Hyperspectral Images Terrain Classification in Combination Spectrum DLDA Subspace

Jing Liu
School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an, China
Email: zyhalj1975@163.com

Yi Liu and Jin Wu
School of Electronic Engineering, Xidian University, Xi'an, China and School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an, China
Email: {yiliu01@mail.xidian.edu.cn, huatao2000@126.com}

***Abstract*—Hyperspectral images face the problem of high dimensionality and low samples number, which results in unsatisfied recognition efficiency, thus dimensionality reduction is needed before terrain classification. A novel hyperspectral images feature extraction method is presented for dimensionality reduction. Firstly, take discrete Fourier transformation (DFT) of each pixel spectral curve, and combine the amplitude spectrum and corresponding phase spectrum; then direct linear discriminant analysis (DLDA) is performed in the combination spectrum space to extract features. Minimum distance classifier is used to evaluate the feature extraction performance in the achieved combination spectrum DLDA subspace. The experimental results for airborne visible/infrared imaging spectrometer (AVIRIS) hyperspectral image show that, comparing with the spectral DLDA subspace method, the present method can improve the terrain classification efficiency.**

***Index Terms*—Terrain classification, Feature subspace, Feature extraction, Hyperspectral image**

I. INTRODUCTION

Hyperspectral images [1] provide abundant information in spatial domain and spectral domain. Hyperspectral images have high dimensionality because of high spectral resolution, while the number of samples is relatively small due to the expensive and time-consuming sample acquirement. This causes following problems in hyperspectral images terrain classification: (i) comparing with the low dimensionality data space, high dimensionality space has lower linear classification reliability and less classifier generalization ability; (ii) with the increase of bands number, the data amount and calculated amount increase rapidly; (iii) on the condition of fixed data number, with the increase of bands number, classification accuracy will descend after ascending to certain degree [2]. Therefore, dimensionality reduction is needed before terrain classification. Linear discriminant analysis [3–6] (LDA) is an efficient feature extraction method for dimensionality reduction based on Fisher criterion, namely, the ratio of total between-class scatter to average within-class scatter is maximized in the LDA subspace. Recently, LDA is mainly applied in the spectral domain of hyperspectral images [7–10]. Spectral domain means the original hyperspectral bands data space, in the spectral domain, a hyperspectral pixel is a datum in the high dimensional spectral space whose dimensionality is the bands number.

In this paper, a novel hyperspectral images feature extraction method is presented, i.e., direct linear discriminant analysis [11] (DLDA) is used to extract features in the combination spectrum space of amplitude spectrum and phase spectrum. Firstly, take discrete Fourier transformation (DFT) of each pixel spectral curve, then combine the amplitude spectrum and corresponding phase spectrum; secondly, DLDA is performed in the combination spectrum space to extract features. DLDA is an improvement of traditional LDA, which is presented for solving the small sample size (SSS) problem [12] in face recognition. In this paper, we apply LDA to the Fourier frequency domain of hyperspectral image for the first time. An amplitude spectrum is a vector whose components determine the intensities in a pixel spectral curve, and the corresponding phase spectrum is a vector of angles that contains important structure characteristics of a spectral curve. Therefore, the combination spectrum which combines amplitude spectrum and phase spectrum contains both intensities and structure information of a pixel spectral curve. From the viewpoint of feature extraction, the combination spectrum itself is a feature of hyperspectral images which is obtained by DFT. Minimum distance classifier is used to evaluate the feature extraction performance in the achieved combination spectrum DLDA subspace. The experimental results for airborne visible/infrared imaging spectrometer (AVIRIS) hyperspectral image data show that, comparing with the spectral DLDA subspace method, the present method can improve the terrain classification accuracy.

II. COMBINATION SPECTRUM DLDA SUBSPACE TERRAIN CLASSIFICATION

doi:10.4304/jcp.9.1.28-33

A new hyperspectral image terrain classification method is presented in this section. Firstly, the combination spectrum of each pixel spectral curve is achieved by combining the amplitude spectrum and the corresponding phase spectrum; secondly, the combination spectrum DLDA subspace is obtained by performing DLDA in the combination spectrum space; finally, minimum distance classifier is used in the combination spectrum DLDA subspace for evaluating the feature extraction performance.

