Dimensionality Reduction for Hyperspectral Image Based on Manifold Learning

Yiting Wang[✉], Shiqi Huang, Hongxia Wang, Daizhi Liu,
and Zhigang Liu

Department Xi'an Research Institute of Hi-Tech,
Xi'an, People's Republic of China
appie5744@sina.com

Abstract. A novel dimensionality reduction method named spectral angle and geodesic distance-based locality preserving projection (SAGD-LPP) was proposed in this paper. Considering the physical characters of hyperspectral imagery, the proposed method primarily select neighbor pixels in the image based on spectral angle distance. Then, using the geodesic distance matrix construct a weighted matrix between pixels. Finally, based on this weighted matrix, the idea of locality preserving projection algorithm is applied to reduce the dimensions of hyperspectral image data. The use of spectral angle to measure the distance between pixels can effectively overcome the spectral amplitude error caused by the uncertainty. At the same time, the use of geodesic distance to construct weight matrix can better reflect the internal structure of the data manifold than the use of Euclidean distance. Therefore, the proposed methods can reserve effectively the original characters of dataset with less loss in the useful information and less distortion on the data structure. Experimental results on real hyperspectral data demonstrate that the proposed methods have higher detection accuracy than the other methods when applied to the target detection of hyperspectral imagery after dimensionality reduction.

Keywords: Hyperspectral · Manifold learning · Target detection

1 Introduction

Hyperspectral sensors measure the radiance of the materials within each pixel area at a very large number of contiguous spectral bands and provide image data containing both spatial and spectral information. The resulting "image cube" is a stack of images in which each pixel has an associated spectral signature or fingerprint that uniquely characterizes the underlying objects. And due to its "one map" and high spectral resolution, hyperspectral remote sensing has opened up new opportunities for analyzing a variety of land cover materials [1].

Although this spectral feature provides sufficient discriminative information of the objects, hyperspectral target detection is always a great challenge due to its high dimensionality. Meanwhile, due to the nonlinear changes of solar radiation and nonlinear propagation of electromagnetic waves in the atmosphere, hyperspectral data has a typical non-linear characteristic. This further increases the difficulty of the

hyperspectral data processing. So how to effectively learn and discover nonlinear structure of hyperspectral data and reasonably reduce the dimensionality of data has important implications for hyperspectral image processing and application [2].

Manifold learning is a kind of common nonlinear dimensionality reduction algorithm. Among them, the typical manifold learning algorithms are isometric feature mapping (ISOMAP) [3], locally linear embedding (LLE) [4], Laplacian eigenmaps (LE) [5], etc. Manifold learning pursuits the goal to embed data that originally lies in a high dimensional space into a lower dimensional space, while preserving characteristic properties. Generally, it is difficult to know the geometry of the data manifold. ISOMAP is a technique that attempts to preserve pairwise geodesic distances between data points to keep the geometry of the data. In LLE, the local properties of the data manifold are constructed by writing the high-dimensional data points as a linear combination of their nearest neighbors. In the low-dimensional representation of the data, LLE attempts to retain the reconstruction weights in the linear combinations as good as possible. Similar to LLE, LE find a low-dimensional data representation by preserving local properties of the manifold. In LE, the local properties are based on the pairwise distances between near neighbors. The dimensionality reduction data obtained from the above manifold learning methods are all able to maintain a good global or local geometry of the original data.

While, an important requirement for dimensionality reduction techniques is the ability to embed new high-dimensional data points into an existing low-dimensional data representation, that is so-called out-of-sample extension. For the above nonlinear dimensionality reduction techniques, they yield mappings that are defined only on the training data points and it remains unclear how to naturally evaluate the maps on novel testing points. Therefore, some approximate out-of-sample extensions have been proposed that is based on computing a linear transformation from a set of landmark points to the complete dataset, in which neighborhood preserving embedding (NPE) algorithm is a linear approximation to the LLE [6], and locality preserving projection (LPP) algorithm is a linear approximation to the LE [7]. At present, these dimensionality reduction techniques are most conducted in the field of hyperspectral image classification but few in the field of target detection. So in this paper, we propose a new linear dimensionality reduction algorithm, called spectral angle and geodesic distance-based locality preserving projection (SAGD-LPP), so as to achieve the purpose of dimensionality reduction and improving target detection performance.

