Discriminant Tensor Spectral–Spatial Feature Extraction for Hyperspectral Image Classification

Zisha Zhong, Bin Fan, Jiangyong Duan, Lingfeng Wang, Kun Ding, Shiming Xiang, and Chunhong Pan

***Abstract*—We propose to integrate spectral–spatial feature extraction and tensor discriminant analysis for hyperspectral image classification. First, we apply remarkable spectral–spatial feature extraction approaches in the hyperspectral cube to extract a feature tensor for each pixel. Then, based on class label information, local tensor discriminant analysis is used to remove redundant information for subsequent classification procedure. The approach not only extracts sufficient spectral–spatial features from original hyperspectral images but also gets better feature representation owing to tensor framework. Comparative results on two benchmarks demonstrate the effectiveness of our method.**

***Index Terms*—Discriminative tensor representation, hyperspectral classification, spectral–spatial feature extraction.**

I. Introduction

HYPERSPECTRAL image classification remains an active topic for many years with great achievements [1], [2]. Traditional pixel-wise classification takes the spectral signatures as inputs. It does not consider the spatial information from neighbor pixels. Thus, the classification result may be noisy like salt-and-pepper noise. Many studies have led pixel-wise classification into spectral–spatial classification, which has been proved very effective for improving classification accuracy and visual effect. Representative works are extended morphological profiles (EMPs) [3], extended attribute profiles (EAPs) [4], and Gabor filtering [5]. Generally speaking, these remarkable spectral–spatial feature extraction methods extract abundant features using well-designed morphological profiles, attribute profiles, or Gabor wavelet filters from full or several top components of the hyperspectral cube with different parameter configurations. However, they usually consider the output of these efficient filters as a long vector, which may cause the problem of curse of dimensionality under the high-dimensional and limited available training samples.

Manuscript received January 22, 2014; revised May 15, 2014, September 9, 2014, and October 9, 2014; accepted November 15, 2014. Date of publication December 18, 2014; date of current version February 5, 2015. This work was supported by the Natural Science Foundation of China under Grants 91338202, 61331018, 61305049, and 61375024.

Z. Zhong, B. Fan, L. Wang, K. Ding, S. Xiang, and C. Pan are with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: zszhong@nlpr.ia.ac.cn; bfan@nlpr.ia.ac.cn; lfwang@nlpr.ia.ac.cn; kding@nlpr.ia.ac.cn; smxiang@nlpr.ia.ac.cn; chpan@nlpr.ia.ac.cn).

J. Duan is with the Payload Operation and Application Center, Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Beijing 100094, China (e-mail: duanjy@csu.ac.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2014.2375188

To overcome the curse of dimensionality, dimension reduction can be achieved with feature selection or feature extraction. Representative feature selection techniques in remote sensing are sequential forward floating selection [6] and extended Jeffreys–Matnsita criterion [7]. Additionally, Bruzzone and Persello [8] have proposed a novel technique based on multiobjective optimization to spatially invariant feature selection for hyperspectral image classification. Considering the feature extraction pipeline, two basic approaches are principal component (PC) analysis [9] and linear discriminant analysis (LDA) [10]. Due to the inherent Gaussian assumption, they may be not suitable for multimodal high-dimensional data with complex structures. Other widely used approaches in hyperspectral communities are maximum noise fraction [11] and nonparametric weighted feature extraction (NWFE) [12].

Recently, tensor representation has attracted great interest and has been widely applied to problems with tensorial data. In remote sensing research communities, Renard and Bourennane [13] have introduced a dimension-reduction strategy based on tensor modeling. Zhang *et al.* [14] have developed a multifeature tensor representation method for target recognition. Bourennane *et al.* [15] have proposed multidimensional Wiener filtering to jointly achieve denoising and dimension reduction. Lin and Bourennane [16] have surveyed two tensor-based denoising methods and proposed a novel combination to preserve rare signals during the denoising procedure. Velasco-Forero and Angulo [17] have integrated morphological decomposition and tensor PC analysis to improve hyperspectral pixel-wise classification. Zhang *et al.* [18] have developed a tensor discriminative locality alignment method for removing redundant information in hyperspectral image classification.

