the natural extension of RXD from second-order to third-order statistics. Essentially, the COSD take full advantage of angle information, which ensures the validity of COSD.

Methods

RXD. The RXD is a detector proposed by Reed and Yu⁴. For each pixel vector \mathbf{r} in an image, RXD can be implemented by an operator specified by

$$\delta_{RXD}(\mathbf{r}) = (\mathbf{r} - \boldsymbol{\mu})^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}), \tag{1}$$

where $\boldsymbol{\mu} = \frac{1}{N} \sum_{i=1}^N \mathbf{r}_i$ is the mean vector of the image, \mathbf{K} is the sample covariance matrix of the image and N is the number of pixels. $\delta_{RXD}(\mathbf{r})$ in Eq.(1) has the same form as Mahalanobis distance. The covariance matrix \mathbf{K} can be decomposed as: $\mathbf{K} = \mathbf{E} \mathbf{D} \mathbf{E}^T$, where $\mathbf{D} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_L\}$; \mathbf{E} is the eigenvectors matrix of \mathbf{K} . We denotes

$\mathbf{F} = \mathbf{E} \mathbf{D}^{-1/2}$ as a whitening operator of the image. Then eq.(1) can be transformed as

$$\delta_{RXD}(\mathbf{r}) = (\mathbf{r} - \boldsymbol{\mu})^T \mathbf{F} \mathbf{F}^T (\mathbf{r} - \boldsymbol{\mu}) = (\mathbf{F}^T (\mathbf{r} - \boldsymbol{\mu}))^T (\mathbf{F}^T (\mathbf{r} - \boldsymbol{\mu})). \tag{2}$$

From Eq. (2) we can see that, $\delta_{RXD}(\mathbf{r})$ is actually the Euclidean distance of \mathbf{r} and $\boldsymbol{\mu}$ of the whitened image.

Tensor introduction. A real m -order n -dimensional tensor $\underline{\mathbf{A}}$ consists of n^m real entries¹⁹, represented as a_{i_1, \dots, i_m} where $i_j = 1, \dots, n$ for $j = 1, \dots, m$. Fig. 8 shows an example of third-order 4-dimensional tensor. The tensor $\underline{\mathbf{A}}$ is supersymmetric if its entries are invariant under any permutation of their indices²⁰, or mathematically, $a_{ijk} = a_{ikj} = a_{jik} = a_{jki} = a_{kji} = a_{kij}$.

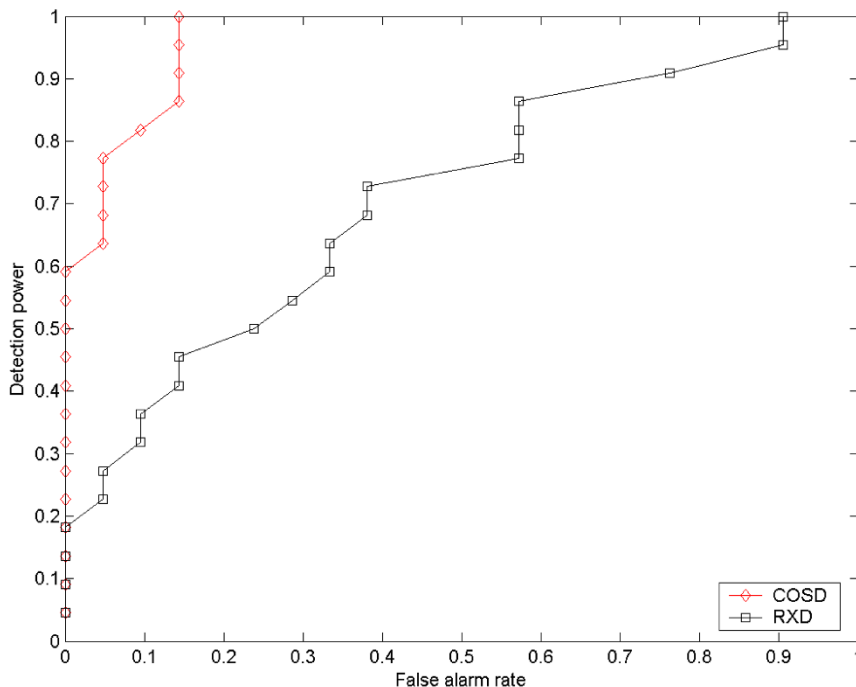


Figure 5 | ROC curves.