A. Combination Spectrum Space

Suppose the spectral resolution of a hyperspectral image is N, then an arbitrary pixel $\boldsymbol{x}=[x_0,x_1,\dots,x_{N-1}]^{\mathrm{T}}$ is a datum in N-dimensional data space, the components of $\boldsymbol{x}$ compose a spectral curve. Shown in Fig. 1 are two pixel spectral curves of two types of terrain objects: corn-notill and grass/trees.

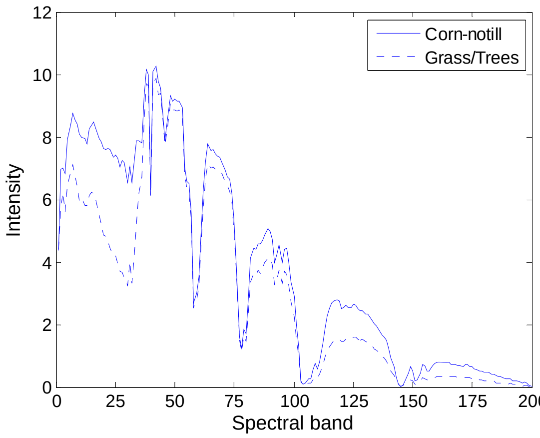

Figure 1. Two pixel spectral curves of corn-notill and grass/trees.

From the viewpoint of signal processing, $\boldsymbol{x}$ is a discrete signal, taking DFT of $\boldsymbol{x}$ yields the spectrum

$$\boldsymbol{X}=[X_0,X_1,\dots,X_{N-1}]^{\mathrm{T}}, \tag{1}$$

where

$$X_u=\sum_{n=0}^{N-1}x_n e^{-j2\pi un/N} \qquad u=0,1,\dots,N-1, \tag{2}$$

and u is the discrete frequency variable. Because DFT is complex in general, it can be expressed in polar form

$$X_u=|X_u|e^{j\phi_u}, \tag{3}$$

where $|X_u|$ and ϕ_u are the magnitude and angle of complex function X_u. The amplitude spectrum of $\boldsymbol{x}$ is

$$\boldsymbol{A}=[|X_0|,|X_1|,\dots,|X_{N-1}|]^{\mathrm{T}}, \tag{4}$$

and the phase spectrum of $\boldsymbol{x}$ is

$$\boldsymbol{\Phi}=[\phi_0,\phi_1,\dots,\phi_{N-1}]^{\mathrm{T}}. \tag{5}$$

The amplitude spectrum of each pixel is even symmetry because the components of a spectral curve are real numbers, thus we combine the former half of amplitude spectrum with corresponding phase spectrum. For an arbitrary pixel $\boldsymbol{x}$, the combination spectrum of amplitude spectrum and phase spectrum is

$$\boldsymbol{C}=[|X_0|,|X_1|,\dots,|X_{(N/2-1)}|,\phi_0,\phi_1,\dots,\phi_{N-1}]^{\mathrm{T}}. \tag{6}$$

Therefore, the dimensionality of the combination spectrum space is $3N/2$. In the combination spectrum space, datum $\boldsymbol{C}$ corresponds to the original pixel $\boldsymbol{x}$.

The components of the amplitude spectrum determine the amplitudes of the sinusoids that combine to form the resulting pixel spectral curve. At any given frequency in the DFT of a spectral curve, a larger amplitude implies a greater prominence of a sinusoid is present in the spectral curve. The phase is a measure of displacement of the various sinusoids with respect to their origin. An amplitude spectrum determines the intensities in a spectral curve, and the corresponding phase spectrum carries important structure characteristics of the spectral curve. Therefore, a combination spectrum contains both intensities and structure information of a spectral curve.

From the viewpoint of feature extraction, the combination spectrum itself is a feature of hyperspectral images which is obtained by DFT. DLDA is performed in the combination spectrum space for further feature extraction and dimensionality reduction. Therefore, the presented hyperspectral images feature extraction method actually includes two successive feature extraction steps, namely, DFT and DLDA.