2 LPP Algorithm

LPP is designed for preserving local structure of high-dimensional data. It is likely that a nearest neighbor search in the low dimensional space will yield similar results to that in the high dimensional space. As a linear approximation of the LE, LPP suppose there exist a linear transformation between the high dimensional data point $\mathbf{x}_i$ and low dimensional data point $\mathbf{y}_i$, i.e. $\mathbf{y}_i = \mathbf{a}^T\mathbf{x}_i$, where $\mathbf{a}$ is a transformation vector. Then the algorithmic procedure of LPP is formally stated below:

1. Select neighbor pixels and construct the adjacency graph G: in graph G every data point $\mathbf{x}_i$ is connected to its k nearest neighbors.
2. Construct the weight matrix: the weight of the edge in the graph G is computed using the Gaussian kernel function. If nodes i and j are connected, put

$$w_{ij} = e^{-\|\mathbf{x}_i - \mathbf{x}_j\|\sigma^2} \tag{1}$$

   and, $w_{ij} = 0$ if there is no such edge.
3. Compute the low-dimensional representations $\mathbf{Y}$: a reasonable criterion for choosing a "good" map $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_1, \ldots, \mathbf{y}_m]$ is to minimize the following objective function

$$\sum_{i,j} \|\mathbf{y}_i - \mathbf{y}_j\|^2 w_{ij} \tag{2}$$

Because $\mathbf{y}_i = \mathbf{a}^T \mathbf{x}_i$, the objective function can be reduced to

$$\sum_{i,j} (\mathbf{y}_i - \mathbf{y}_j)^2 w_{ij} = \sum_{i,j} (\mathbf{a}^T \mathbf{x}_i - \mathbf{a}^T \mathbf{x}_j)^2 w_{ij} = 2\mathbf{a}^T \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{a} - 2\mathbf{a}^T \mathbf{X} \mathbf{W} \mathbf{X}^T \mathbf{a} = 2\mathbf{a}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{a} \tag{3}$$

where $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]$, and $\mathbf{D}$ is a diagonal matrix, $d_{ii} = \sum_j w_{ij}$. $\mathbf{L} = \mathbf{D} - \mathbf{W}$ is the Laplacian matrix of graph G. And a constraint is imposed as follows:

$$\mathbf{a}^T \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{a} = 1 \tag{4}$$

Finally, the minimization problem reduces to finding:

$$\underset{\mathbf{a}^T \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{a} = 1}{\arg\min} \ \mathbf{a}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{a} \tag{5}$$

It is a generalized eigenvector problem:

$$\mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{a} = \lambda \mathbf{X} \mathbf{D} \mathbf{X}^T \mathbf{a} \tag{6}$$

Suppose the eigenvectors $\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_d$ are the solutions of Eq. (6), and their corresponding eigenvalues $\lambda_0 < \lambda_1 < \ldots < \lambda_d$, Thus, the embedding is

$$\mathbf{y}_i = \mathbf{a}^T \mathbf{x}_i, \mathbf{a} = [\mathbf{a}_0, \ldots, \mathbf{a}_d] \tag{7}$$

When the transformation vector $\mathbf{a}$ is computed out based on the training data, an explicit expression of linear maps can be obtained, so the new testing data could be embedded into the existing low-dimensional data representation.

3   SAGD-LPP Algorithm

In hyperspectral image, due to the widespread of the uncertainty, the spectral radiant intensity of the same class feature show large changes. At the same time, the spectral radiation of the object in shaded areas is greatly different from that of the same object in non-shaded areas. However, regardless of how changes in the amplitude of the spectral curve, the spectral shape of the same class feature is substantially similar. According to this physical characteristic of hyperspectral image, we can know that if we select the neighbor pixels based on the Euclidean distance, there will be a large errors in constructing the adjacency graph G. And this may lead to the physical neighbor pixels in the hyperspectral image extend away from each other in the low-dimensional data, reducing the accuracy of target identification.

Therefore, in the first step of our proposed algorithm SAGD-LPP, we select the neighbor pixels based on the spectral angle distance to construct the adjacency graph G. Spectral angle distance can overcome errors caused by changes in the spectrum amplitude, making the physical neighbor pixels similar with each other. The spectral angular distance between two pixels can be expressed as:

$$d(\mathbf{x}_i, \mathbf{x}_j) = arc\cos\left[\frac{\sum\limits_{k=1}^{p} x_{ik}x_{jk}}{\left(\sum\limits_{k=1}^{p} x_{ik}^2\right)^{\frac{1}{2}}\left(\sum\limits_{k=1}^{p} x_{jk}^2\right)^{\frac{1}{2}}}\right] \tag{8}$$

where, P is the number of bands.