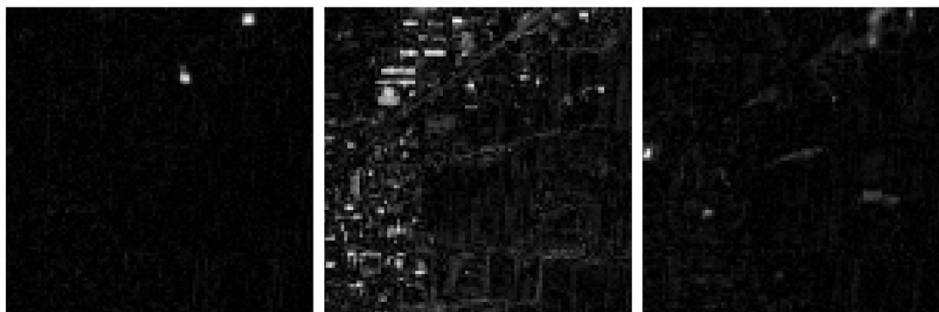


Figure 6 | The different results of the first independent component by FastICA.

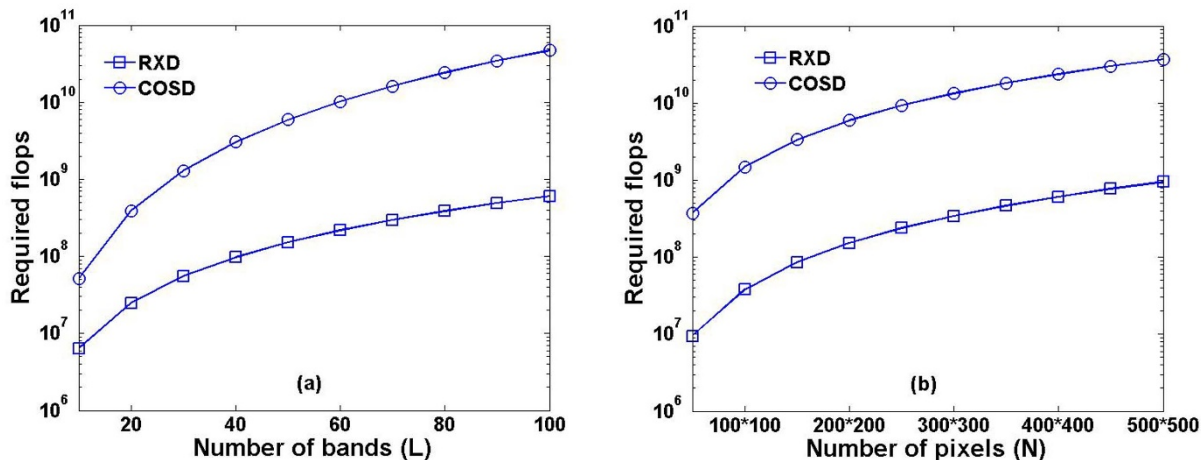


Figure 7 | The computational complexity of RXD and COSD versus (a) the number of bands ($N = 200 \times 200$) and (b) the number of pixels ($L = 50$).

