FUSION OF SPECTRAL AND SPATIAL INFORMATION FOR LAND COVER CLASSIFICATION

Wenzhi Liao[1], *Hongyan Zhang*[1,2], *Jie Li*[1,3], *Shaoguang Huang*[1], *Rui Wang*[1,4], *Renbo Luo*[1,5], *Aleksandra Pižurica*[1]

[1]Ghent University-TELIN-IPI-iMinds, Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
[2]The State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, China
[3]School of Electronics & Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[4]Sichuan Provincial Key Laboratory of Information Coding and Transmission, Southwest Jiaotong University, Chengdu 610031, China
[5]School of Automation Science and Engineering, South China University of Technology, 510640 Guangzhou, China
Email: {wliao, hozhang, jiejli, shhuang, ruiwan, reluo and sanja}@telin.ugent.be

ABSTRACT

Hyperspectral imagery contains a wealth of spectral and spatial information that can improve target detection and recognition performance. Existing feature extraction methods cannot fully utilize both the spectral and spatial information. Data fusion by simply stacking different feature sources together does not work well either, as it does not take into account the differences between feature sources. In this paper, we present our recent graph-based approach for fusing spectral information and spatial information of hyperspectral imagery, and we show a case study on how our graph-based fusion method combines multiple features, which can be applied in other applications. Our approach takes into account the properties of different data sources, and makes full advantage of both the spectral and the spatial features through the fusion graph. Experimental results on the classification of fusing real hyperspectral images are very encouraging.

Index Terms— Hyperspectral images, remote sensing, classification, data fusion

1. INTRODUCTION

Recent advances in sensors technology have led to an increased availability of hyperspectral data at very high both spatial and spectral resolutions. Many approaches have been developed to exploit the spectral and the spatial information of hyperspectral imagery for classification. Some of these approaches focus on increasing the spectral discrimination through dimension reduction [1]. Others explore the spatial information of HS data through morphological features [2, 3, 4, 5].

A limitation of the above approaches is that they rely mainly on a single type of features (spectral or geometrical features) and do not fully utilize the wealth of information available in the HS data. The spatial information, once combined with the spectral information, can contribute to a more comprehensive interpretation of objects on the ground. For example, spectral signatures cannot differentiate between objects made of the same material (e.g. roofs and roads made with the same asphalt), while they can often be easily distinguished by their geometry. On the other hand, spatial features alone may fail to discriminate between objects that are quite different in nature (e.g. grass field, parking or a swimming pool), if their shape and size are similar. Many approaches have been developed to fuse the spectral and spatial information for the classification of remote sensing data [6, 8, 9, 10, 11]. Some of these approaches employ the so-called composite kernel methods [6, 7] or their generalization [9]. Others define spatial information through morphological profiles, and concatenate spectral and spatial features in a stacked architecture for classification [10, 11].

Despite their simplicity, the feature fusion methods that simply concatenate several kinds of features together are rarely useful in practice. These simple stacking methods can perform even worse than using a single feature, because the information contained by different features is not equally represented or measured. The element values of different features can be significantly unbalanced. Furthermore, the resulting data by stacking several kinds of features may contain redundant information. Last, but not least, the increase

This work was supported by the FWO (Fund for Scientific Research in Flanders) project G037115N "Data fusion for image analysis in remote sensing."

in the dimensionality of the stacked features, as well as the limited number of labeled samples in many real applications may pose the problem of the curse of dimensionality and, as a consequence, result in the risk of overfitting the training data.

In this paper, we present a method for graph-based fusion of spectral and spatial information, abbreviated by GFSS. This method couples dimension reduction and data fusion of spectral features (from the original HS image) and spatial features contained in the morphological features (computed from the HS image). Variants of our GFSS method were very successful in Data Fusion Contests of the IEEE Geoscience and Remote Sensing Society (GRSS) in 2013[1], focusing on the fusion of hyperspectral and LiDAR data [15] and in 2014, focusing on the fusion of thermal hyperspectral and visible images [16]. Here we present the essence of GFSS in a comprehensive way and we focus on fusing the spectral and spatial information from a hyperspectral imagery as a prerequisite for classification. We evaluate the performance of GFSS in combination with support vector machines (SVM) classifier in a case study on AVIRIS hyperspectral data set. Specifically, our method first generates morphological profiles (MPs) on the first few principal components (PCs) of the original HS image. Then, we build a fusion graph where only the feature points with similar both spectral and spatial characteristics are connected. Finally, we solve the problem of data fusion by projecting all the features onto a linear subspace, on which neighboring data points (i.e., with similar both spectral and spatial characteristics) in the high-dimensional feature space are kept on neighborhood in the low-dimensional projected subspace as well. The organization of this paper is as follows. Section 2 provides a brief review of morphological features. In Section 3, we present the proposed graph-based feature fusion method. The experimental results on real urban hyperspectral images are presented and discussed in Section 4. Finally, the conclusions of the paper are drawn in Section 5.