In this letter, we propose a spectral–spatial discriminative feature extraction method based on 2-D feature tensor representation to improve hyperspectral image classification. First, we apply spectral–spatial feature extraction methods to extract abundant spectral–spatial features and represent the features for each pixel as a second-order feature tensor (i.e., a matrix). Then based on class label information, the local tensor discriminant analysis (LTDA) [19] technique is adopted to remove redundant information and extract discriminative representation for the subsequent classification procedure.

Our contributions mainly lie in three points: 1) We represent spectral–spatial features as feature tensors, which combine the advantage of spectral–spatial feature extraction and tensor representation. 2) We use a tensor discriminant analysis method, i.e., LTDA [19], to remove redundant information and extract discriminative representation from feature tensors. Meanwhile, the optimal reduced dimensions are automatically

extracted. 3) We conduct a comparative analysis on three well-known spectral–spatial feature extraction approaches and their corresponding discriminative improvements based on second-order tensor representation. Our work is a natural extension as [18]. First, we extend their pixel-wise feature extraction to the spectral–spatial counterpart, which can enhance descriptive power of features and achieve smoother classification maps. Second, tensor discriminative locality alignment in [18] is based on the patch alignment framework, whereas the LTDA is easily derived with a well-defined criterion similar to LDA and could automatically obtain the optimal reduced dimensions during the optimization procedure.

This letter is organized as follows. Section II introduces the LTDA. Section III introduces the proposed discriminative tensor feature extraction of spectral–spatial information for hyperspectral image classification. Section IV presents the experiments and analysis. Section V concludes this letter.

II. LTDA

In this section, we briefly introduce some background on tensor and the LTDA method.

A. Tensor Basics

A tensor with an order m is defined as $\mathsf{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_m}$ with its element denoted by $\mathsf{A}_{i_1,i_2,\ldots,i_m}$. The *inner product* of two tensors $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_m}$ is defined by $\langle \mathsf{A}, \mathsf{B} \rangle = \sum_{i_1}^{I_1} \sum_{i_2}^{I_2} \cdots \sum_{i_m}^{I_m} \mathsf{A}_{i_1,i_2,\ldots,i_m} \mathsf{B}_{i_1,i_2,\ldots,i_m}$. Then, the *Frobenius norm* of a tensor A is defined as $\|\mathsf{A}\|_F = \langle \mathsf{A}, \mathsf{A} \rangle = \sqrt{\sum_{i_1}^{I_1} \cdots \sum_{i_m}^{I_m} \mathsf{A}^2_{i_1,\ldots,i_m}}$. Please refer to [19] for more details.

B. LTDA

Given n high-dimensional tensors $\mathsf{A}_i \in \mathbb{R}^{H_1 \times H_2 \times \cdots \times H_m}$ and its class label from $\{1, 2, \ldots, c\}$, LTDA [19] is used to find m multilinear transformation matrices $\mathbf{U}_k \in \mathbb{R}^{H_k \times L_k} (L_k < H_k, k = 1, 2, \ldots, m)$ to obtain n low-dimensional tensors $\mathsf{B}_i \in \mathbb{R}^{L_1 \times L_2 \times \cdots \times L_m}$, i.e.,

$$\mathsf{B}_i = \mathsf{A}_i \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times \cdots \times_m \mathbf{U}_m \tag{1}$$

such that the separability in the reduced space is maximized. Here, $\times_k$ $(k = 1, 2, \ldots, m)$ is *k-mode product* in tensor algebra and defined as

$$\mathsf{B}_i = \mathsf{A}_i \times_k \mathbf{U}_k \tag{2}$$

$$(\mathsf{B}_i)_{i_1,\cdots,i_{k-1},h,i_{k+1},\cdots,i_m} = \sum_{i_k=1}^{I_k} (\mathsf{A}_i)_{i_1,\ldots,i_{k-1},i_k,i_{k+1},\ldots,i_m} \mathbf{U}_{i_k,h}. \tag{3}$$