The tensor $\underline{\mathbf{A}}$ defines an m th-degree homogeneous polynomial

$$f(\mathbf{x}) = \underline{\mathbf{A}}\mathbf{x}^m, \tag{3}$$

where $\mathbf{x} = [x_1, \dots, x_n]^T$, \mathbf{x}^m is a tensor with m orders, n dimension and rank being 1²⁰ and its elements are respectively $x_{i_1} \dots x_{i_m}$ where $i_j = 1, \dots, n$ for $j = 1, \dots, m$. $\underline{\mathbf{A}}\mathbf{x}^m$ is the tensor product of $\underline{\mathbf{A}}$ and \mathbf{x}^m ¹⁹. For example, when $m = 2$, $\underline{\mathbf{A}}$ is a matrix of $n \times n$, and $\mathbf{x}^2 = \mathbf{x}\mathbf{x}^T$, $f(\mathbf{x}) = \underline{\mathbf{A}}\mathbf{x}^2 = \mathbf{x}^T \underline{\mathbf{A}}\mathbf{x}$. For m -order tensors, $\underline{\mathbf{A}}\mathbf{x}^m$ can be decomposed in m steps as following:

$$f(\mathbf{x}) = \underline{\mathbf{A}}\mathbf{x}^m = \underline{\mathbf{A}} \times_1 \mathbf{x} \times_2 \mathbf{x} \times_3 \dots \times_m \mathbf{x}, \tag{4}$$

where \times_i denotes the i -mode product operator. Fig. 9 shows the explanation of the multiplication of a third-way tensor and a vector which yields a scalar. As will be seen later, that scalar is the corresponding skewness in the direction \mathbf{x} if $\underline{\mathbf{A}}$ is the coskewness tensor.

For a hyperspectral image data set $S = \{\mathbf{r}_1, \dots, \mathbf{r}_N\}$, its m -order cumulant matrix (tensor) is defined as:

$$\underline{\mathbf{R}} = \frac{1}{N} \sum_{i=1}^N \mathbf{r}_i^m, \tag{5}$$

where \mathbf{r}_i is the spectral column vector of an image; \mathbf{r}_i^m is a m -order L -dimensional tensor (where L is the number of bands in the image) with rank being 1. Obviously, the m -order statistical tensor $\underline{\mathbf{R}}$ is a supersymmetric tensor. This paper will focus on the research of anomaly extraction for the hyperspectral image by using the high-order statistical tensor.

coskewness tensor detector (COSD). For a hyperspectral image data set

$\tilde{S} = \{\tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_N\}$, suppose that $\boldsymbol{\mu} = \frac{1}{N} \sum_{i=1}^N \tilde{\mathbf{r}}_i = \mathbf{0}$ and $E(\tilde{S}\tilde{S}^T) = \mathbf{I}$, where \mathbf{I} is $L \times L$ unit

matrix. It means that each band of the image has a variance of 1, and the correlation coefficient between the bands is zero. That is to say, the hyperspectral data has been normalized. It is not difficult to reach these two conditions. If the mean vector of the image is not zero, it can be achieved by moving the origin of the image to the mean vector. Besides, the real hyperspectral image can meet the second condition by data whitening.

Here, we propose a new anomaly detector, named high-order statistic detector (HOSD), which is defined as follows:

$$\delta_{HOSD}(\bar{\mathbf{r}}) = \underline{\mathbf{K}}\bar{\mathbf{r}}^m, \tag{6}$$

where $\underline{\mathbf{K}} = \frac{1}{N} \sum_{i=1}^N \mathbf{r}_i^m$ is a high-order statistic tensor of the image defined as Eq.(5); \mathbf{r} is

the pixel vector of the image. Similar to Eq.(3), $\underline{\mathbf{K}}\bar{\mathbf{r}}^m$ is a scalar which is the tensor product of the m -order statistic tensor $\underline{\mathbf{K}}$ and $\bar{\mathbf{r}}^m$. In this paper, we will discuss a case where $m = 3$. And then Eq.(5) can be transformed as

$$\delta_{COSD}(\bar{\mathbf{r}}) = \underline{\mathbf{K}}\bar{\mathbf{r}}^3, \tag{7}$$

which can be called as coskewness tensor detector (COSD), and here $\underline{\mathbf{K}}$ is the coskewness tensor. Apparently the coskewness tensor is a supersymmetric third-order tensor with the dimension of $L \times L \times L$.

