LEARNING A JOINT MANIFOLD WITH GLOBAL-LOCAL PRESERVATION FOR MULTITEMPORAL HYPERSPECTRAL IMAGE CLASSIFICATION

Hsiuhan Lexie Yang, Student Member, IEEE and Melba M. Crawford, Fellow, IEEE

School of Civil Engineering
Purdue University, West Lafayette, IN 47907, USA

ABSTRACT

Adapting a pre-trained classifier with labeled samples from an image for classification of another temporally related image is a common multitemporal image classification strategy. However, the adaptation is not effective when the spectral drift exhibited in temporal data is significant. Instead of iteratively redefining classifier parameters, we exploit similar data geometries of temporal data and project temporal data into a joint manifold space where similar samples are clustered. The proposed classification framework is based on aligning global temporal data manifolds. In addition to global structures, we also consider the local scale by incorporating local point relations into the alignment process. In experiments with challenging temporal hyperspectral data, the proposed framework provides favorable classification results, compared to the baseline.

Index Terms— Hyperspectral, multitemporal, manifold learning, manifold alignment

1. INTRODUCTION

Manifold learning has been demonstrated as a successful nonlinear dimension reduction technique for analyzing high dimensional data. Researchers are able to exploit essential features in a low dimensional manifold space that represents some aspects of nonlinear scattering. Similarity is the most utilized concept for representing geometric structures for developing manifolds, largely because highly similar samples are clustered, and dissimilar samples are separated. In general, similarity characteristics used in manifold learning are based on some distance measure, which is typically categorized as either a local or global based distance. Local based methods [1, 2] consider only locally clustered samples, often within Euclidean distance or the spectral angle of remotely sensed data, whereas global focused methods [3, 4] seek to preserve global structures using geodesic distance. Many studies have demonstrated specific advantages of these two categories for various applications and data sets. For example, the global approaches have a major advantage in data representation, where it is desirable to retain relationships between different classes. The local approaches are more favored in discriminating difficult classes and are computationally advantageous. In the remote sensing community, both local and global methods have been successfully used to analyze hyperspectral imagery, particularly for classification [5, 6, 7]. In recent years, researchers have investigated linking these two general approaches [8, 9]. However, work related to combining characteristics of local and global manifolds for the classification task has not be explored to date.

Classification of a single hyperspectral image in the manifold space has been investigated in several papers [6, 7]. However, it is challenging to extend manifold learning intuitively for classifying temporal hyperspectral images. Analyzing manifolds developed for individual images is not viable, unless conditions do not vary between dates. Learning a proper joint manifold by exploiting spectral similarity across temporal images is difficult due to spectra shifts that occur in many classes over time. Assuming classes are the same in temporal images, Manifold Alignment (MA) is a possible strategy for analyzing multiple similar manifolds [10, 11]. With MA, a joint manifold is obtained by aligning similar geometries. Thus, it is unnecessary to adapt classifiers and therefore avoid the poor initial condition that occurs when spectral drifting is severe. In our previous MA work [12], we concluded that two temporal data manifolds developed with the global structure preserving concept could be similar, although some local changes in the multitemporal data manifolds may also occur.

In this paper, we propose a framework to learn an adequate joint manifold that combines the benefits of global based structures and local clusters for MA (GLMA). We first extract global structures and align two global manifolds using some corresponding pairs of hyperspectral images acquired on two different dates. We also explored geographic proximity for selecting pairs that can be used to confidently bridge temporal manifolds. With the proposed approach, we are able to predict class labels of one image (target image) with training data only from another image (source image). Three temporal related hyperspectral images are used to validate the proposed method, and results are compared to a baseline.

2. PROPOSED METHOD

Let $X^S = \{x_1^S, x_2^S, x_3^S, \cdots, x_m^S\} \in R^d$ and $X^T = \{x_1^T, x_2^T, x_3^T, \cdots, x_n^T\} \in R^d$ denote m and n samples from a source and a target d-dimension hyperspectral image, respectively. $Y^S = \{y_1^S, y_2^S, y_3^S, \cdots, y_m^S\}$ is the corresponding class label set of X^S. We want to predict the labels of the samples X^T, denoted as $\hat{Y}^T = \{\hat{y}_1^T, \hat{y}_2^T, \hat{y}_3^T, \cdots, \hat{y}_n^T\}$ in a joint manifold space based on the assumption that samples from the same class should be clustered in a properly learned joint manifold space.