Similar to LDA, LTDA then minimizes the distances within a class and maximizes the distances between classes. It is formulated as maximizing the following criterion:

$$L(\mathbf{U}) = \mathbf{S}_b^{\mathbf{U}} - \gamma \mathbf{S}_w^{\mathbf{U}} \tag{4}$$

where $\mathbf{S}_w$ and $\mathbf{S}_b$ are the local within-class scatter and the local between-class scatter matrix, respectively, and $\mathbf{U} = \{\mathbf{U}_1, \ldots, \mathbf{U}_m\}$ is the set of multilinear projection matrices. It is noted that to construct $\mathbf{S}_w$ and $\mathbf{S}_b$ scatter matrices, we consider k_w and k_b neighbors, respectively. For detailed explanation, refer to [19] and [20].

Based on (4), the optimization problem for LTDA can be formulated as

$$\mathbf{U}^* = \arg \max_{1 \le L_k \le H_k \ (k=1,2,\ldots,m)} \max_{\mathbf{U}_k^T \mathbf{U}_k = \mathbf{I}_k \ (k=1,2,\ldots,m)} L(\mathbf{U}). \tag{5}$$

Generally, (5) has no closed solution. Nevertheless, we can obtain a local optimal solution to this problem by alternative optimization. This optimization problem is finally reformulated as a generalized eigenvalue problem, and it is derived that the optimal reduced dimensions L_k $(k = 1, 2, \ldots, m)$ is automatically adapted to the data [19].

In this letter, we extend their pixel-wise feature extraction of the TDLA [18] to the spectral–spatial counterpart, which can enhance descriptive power of features and achieve smoother classification maps. Meanwhile, the LTDA is adopted to remove the redundant information in feature tensors. Furthermore, compared with those conventional vector-based supervised dimension-reduction methods, such as LDA [10] and NWFE [12], there is a large reduction of unknown parameters when using the LTDA, which helps prevent the overfitting problem and consequently improve generalization performance. In Section III, we will describe the proposed method in detail.

III. Proposed Approach

The proposed approach contains three steps: 1) perform spectral–spatial feature extraction via efficient approaches in the hyperspectral image to generate second-order feature tensors; 2) perform LTDA on these feature tensors based on class label; and 3) perform classification with the reduced feature representation. The flowchart is shown in Fig. 1. In the following, we will describe these three steps in detail.

A. Spectral–Spatial Feature Extraction

The well-known spectral–spatial feature extraction methods, such as EMP [3], EAP [4], and Gabor filtering [5], are generally computed with different parameter settings on several top PCs of original hyperspectral data. On each component, they extract morphological profiles, attribute profiles, or filtering responses and then stack them into a vector. To better preserve structure information within these features along PCs and along different parameter settings, we model the extracted features as a second-order tensor (i.e., a matrix). The row-wise of the feature tensor is the view of PCs, and the col-wise is the view of different parameter settings of profiles or responses.

Take the Gabor filtering method as an example, we first design some Gabor filters with five scales and eight orientations under the wavelet framework. With these filters, we convolve each band or the several top PCs of a 3-D hyperspectral cube, which produces 40 3-D Gabor features. Then, from these Gabor