Like RXD, the coskewness tensor based anomaly detector also uses all pixels of the image by eq. (7) to get a gray image, where the anomalies of the image will appear to be very bright or dark. The dark pixels in the gray image are caused by the negative values from Eq. (7). Therefore, we usually determine anomaly pixel by using the absolute value of the gray image.

In Eq. (7), if $\bar{\mathbf{r}}$ is a unit vector, then $\delta_{COSD}(\bar{\mathbf{r}})$ is the skewness of the image against the direction of $\bar{\mathbf{r}}$. Eq.(7) can be transformed as follows

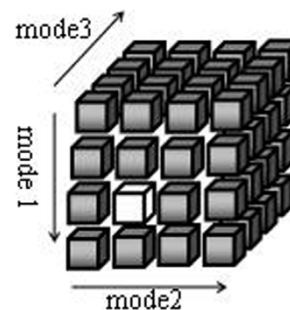


Figure 8 | A sketch map of third-order tensor with size of $4 \times 4 \times 4$.

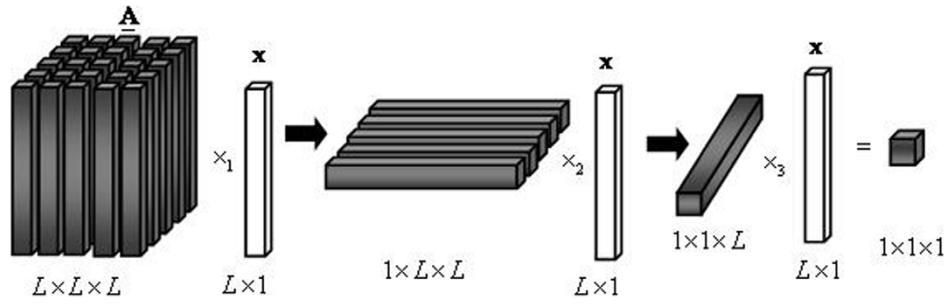


Figure 9 | Illustration of n-mode product of a 3-way tensor and a vector.

$$\delta_{\text{COSD}}(\bar{\mathbf{r}}) = (\|\bar{\mathbf{r}}\|_2)^3 \mathbf{K}(\bar{\mathbf{r}}/\|\bar{\mathbf{r}}\|_2)^3. \quad (8)$$

From Eq. (8), we can see that the anomaly extraction using Eq. (7) is mainly dependent on two indices: One is the skewness and the other is the cube of the 2-norms. If we eliminate the item $(\|\bar{\mathbf{r}}\|_2)^3$ from Eq.(8), the COSD operator becomes a normalized COSD (NCOSD) operator

$$\delta_{\text{NCOSD}}(\bar{\mathbf{r}}) = \mathbf{K}(\bar{\mathbf{r}}/\|\bar{\mathbf{r}}\|_2)^3. \quad (9)$$

Practically, $\delta_{\text{NCOSD}}(\bar{\mathbf{r}})$ is the skewness of image data against the direction of $\bar{\mathbf{r}}$.

Let us assume that hyperspectral image is composed of background and abnormal pixels,

$$\tilde{\mathbf{S}} = \tilde{\mathbf{S}}_b + \tilde{\mathbf{S}}_a, \quad (10)$$

where $\tilde{\mathbf{S}} = \{\tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_N\}$ is the whitened hyperspectral image; $\tilde{\mathbf{S}}_b = \{\tilde{\mathbf{r}}_{b_1}, \dots, \tilde{\mathbf{r}}_{b_{N_b}}\}$ is background (N_b is the number of background pixels); and $\tilde{\mathbf{S}}_a = \{\tilde{\mathbf{r}}_{a_1}, \dots, \tilde{\mathbf{r}}_{a_{N_a}}\}$ is anomaly (N_a is the number of abnormal pixels). It is obvious that $N = N_b + N_a$. The coskewness tensor of the image can be transformed as