To obtain a representative joint manifold, we define an $(m+n)$ by $(m+n)$ distance matrix D_G to preserve similarities within the same data set, and identify clusters of spectrally similar samples across images:

$$D_G = \begin{bmatrix} D_{x^S,x^S} & D_{x^S,x^T} \\ D_{x^T,x^S} & D_{x^T,x^T} \end{bmatrix} \quad (1)$$

where D_{x^S,x^S} and D_{x^T,X^T} are the geodesic distance matrices of the source and target images, respectively. These two components retain the global geometry within the two data sets. $D_{x^T,x^S} = D^T_{x^S,x^T}$ describes the connections between two data sets. The method which is used to construct $D_{x^T,x^S} = D^T_{x^S,x^T}$ considers the unique characteristics of multitemporal remote sensing data, as discussed subsequently.

Spectral signatures may change over time due to seasonal changes in targets and environmental conditions. The geodesic measure within these two images could have different scales, so D_{x^S,x^S} and D_{x^T,x^T} should be rescaled so that the two distance measures are comparable. In order to find a rescaling factor, denoted as μ, we exploit the concept of corresponding pairs discussed in our previous work [12] to search the optimal μ. The correspondence pairs $(x_{c^i}^S, x_{c^i}^T), i \in [1, u]$ serve as matching samples between two temporal images while aligning similar data manifolds. That is, two samples in each pair from the source and the target image, respectively, are highly likely to be the same class labels. The optimal μ can be given by minimizing $\|D_{x_c^S,x_c^S} - \mu D_{x_c^T,x_c^T}\|$ where $D_{x_c^T,x_c^T}$ and $D_{x_c^T,x_c^T}$ are distance matrices of u corresponding pairs.

The corresponding pairs could be simply defined as the nearest spectral neighbor across temporal images. However, spectral changes may compromise the confidence of the resulting pairs. In the paper, we consider both spectral proximities and spatial locations to obtain a robust pairing condition. Let $Z^S = \{z_1^S, z_2^S, z_3^S, \cdots, z_m^S\} \in R^2$ and $Z^T = \{z_1^T, z_2^T, z_3^T, \cdots, z_n^T\} \in R^2$ represent geographical coordinates of m and n pixels, respectively. The i-th corresponding pair $(x_{c_p^i}^S, x_{c_q^i}^T)$ is defined as:

$$\arg\min_{p \le m, q \le n} (\|x_p^S - x_q^T\| + a \|z_p^S - z_q^T\|) \quad (2)$$

where a is a balancing coefficient. The two samples in a corresponding pair should be spectrally similar and spatially adjacent. Therefore, these pairs should thus be able to bridge the two data manifolds.

Table 1. Botswana Class Label and Data Sample Size

ID	Class	May	June	July
1	Water	297	361	185
2	Primary Floodplain	437	308	96
3	Riparian	448	303	164
4	Firescar	354	335	186
5	Island Interior	337	370	131
6	Woodlands	357	324	169
7	Savanna	330	342	171
8	Short Mopane	239	299	152
9	Exposed Soils	215	229	96

The corresponding pair is also a key to preserving local geometry. By defining $D_{x_i^S,x_j^T}(i,j) = \min(D_{x_i^S,x_{c_p^i}^S} + D_{x_j^T,x_{c_q^i}^T})$, we note that the local spectral neighbor of x_i^S and x_j^T will be one of the corresponding pairs $(x_{c_i}^S, x_{c_i}^T)$. Further, the distance $D_{x_i^S,x_j^T}(i,j) = 0$ when $x_i^S = x_{c_i}^S$ and $x_j^T = x_{c_i}^T$. This implies that the samples in a corresponding pairs are forced to be aligned in the joint manifold space.

With the definition of D_G in (2), an optimal joint manifold feature space $F \in R^{D_{D<d}}$ is computed by minimizing the cost function $E = \|\tau(D_G) - \tau(D_F)\|$ where D_F is a distance matrix where elements $D_F(i,j) = \sqrt{f_i - f_j}$ and the τ operator converts distance to inner products, which uniquely characterize the geometry [3]. The optimal solution is obtained by applying multidimensional scaling. That is, the top eigenvectors (in decreasing order) of D_G are the sample configurations in the joint manifold space. In this joint manifold space, the samples whose spectra are similar will be clustered. Therefore, we are able to predict class labels of X^T (target) in the joint manifold space using only labeled data Y^S from the source image.