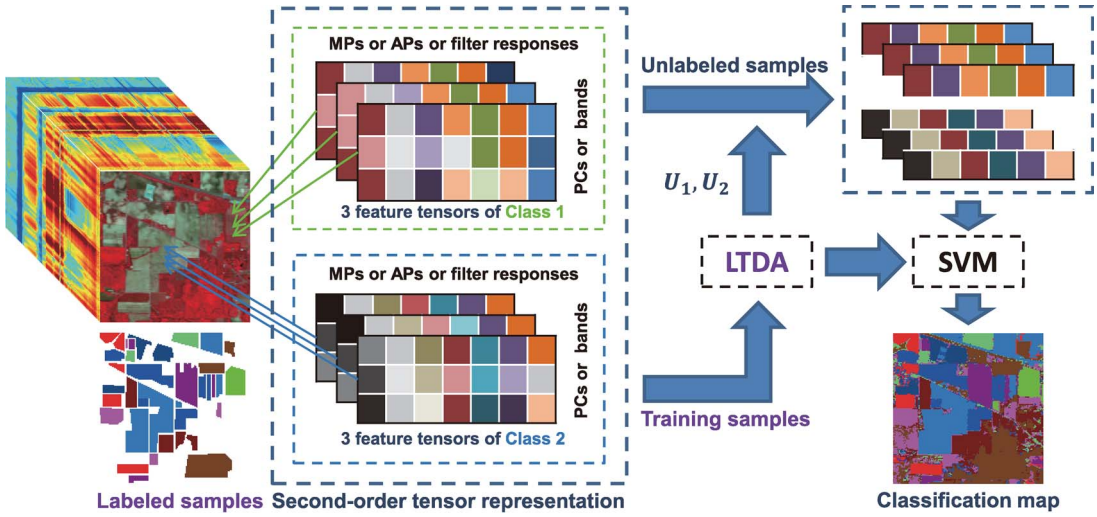

Fig. 1. Flowchart of the proposed method.

features, we can obtain $40 \times P$ (P is the number of total bands or top PCs) features for one pixel. Traditional methods based on Gabor filtering treat the features as a long vector, which loses the structural correlated information among these different filter parameters. In addition, considering the high redundancy of these features, we take advantage of tensor representation and rearrange them into a second-order tensor $\mathsf{A} \in \mathbb{R}^{40\times P}$. A similar strategy can also be adopted for EMP, EAP, or others.

B. Discriminative Feature Extraction

With the LTDA introduced in Section II, we can transform a second-order feature tensor $\mathsf{A} \in \mathbb{R}^{H_1\times H_2}$ into a low-dimensional tensor $\mathsf{B} \in \mathbb{R}^{L_1\times L_2}$. Generally, B has much smaller size than A but can still keep most of its information. LTDA can effectively overcome the limitation of the Gaussian assumption such as LDA. Meanwhile, the optimal reduced dimensions are automatically obtained during the optimization procedure [19]. In this letter, we reshape the reduced feature tensor back into vector representation and feed it into the subsequent classification stage.

C. Classification With Tensor Representation

To sum up, we first extract abundant spectral–spatial features from the total bands or several top components of the 3-D hyperspectral. Then, we rearrange these features for one pixel into a second-order tensor. Based on class label, we take the LTDA technique to effectively reduce the redundant information among these feature tensors as well as extract discriminative features. After that, the extracted feature representations are rearranged back into vector representation, and later, we train support vector machine (SVM) classifiers with them. To validate the effectiveness of our method, we have conducted several comparative experiments. The detailed information will be described in the following section.

IV. Experiment

In this section, we first overview the two standard hyperspectral benchmark data sets, and then, quantitative results and visual classification maps are illustrated to evaluate the effectiveness of our method.

TABLE I
Average (on Five Trials) Overall Classification Accuracy (%), Kappa Coefficient (κ), and Standard Deviation Values Obtained for Both Data Sets

Indian Pines Image

	Original	LDA	NWFE	TDLA	LTDA
Spec.	78.12±0.98	74.34±0.57	84.64±1.03	84.25±0.55	**86.90±1.23**
	0.7444±0.0116	0.7002±0.0067	0.8212±0.0120	0.8168±0.0062	**0.8473±0.0145**
EMP	94.09±0.30	94.01±0.22	94.37±0.31	94.48±0.14	**94.98±0.56**
	0.9314±0.0035	0.9304±0.0026	0.9346±0.0037	0.9358±0.0015	**0.9417±0.0065**
EAP	95.37±0.53	95.49±0.17	95.90±0.21	**96.77±0.32**	96.73±0.35
	0.9462±0.0062	0.9476±0.0020	0.9524±0.0024	**0.9624±0.0037**	0.9620±0.0041
Gabor	98.32±0.15	—	98.64±0.18	**98.72±0.08**	98.72±0.18
	0.9805±0.0018	—	0.9842±0.0021	**0.9851±0.0009**	0.9851±0.0020

