Robust Metric based Anomaly Detection in Kernel Feature Space

Bo Du[1,2], Liangpei Zhang[2], Huang Xin[2]
1 School of Computer Science, Wuhan University
2 The State Key Laboratory of Information Engineering in
Surveying, Mapping, and Remote Sensing
Wuhan University, P.R. China.

Abstract: This thesis analyzes the anomalous measurement metric in high dimension feature space, where it is supposed the Gaussian assumption for state-of-art mahanlanobis algorithms is reasonable. The realization of the detector in high dimension feature space is by kernel trick. Besides, the masking and swamping effect is further inhibited by an iterative approach in the feature space. The proposed robust metric based anomaly detection presents promising performance in hyperspectral remote sensing images: the separability between anomalies and background is enlarged; background statistics is more concentrated, and immune to the contamination by anomalies.

Keywords: anomaly detection, hyperspectral images, Manhanlobis distance

Introduction

Anomaly targets in hyperspectal images (HSI) refer to those deviating obviously from the other background pixels, especially by means of the spectral feature [1]. Typical ones are the man-made objects in nature scene, such as the vehicles in a grass field. State-of-arts methods mainly evaluate it by exploiting the distance of an observing pixel to the background statistics center. So the key is the background statistics, or the anomalous metric. RX and its variants take use of a Manhanlobis distance from background statistics [2]. In spite of their effectiveness, they are proved to be susceptible to the masking and swamping effect, due to the contaminated background statistics [3]. Multivariate outlier detection methods, focusing to alleviate this effect, figure out a more robust metric by eliminating the probable background pixels or a contracting iteration procedure to obtain a new covariance matrix [3, 4]. Traditional ways include iterative exclusion algorithm, with

each iteration excluding the most anomalous samples until the rest samples unchanged. Then the metric by the rest samples is believed to be immune to the anomalies, or a robust one. However, the robust metric anomaly detection methods don't take into consideration the nonlinear relationship between different bands of the hyperspectral images. The Gaussian assumption of the target present hypothesis and target absent hypothesis may not be valid either [5]. Besides, on the boundary between background and anomaly it is very common to find the mixed pixels. So does it when the size of anomaly is smaller than the images' spatial resolution. The cases become much worse when the pixels are seriously mixed, or the nonlinear mixed pixels. Kernel based anomaly detectors, such as kernel-RX, have been developed to solve the above problem, while the metric in high dimension feature space is not robust since the anomaly pixels may be contained in the background gram matrix.
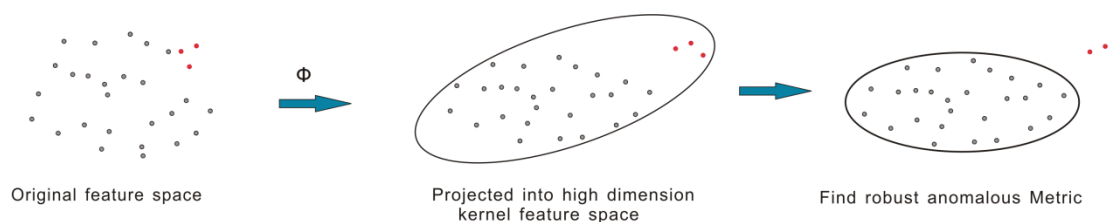

Fig. 1   The schematic flow sheet of our method. The black dots represent the background pixels, and the red ones corresponding to the anomaly targets.

This thesis proposes a new anomaly detector by exploiting a robust metric in the kernelized feature space. The idea is shown in Fig. 1, where vectored pixels in original feature space may not be fit for Gaussian distribution, but it is the case for some high dimension feature space like the middle picture with the counter corresponding to the Manhanlobis metric. But only with the metric excluding the anomaly pixels can the real anomalous degree of the anomalies be presented, shown in the last picture.