3. EXPERIMENTS AND RESULTS

Three Hyperion images, collected over the Okavango Delta, Botswana in May, June and July of 2001, were used to validate the proposed method. The Hyperion on the NASA EO-1 satellite acquires 224 bands at 30m spatial resolution in 10 nm bands from 400 - 2500 nm portion of the spectrum. Removal of uncalibrated and noisy bands resulted in 145 bands that were used in experiments. The images have nine common identified classes, as shown in Table 1. Two data pairs comprise a source image (May or June) and a target image (June or July). Among these three images, June and July data cover the same area. May imagery is not exactly co-located with the other two images due to a change in the satellite pointing. Therefore, the classification with May/June data pair is a more challenging problem, considering spectral drift in both

Table 2. Classification Accuracy

Data Pair		May/June		June/July	
		Baseline	GLMA	Baseline	GLMA
Overall Acc.		76.07	84.24	86.59	86.44
Cls. ID	1	100	100	100	100
	2	99.4	96.7	72.9	76.0
	3	62.4	53.8	63.4	54.3
	4	99.7	98.8	98.4	96.2
	5	93.2	88.9	96.9	97.7
	6	36.1	51.9	54.1	62.1
	7	71.3	71.9	97.7	97.1
	8	59.2	99.0	100	100
	9	48.5	87.8	97.9	93.8

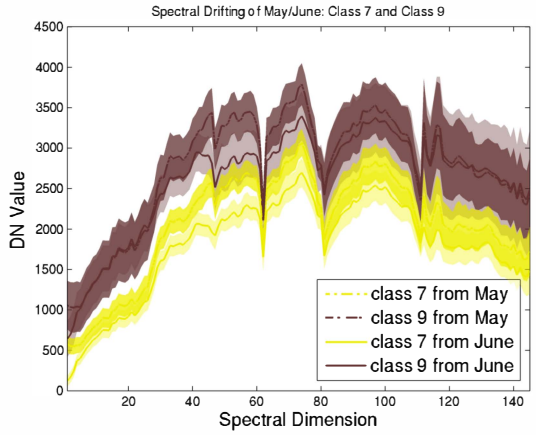

Fig. 1. Spectral drifting example

the spatial and the temporal domain. Evaluations were performed based on the k-nearest neighbor classifier in which k is determined by grid search. The labels from the source image were used as training data, and the labeled samples from the target image were used to validate the predicted labels. That is, unsupervised classifications is performed on the target image.

We also created a baseline based on ISOMAP [1], a popular global based manifold learning approach. The baseline results are obtained by performing classifications in a joint manifold learned with ISOMAP without performing MA. The results obtained with the proposed method and the baseline are shown in Table 2.

The overall accuracy achieved by the alignment method was higher than the baseline method for the May/June pair. The gain was achieved by classes that exhibit significant spectra shifts between the dates (e.g. Class 9 (Exposed Soils) and Class 8 (Short Mopane)). These changes in spectra may result in inaccurate spectral neighbors in the nearest neighbor search for while developing manifolds with ISOMAP. As Figure 1 shows, the spectral signatures of Class 7 from the May data and Class 9 from the June data overlap in some spectral bands. The class accuracy of Class 9 is low because some samples of Class 9 are misclassified as Class 7, as shown in Table 3 and Table 4. The confusion matrices also show improvements in class accuracies in Class 5 and Class 8 resulted from MA.

Table 3. Confusion Matrix: Baseline of May/June pair

ID	1	2	3	4	5	6	7	8	9
1	**361**	0	0	0	0	0	0	0	0
2	0	**306**	1	0	0	0	1	0	0
3	0	15	**189**	1	0	98	0	0	0
4	1	0	0	**334**	0	0	0	0	0
5	0	2	0	1	**345**	0	17	0	5
6	1	97	106	2	0	**117**	1	0	0
7	0	61	0	26	11	0	**244**	0	0
8	0	0	0	0	107	7	8	**177**	0
9	0	0	0	0	36	0	82	0	**111**

Table 4. Confusion Matrix: GLMA of May/June pair

ID	1	2	3	4	5	6	7	8	9
1	**361**	0	0	0	0	0	0	0	0
2	0	**298**	0	3	0	6	1	0	0
3	0	4	**163**	1	0	135	0	0	0
4	0	0	0	**331**	0	4	0	0	0
5	0	0	0	4	**329**	0	28	0	9
6	0	73	75	5	0	**168**	2	1	0
7	0	37	0	35	22	0	**246**	0	2
8	0	0	0	0	1	0	1	**296**	1
9	0	0	0	0	23	0	5	0	**201**

For June/July data pair, the baseline and the proposed method yield similar overall accuracies. The impact of spectral shifts is relatively small on this data pair, so the benefit of MA is limited.

