Kernel-based Invariant Subspace Method for Hyperspectral Target Detection

Ye ZHANG, Yanfeng GU

Department of Information Engineering, Harbin Institute of Technology, Harbin 150001, China

ABSTRACT

In this paper, a kernel-based invariant subspace detection method is proposed for small target detection of hyperspectral images. The method combines kernel principal component analysis (KPCA) and linear mixture model (LMM). The LMM is used to describe each pixel in the hyperspectral images as mixture of target, background and noise. The KPCA is used to build subspaces of target and background. A generalized likelihood ratio test is used to detect whether each pixel in hyperspectral image includes target. The numerical experiments are performed on AVIRIS hyperspectral data with 126 bands. The experimental results show the effectiveness of the proposed method and prove that this method can commendably overcome spectral variability in the hyperspectral target detection, and it has good ability to separate target from background.

1. INTRODUCTION

Relative to multispectral sensing, hyperspectral sensing can increase the detectability of pixel size targets by exploiting finer detail in the spectral signatures of targets and natural backgrounds [1,2]. Recently, hyperspectral target detection algorithms mainly include statistical classification and spectra-based matched filtering approaches. Spectral Angle Mapper (SAM) and Matched Filtering Detector (MFD) are two classical approaches to detect full pixel targets from multispectral data [5], but they are usually effectless for processing the scene of the mixed pixels. Chein-I Chang firstly introduced the concept of the subspace projection into multispectral image classification and proposed the orthogonal subspace projection (OSP) method. Manolakis *et al* proposed a hyperspectral subpixel target detection algorithm, based on the linear mixture model (LMM) and generalized likelihood ratio test (GLRT) [1,2,6]. This algorithm uses the concept of subspace to describe target and background respectively and defines that a pixel is composed of target, background and noise. This subspace model can better describe pixel information beyond subpixel level. To overcome the effect of the spectral variability in terms of target detection, Thai *et al* investigated an invariant subpixel material detection (ISMD) algorithm [4,7]. The algorithm uses a small number of target basis vectors to define a target material subspace, and the defined target subspace contains material information over illumination and atmospheric conditions.

In this paper, a new hyperspectral target detection method is proposed and is called kernel-based invariant subspace detection (KISD), which combines kernel principal component analysis (KPCA) with linear mixture model (LMM). In terms of the current problems of the mixed pixel and spectral variability in the hyperspectral target detection, the LMM is used to describe each pixel in the hyperspectral image, and each pixel in the scene is considered as the combination of target, background and noise. The KPCA is used to construct subspace of target and background, and hyperspectral target detection is performed by introducing a generalized likelihood ratio test (GLRT).

This work is supported by Natural Science Foundation of China, project No. 60272073

2. Kernel-based Invariant Subspace Detection

2.1. Linear mixture model

In hyperspectral target detection, the main task is to detect, discriminate and identify materials. The concrete processing includes two basic approaches: region-by-region and pixel-by-pixel. The pixel-by-pixel processing is necessary when the spatial resolution of images is low and mixed pixels exist in the hyperspectral images. Thus, identifying target pixel by pixel is a dependable approach in most case, and the detection algorithm is to estimate whether a pixel is target or it includes target.

In the LMM, the spectrum of a mixed pixel is represented as a linear combination of component spectra (endmembers). The weight of each endmember spectrum (abundance) is proportial to the fraction of the pixel area covered by the endmember [2,3]. Assume that there are hyperspectral images with N bands. Each pixel in the hyperspectral images can be represented by N-dimensional vector. Under such a circumstance that a pixel in the hyperspectral images is considered in terms of the task of pixel target detection, it is assumed that each pixel includes background and noise, and possibly includes target. Thus, the linear spectral mixing of the pixel can be expressed by

$$\mathbf{y} = \boldsymbol{\alpha}\mathbf{T} + \boldsymbol{\beta}\mathbf{B} + \mathbf{n} \quad (1)$$

where T and B represent the target and the background respectively, the $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ represent the abundance of target and background respectively, and the **n** is noise fraction. Generally, it is assumed that the number of the target is less than or equal to one in each pixel, and the number of the background exceeds one.