$$\mathbf{K} = \frac{1}{N} \sum_{i=b_1}^{b_{N_b}} \tilde{\mathbf{r}}_i^3 + \frac{1}{N} \sum_{i=a_1}^{a_{N_a}} \tilde{\mathbf{r}}_i^3 = \mathbf{K}_b + \mathbf{K}_a, \quad (11)$$

where \mathbf{K}_b is composed of background pixels and \mathbf{K}_a is composed of abnormal pixels. In this study, we just discuss the case that only one class of anomaly lies in the image and denote the spectrum of anomaly as $\tilde{\mathbf{r}}_a$. Then we have

$$\mathbf{K}_a = \frac{N_a}{N} \tilde{\mathbf{r}}_a^3. \quad (12)$$

and the coskewness tensor of the image can be transformed as:

$$\mathbf{K} = \mathbf{K}_b + \frac{N_a}{N} \tilde{\mathbf{r}}_a^3. \quad (13)$$

In general, the number of abnormal pixels in an image is very small, thus $\mu_b \approx \mu = 0$, where $\mu_b = \frac{1}{N_b} \sum_{i=b_1}^{b_{N_b}} \tilde{\mathbf{r}}_i$ is the mean vector of the background image. Accordingly, \mathbf{K}_b can be approximately considered as the coskewness tensor of the background image. If we assume that the background image fits a Gaussian distribution, we can get

$$\mathbf{K}_b \approx \mathbf{0}. \quad (14)$$

It means that all the elements of \mathbf{K}_b are close to zero. Thus the skewness of the image in the direction of $\bar{\mathbf{r}}$ can be expressed as

$$\delta_{\text{NCOSD}}(\bar{\mathbf{r}}) = \mathbf{K}(\bar{\mathbf{r}}/\|\bar{\mathbf{r}}\|_2)^3 \approx \frac{N_a}{N} \tilde{\mathbf{r}}_a^3 (\bar{\mathbf{r}}/\|\bar{\mathbf{r}}\|_2)^3 = \frac{N_a}{N} (\bar{\mathbf{r}}^T \tilde{\mathbf{r}}_a / \|\bar{\mathbf{r}}\|_2)^3 = \frac{N_a}{N} \|\tilde{\mathbf{r}}_a\|_2^3 \cos^3 \theta, \quad (15)$$

where θ is the angle between the vector $\bar{\mathbf{r}}$ and $\tilde{\mathbf{r}}_a$. Considering eq.(15), eq.(8) can be rewritten as

$$\delta_{\text{COSD}}(\bar{\mathbf{r}}) \approx \frac{N_a}{N} \|\bar{\mathbf{r}}\|_2^3 \|\tilde{\mathbf{r}}_a\|_2^3 \cos^3 \theta. \quad (16)$$

Since $\tilde{\mathbf{S}} = \{\tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_N\}$ is an image after centering and whitening, the RX operator can be expressed as

$$\delta_{\text{RXD}}(\bar{\mathbf{r}}) = \|\bar{\mathbf{r}}\|_2^2. \quad (17)$$

From eq.(15–17), it can be seen that there are distinguished differences among RX, NCOSD and COSD. Specifically, RXD depends only on the distance between a pixel

and the origin in the feature space of the whitened image. Only when abnormal pixels are far away from the origin and all the background pixels are relatively close to the origin, RXD can achieve a good anomaly detection result. NCOSD is based on the skewness of the image, and its detection performance is dependent on angles between an abnormal pixel and all the background pixels. When all the angles are large, NCOSD can get a good detection result. As for COSD, it does not only take the distance into account, but also the angle. So it can overcome shortcomings of both RXD and NCOSD, both of which focus only on one single index.