We also compared to the alignment results in [12], where only local geometry is considered. The overall accuracy in [12] is 79.21% (May/June) and 86.15% (June/July). GLMA provides higher classification results for the May/June data (84.24%) and essentially the same accuracy for June/July data (86.44%). The local structure preserving based alignment results in [12] are useful to retain local clusters for classifications, compared to our baseline in this paper (as in Table 2). However, it is difficult to interpret the data sets globally since this type of method considers only nearby samples. The results obtained by the proposed global-local preserving method present a more illustrative feature space because of the isometric-preserving property. We can better understand the relationships among different classes in the joint manifold, as shown in Figure 2. Most of samples that from the same class, either from May or from June data, reside in the same cluster, except samples from Class 1 (Water) and Class 4 (Firescar). For these two classes, the spectral signatures of two temporal data are quite different for the following reasons. During May and June, the extent a of the flooding was different. As for the Firescar class, re-growth of vegetation was quite rapid, and the spectral signatures changed between

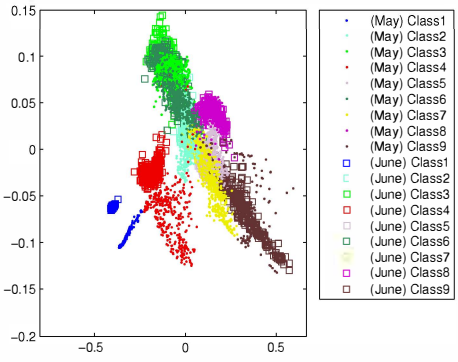

Fig. 2. GLMA alignment results of May/June

the two acquisition dates. Those samples can still be correctly classified in the MA space, not in exactly the same cluster, but the clusters are close each other. Another implication is that we can observe such abrupt spectral changes over time in the resulted joint MA space.

4. CONCLUSION

In this paper, we propose a new approach to learn a joint manifold for multitemporal hyperspectral image classifications. The joint manifold is created by aligning two global preserving manifolds with local clusters that are retained from correspondence pairs between two data sets. The proposed framework is the first work to discuss both global and local characteristics of temporal data manifolds. The preliminary results show the potential usefulness of combining global and local geometry for learning a representative data manifold of temporal hyperspectral images impacted by spectral drift. We are currently investigating alternative local geometry preserving approaches (the current approach only considers one local spectral neighbor). Including additional spectral neighbors or considering the spatial proximity of samples may be beneficial to retain better local clusters.

ACKNOWLEDGMENTS

This research was supported by NSF Grant No. 0705836, NASA AIST Grant No. 11-0077, and the Purdue Laboratory for Applications of Remote Sensing.

5. REFERENCES

[1] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.

[2] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data," *Neural Computation*, vol. 1396, no. 6, pp. 1373–1396, 2003.

[3] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[4] C. K. I. Williams, "On a connection between kernel PCA and metric multidimensional scaling," *Machine Learning*, vol. 46, no. 1, pp. 11–19, 2002.

[5] C. M. Bachmann, T. L. Ainsworth, and R. A. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 441–454, 2005.

[6] M. M. Crawford, L. Ma, and W. Kim, "Exploring nonlinear manifold learning for classification of hyperspectral data," in *Optical Remote Sensing*, vol. 3, pp. 207–234. Springer Berlin Heidelberg, 2011.

[7] L. Ma, M. M. Crawford, and J. Tian, "Local manifold learning-based k-nearest-neighbor for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 11, pp. 4099–4109, 2010.

[8] J. Verbeek, "Learning nonlinear image manifolds by global alignment of local linear models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1236–1250, 2006.

[9] R. Wang, S. Shan, X. Chen, J. Chen, and W. Gao, "Maximal linear embedding for dimensionality reduction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pp. 1776–1792, 2011.

[10] C. Wang and S. Mahadevan, "A general framework for manifold alignment," in *AAAI Fall Symposium on Manifold Learning and its Applications*, 2009.

[11] S. Lafon, Y. Keller, and R. R. Coifman, "Data fusion and multicue data matching by diffusion maps," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1784–1797, 2006.

[12] H. L. Yang and M. Crawford, "Manifold alignment for multitemporal hyperspectral image classification," in *Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2011, pp. 4332–4335.