2. MORPHOLOGICAL FEATURES

Morphological features are generated by either applying morphological openings or closings by reconstruction [2] on the image, using a structural element (SE) of predefined size and shape. An opening acts on bright objects compared with their surrounding, while closings act on dark objects. For example, an opening deletes (this means the pixels in the object take on the value of their surrounding) bright objects that are smaller than the SE. By increasing the size of the SE and repeating the previous operation, a complete morphological profile (MP) is built, carrying information about the size and the shape of objects in the image.

A morphological profile (MP) consists of the opening profile (OP) and the closing profile (CP). For the panchromatic image, MP is built on the original single band image directly.

[1] $http://hyperspectral.ee.uh.edu/?page\_id=795$.

The OP with M scales at pixel $\mathbf{x}$ forms M-dimensional vector, and so does the CP. By incorporating the OP, the CP and original image, the morphological profile of pixel $\mathbf{x}$ is defined as $(2M+1)$-dimensional vector. When applying MP to the hyperspectral data, principal component analysis (PCA) is widely used as a pre-processing step to reduce the dimensionality of the high-dimensional original data, as well as to reduce the redundancy within the bands. Then one applies morphological openings and closings with reconstruction to construct morphological profile on each PC independently. An extended MP (EMP) is formed as a stacked vector consisting of all the morphological profiles. Suppose p PCs are extracted from the original hyperspectral data, then the EMP of pixel $\mathbf{x}$ is a $p(2M+1)$-dimensional vector. Fig. 1 shows a OP built on the first PC. The effect of using morphological features for classification of remote sensing data from urban areas has been discussed in numerous studies [2, 3, 4, 5, 8, 9, 10, 11].

3. THE GRAPH-BASED FUSION METHOD

Let $\mathbf{X}^{Spe} = \{\mathbf{x}_i^{Spe}\}_{i=1}^n$ and $\mathbf{X}^{Spa} = \{\mathbf{x}_i^{Spa}\}_{i=1}^n$ denote the spectral and spatial features after normalization to the same dimension, where $\mathbf{x}_i^{Spe} \in \Re^D$ and $\mathbf{x}_i^{Spa} \in \Re^D$. $\mathbf{X}^{Sta} = \{\mathbf{x}_i^{Sta}\}_{i=1}^n = [\mathbf{X}^{Spe}; \mathbf{X}^{Spa}]$, and $\mathbf{x}_i^{Sta} = [\mathbf{x}_i^{Spe}; \mathbf{x}_i^{Spa}] \in \Re^{2D}$ denotes the vector stacked by the spectral and spatial features. $\{\mathbf{z}_i\}_{i=1}^n$, and $\mathbf{z}_i \in \Re^d$ denote the fusion features in a lower dimensional feature space with $d \leq 2D$.

The goal of this paper is to find a transformation matrix $\mathbf{W} \in \Re^{2D \times d}$, which can couple dimensionality reduction and feature fusion in a way of:

$$\mathbf{z}_i = \mathbf{W}^T \mathbf{x}_i \tag{1}$$

where $\mathbf{x}_i$ is a variable, which can set to be $\mathbf{x}_i^{Sta}$. The transformation matrix $\mathbf{W}$ should not only fuse different features in a lower dimensional feature space, but also preserve local neighborhood information and detect the manifold embedded in the high-dimensional feature space. A reasonable way [12] to find the transformation matrix $\mathbf{W}$ can be defined as follows:

$$\arg \min_{\mathbf{W} \in \Re^{2D \times d}} \left(\sum_{i,j=1}^{n} ||\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j||^2 A_{ij} \right) \tag{2}$$

where the matrix $\mathbf{A}$ is the edge of the graph $\mathbf{G} = (\mathbf{X}, \mathbf{A})$. We assume that the edge (between data point $\mathbf{x}_i$ and $\mathbf{x}_j$) $A_{ij} \in \{0, 1\}$; $A_{ij} = 1$ if $\mathbf{x}_i$ and $\mathbf{x}_j$ are "close" and $A_{ij} = 0$ if $\mathbf{x}_i$ and $\mathbf{x}_j$ are "far apart". The "close" here is defined by finding the k nearest neighbors (kNN) of the data point $\mathbf{x}_i$. The kNN is determined first by calculating the distance (we use Euclidean distance here) between data point $\mathbf{x}_i$ and all the data points, then sorting the distance and determining nearest neighbors based on the k-th minimum distance.

