A New Nonlinear Dimensionality Reduction Method with Application to Hyperspectral Image Analysis

Shen-En Qian and Guangyi Chen
Canadian Space Agency
6767, Route de l'Aéroport, St-Hubert, Quebec, Canada, J3Y 8Y9.
Email: {Shen-en.Qian, Guangyi.Chen}@space.gc.ca.

***Abstract*—In this paper, we propose a new nonlinear dimensionality reduction method by combining Locally Linear Embedding (LLE) with Laplacian Eigenmaps, and apply it to hyperspectral data. LLE projects high dimensional data into a low-dimensional Euclidean space while preserving local topological structures. However, it may not keep the relative distance between data points in the dimension-reduced space as in the original data space. Laplacian Eigenmaps, on the other hand, can preserve the locality characteristics in terms of distances between data points. By combining these two methods, a better locality preserving method is created for nonlinear dimensionality reduction. Experiments conducted in this paper confirms the feasibility of the new method for hyperspectral dimensionality reduction. The new method can find the same number of endmembers as PCA and LLE, but it is more accurate than them in terms of endmember location. Moreover, the new method is better than Laplacian Eigenmap alone because it identifies more pure mineral endmembers.**

I. Introduction

Hyperspectral imaging has recently become one of the most active research areas in remote sensing. Hyperspectral imagery possesses much richer spectral information than multispectral imagery because the number of spectral bands in hyperspectral imagery is in the hundreds instead of in the tens. This larger data volumes produced by hyperspectral sensors present a challenge to traditional data processing techniques. Conventional classification methods may not be used without dimension reduction as a preprocessing step. This is due to the 'curse' of dimensionality. A number of methods have been developed to mitigate the effects of dimensionality on information extraction from hyperspectral data, such as Principal Component Analysis (PCA) [1], Minimum Noise Fraction (MNF) [2], and Linear Discriminate Analysis (LDA) [3]. They all depend on linear projection and can result in lose of nonlinear properties of the original data after reduction of dimensionality.

In hyperspectral remote sensing, nonlinear properties are originated from the multi-scattering between photons and ground targets, within-pixel spectral mixing, and scene heterogeneity. We briefly review a few methods for hyperspectral dimensionality reduction here. Plaza et al. [4] described sequences of extended morphological transformations for dimensionality reduction and classification of high-dimensional remote sensed hyperspectral datasets. Harsanyi and Chang [5] investigated hyperspectral image classification and dimensionality reduction by using an orthogonal subspace projection approach. Wavelet transforms have been used in hyperspectral data dimensionality reduction ([6], [7]). Wavelet transforms can preserve the high and low frequency features during the signal decomposition, hence preserving the peaks and valleys found in typical spectra. Wang and Chang [8] proposed three ICA-based dimensionality reduction methods for hyperspectral data. Their methods are better than PCA and MNF in their experiments. Locally Linear Embedding (LLE) proposed by Roweis and Saul ([9], [10]) is a nonlinear feature extraction method that projects high dimensional data into a low-dimensional Euclidean space while preserving local topological structures. However, the computational complexity of this method is very intensive in computation and memory consumption. Belkin and Niyogi [11] developed the Laplacian Eigenmap for dimensionality reduction, and it preserves the relative distance between data points. More recently, Chang and Yeung [12] proposed robust locally linear embedding for nonlinear dimensionality reduction, and they demonstrated that the method is better suited for outlier problem. Chen and Qian [13] improved the existing LLE by introducing a spatial neighbourhood window for hyperspectral dimensionality reduction.

In this paper, we develop a novel nonlinear dimensionality reduction method by combining LLE with Laplacian Eigenmaps. LLE projects high dimensional data into a low-dimensional Euclidean space while preserving local topological structures. However, it may not maintain the distance between data points in the dimension-reduced space as in the original space. We know that Laplacian Eigenmaps can preserve the relative distances between data points. Therefore, the combination of these two methods leads to a better locality preserving method for nonlinear dimensionality reduction. We conducted experiments for a hyperspectral datacube and we found that the new method is promising in dimensionality reduction. The new method is better than PCA, LLE, and Laplacian Eigenmaps in terms of endmember extraction.

1-4244-1212-9/07/$25.00 ©2007 IEEE.