2.2. Constructing subspace by using KPCA

In practice, the spectrum of the same material changes with atmospheric conditions, illumination, and other factors. This variability can be described using a statistical distribution or a linear subspace with a dimension less than the number of bands. One way to obtain such a subspace is to use principal component analysis (PCA) [4,7]. But PCA is linear transform and cannot better extract useful features from the hyperspectral images with nonlinear characteristics. Therefore, in the KISD method, KPCA[8] is introduced to construct linear subspace of the target and background in high dimensional feature space.

Given a set of centered random samples $\mathbf{x}_k\,(k=1,2,\cdots,n)$, $\mathbf{x}_k \in R^d$, $\sum_{k=1}^{n}\mathbf{x}_k = 0$. Firstly, the sample set $\{x_k\}$ is mapped into a feature space R^χ by ϕ $(\phi : R^d \to R^\chi)$ and the covariance matrix Σ_ϕ is computed in the feature space R^χ ,

$$\Sigma_\phi = \frac{1}{n}\sum_{j=1}^{n}\phi(x_j)\phi(x_j)^T \tag{2}$$

Let $\mathbf{V}$ $(\mathbf{V}\neq 0)$ be an eigenvector of Σ_ϕ that corresponds to positive eigenvalue λ of Σ_ϕ . So the eigenvector is in the space spanned by the mapped samples, i.e. $\mathbf{V} \in span\{\phi(x_1),\phi(x_2),\cdots,\phi(x_n)\}$. This can be described as

$$\mathbf{V} = \sum_{i=1}^{n}\beta_i\phi(x_i) \tag{3}$$

Thus, eigenvalue decomposition can be written as

$$\lambda\cdot\mathbf{V} = \Sigma_\phi\cdot\mathbf{V} \tag{4}$$

Furthermore, we multiply by $\phi(x)$ at both sides of equation (4) and obtain the expression as

$$\lambda(\phi(x)\cdot\mathbf{V}) = (\phi(x)\cdot\Sigma_\phi\cdot\mathbf{V}) \tag{5}$$

For the all n eigenvectors, expression (5) can be also written as

$$\lambda(\phi(x)\cdot\mathbf{V}^k) = (\phi(x)\cdot\Sigma_\phi\cdot\mathbf{V}^k) \tag{6}$$

where $k=1,2,\cdots,n$, $\mathbf{V}^k = \sum_{i=1}^{n}\beta_i^k\phi(x_i)$.

According to the kernel function constructing method, define a $n\times n$ dimensionality of kernel matrix $\mathbf{K}$ in feature space as

$$(\mathbf{K})_{i,j} = (\phi(x_i)\cdot\phi(x_j)) = k(x_i,x_j) \tag{7}$$

and consider an eigenvalue decomposition for the expansion coefficients β^k by using kernel matrix $\mathbf{K}$ as

$$\lambda\beta^k = \mathbf{K}\beta^k\,,\ \left(\beta^k = (\beta_1^k,\beta_2^k,\cdots,\beta_n^k)^T\right) \tag{8}$$

The obtained solution (λ_k,β^k) needs to be normalized by imposing $\lambda_k(\beta^k\cdot\beta^k)=1$ and to be centered by substituting centered $\mathbf{K}_c$ for the $\mathbf{K}$. The $\mathbf{K}_c$ is given by

$$\mathbf{K}_c = \mathbf{K} - 1_n\cdot\mathbf{K} - \mathbf{K}\cdot 1_n + 1_n\cdot\mathbf{K}\cdot 1_n \tag{9}$$

where 1_n is $n\times n$ matrix, of which all elements are equal to $1/n$.

To obtain a new feature of the samples, which projects the mapped sample $\phi(x)$ on vector $\mathbf{V}^k$ is necessary, namely,

$$(\mathbf{V}^k\cdot\phi(x)) = \sum_{i=1}^{n}\beta_i^k(\phi(x_i)\cdot\phi(x)) = \sum_{i=1}^{n}\beta_i^k k(x_i,x) \tag{10}$$

Because the expansion coefficient $\beta^k(k=1,2,\cdots n)$ is obtained by solving the eigenvalue of matrix $\mathbf{K}_c$, so such expansion coefficients are useless in the weighting mapped samples if $\beta^k(k=1,2,\cdots,n)$ is corresponding to zero eigenvalues of matrix $\mathbf{K}_c$. To avoid the occurrence of this situation, a matrix decomposition is introduced to further reduce computational loads.