Moreover, compared to the Euclidean distance, geodesic distance can better reflect the internal structure of the high dimensional manifold. Therefore, in the second step of the algorithm, we use the geodesic distance to construct the weight matrix $\tilde{\mathbf{W}}$. And it can be expressed as:

$$\tilde{w}_{ij} = e^{-d_G(\mathbf{x}_i,\mathbf{x}_j)\sigma^2} \tag{9}$$

where, $d_G(\mathbf{x}_i, \mathbf{x}_j)$ is the geodesic distance between pixel $\mathbf{x}_i$ and $\mathbf{x}_j$.

Then, the objective function in the third step of proposed algorithm is translated to

$$\arg\min \sum\nolimits_{i,j} \left\|\mathbf{y}_i - \mathbf{y}_j\right\|^2 \tilde{w}_{ij} \tag{10}$$

i.e.

$$\underset{\mathbf{a}^T\mathbf{X}\tilde{\mathbf{D}}\mathbf{X}^T\mathbf{a}=1}{\arg\min} \ \mathbf{a}^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\mathbf{a} \tag{11}$$

where $\mathbf{D}$ is a diagonal matrix, $\tilde{d}_{ii} = \sum\limits_{j} \tilde{w}_{ij}$. $\tilde{\mathbf{L}} = \tilde{\mathbf{D}} - \tilde{\mathbf{W}}$ is the Laplacian matrix.

4   Experimental Validation

In this section, we first use the proposed method to obtain a dimension reduction data based on the real hyperspectral image data. Then, based on the obtained low-dimensional data, two detection methods constrained energy minimization (CEM) [8] algorithm and adaptive coherence estimator (ACE) [9] will be applied to target detection. Finally, the detection results are used to validate the effectiveness of the proposed method. In this experiment, the proposed method is compared with three classical dimensionality reduction methods that are PCA, NPE, LPP. The ROC curve is adopted to quantitatively measure the effect of target detection [10]. If the target is more similar to the background and it is hard to be detected, the ROC curve will become straighter, and the area under the curve (AUC) will be smaller. While, if the target is less similar to the background, that is, the target is more easily detected, the curve will bend to the left, and the AUC will be larger.

4.1   Experimental Data

The experimental data is obtained from AVIRIS hyperspectral image of the United States Santiago North Island Naval airport. The original image size is 400×400 with a total of 224 bands, and its spatial resolution is 3.5 m. Remove the invalid bands and left 189 effective bands. The image data used in this experiment are two interceptions from the original image; the size of the two images respectively is 100×100. Figure 1 shows a diagram of the experimental data. The image is an airport tarmac. The aircraft is the goal of detection. Figure 1(a) is the gray image of sub image I on band 10. Figure 1(b) is the gray image of sub image II on band 10.
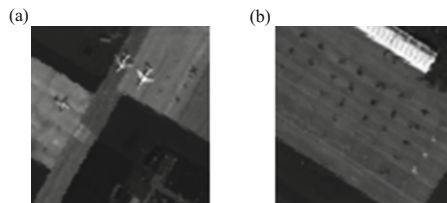

Fig. 1.   Experimental image of AVIRIS data (a) Sub image I (b) Sub image II

4.2   Experimental Results

Because of the variation of spectral characteristics between different objects, it is difficult to know the optimal dimensionality of the data. So we choose the optimal detection result that can be achieved by each of the algorithm for the comparison. Figure 2 show the detection results based on sub image I by using CEM detector. Figure 2(a) shows the detection result based on the original image data and Fig. 2(b)–(e) respectively show the detection result based on the new obtained low dimensionality data by using the NPE, PCA, LPP and SAGD-LPP. Similarly, Fig. 3(a)–(e) show

the corresponding results of sub image I obtained by ACE detector. Figure 4(a)–(b) show the corresponding ROC comparison figure under the two detection methods. Table 1 lists the corresponding AUC value of ROC curve and the optimal dimensionality of each method.