The organization of the paper is as follows. Section II reviews the basic concept of LLE and Laplacian Eigenmaps. Section III proposes a new technique for reducing the dimensionality by combining LLE with Laplacian Eigenmaps. Section IV conducts some experiments for hyperspectral data dimensionality reduction and endmember extraction. Finally, Section V draws the conclusions of the paper and gives future work.

II. Review of LLE and Laplacian Eigenmaps

In this section, we give an overview of two existing nonlinear dimensionality reduction methods, namely, LLE and Laplacian Eigenmaps.

A. LLE

LLE ([9], [10]) is a nonlinear dimensionality reduction method that maps high dimensional data into a low-dimensional Euclidean space while preserving local topological structures. LLE assumes that the manifold is well sampled, i.e., there are enough data. Each data point and its nearest neighbours lie on or close to a locally linear patch of the manifold. Therefore, we can approximate a data point x_i by a linear combination of its neighbours. The constrained weights minimizing the errors satisfy the following property: for any data point, they are invariant to rotation, scaling and translation of that data point and its neighbours. LLE is an unsupervised and non-iterative method, and it avoids the local minima problems plaguing many competing methods. LLE approximates high dimensional space with small patches, each of which can be considered as almost flat. These small patches are stitched together in the low dimensional space so that nonlinear structures in the high dimensional space are well preserved. LLE consists of the following three steps:

1) For each point in the original space x_i, find its K nearest neighbours.
2) Measure reconstruction error resulting from the approximation of each x_i by its nearest neighbours and calculate the reconstruction weights $w_{i,j}$ such that

   $$\begin{aligned} &\textit{min } (||x_i - \textstyle\sum_j w_{i,j} x_j||) \\ &\textit{st: } \textstyle\sum_j w_{i,j} = 1. \end{aligned}$$

3) Determine the low-dimensional embedding y_i that best preserves the local geometry represented by the reconstruction weights

   $$\begin{aligned} &\textit{min } \textstyle\sum_{i=1}^{n} (||y_i - \sum_j w_{i,j} y_j||) \\ &\textit{st: } 1/n \textstyle\sum_i y_i y_i^T = I, \text{ and } \sum_i y_i = 0. \end{aligned}$$

The computational complexity of LLE for the three steps are of order $O(dn^2)$, $O(dnK^3)$ and $O(rn^2)$, respectively, where d is the input dimensionality, K the number of nearest neighbours, n the number of data points, r the output dimensionality.

B. Laplacian Eigenmaps

The Laplacian Eigenmap was proposed by Belkin and Niyogi [11]. They construct a weighted graph with n nodes and a set of edges connecting the neighbouring points. The algorithm has three steps:

1) Constructing the adjacency graph. Put an edge between nodes i and j if x_i and x_j are close. It can be based on ϵ - neighbourhoods or K nearest neighbours.
2) Choosing the weights. There are two ways to choose the weights:
   a) Heat kernel. Set $v_{ij} = e^{-\frac{||x_i - x_j||^2}{t}}$ if nodes i and j are connected. Otherwise, put $v_{ij} = 0$.
   b) Set $v_{ij} = 1$ if nodes i and j are connected. Otherwise, put $v_{ij} = 0$.
3) Compute the eigenvalues and eigenvectors for the generalized eigenvector problem

   $$Lf = \lambda Df$$

   where $D = diag\{D_{11}, D_{22}, \cdots, D_{nn}\}$ is a diagonal matrix, $D_{ii} = \sum_j v_{ij}$, and $L = D - V$ is the Laplacian matrix.

We leave out the eigenvector corresponding to eigenvalue 0, and use the next bottom r eigenvectors for the embedding in the reduced dimension space. Laplacian eigenmap is relatively insensitive to outliers and noise because of its locality-preserving property. It is not prone to short circuiting as only the local distances are used in the algorithm.