1. Tan, K., Li, E., Du, Q. & Du, P. Hyperspectral Image Classification Using Band Selection and Morphological Profiles. *IEEE J. Sel. Topics Appl. Earth Observ.* **7**, 40–48 (2014).
2. Chang, C.-I. *Hyperspectral Imaging: Techniques for Spectral Detection and Classification* (Plenum Press, New York, 2003).
3. Manolakis, D. & Shaw, G. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Process Mag* **19**, 29–43 (2002).
4. Reed, I. S. & Yu, X. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Trans. Acoust Speech Signal Process* **38**, 1760–1770 (1990).
5. Yu, X., Reed, I. S. & Stocker, A. D. Comparative performance analysis of adaptive multispectral detectors. *IEEE Trans. Signal Process* **41**, 2639–2656 (1993).
6. Yu, X., Hoff, L. E., Reed, I. S., Chen, A. M. & Stotts, L. B. Automatic target detection and recognition in multiband imagery: A unified ML detection and estimation approach. *IEEE Trans. Image Process* **6**, 143–156 (1997).
7. Chang, C. I. & Chiang, S. S. Anomaly detection and classification for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens* **40**, 1314–1325 (2002).
8. Hsueh, M. & Chang, C. I. Adaptive causal anomaly detection for hyperspectral imagery. In *IEEE International Geoscience & Remote Sensing Symposium* **5**, 3222–3224 (2004).
9. Harsanyi, J. C. *Detection and classification of subpixel spectral signatures in hyperspectral image sequences* Ph.D thesis, Baltimore, (1993).
10. Kwon, H. Adaptive anomaly detection using subspace separation for hyperspectral imagery. *Opt Eng* **42**, 3342–3351 (2003).
11. Liu, W. & Chang, C. I. A nested spatial window-based approach to target detection for hyperspectral imagery. In *IEEE International Geoscience and Remote Sensing Symposium* 20–24 (Alaska, 2004).
12. Rosario, D. A nonparametric F-distribution anomaly detector for hyperspectral imagery. In *Aerospace Conference, IEEE 2022–2029* (2005).
13. Schwerizer, S. M. & Moura, J. M. F. Efficient detection in hyperspectral imagery. *IEEE Trans. Image Process* **10**, 584–597 (2001).
14. Gu, Y., Liu, Y. & Zhang, Y. A selective KPCA algorithm based on high-order statistics for anomaly detection in hyperspectral imagery. *IEEE Geosci. Remote Sens Lett* **5**, 43–47 (2008).
15. Chiang, S.-S., Chang, C.-I. & Ginsberg, I. W. Unsupervised target detection in hyperspectral images using projection pursuit. *IEEE Trans. Geosci. Remote Sens* **39**, 1380–1391 (2001).
16. Xun, L. & Fang, Y. Anomaly Detection Based on High-order Statistics in Hyperspectral Imagery. In *The Sixth World Congress on Intelligent Control and Automation* **2**, 10416–10419 (2006).
17. Ren, H. & Chang, Y.-L. A Parallel Approach for Initialization of High-Order Statistics Anomaly Detection in hyperspectral imagery. In *IEEE International Geoscience & Remote Sensing Symposium* **2**, II-1017-II-1020 (2008).
18. Geng, X., Ji, L., Zhao, Y. & Wang, F. A small target detection method for the hyperpectral image based on high-order singular value decomposition (HOSVD). *IEEE Geosci. Remote Sens Lett* **10**, 1305–1308 (2013).
19. Qi, L. Eigenvalues of a real supersymmetric tensor. *J. Symbolic Comput.* **40**, 1302–1324 (2005).
20. Kofidis, E. & Regalia, P. On the best rank-1 approximation of higher-order supersymmetric tensors. *SIAM J Matrix Anal A* **23**, 863–884 (2002).

Acknowledgments

We are very grateful to Dr. Suhong Liu from Beijing Normal University, China for her great help in language.



Author contributions

X.G. conceived the idea. X.G., K.S., L.J. and Y.Z. designed and performed the experiments and analyzed the data. X.G., K.S. and L.J. wrote the main manuscript text. All authors reviewed the manuscript.

Additional information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Geng, X., Sun, K., Ji, L. & Zhao, Y. A High-Order Statistical Tensor

Based Algorithm for Anomaly Detection in Hyperspectral Imagery. *Sci. Rep.* **4**, 6869; DOI:10.1038/srep06869 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>