When the graph is constructed by spectral features (i.e. $\mathbf{G} = \mathbf{G}^{Spe} = (\mathbf{X}^{Spe}, \mathbf{A}^{Spe})$), the k nearest neighbors (i.e.

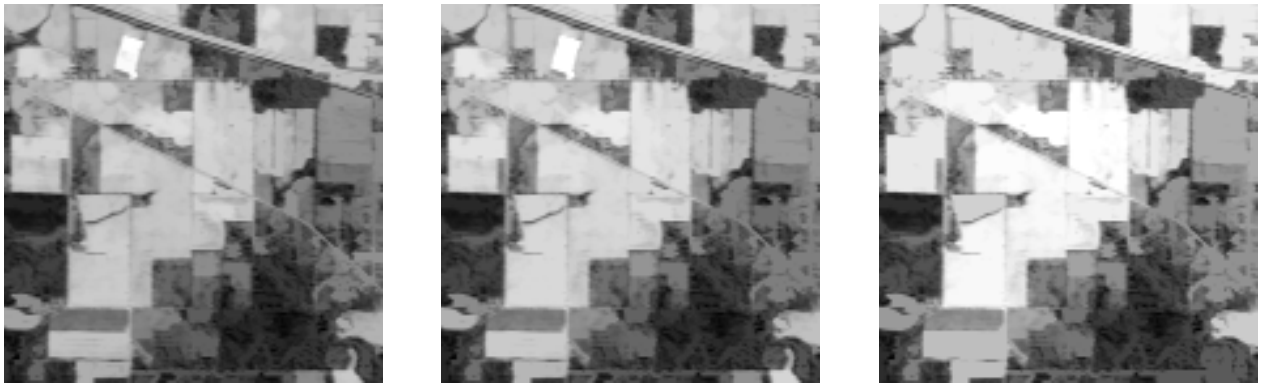


Fig. 1: Morphological opening profile built on the first PC of hyperspectral image. The scale of circular SE varies from 2 to 6, with step size increment of 2. As the size of the SE increases in openings, more small bright objects disappear in the dark background.

$A_{i,j} = A_{i,j}^{Spe} = 1$, $j \in \{1, 2, \cdots, k\}$) of the data point $\mathbf{x}_i^{Spe}$ indicate the spectral signatures of these kNN data points $\mathbf{x}_j^{Spe}$ are more similar in terms of Euclidean distance. On the other hand, when the graph is constructed by spatial features (i.e. $\mathbf{G} = \mathbf{G}^{Spa} = (\mathbf{X}^{Spa}, \mathbf{A}^{Spa})$), the k nearest neighbors (i.e. $A_{i,j} = A_{i,j}^{Spa} = 1$, $j \in \{1, 2, \cdots, k\}$) of the data point $\mathbf{x}_i^{Spa}$ mean that they are more similar in term of the spatial characteristic. We propose a fusion graph which we define $\mathbf{G}^{Fus} = (\mathbf{X}^{Sta}, \mathbf{A}^{Fus})$ as follows:

$$\mathbf{A}^{Fus} = \mathbf{A}^{Spe} \odot \mathbf{A}^{Spa} \tag{3}$$

where the operator '$\odot$' denotes element-wise multiplication, i.e. $A_{i,j}^{Fus} = A_{i,j}^{Spe} A_{i,j}^{Spa}$. Note that $A_{ij}^{Fus} = 1$ only if $A_{ij}^{Spe} = 1$ and $A_{ij}^{Spa} = 1$. This means that the stacked data point $\mathbf{x}_i^{Sta}$ is "close" to $\mathbf{x}_j^{Sta}$ only if all individual feature points $\mathbf{x}_i^{Ind}$ ($Ind \in Spe, Spa$) is "close" to $\mathbf{x}_j^{Ind}$. The connected data points $\mathbf{x}_i^{Sta}$ and $\mathbf{x}_j^{Sta}$ have similar spectral and spatial characteristics. If any individual feature point $\mathbf{x}_i^{Ind}$ is "far apart" from $\mathbf{x}_j^{Ind}$, then $A_{ij}^{Fus} = 0$. In real data, the data points from the roofs ($\mathbf{x}_i^{Sta}$) and roads ($\mathbf{x}_j^{Sta}$) are both made with the same materials (e.g. asphalt) and have similar spectral characteristics ($\mathbf{A}_{i,j}^{Spe} = 1$), but different spatial information (i.e. shape and size) ($\mathbf{A}_{i,j}^{Spa} = 0$), so these two data points are not "close" (i.e. $\mathbf{A}_{i,j}^{Fus} = 0$). On the other hand, the data points from the grassy areas ($\mathbf{x}_i^{Sta}$) and soil areas ($\mathbf{x}_j^{Sta}$) have different spectral characteristics ($\mathbf{A}_{i,j}^{Spe} = 0$), but similar spatial information ($\mathbf{A}_{i,j}^{Spa} = 1$), so $\mathbf{A}_{i,j}^{Fus} = 0$ and these two data points are "far apart". When using the constraint in [13] for avoiding degeneracy:

$$\mathbf{W}^T (\mathbf{X}^{Sta}) \mathbf{D}^{Fus} (\mathbf{X}^{Sta})^T \mathbf{W} = \mathbf{I} \tag{4}$$

where $\mathbf{D}^{Fus}$ is a diagonal matrix with $D_{i,i}^{Fus} = \sum_{j=1}^{n} A_{i,j}^{Fus}$ and $\mathbf{I}$ the identity matrix, we can obtain the transformation matrix $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_r)$ which is made up by r eigenvectors associated with the least r eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$ of the following generalized eigenvalue problem:

$$(\mathbf{X}^{Sta}) \mathbf{L}^{Fus} (\mathbf{X}^{Sta})^T \mathbf{w} = \lambda (\mathbf{X}^{Sta}) \mathbf{D}^{Fus} (\mathbf{X}^{Sta})^T \mathbf{w} \tag{5}$$

where $\mathbf{L}^{Fus} = \mathbf{D}^{Fus} - \mathbf{A}^{Fus}$ is the fusion Laplacian matrix.

4. EXPERIMENTAL RESULTS

The dataset that we used in the experiments was captured by Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over northwestern Indiana in June 1992, with 220 spectral bands in the wavelength range from 0.4 to 2.5μm and spatial resolution of 20 meters by pixel. The calibrated data are available online (along with detailed ground-truth information) from http://cobweb.ecn.purdue.edu/~biehl/. The whole scene, consisting of the full 145×145 pixels, which contains 16 classes, ranging in size from 20 to 2468 pixels, see Fig. 2. Table 1 shows the number of labeled samples in each class. Note that the color in the cell denotes different classes in the classification maps (Fig. 2).

To apply the morphological profiles to hyperspectral images, principal component analysis (PCA) was first applied to the original hyperspectral data set, and the first 3 principal components (PCs) were selected (representing 99% of the cumulative variance) to construct the EMP. A circular SE ranging from 1 to 10 with step size increment of 1 was used. 10 openings and closings were computed for each PC, resulting in an EMP of dimension 63.

The SVM classifier with radial basis function (RBF) [14] kernels is applied in our experiments. SVM with RBF kernels has two parameters: the penalty factor C and the RBF kernel width γ. We apply a grid-search on C and γ using 5-fold cross-validation to find the best C within the given set $\{10^{-1}, 10^0, 10^1, 10^2, 10^3\}$ and the best γ within the given set $\{10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1\}$. We compare our approach