According to matrix theory, any real symmetric and positive semidefinite matrix $\mathbf{H}$ with rank q can be diagonalized as

$$\mathbf{H}_{n\times n} = \mathbf{U}_{n\times q}^T\mathbf{Q}_{q\times q}\mathbf{U}_{q\times n} \tag{11}$$

where $\mathbf{Q}$ is a nonsingular diagonal matrix with rank $q(q\leq n)$ containing only the positive eigenvalues of $\mathbf{H}$. $\mathbf{U}$ is a row orthogonal matrix, i.e. $\mathbf{U}\mathbf{U}^T = \mathbf{I}_q$ and q row vectors of $\mathbf{U}$ are corresponding eigenvectors.

Without loss of generality, it is supposed that all eigenvalues of kernel matrix $\mathbf{K}_c$ are $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_n\geq 0$. Substituting the spectral decomposition for the eigenvalue decomposition of kernel matrix $\mathbf{K}_c$ in the centered PCA method, we can obtain nonsingular diagonal matrix $\mathbf{K}_q$ with rank $q(q\leq n)$ and the expansion coefficients $\beta^k(k=1,2,\cdots,q)$ that are eigenvectors of $\mathbf{K}_q$. Thus the equation (10) can be simplified as

$$S = (\mathbf{V}^k\cdot\phi(x)) = \sum_{i=1}^{q}\beta_i^k k(x_i,x) \tag{12}$$

where $k=1,2,\cdots,q(q\leq n)$.

Thus linear kernel subspace in high dimensional feature space is obtained by KPCA. Using such a processing can obtain the linear kernel subspace of the target and background respectively.

2.3. Kernel-based invariant subspace detection

Using the linear mixture model (1) and substituting the kernel subspace of target and background for the original subspace, the hyperspectral target detection can be described as the following hypothesis test

$$\begin{aligned} H_0&: \ \mathbf{y} = \boldsymbol{\beta}_k \mathbf{B}_k + \mathbf{n} \\ H_1&: \ \mathbf{y} = \boldsymbol{\alpha}_k \mathbf{T}_k + \boldsymbol{\beta}_k \mathbf{B}_k + \mathbf{n} \end{aligned} \quad (13)$$

where T_k and B_k represent the target and the background respectively, and the $\boldsymbol{\alpha}_k$, $\boldsymbol{\beta}_k$ represent the abundance of target and background respectively, and the $\mathbf{n}$ is Gaussian white noise fraction. Thus the likelihood ratio is given

$$L(y) = \frac{p_1(\mathbf{y}/H_1)}{p_0(\mathbf{y}/H_0)} \quad (14)$$

where $p_0(y/H_0)$ and $p_1(y/H_1)$ are the conditional probability density function of the observation y under two hypothesis tests of H_1 and H_0 respectively [4,7].

The reference [4] provides a concrete expression of the generalized likelihood ratio as the following

$$\widetilde{L}(y) = \frac{\mathbf{y}^T(\mathbf{I} - \mathbf{B}\mathbf{B}^T)\mathbf{y}}{\mathbf{y}^T(\mathbf{I} - \mathbf{A}\mathbf{A}^+)\mathbf{y}} \quad (15)$$

where $\mathbf{A}_k = [\mathbf{B}_k \ \ \mathbf{T}_k]$, A_k^+ is the pseudo-inverse of A_k, and I is the identity matrix. Using the Gram-Schmidt process can extract matrix Q_k from A_k, of which orthonormal columns span the same subspace as the columns of A_k. Thus expression (15) can be modified as the following

$$\widetilde{L}(y) = \frac{\mathbf{y}^T(\mathbf{I} - \mathbf{B}_k\mathbf{B}_k^{\ T})\mathbf{y}}{\mathbf{y}^T(\mathbf{I} - \mathbf{Q}_k\mathbf{Q}_k^{\ T})\mathbf{y}} \quad (16)$$

The expression (16) can be used as a discriminant to detect hyperspectral small target pixel by pixel. If the likelihood ratio in the expression (16) exceeds the detection threshold in terms of a pixel, it can be concluded that the pixel includes target. Otherwise, it can be considered that there is no target in the pixel. So the kernel-based invariant subspace target detection is realized.