B. DLDA

DLDA is an improvement of traditional LDA, which is presented for solving the small sample size (SSS) problem in face recognition [11]. Suppose there are c classes training data in an N-dimensional data space, the within-class scatter matrix is defined as

$$\boldsymbol{S}_w=\sum_{j=1}^{c}P_j\Big[\frac{1}{n_j}\sum_{k=1}^{n_j}(\boldsymbol{x}_j^{(k)}-\boldsymbol{m}_j)(\boldsymbol{x}_j^{(k)}-\boldsymbol{m}_j)^{\mathrm{T}}\Big], \tag{7}$$

where n_j is the sample number of class j, $\boldsymbol{x}_j^{(k)}$ is the k-th sample of class j. $\boldsymbol{m}_j$ and P_j are the mean and a priori probability of class j, respectively. The between-class scatter matrix is defined as

$$\boldsymbol{S}_b=\sum_{j=1}^{c}P_j(\boldsymbol{m}_j-\boldsymbol{m})(\boldsymbol{m}_j-\boldsymbol{m})^{\mathrm{T}}, \tag{8}$$

where $\boldsymbol{m}=\sum_{j=1}^{c}P_j\boldsymbol{m}_j$ is the overall mean. LDA finds a linear transformation matrix $\boldsymbol{W}\in\Re^{N\times d}$, such that after the N-dimensional original data $\boldsymbol{x}$ are mapped to a d-dimensional feature subspace by

$\boldsymbol{y} = \boldsymbol{W}^{\mathrm{T}}\boldsymbol{x}$, Fisher criterion

$$J_{\mathrm{F}}(\boldsymbol{W}) = \max_{\boldsymbol{W}}\{tr[(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{S}_w\boldsymbol{W})^{-1}(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{S}_b\boldsymbol{W})]\} \quad (9)$$

is maximized. The solution of (9) consists of the eigenvectors corresponding to the d largest eigenvalues of $\boldsymbol{S}_w^{-1}\boldsymbol{S}_b$ [13,14]. The eigenvalue decomposition of $\boldsymbol{S}_w^{-1}\boldsymbol{S}_b$ is equivalent to diagonalizing $\boldsymbol{S}_w$ and $\boldsymbol{S}_b$ simultaneously. For solving $\boldsymbol{W}$, LDA whitens $\boldsymbol{S}_w$ first, and then diagonalizes $\boldsymbol{S}_b$; while DLDA whithens $\boldsymbol{S}_b$ first, and then diagonalizes $\boldsymbol{S}_w$. The solving steps of DLDA are as follows:

(1) Whiten $\boldsymbol{S}_b$. Let $\boldsymbol{U}_b$ and $\boldsymbol{\Lambda}_b$ be the eigenvector matrix and eigenvalue matrix of $\boldsymbol{S}_b$ respectively, i.e., $\boldsymbol{U}_b^{\mathrm{T}}\boldsymbol{S}_b\boldsymbol{U}_b = \boldsymbol{\Lambda}_b$, $\boldsymbol{\Lambda}_b$ is a diagonal matrix containing the nonzero eigenvalues of $\boldsymbol{S}_b$ in descending order. There exists a matrix $\boldsymbol{W}_1 = \boldsymbol{U}_b\boldsymbol{\Lambda}_b^{-1/2}$ which satisfies $\boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{S}_b\boldsymbol{W}_1 = \boldsymbol{I}$, $\boldsymbol{I}$ is an identity matrix.

(2) Diagonalize $\boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{S}_w\boldsymbol{W}_1$. There exists a matrix $\boldsymbol{U}'_w$ satisfies $(\boldsymbol{U}'_w)^{\mathrm{T}}\boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{S}_w\boldsymbol{W}_1\boldsymbol{U}'_w = \boldsymbol{\Lambda}'_w$, $\boldsymbol{\Lambda}'_w$ is a diagonal matrix containing the d smallest eigenvalues of $\boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{S}_w\boldsymbol{W}_1$ in ascending order, the columns of $\boldsymbol{U}'_w$ are the corresponding eigenvectors.

(3) Sphere the data in feature subspace. There exists a matrix $\boldsymbol{W}_2 = (\boldsymbol{U}'_w)(\boldsymbol{\Lambda}'_w)^{-1/2}$ satisfies $\boldsymbol{W}_2^{\mathrm{T}}\boldsymbol{W}_1^{\mathrm{T}}\boldsymbol{S}_w\boldsymbol{W}_1\boldsymbol{W}_2 = \boldsymbol{I}$.