It can be found from Figs. 2 and 3, the detection results are unsatisfactory, when apply the both detection method to the original data and the low dimensionality data obtained by NPE. And based on the low dimensionality data obtained by PCA and LPP, all of three planes are detected, but the background information is not suppressed enough having a high false alarm rate. Based on the data obtained by SAGD-LPP algorithm not only all of the three planes are detected, but also good background suppression is get which having a low false alarm rate. It can be seen from Fig. 4 and Table 1, the SAGD-LPP performs outperforms other three algorithms, followed by LPP and PCA. The NPE performs poorly.
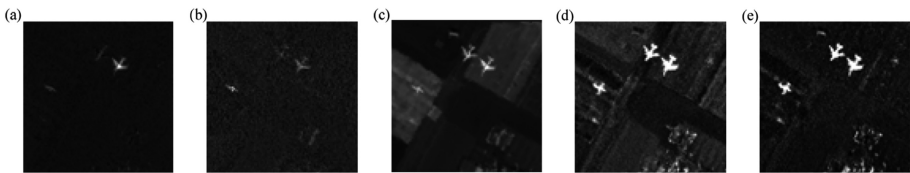

Fig. 2. Detection result based on sub image I by using CEM detector (a) Original data (b) NPE (c) PCA (d) LPP (e) SAGD-LPP
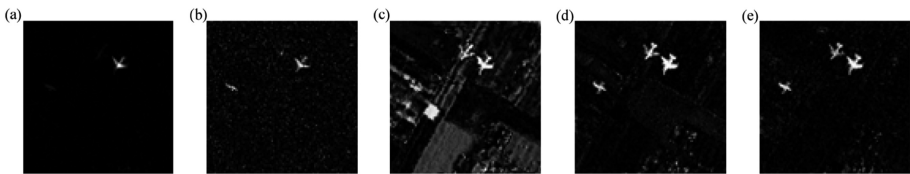

Fig. 3. Detection result based on sub image I by using ACE detector (a) Orignal data (b) NPE (c) PCA (d) LPP (e) SAGD-LPP
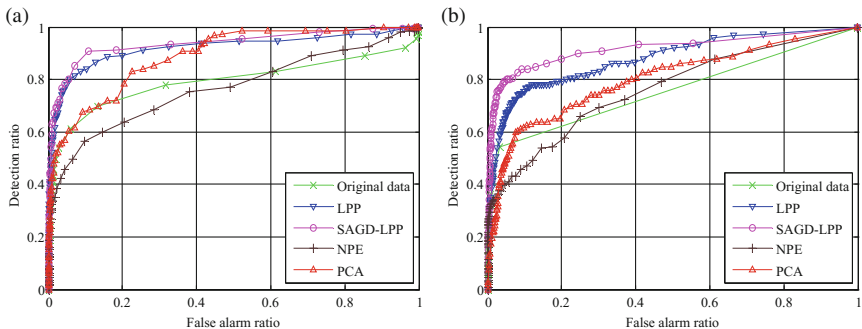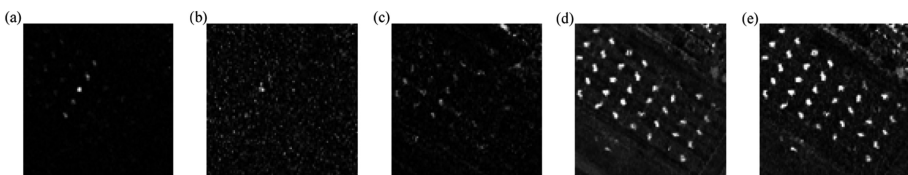

Fig. 4. ROC comparison figure obtained from sub image I (a) CEM detector (b) ACE detector

Table 1. Target detection results obtained from sub image I via different dimensionality reduction methods

Dimensionality reduction method	CEM detector		ACE detector	
	AUC	Dimensionality	AUC	Dimensionality
Original image	0.7914	198	0.7600	198
NPE	0.7666	30	0.7366	32
PCA	0.8019	4	0.8903	2
LPP	0.9314	8	0.9100	8
SAGD-LPP	0.9361	12	0.9275	12

Figures 5(a)–(e) and 6(a)–(e) respectively show the detection results of sub image II by using CEM and ACE detector based on the original image data and the new obtained low dimensionality data by NPE, PCA, LPP and SAGD-LPP. Figure 7(a)–(b) show the corresponding ROC comparison figure of sub image II. Table 2 lists the AUC value of ROC curve and the optimal dimensionality of each method corresponding to Fig. 7.
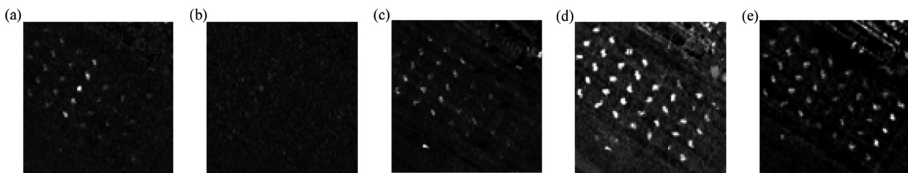
Fig. 5. Detection result based on sub image II by using CEM detector (a) Orignal data (b) NPE (c) PCA (d) LPP (e) SAGD-LPP