III. Dimensionality Reduction by Combining LLE with Laplacian Eigenmaps

LLE characterizes the local geometry by linear coefficients w_{ij} that reconstruct each data point x_i from its K nearest neighbours. However, it may not map close data points in the original space into close data points in the dimension-reduced space. Laplacian Eigenmap can preserve the relative distance between data points by incurring a heavy penalty if neighbouring points x_i and x_j are mapped far apart. By combining LLE with Laplacian Eigenmaps, we obtain a new nonlinear dimensionality reduction method which keeps both the local geometry and the relative distance between the data points. The new method has three steps as well:

1) For each point in the original space x_i, find its K nearest neighbours. Also, Put an edge between nodes i and j if x_i and x_j are close.
2) Measure reconstruction error resulting from the approximation of each x_i by its nearest neighbours and calculate the reconstruction weights $w_{i,j}$ according to LLE. Also, choose the weight v_{ij} according to Laplacian Eigenmaps.
3) Determine the low-dimensional embedding y_i that best preserves the local geometry and the relative distance between data points by solving the following optimization problem:

   $$\begin{aligned} min \quad & \textstyle\sum_{i=1}^{n} ||y_i - \sum_{j=1}^{K} w_{ij} y_j||^2 + \frac{1}{2} \sum_{i,j=1}^{n} ||y_i - y_j||^2 v_{ij} \\ st: \quad & \tfrac{1}{n} Y Y^T = I, \\ & Yl = 0. \end{aligned}$$

where $Y = [y_1, y_2, \cdots, y_n]_{r\times n}$, I is the identity matrix, and $l = [1, 1, \cdots, 1]^T$. The weights v_{ij} incurs a heavy penalty if neighbouring points x_i and x_j are mapped far apart.

We know that

$$\begin{aligned}
&\textstyle\sum_{i=1}^{n} ||y_i - \sum_{j=1}^{K} w_{ij} y_j||^2 \\
&= \textstyle\sum_{ij}(\delta_{ij} - w_{ij} - w_{ji} + \sum_k w_{ki} w_{kj}) < y_i, y_j > \\
&= tr(Y(I-W)(I-W)^T Y^T) \\
&= tr(YMY^T)
\end{aligned}$$

where $M = (I-W)(I-W)^T$, $< y_i, y_j >$ is the dot product of y_i and y_j, and $W = [w_1, w_2, \cdots, w_n]$. also,

$$\begin{aligned}
&\textstyle\sum_{i,j} ||y_i - y_j||^2 v_{ij} \\
&= \textstyle\sum_{i,j}\{||y_i||^2 + ||y_j||^2 - 2 < y_i, y_j >\} v_{ij} \\
&= \textstyle\sum_i ||y_i||^2 D_{ii} + \sum_j ||y_j||^2 D_{jj} - 2\sum_{i,j} v_{ij} < y_i, y_j > \\
&= \textstyle 2\sum_i ||y_i||^2 D_{ii} - 2\sum_{i,j} v_{ij} < y_i, y_j > \\
&= \textstyle 2\sum_k \sum_i (y_i^k)^2 D_{ii} - 2\sum_k \sum_{i,j} v_{ij} y_i^k y_j^k \\
&= 2\{tr(YDY^T) - tr(YVY^T)\} \\
&= 2\{tr(Y(D-V)Y^T)\}
\end{aligned}$$

where $D = diag\{D_{11}, D_{22}, \cdots, D_{nn}\}$ and $D_{ii} = \sum_j v_{ij}$.

The constrained minimization problem can be done by using the Lagrange multipliers

$$L(Y) = YMY^T + Y(D-V)Y^T + (NI - YY^T)\Lambda.$$

Setting the gradients with respect to Y to zero, we have

$$2MY^T + 2(D-V)Y^T - 2Y^T\Lambda = 0.$$

therefore we need to solve the following symmetric eignevalue problem:

$$(M + D - V)Y^T = Y^T\Lambda.$$

It is easy to show that $l = [1, 1, \cdots, 1]^T$ is an eigenvector corresponding to eigenvalue 0. Discarding this eigenvector enforces the constraint that the embeddings have zero mean by virtue of orthogonality. The remaining r bottom eigenvectors form the embedding in the reduced dimension space. Since $M + D - V$ is a large sparse symmetric matrix, we can use any well known eigensolver to solve it. In this paper, we use the $eigs()$ function in Matlab to calculate a few smallest eigenvalues (in magnitude) and their corresponding eigenvectors.