(4) The resulting linear transformation matrix $\boldsymbol{W} = \boldsymbol{W}_1\boldsymbol{W}_2$, and an N-dimensional original datum $\boldsymbol{x}$ is mapped to the d-dimensional feature subspace by

$$\boldsymbol{y} = \boldsymbol{W}^{\mathrm{T}}\boldsymbol{x}. \quad (10)$$

Comparing with LDA, DLDA is more suitable for feature extraction in the combination spectrum space of hyperspectral images, because: (i) DLDA whitens between-class scatter matrix first, and the null space of between-class scatter matrix which does not contain classification information is discarded in this step. (ii) The combination spectrum space also has SSS problem, under the small sample condition, within-class scatter matrix is singular in general. LDA whitens the within-class scatter matrix firstly, and discards the null space of within-class scatter matrix which contains important classification information; while DLDA diagonalizes the within-class scatter matrix latterly, which can avoid discarding the null space of within-class scatter matrix. (iii) The dimensionality of LDA subspace is limited by classes number, while the dimensionality of DLDA subspace is not limited by classes number.

C. Combination Spectrum DLDA Subspace Terrain Classification

The procedure of the presented terrain classification method is: firstly, calculate the combination spectrum for each pixel; secondly, DLDA is performed in the combination spectrum space to extract features; lastly, minimum distance classifier is designed in the achieved combination spectrum DLDA subspace for recognition. The principle of the presented method is shown in Fig. 2.

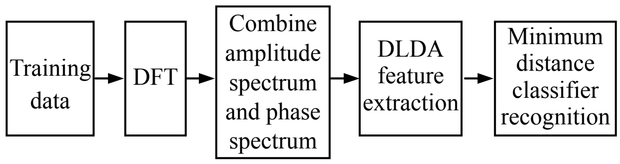

Figure 2. Principle of terrain classification in combination spectrum DLDA subspace.

The detailed steps of the presented method are as follows:

(i) *Feature extraction*.

① Take DFT of each training datum, and the combination spectrum of each training datum is obtained according to (6).

② Perform DLDA in the combination spectrum space according to 2.2, and the linear transformation matrix $\boldsymbol{W}$ is obtained.

(ii) *Set up the templates database*. Using (10), the combination spectrum of each training datum is mapped to the combination spectrum DLDA subspace. The mean vectors $\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \cdots, \boldsymbol{\mu}_c\}$ of each class in the subspace are saved as templates.

(iii) *Recognition*.

① Perform feature extraction to testing data. Take DFT of a testing datum, and the combination spectrum of the testing datum is obtained according to (6), then the combination spectrum of the testing datum is mapped to the combination spectrum DLDA subspace according to (10), and is denoted as $\boldsymbol{y}$.

② Design minimum distance classifier. $\boldsymbol{y}$ belongs to class $\arg\min_{i=1,\ldots,c}\|\boldsymbol{\mu}_i - \boldsymbol{y}\|_2$, where $\|\boldsymbol{\mu}_i - \boldsymbol{y}\|_2$ is the Euclidean distance between $\boldsymbol{y}$ and the i-th template.

III. Experiments

For verifying the efficiency of the presented method, the experimental results of the presented combination spectrum DLDA subspace method are compared with those of the original spectral space method, original combination spectrum space method, and spectral DLDA subspace method, under the condition of same training data and testing data.

A. Experimental Data

We use the AVIRIS hyperspectral image which is taken over Northwest Indiana's Indian Pine test site in

June 1992. The Indian Pines image contains 16 land-cover classes, 145×145 pixels, and 220 spectral bands in the 400–2450-nm range. We remove 20 noisy bands corresponding the region of water absorption, so the final dimensionality of each pixel is 200. Besides the whole-image, we also use a part of the scene, called the sub-image, consisting of pixels [27–94]×[31–116] for a size of 68×86. The sub-image contains 4 classes: Corn-notill, Grass/Trees, Soybeans-notill, and Soybeans-min. Shown in Fig. 3 are the RGB composition maps of the whole-image and the sub-image corresponding to bands 35, 17, and 6.

(a) (b)

Figure 3. RGB composition maps of Indian Pines image corresponding to bands 35, 17, and 6. (a) Whole-image (b) Sub-image.