Pavia University Image

	Original	LDA	NWFE	TDLA	LTDA
Spec.	94.07±0.25	89.08±0.19	91.48±0.26	**96.24±0.26**	94.81±0.19
	0.9218±0.0034	0.8558±0.0025	0.8875±0.0034	**0.9506±0.0034**	0.9316±0.0026
EMP	97.36±0.11	96.92±0.20	97.26±0.08	96.99±0.23	**97.54±0.20**
	0.9653±0.0015	0.9595±0.0027	0.9641±0.0011	0.9605±0.0030	**0.9677±0.0027**
EAP	99.47±0.02	99.30±0.07	**99.60±0.06**	99.51±0.03	99.53±0.03
	0.9931±0.0003	0.9908±0.0010	**0.9947±0.0008**	0.9936±0.0004	0.9939±0.0004
Gabor	98.96±0.41	—	99.34±0.12	**99.36±0.12**	99.25±0.10
	0.9863±0.0054	—	0.9913±0.0015	**0.9916±0.0016**	0.9902±0.0013

A. Datasets

Indian Pines Image: The original ground truth has 16 labeled classes, we discard six classes due to too less training samples, and finally, the selected classes are as follows: Soybean–Mintill (2468), Corn–Notill (1434), Grass–Trees (747), Soybean–Notill (968), Corn–Mintil (834), Hay–Windrowed (489), Soybean–Clean (614), Grass–Pasture (497), Woods (1294), and Buildings–Grass–Trees (380).

Pavia University Image: The ground truth has nine labeled classes: Meadows (18 686), Asphalt (6852), Bare Soil (5104), Self-Blocking Bricks (3878), Trees (3436), Gravel (2207), Painted metal sheets (1378), Bitumen (1356), and Shadows (1026).

In both data sets, 10% of samples from each class are randomly chosen as the training set and the remaining ones as the testing set, and this is repeated five times for evaluation.

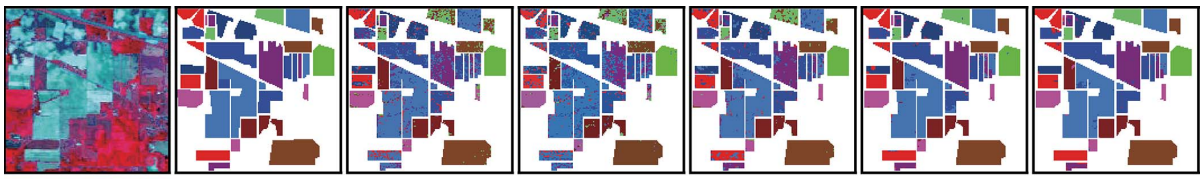

Fig. 2. (a) Indian Pines Image with size of 145 × 145 in pixels and (b) its ground truth image and classification maps obtained by (c) Spectral-Original, (d) Spectral-LDA, (e) Spectral-LTDA, (f) EAP-Original, and (g) EAP-LTDA.

B. Experimental Settings

For comparison, the original spectral signatures (denoted **Original**, also as the baseline method) and three spectral–spatial feature extraction methods are tested, i.e., **EMP** [3], **EAP** [4], and **Gabor** [5]. They are listed as follows. For **Original**, the spectral signature is directly used. For **EMP** and **EAP**, we use five and two PCs in the two data sets, respectively. Nine MP features were computed for each component with disk-shaped structural elements of radius increased from 1 with a step size of 2. For **EAP**, four attributes were computed for each component with the same parameters as [21]. For **Gabor** (denoted **G**), Gabor features with five spatial scales and eight frequency orientations are extracted for each band and then stacked into a long vector. For the second-order tensor representation of **EMP** and **EAP**, we concatenated the profiles or filtering responses for each PC row by row.