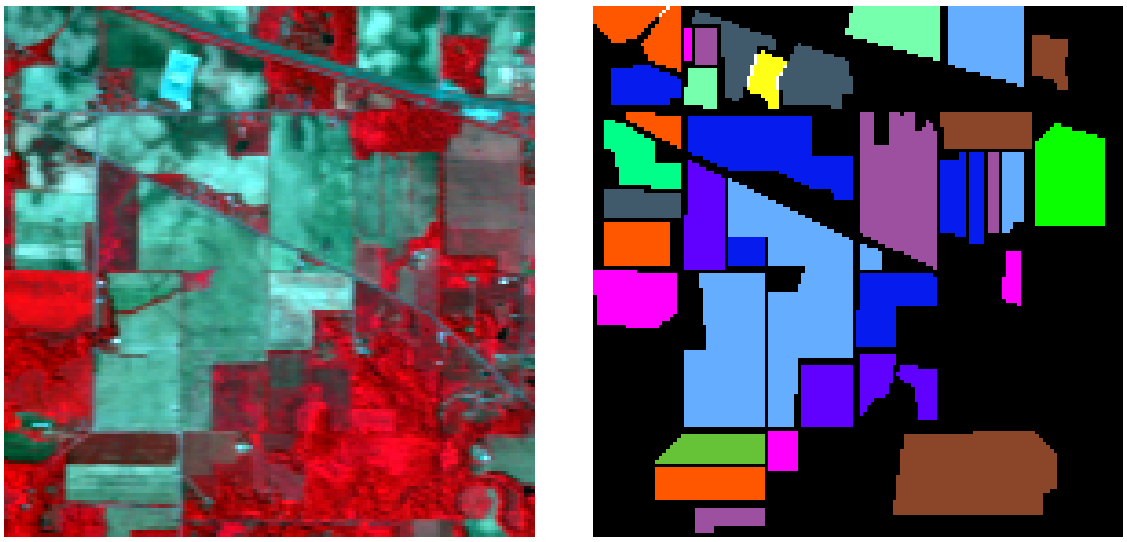


Fig. 2: HSI data sets used in our experiments. Left: false color image of Indian Pines; Right: ground truth of the area with 13 classes

Table 1: Data Sets Used in The Experiments

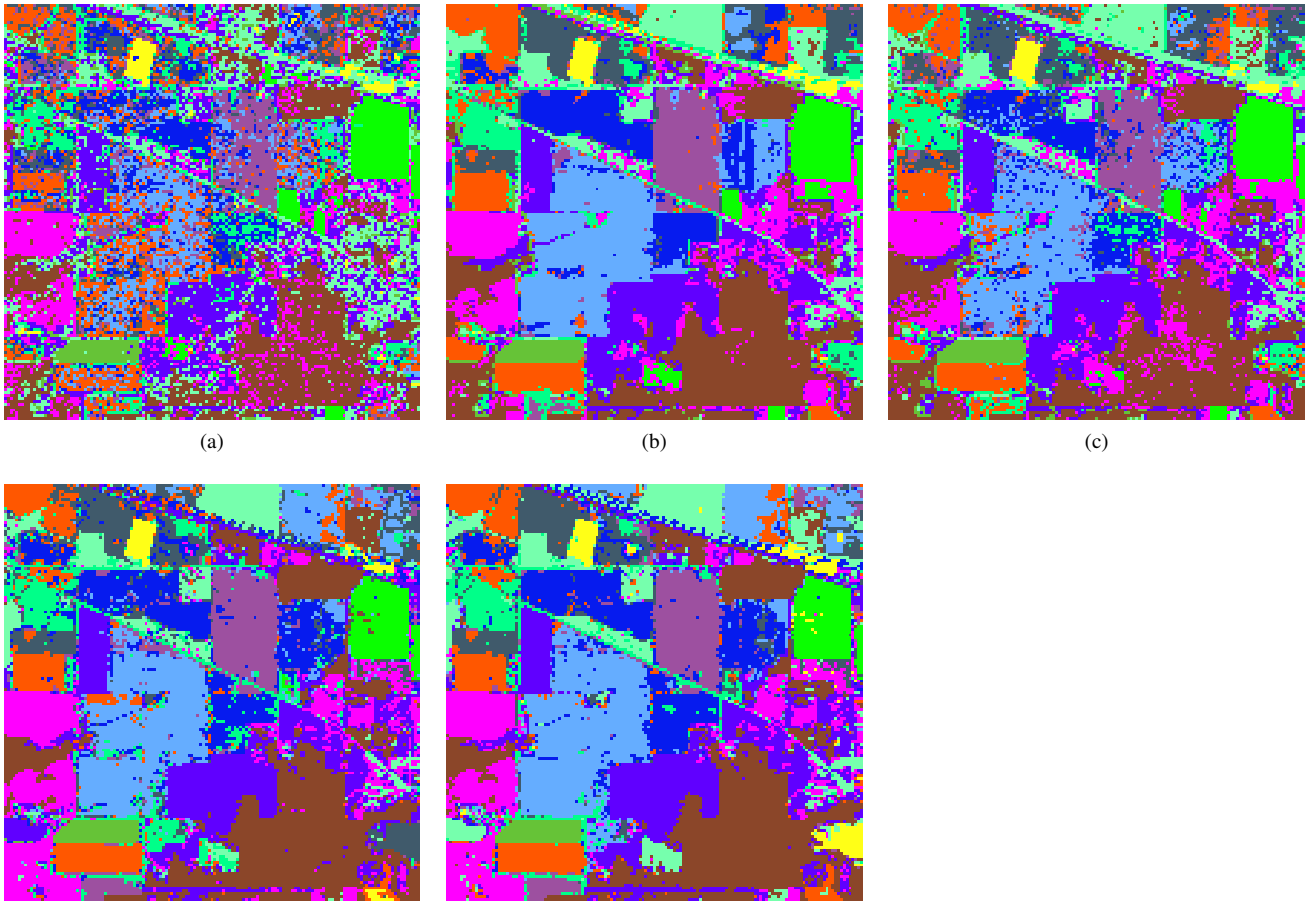
Class No.	Class Name	# Samples	Class No.	Class Name	# Samples
1	Corn-notill	1434	2	Corn-min	834
3	Corn	234	4	Grass/Pasture	497
5	Grass/Trees	747	6	Hay-windrowed	489
7	Soybeans-notill	968	8	Soybeans-min	2468
9	Soybeans-clean	614	10	Wheat	212
11	Woods	1294	12	Bldg-Grass-Trees	380
13	Stone-steel towers	95			