Robust Anomalous Metric in High Dimension Feature Space

Traditional target detection methods exploit the linear separablity between targets and backgrounds signals [6]. Classic approaches include subspace model and linear mixture model [6]. Due to the lack of prior information on targets, anomaly detection methods depend on the

measurement metric from background pixels, where the background pixels dataset is usually contaminated since it is usually composed of all the pixels in the image [7]. Robust Mahanlobis distance based methods construct an iterative procedure [4]. In each iteration the first and second order statistics are computed to figure out the anomalous distance, then the pixels presenting distance larger than a predefined threshold would be excluded and the statistics are updated by the rest dataset. The iteration is done until the rest dataset wouldn't change. The underlying idea is that the contamination by anomalies can be gradually eliminated by a dataset shrinking procedure, and the anomalies can be detected at the same time.

Hyperspectral images contain a large number of spectral bands. Though state-of-art methods prove promising in separating anomalies from backgrounds, the nonlinear correlations between different spectral bands are not considered. Different materials present spectrally absorption at different spectral position, so that the nonlinear correlations are not evitable. Kernel based anomaly detection methods have made great success, the typical ones are kernel-RX. Another factor that needs further investigation is that the mixed manner of each pixel is much more complex than linear mixture model. As the spatial resolution is limited, intimate mixture, instead of linear mixture, is more widespread and reasonable [8]. In intimate mixtures, the photons are interacting with all the materials simultaneously. In linear mixtures, the assumption was the photons scattered off one material at a time. Since intimate mixtures have multiple different materials in close relation to one another, the photons can bounce from one particle to another causing different absorption and reflection effects. The result is mixing that cannot be well captured by simple linear models [8]. Inspired by kernel-RX and robust anomaly detection methods, we proposed the new robust one, with the detailed steps are presented as following:

Step 1: Since the gram matrix is usually N×N with N being the number of background pixels, it is not possible to consider all the pixels at one time otherwise it would exceed the computing capacity very easily. So a k-means clustering method is employed to segment the dataset into k classes.

Step 2: For each clustered class, all the pixels are projected into the high feature space $x \rightarrow \phi(x)$, constituting a new dataset D. It is assumed that

Step 3: The statistics of these projected pixels from D are figured out, including mean m and covariance C.

$$m_\Phi = \frac{1}{M} \sum_{i=1}^{M} \phi(\mathbf{x}_i) \qquad (1)$$

$$\mathbf{C}_\Phi = \begin{bmatrix} \sigma_{11}\sigma_{12}...\sigma_{1b'} \\ ... \\ \sigma_{b'1}\sigma_{b'2}...\sigma_{b'b'} \end{bmatrix} \qquad (2)$$

where $\sigma_{ij} = \frac{1}{M} \sum_{k} (\phi(\mathbf{x}_{ik}) - \mathbf{m}_i)(\phi(\mathbf{x}_{jk}) - \mathbf{m}_j), k=1,...,M$, M is the number of pixels in the dataset, m_i and m_j are the means of the ith band and the jth band, respectively. $\mathbf{x}_{ik}$ is the value in the ith band of the kth pixel.

Step 4: Define the Manhalobis distance metric of each pixel in D by the mean and covariance.

The numerical expression of the anomaly detector in each iteration is given as:

$$d(\phi(x_i)) = (\phi(x_i) - m_\Phi)C_\Phi(\phi(x_i) - m_\Phi)^T \qquad (3)$$

With the algebra computation as well as the kernelized trick, the distance can be done by dot product of pixels in original low dimension feature space. By spectrum decomposition, the background covariance matrix can be transferred into:

$$C_\Phi = V_\Phi \Lambda V_\Phi{}^T \qquad (4)$$

where V_Φ and Λ are the eigenvectors and eigenvalues matrix, respectively. It is proved that each eigenvectors in the feature space can be expressed as the linear composition of the centralized input samples in the feature space [5].

$$v_\Phi{}^j = \sum_{i=1}^{M} \beta_i{}^j \Phi(x_i) = X_\Phi \beta^j \qquad (5)$$

where X_Φ is the composed of the kernelized input dataset samples and β^j is the eigenvectors of the centered kernel matrix (Gram matrix). Due to the physical structure in high dimension feature space and some formulas computation, all the eigenvectors with nonzero eigenvalues:

$$V_\Phi = X_\Phi \beta \qquad (6)$$

By substitute and into , and some similar algebra computation, the final expression of the detector is:

$$d(\phi(x_i)) = (K_r^T - K_{m_\Phi}^T)^T \widehat{K}_b^{-1}(K_r^T - K_{m_\Phi}^T) \qquad (7)$$

where each item is a centralized Gram matrix [5], which can be figured out by the kernel function on original samples.