The main procedure of the KISD can be summarized in the following steps. Firstly, the KPCA is used to calculate the kernel subspace T_k and B_k of the target and background on their sample data. Secondly, T_k and B_k are used to construct full subspace matrix A_k, and Q_k is obtained by using the Gram-Schmidt processing on A_k. Thirdly, the likelihood ratio is calculated by using the expression (16). Finally, the target detection is performed, based on the likelihood ratio calculated.

Generally, the sample data of the target are known, and the sample data of the background are uniformly selected from the hyperspectral images.

3. EXPERIMENTS AND RESULTS

3.1. Data description

In order to test the effectiveness of the proposed method for HTR, the numerical experiments are performed on AVIRIS data, which is an AVIRIS data set of naval military base acquired in San Diego, USA. After removing those bands that are corresponding to the water absorption regions, low SNR and bad bands, we remain 126 bands available in the 0.4–1.8 μm wavelength range. A scene of 100×100 pixels was selected for our experiments, in which there are three small planes as targets for our detection. The ground sampling distance of those hyperspectral images is 3.5*m*.

3.2. Experiments and discussion of the results

In the experiments, in order to test the ability of the proposed algorithm in terms of spectral variability, 3 classes of different ground covers are randomly selected as the training samples of the background (non-target), and the training samples of target and background are obtained from spectral library. Figure.1 provides spectrum of two target samples that are randomly selected from the target training set. By comparing the spectrum of two targets, it can be found that the spectral difference is great.

In order to prove the effectiveness of the proposed algorithm, ISMD, MFD and OSP are realized in the experiment to be compared with the proposed method. In proposed kernel-based methods, a radial basis function kernel is used, which is defined as

$$K(x, y) = \exp\left(\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (17)$$

where σ set to 200 in the experiments.

The detection results of the four algorithms show in the Figure.2. The dimensionality of the target subspace in four algorithms is one and the dimensionality of the background subspace is selected as 1, 5, and 10 in each algorithm respectively. From the corresponding detection result, it can be found that the separability between targets and background of the proposed algorithm and ISMD is better than the MFD and the OSP. This indicates detection ratios of the proposed algorithm and ISMD are higher than the MFD and the OSP. In addition, it also can be found that the detection result obtained by the proposed algorithm under different numbers of the background subspace is almost identical. This indicates the construction of the background subspace in the proposed algorithm is more effective than the other algorithms.