GFSS with the schemes of (1) Using original HSI (Raw); (2) Using the MPs computed on the first 4 PCs of original HSI (EMP) [2]; (3) Stacking all spectral and MPs together (Sta), similar as [11]; (4) Features fused by using the graph constructed by stacked features $\mathbf{X}^{Sta}$(i.e. LPP [13]) (LPP). We select 20 labeled samples per class for training. The classifiers were evaluated against the remaining labeled samples by measuring the Overall Accuracy (OA), the Average Accuracy (AA) and the Kappa coefficient (κ). Table 2 shows the accuracies obtained from the experiments, and Fig. 3 shows the classification maps.

It can be found that only using single spectral/spatial feature is not enough for a reliable classification. It is better sometimes to use single feature source than simply stacking many of them for classification. Compared to the situation with single spatial features (EMP), the OA of simply stacking original spectral and spatial features (Sta) decreases more than 10 percentage points, while increasing the dimensionality. This indicates that the spatial information contained in the original EMP was not well exploited in such a stacked architecture. Indeed, when stacking all features together, the element values of different features can be significantly unbalanced, and the information contained by different features is not equally represented. The same problems happen when using the stacked features to build a graph in LPP method. Our approach GFSS produced the best results, with OA improvements of 4-30% over only using the single spectral/spatial feature source, with improvements of 14% over stacking both the spectral and the spatial features by Sta, and with 3% improvement over the LPP.

Table 2: Classification accuracies obtained by different approaches. The numbers in brackets are the numbers of extracted spectral and spatial features, respectively.

	Raw	EMP	Sta	LPP	GFSS
No. of Features	220	83	303	30	36
OA (%)	55.64	83.67	73.01	84.36	**87.27**
AA (%)	65.65	87.82	81.55	87.86	**89.85**
κ	0.506	0.815	0.697	0.823	**0.855**
Corn-notill	40.38	77.75	53.14	76.36	83.33
Corn-min	54.44	93.29	87.53	90.89	**93.17**
Corn	66.67	87.18	87.18	82.05	**90.60**
Grass/Pasture	72.43	77.26	75.86	77.06	**79.07**
Grass/Trees	81.39	92.37	93.17	97.99	**98.26**
Hay-windrowed	98.77	99.59	99.59	95.71	**96.73**
Soybeans-notill	52.17	**78.31**	62.60	81.82	83.06
Soybeans-min	34.40	**77.07**	61.10	75.24	82.13
Soybeans-clean	36.97	**76.38**	69.38	71.17	74.59
Wheat	96.23	**99.53**	99.53	99.53	99.53
Woods	79.06	**87.09**	80.60	97.99	93.59
Bldg-Grass-Trees	45.97	**97.89**	92.63	98.42	98.16
Stone-steel towers	94.74	**97.89**	97.89	97.89	95.79



(a) (b) (c) (d) (e)

Fig. 3: Classification maps produced by the described schemes. Thematic map using (a) original HS data; (b) EMP of HS data; (d) the LPP; (e) GFSS.