Step 5: Eliminate the pixel with the distance larger than a predefined threshold to update D. Some adaptive thresholding approach can be used. For simplicity, a percentage of the number of the samples is used. For example, the largest 1% samples can be excluded.

Step 6: Iterate Step 2. to Step 4. until the dataset D doesn't change.

Step 7: After all the clustered classes are performed, the pixels excluded in each clustered class are labeled as the anomaly targets.

Experiments and anslysis

Experiments on real-world hyperspectral images have been done to evaluate our proposed method, including air-born hyperspectral remote sensing images and near scene hyperspectral images. Five rows of panels distribute in the scene and considered as anomalies, shown in Fig.2 . Several state of art methods are used as comparison ones. Our method iterates 5 times until the results keep stable. Preliminary experiments results with AVIRIS hyperspectral images are shown in Fig. 3. It is obvious that our proposed method did best among all the methods. Considering the mixed boundary anomaly pixels, the performances of different methods for different kinds of anomalies are also presented in Table I, which further reveal that the improvement of our method is partly due to the superior performance on the transition boundary pixels.
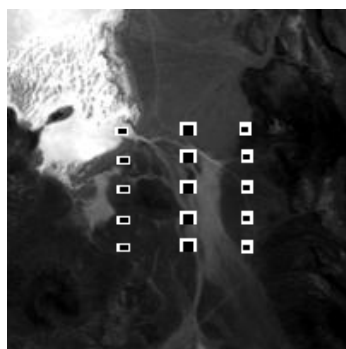

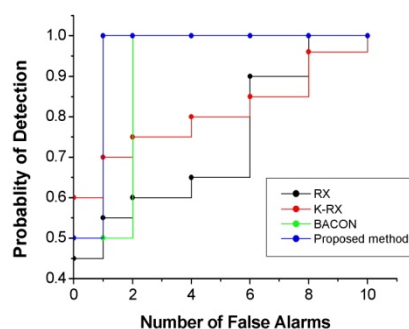Fig. 2. Target panels in the AVIRIS image.

Fig. 3. Performances of different methods.

Table.1 Number of detected anomalies under minimized false alarms

Methods	Proposed Method	RX	K-RX	BACON
Number of Boundary anomaly pixels detected	13	4	6	7

Conclusion

Combining the robust Manhanlobis anomaly detection methods and nonlinear mixture models, a robust metric based anomaly detection method in kernel feature space is proposed. Experiments reveal that the proposed method does provide a more reliable and robust metric for anomaly detection from hyperspectral remote sensing images, especially for detecting the ones on resolved pixels.

Reference

[1] B. Du and L. Zhang, "Random Selection based Anomaly Detector for Hyperspectral Imagery," *IEEE Trans. Geosci. Remote Sens.,* vol. 49, no. 5, pp. 1578 – 1589, 2011.

[2] 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, pp.1760–1770, 1990.

[3] T. E. Smetek and K.W. Bauer, "A Comparison of Multivariate Outlier Detection Methods for Finding Hyperspectral Anomalies," *Military Operations Research*, vol. 13, no. 4, pp. 19–44, 2008.

[4] T. E. Smetek, "Hyperspectra Improved Anomaly Detection and Signature Matching Methods," *Ph.D. dissertation*, Air Univ., Ohio, 2007.

[5] H. Kwon and N. M. 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.

[6] D. Manolakis and G. Shaw, "Detection Algorithms for Hyperspectral Imaging Applications," *IEEE Signal Processing Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[7] 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, 2002.

[8] J. Plaza, P. Martinez, R. Perez, and A. Plaza, "Nonlinear neural network mixture models for fractional abundance estimation in AVIRIS hyperspectral images," in *Proc. XIII JPL Airborne Earth Science Workshop*, Pasadena, CA, 2004.

[9] J. B. Broadwater and A. Banerjee, "A comparison of kernel functions for intimate mixture models," in *Proc. of IEEE WHISPERS '09*, Grenoble, France, Aug. 2009, pp. 1–4.