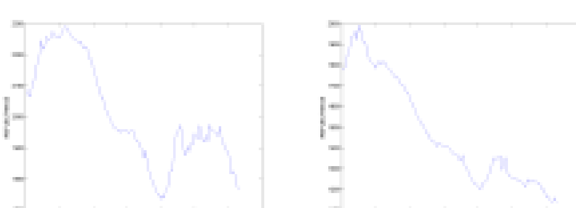

Fig 1. Spectral variability of targets from two different pixels

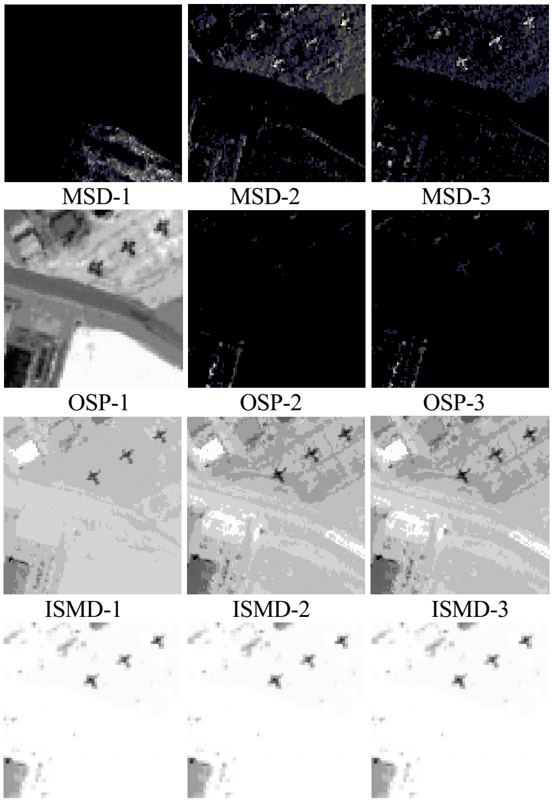

Fig 2. The detection result by using four algorithms. The dimensionality of the background subspace is 1, 5 and 10 respectively.

In order to describe the false detection ratio, an evaluation rule is defined as

$$P_{fdr} = \frac{N_f}{N - N_t} \tag{18}$$

where N_f is the number of the detected non-targets, N is total number of the image pixels and the N_t is the number of the target pixels resident in the testing scene. The Table1 gives the numerical comparison of four algorithms in terms of the false detection ratio where horizontal grid describes the number of the detected pixels that include true and false targets.

According to the experimental data from Table1, it is clearly found that the false detection ratio of the proposed algorithm is less than the other three algorithms whereas same number of pixels that possibly includes true and false targets is detected.

4. CONCLUSION

In this paper, a kernel-based invariant subspace detection (KISD) method is proposed for target detection of hyperspectral images. The method combines KPCA and LMM. In the method, KPCA is used to build the kernel subspace of target and background. The experimental results prove that this method has good ability to construct background subspace and can greatly overcome the spectral variability in the hyperspectral target detection. In terms of the false detection ratio, the proposed KISD method is better than the other.

Table 1 Comparison of the false detection ratio between four algorithms

$P_{fdr}(10^{-3})$		MSD	OSP	ISMD	KISD
The number of the detected pixels (include true and false targets)	70	13.09	11.48	4.431	4.030
	71	13.00	11.58	4.330	4.130
	72	13.09	11.68	4.230	4.030
	73	13.19	11.78	4.330	4.130
	74	13.29	11.88	4.431	4.230
	75	13.39	11.98	4.532	4.330
	76	13.49	12.08	4.632	4.431
	77	13.60	12.18	4.733	4.532
	78	13.70	12.28	4.834	4.431
	79	13.80	12.38	4.935	4.532
	80	13.90	12.48	5.035	4.632

5. REFERENCES

[1] D.Manolakis and G.Shaw. Detection Algorithms for Hyperspectral Imaging Applications. IEEE Signal Processing Magazine, pp. 29-43, January 2002.

[2] D.Manolakis, C.Siracusa and G.Shaw. Hyperspectral Subpixel Target Detection Using the Linear Mixing Model. *IEEE Trans. Geosci. Remote Sensing*, Vol. 39, No.7, pp. 1392-1408, July 2001.

[3] N.Keshava and J.F.Mustard. Spectral Unmixing. IEEE Signal Processing Magazine, pp. 44-57, January 2002.

[4] B.Thai and G.Healey. Invariant Subpixel Material Detection in Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sensing*, Vol. 40, No.3, pp. 599-608, July 2002.

[5] K.Choen and C.Seonghoon. Spectral angle mapper classification and vegetation indices analysis for winter cover monitoring using JERS-1 OPS data. Geoscience and Remote Sensing Symposium, IGARSS '96. Remote Sensing for a Sustainable Future. Vol.4, pp.1977 -1979 May 1996

[6] D.Manolakis, C.Siracusa, D.Marden and G.Shaw. Hyperspectral Adaptive Matched Filter Detectors: Practical Performance Comparison. Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII, Proceedings of SPIE Vol.4381, pp.18-33, 2001.

[7] B.Thai and G.Healey. Invariant Subpixel Material Identification in Hyperspectral Imagery. Part of the SPIE Conference on Algorithms for Multispectral and Hyperspectral Imagery V. SPIE Vol. 3717, pp.14-24, April 1999.

[8] B.Schölkopf, A.J.Smola, and K-R.Müller. Nonliear Component Analysis as A Kernel Eigenvalue Problem. *Neural Comput*, Vol.10, pp.1299-1319, 1998.