For comparison of supervised dimension reduction, three methods are evaluated: **LDA** [10], **NWFE** [12], and **TDLA** [18]. For **LDA**, the reduced dimensions for all methods are $c-1$. For **G-LDA**, the original dimension is $5*8*220 = 8800$, which leads to a singular inverse of S_w in **LDA**; thus, we remove the results. For **NWFE**, the optimal reduced dimension is selected by cross validation. For **LTDA** and **TDLA**, the five spectral signatures in the 4-nearest-neighbor (4-NN) neighborhood are rearranged into a matrix as inputs, as that in [18]. For **TDLA**, there are five adjustable parameters: n_1 is the number of NNs that have the same class label, n_2 is the number of NNs that have different class labels, α is a scaling factor, P_1 is the reduced dimension of the first mode of feature tensors, and P_2 is the reduced dimension of the second mode of feature tensors. As suggested in [18], we set $n_1 = n_2 = 6$, and α is constantly set to 2. P_1 and P_2 are selected according to I_1 and I_2 by cross validation, respectively.

With all the methods, the extracted features are used as the input for multiclass SVM for classification. The Gaussian kernel is adopted in SVM, and the regularization and Gaussian kernel parameters in SVM are selected from $\{2^{-5}, 2^{-4}, \ldots, 2^{15}\}$ and $\{2^{-15}, 2^{-14}, \ldots, 2^{10}\}$ by cross validation. For performance evaluation, two quantitative indexes are used [17]: *overall accuracy* (OA) and *kappa coefficient* (κ), which indicate better results with larger values.

C. Experimental Results

Table I shows the comparative OA and κ with their standard deviations in the two standard benchmark hyperspectral data sets. As shown in the table, the methods with tensor representation achieve overall better performance compared with their vector versions. In particular, the OA accuracy of G-LTDA is about 20% higher compared with the baseline in the Indian Pines Image.

On the whole, the feature extraction methods that incorporate spatial information are better than the baseline. This validates the efficiency of the spectral–spatial feature extraction strategy in hyperspectral image classification. Furthermore, their tensor versions have consistent higher accuracies than their vectorial versions. This also indicates the efficiency of second-order tensor discriminant analysis. In particular, G-LTDA is better than LTDA and Gabor. On the one hand, the better performance over LTDA indicates that spatial feature extraction via Gabor filtering is better than original pixel-wise spectral signatures. On the other hand, the better performance over Gabor gives support on the efficiency of tensor representation. Unfortunately, in some classes, it is slightly worse than others. The main reason is that in the pixels near the edge, the surface texture is complex; thereby, the discriminative capability of the features is weakened.

The classification maps in the Indian Pines Image are illustrated in Fig. 2. As shown in the figure, the baseline has the problem of salt-and-pepper noise. As a result, their classification maps are not smooth. Fortunately, the methods with efficient spectral–spatial feature extraction can overcome this problem, which makes the visual effect better than those of other methods. Such a good performance is attributed to the use of spatial smoothing, which helps suppress noise. Similar results can also be observed in Fig. 3 for the Pavia University Image.

The ANOVA analysis of quantitative results for 20 methods in our experiments is shown in Fig. 4. From these two figures, our method accomplishes comparable or better performance than the other methods. The methods with tensor representation overall achieve higher accuracies than those with vectorial versions. On the other hand, the size of unknown parameters in the projection matrix of LDA or NWFE (which have the vector-based inputs with size of H_1H_2) is $H_1H_2 \times L_1L_2$ (where L_1L_2 is the reduced dimension), whereas that of the projection matrices of the LTDA is $H_1 \times L_1 + H_2 \times L_2$. Thus, owing to the large reduction of unknown parameters of the projection matrices, particularly under the condition of high-dimensional but limited number of training samples, feature extraction with tensor representation has more compact feature representations and lower evaluated variances.

V. Conclusion

In this letter, we have proposed a discriminative tensor representation strategy for spectral–spatial feature extraction in

Fig. 3. (a) Pavia University Image with size of 640 × 310 in pixels and (b) its ground truth image and classification maps obtained by (c) Spectral-TDLA, (d) EMP-Original, (e) EMP-LTDA, (f) EAP-NWFE, and (g) Gabor-LTDA.