IV. Experimental Results

In this section, we conduct some experiments in Matlab for hyperspectral data dimensionality reduction in order to demonstrate the feasibility of the proposed method in this paper. We use a Pentium-4 PC with a $3.20GHz$ CPU and $1G$ RAM for our experiments. Our datacube was acquired using an Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) in the Cuprite mining district, Nevada, by Jet Propulsion Laboratory (JPL) in 1997. The original scene with size of 614×512 pixels and 224 bands is available online at $http://aviris.jpl.nasa.gov/html/aviris.freedata.html$. As in [8], we cut out the upper-right corner of the scene that consists of 350×350 pixels and 224 bands. This scene is well understood mineralogically and it has been made a standard test site for validation and in assessment of endmember extraction methods. Also as in [8] and [14], due to water absorption and low signal-to-noise ratio (SNR), we remove the bands 105-115 and bands 150-170 in our experiments. As a result, a total of 192 bands are used in our experiments. Fig. 1 shows the image of the datacube at wavelength $827nm$ (spectral band $\#50$). As mentioned in [8], in this scene there are five ground-truth endmembers for minerals: alunite at (62,161), buddingtonite at (209,234), calcite at (30,347), kaolinite at (22,298), and muscovite at (33,271). In our experiments, we only extract a region of 64×64 pixels (see the black thick line square in Fig. 1) for testing our hyperspectral dimensionality reduction. We set the number of nearest neighbours for reconstruction to 26, that is, $K = 26$. The output dimensionality is set to $r = 10$.

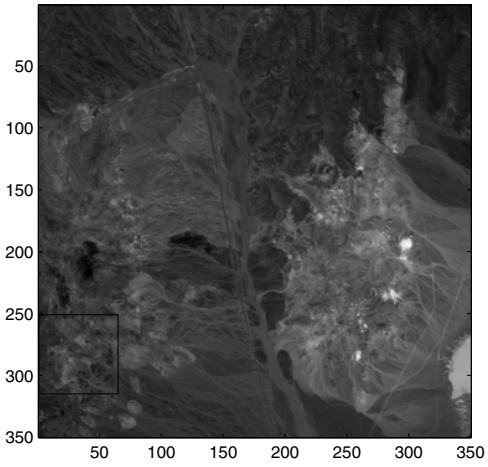

Fig. 1. The AVIRIS Cuprite scene at wave-length $827nm$ (spectral band $\#50$).

We conducted experiments for endmember extraction of a hyperspectral image with a region of 64×64 pixels. Within this region, only two ground-truth endmembers in the original 350×350 image scene are present: kaolinite at (22,298) and muscovite at (33,271). Endmember extraction is one of the fundamental tasks in hyperspectral image analysis. We extract endmembers using Pixel Purity Index (PPI) [15] available in Environment for Visualizing Images (ENVI) 4.2 software system [16] developed by Research Systems, Inc. In our experiments, we set the number of iterations in PPI parameters to 10000, and the threshold factor to 1. We do not generate the final PPI-found endmembers using the cluster mean spectra of the data points in the extreme corners of the data cloud. Instead, we select individual pixels falling into the corners of the data cloud and use those pixel signatures as the final endmember set. Table I tabulates the spectral angles (in degrees) between the ground-truth endmembers and the endmembers extracted from the data sets after the dimensionality being

TABLE I
THE SPECTRAL ANGLES (IN DEGREES) BETWEEN THE GROUND-TRUTH ENDMEMBERS AND THE ENDMEMBERS EXTRACTED WITH PCA, LLE, LAPLACIAN EIGENMAPS, AND THE PROPOSED METHOD BY USING PPI.

Reduction Methods	*Extracted Endmembers*	*Ground-truth Endmembers*				
		kaolinite	muscovite	buddingtonite	calcite	alunite
		(22,298)	(33,271)	(209,234)	(30,347)	(62,161)
PCA	(23,304)	**0.03**	0.14	0.18	0.22	0.10
	(20,281)	0.13	**0.07**	0.09	0.12	0.14
	(56,303)	0.24	0.14	0.12	**0.05**	0.24
LLE	(23,304)	**0.03**	0.14	0.18	0.22	0.10
	(19,281)	0.13	**0.07**	0.08	0.12	0.15
	(56,304)	0.25	0.15	0.12	**0.07**	0.25
Laplacian Eigenmaps	(22,298)	**0.00**	0.13	0.17	0.21	0.10
	(26,278)	0.13	0.08	**0.07**	0.10	0.14
proposed method	(22,298)	**0.00**	0.13	0.17	0.21	0.10
	(33,270)	0.13	**0.04**	0.09	0.11	0.15
	(63,274)	0.16	0.08	**0.07**	0.11	0.16