5. CONCLUSION

In this paper, we give a comprehensive presentation of a graph-based feature fusion method, which enables to include

both spectral and spatial information in the classification process. The morphological features, which carry the spatial information, are first generated on the first few PCs of HS image. Then, we build a fusion graph where only the feature points with both similar spectral and spatial characteristics are connected. Finally, we solve the problem of data fusion by projecting all the features into a linear subspace. This projection guarantees preservation of the local geometry properties. The neighboring relations are kept after the dimension reduction. Experiments on a real hyperspectral image demonstrate that the proposed fusion method can greatly benefit the accuracy of the subsequent classification.

Acknowledgment

The authors would like to thank Prof. Landgrebe for providing the AVIRIS Indian Pines dataset.

6. REFERENCES

[1] B. C. Kuo and D. A. Landgrebe, "Nonparametric weighted feature extraction for classification", *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 5, pp. 1096-1105, May 2004.

[2] J. A. Benediktsson, J. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles", *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480-491, Mar. 2005.

[3] W. Liao, R. Bellens, A. Pižurica, W. Philips and Y. Pi, "Classification of Hyperspectral Data Over Urban Areas Using Directional Morphological Profiles and Semi-Supervised Feature Extraction," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 5, no. 4, pp. 1177-1190, Aug. 2012.

[4] W. Liao, M. Dalla Mura, J. Chanussot, A. Pizurica, Fusion of Spectral and Spatial Information for Classification of Hyperspectral Remote Sensed Imagery by Local Graph, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 2, pp. 583-594, Feb. 2016.

[5] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Extended profiles with morphological attribute filters for the analysis of hyperspectral data," *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5975-5991, Nov. 2010.

[6] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93-97, Jan. 2006.

[7] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial-spectral kernel-based approach for the classification of remote-sensing images," *Pattern Recognit.*, vol. 45, no. 1, pp. 381-392, 2012.

[8] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, J. C. Tilton, "Advances in Spectral-Spatial Classification of Hyperspectral Images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652-675, Mar. 2013.

[9] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias and J. A. Benediktsson, "Generalized Composite Kernel Framework for Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4816-4829, Sep. 2013.

[10] J. A. Palmason, J. A. Benediktsson, J. R. Sveinsson, and J. Chanussot, "Fusion of morphological and spectral information for classification of hyperspectral urban remote sensing data," *in Proc. IGARSS 06*, pp. 2506C2509, Jul. 2006.

[11] M. Fauvel, J. A. Benediktsson, J. Chanussot and J. R. Sveinsson, "Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profile", *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804-3814, Nov. 2008.

[12] M. Belkin and P. Niyogi, "Laplacia Eigenmaps and Spectral Techniques for Embedding and Clustering," *Advances in Neural Information Processing Systems 14*, pp. 585-591, MIT Press, British Columbia, Canada, 2002.

[13] X. F. He, P. Niyogi, "Locality preserving projections," *Advances in Neural Information Processing Systems 16*, pp. 153-160, MIT Press, Cambridge, 2004.

[14] C. C. Chang and C. J. Lin, "LIBSVM: A Library for Support Vector Machines", 2001, *http://www.csie.ntu.edu.tw/ ˜cjlin/libsvm*.

[15] C. Debes, A. Merentitis, R. Heremans, J. Hahn, N. Frangiadakis, T. Kasteren, W. Liao, R. Bellens, A. Pizurica, S. Gautama, W. Philips, S. Prasad, Q. Du, F. Pacifici, "Hyperspectral and LiDAR Data Fusion: Outcome of the 2013 GRSS Data Fusion Contest," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol.7, no. 6, pp. 2405-2418, March 2014.

[16] W. Liao, X. Huang, F. Coillie, S. Gautama, A. Pizurica, W. Philips, H. Liu, T. Zhu, M. Shimoni, G. Moser, and D. Tuia, "Processing of Multiresolution Thermal Hyperspectral and Digital Color Data: Outcome of the 2014 IEEE GRSS Data Fusion Contest," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2984-2996, June 2015..