Fig. 6. Detection result based on sub image II by using ACE detector (a) Orignal data (b) NPE (c) PCA (d) LPP (e) SAGD-LPP

As can be seen from Figs. 5 and 6, based on the low dimensionality data obtained by NPE, no target can be detected. And the detection performance based on the data transformed by the PAC is worse than that of original data. Compared to the other algorithms, more planes are detected based on the data transformed by LPP, and the largest numbers of planes are detected based on the data transformed by SAGD-LPP. On Sub image II, SAGD-LPP algorithm still leads to the best performance. From Fig. 7 and Table 2, one can observe that, SAGD-LPP show the better detection performance
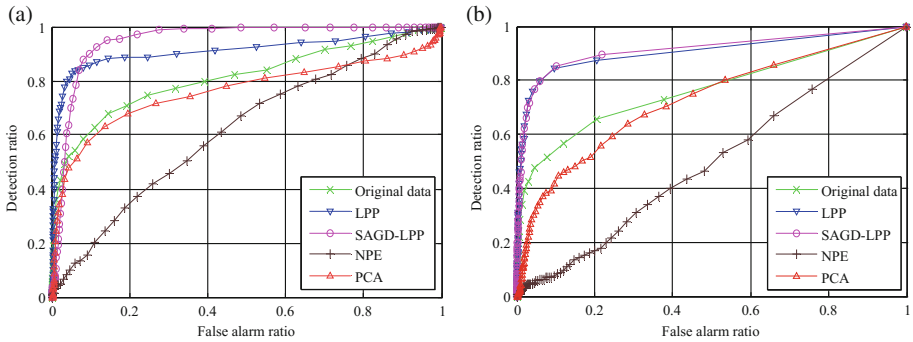
Fig. 7. ROC comparison figure obtained from sub image II (a) CEM detector (b) ACE detector

Table 2. Target detection results obtained from sub image II via different dimensionality reduction methods

Dimensionality reduction method	CEM detector		ACE detector	
	AUC	Dimensionality	AUC	Dimensionality
Original image	0.8115	198	0.7638	198
NPE	0.6813	39	0.4973	39
PCA	0.7610	30	0.7255	33
LPP	0.9112	13	0.9062	11
SAGD-LPP	0.9502	9	0.9151	10

than other algorithms in terms of AUC. And the AUC of SAGD-LPP increased by approximately 3 % compared with the LPP, and increased by approximately 18 % compared with other algorithm.

5   Conclusion

This paper presents a novel dimensionality reduction algorithm called spectral angle and geodesic distance-based locality preserving projection (SAGD-LPP). The proposed algorithm is based on the physical characteristics of hyperspectral data, so the structure of low-dimensional manifold can be better identified and the redundant information of hyperspectral data can be effectively removed during the dimensionality reduction. Real hyperspectral data experimental results show that the proposed SAGD-LPP method significantly outperformed the other methods dimensionality reduction applying to the target detection area.

References

1. Wang, Y.T., Huang, S.Q., Liu, D.Z., Wang, B.H.: A new band removed selection method for target detection in hyperspectral image. J. Opt. **42**(3), 208–213 (2013)
2. Pu, H.Y., Wang, B., Zhang, L.M.: New dimensionality reduction algorithms for hyperspectral imagery based on manifold learning. Infrared Laser Eng. **43**(1), 232–237 (2014). (in Chinese)
3. Tenenbaum, J.B., Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290**(5500), 2319–2323 (2000)
4. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290**(5500), 2323–2326 (2000)
5. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. Adv. Neural Inf. Process. Syst. **14**, 585–591 (2001)
6. He, X.F., Cai, D., Yan, S.C., et al.: Neighborhood preserving embedding. In: Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV 2005), pp. 1208–1213. Beijing (2005)
7. He, X.F., Niyogi, P.: Locality preserving projections. In: Proceedings of Neural Information Processing System. MIT, Vancouver (2003)
8. Gholizadeh, H., et al.: A decision fusion framework for hyperspectral subpixel target detection. Photogrammetrie Fernerkundung Geoinformation **3**, 267–280 (2012)
9. Kraut, S., Scharf, L.L.: The CFAR adaptive subspace detector is a scale-invariant GLRT. IEEE Trans. Signal Process. **47**(9), 2538–2541 (1999)
10. Manolakis, D., Marden, D., Shaw, G.A.: Hyperspectral image processing for automatic target detection applications. Lincoln Lab. J. **14**(1), 79–116 (2003)