reduced to $r = 10$ from 192 using PCA, LLE, Laplacian Eigenmaps, and the proposed method, respectively. From the table we can see that PCA, LLE, Laplacian Eigenmaps, and the new method find three, three, two, and three endmembers, respectively. From the number of endmembers that the new method can find we know that the new method is comparable to PCA and LLE, and it is better than Laplacian Eigenmaps. In addition, by looking at the location of the endmembers extracted by the new method, we know that it finds one endmember (kaolinite) at exactly the same location as the ground-truth endmember at (22,298). Also, it finds another endmember (muscovite) that has only one pixel shift from the ground-truth endmember at (33,271). On the other hand, both PCA and LLE find endmembers at locations that are far away from the ground-truth endmembers. Laplacian Eigenmaps identifies one exact ground-truth endmember at (22,298) as well. However, it only finds two endmembers in total which are less than other methods tested in this paper. Therefore, the new method proposed in this paper is better than PCA, LLE, and Laplacian Eigenmap in endmember extraction.

V. CONCLUSIONS

A new nonlinear dimensionality reduction method is proposed in this paper by combining LLE with Laplacian Eigenmaps. The new method overcomes the drawback of LLE, i.e., it may map close data points in the original space into non-close data points in the dimension-reduced space. This is because Laplacian Eigenmap keeps the relative distance between neighbouring points. We conducted experiments for hyperspectral data analysis and found that the new method is better than PCA, LLE, and Laplacian Eigenmap in identifying pure mineral endmembers.

ACKNOWLEDGMENTS

This work was partially supported by the Visiting Fellowship in Canadian Government Laboratory coordinated by Natural Sciences and Engineering Research Council of Canada (NSERC).

REFERENCES

[1] I. T. Jolliffe, Principal Component Analysis, Springer, New York, 2002.
[2] J. W. Boardman and F. A. Kruse, "Automated spectral analysis: A geologic example using AVIRIS data, north Grapevine Mountains, Nevada," *Proceedings of the Tenth Thematic Conference on Geologic Remote Sensing*, Environmental Research Institute of Michigan, Ann Arbor, MI, USA, pp. I-407, 1994.
[3] K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press, San Diego, California, 1990.
[4] A. Plaza, P. Martinez, J. Plaza and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 466-479, 2005.
[5] J. C. Harsanyi and C. I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 4, pp. 779-785, 1994.
[6] S. Kaewpijit, J. L. Moigne and T. El-Ghazawi, "Automatic reduction of hyperspectral imagery using wavelet spectral analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 4, pp. 863-871, 2003.
[7] L. M. Bruce, C. H. Koger and J. Li, "Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2331-2338, 2002.
[8] J. Wang and C. I. Chang, "Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1586-1600, 2006.
[9] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323-2326, 2000.
[10] L. K. Saul and S. T. Roweis, "Think globally, fit locally: unsupervised learning of low dimensional manifolds," *Journal of Machine Learning Research*, vol. 4, pp. 119-155, 2003.
[11] M. Belkin and P. Niyogi, "Laplacian Eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, pp. 1373-1396, 2003.
[12] H. Chang and D. Y. Yeung, "Robust locally linear embedding," *Pattern Recognition*, vol. 39, no. 6, pp. 1053-1065, 2006.
[13] G. Y. Chen and S.-E. Qian, "Dimensionality reduction of hyperspectral imagery using improved locally linear embedding," *Journal of Applied Remote Sensing*, vol. 1, 013509, 2007.
[14] A. Plaza and C.-I Chang. "Impact of initialization on design of endmember extraction algorithms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3397-3407, 2006.
[15] J. W. Boardman, F. A. Kruse and R. O. Green,"Mapping target signatures via partial unmixing of AVIRIS data," *Summaries of JPL Airborne Earth Science Workshop*, Pasadena, CA, 1995.
[16] Research Systems Inc., *ENVI User's Guide*. Boulder, CO:Research Systems, Inc., 2001.