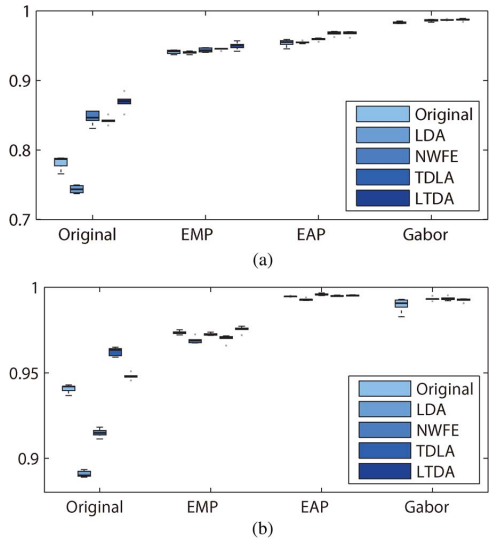

Fig. 4. ANOVA. Black line in each box is the median value. (a) Indian Pines Image. (b) Pavia University Image.

hyperspectral images. First, we extract abundant spectral–spatial features of a pixel to compose a second-order tensor. Then, LTDA is adopted to preserve the discriminativity between classes for classification. Compared with several remarkable methods on two standard data sets, the methods with tensor representation achieve higher classification accuracies and better visual results.

REFERENCES

[1] J. Bioucas-Dias *et al.*, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[2] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, "Advances in hyperspectral image classification: Earth monitoring with statistical learning methods," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 45–54, Jan. 2014.

[3] J. A. Benediktsson, J. A. 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.

[4] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3747–3762, Mar. 2010.

[5] O. Rajadell, P. Garcia-Sevilla, and F. Pla, "Spectral–spatial pixel characterization using Gabor filters for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 860–864, Jul. 2013.

[6] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 153–158, Feb. 1997.

[7] L. Bruzzone, F. Roli, and S. B. Serpico, "An extension of the Jeffreys–Matusita distance to multiclass cases for feature selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 6, pp. 1318–1321, Nov. 1995.

[8] L. Bruzzone and C. Persello, "A novel approach to the selection of spatially invariant features for the classification of hyperspectral images with improved generalization capability," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 9, pp. 3180–3191, Sep. 2009.

[9] I. Jolliffe, *Principal Component Analysis*. Hoboken, NJ, USA: Wiley Online Library, 2005.

[10] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. Mullers, "Fisher discriminant analysis with kernels," in *Proc. IEEE Signal Process. Soc. Workshop*, Madison, WI, USA, Aug. 1999, pp. 41–48.

[11] C.-I. Chang and Q. Du, "Interference and noise-adjusted principal components analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2387–2396, Sep. 1999.

[12] 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.

[13] N. Renard and S. Bourennane, "Dimensionality reduction based on tensor modeling for classification methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1123–1131, Apr. 2009.

[14] L. Zhang, L. Zhang, D. Tao, and X. Huang, "A multifeature tensor for remote-sensing target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 374–378, Mar. 2011.

[15] S. Bourennane, C. Fossati, and A. Cailly, "Improvement of classification for hyperspectral images based on tensor modeling," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 801–805, Oct. 2010.

[16] T. Lin and S. Bourennane, "Survey of hyperspectral image denoising methods based on tensor decompositions," *EURASIP J. Adv. Signal Process.*, vol. 2013, no. 1, pp. 186–196, 2013.

[17] S. Velasco-Forero and J. Angulo, "Classification of hyperspectral images by tensor modeling and additive morphological decomposition," *Pattern Recognit.*, vol. 46, no. 2, pp. 566–577, Feb. 2013.

[18] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Tensor discriminative locality alignment for hyperspectral image spectral–spatial feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 242–256, Jan. 2013.

[19] F. Nie, S. Xiang, Y. Song, and C. Zhang, "Extracting the optimal dimensionality for local tensor discriminant analysis," *Pattern Recognit.*, vol. 42, no. 1, pp. 105–114, Jan. 2009.

[20] S. Xiang, G. Meng, Y. Wang, C. Pan, and C. Zhang, "Image deblurring with matrix regression and gradient evolution," *Pattern Recognit.*, vol. 45, no. 6, pp. 2164–2179, Jun. 2012.

[21] 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, Dec. 2010.