B. *Experimental Results*

For both the whole-image and the sub-image, we randomly select 20% of each class samples as training data, and the rest as testing data. Table I shows the class label and the testing data number of each class. The total testing data number is 8292.

TABLE I.
CLASS LABELS AND TESTING DATA NUMBERS OF WHOLE-IMAGE

Class names	Class labels	Testing data numbers	Class names	Class labels	Testing data numbers
Alfalfa	1	43	Oats	9	16
Corn-notill	2	1147	Soybeans-notill	10	774
Corn-min	3	667	Soybeans-min	11	1974
Corn	4	187	Soybean-clean	12	491
Grass/Pasture	5	398	Wheat	13	170
Grass/Trees	6	598	Woods	14	1035
Grass/pasture-mowed	7	21	Bldg-Grass-Tree-Drives	15	304
Hay-windrowed	8	391	Stone-steel towers	16	76

Table II shows the average recognition rates of the whole-image. It can be seen from Table II that, in original spectral space and original combination spectrum space, the average recognition rates are both very low, because the original spectral space and original combination spectrum space both have high dimensionalities among which are a lot of redundant dimensions. After DLDA feature extraction, the average recognition rates in spectral DLDA subspace and combination spectrum DLDA subspace are both dramatically improved. The average recognition rate in combination spectrum DLDA subspace is 1.11 percentage points larger than that in spectral DLDA subspace, which shows that the combination spectrum DLDA subspace contains more separability information than spectral DLDA subspace.

TABLE II.
AVERAGE RECOGNITION RATES OF WHOLE-IMAGE IN ORIGINAL SPECTRAL SPACE, SPECTRAL DLDA SUBSPACE, ORIGINAL COMBINATION SPECTRUM SPACE, AND COMBINATION SPECTRUM DLDA SUBSPACE

	Spectral space (200 dimensions)	Spectral DLDA subspace (10 dimensions)	Combination spectrum space (300 dimensions)	Combination spectrum DLDA subspace (16 dimensions)
Average recognition rates [%]	48.38	77.98	43.27	79.09

Table III shows the confusion matrixes for the sub-image in original spectral space and spectral DLDA subspace. The spectral DLDA subspace is achieved by performing DLDA in the original spectral space. It can be seen that, after DLDA feature extraction, the classification accuracy is greatly improved, and the data dimensionality is reduced from 200 to 3.

TABLE III.
CONFUSION MATRIXES FOR SUB-IMAGE IN ORIGINAL SPECTRAL SPACE AND SPECTRAL DLDA SUBSPACE

	Original spectral space (200 dimensions)				Spectral DLDA subspace (3 dimensions)			
Class labels	2	6	10	11	2	6	10	11
2	69.98	0	23.54	36.47	92.0	0	0.7	8.2
6	0.74	100.0	0.17	1.56	0	100.0	0.3	1.0
10	15.63	0	67.53	26.02	3.5	0	94.2	7.1
11	13.65	0	8.76	35.95	4.50	0	4.8	83.7
Average recognition rates [%]	68.37				92.48			

Table IV shows the confusion matrixes for the sub-image in original combination spectrum space and combination spectrum DLDA subspace. Comparing Table IV with Table III we can see that, the average recognition rate in original spectral space is 5.2 percentage points larger than that in original combination spectrum space. In both the original spectral space and original combination spectrum space, the recognition rates of class 2, class 10, and class 11 are very low, this is because the Indian Pines image was collected in June, these crops were very early in their growth cycle with about 5% coverage, discriminating them under this condition can be very difficult.

The average recognition rate in combination spectrum DLDA subspace is 1.82 percentage points larger than that in spectral DLDA subspace, and the recognition rates of class 2, class 10, and class 11 are improved 0.9 percentage point, 1 percentage point, and 6.2 percentage points respectively. These experimental results verify the efficiency of the present method.

TABLE IV.
CONFUSION MATRIXES FOR SUB-IMAGE IN ORIGINAL COMBINATION SPECTRUM SPACE AND COMBINATION SPECTRUM DLDA SUBSPACE

	Combination spectrum space (300 dimensions)				Combination spectrum DLDA subspace (3 dimensions)			
Class labels	2	6	10	11	2	6	10	11
2	70.35	0.51	25.43	36.47	92.9	0	1.2	5.1
6	11.04	99.49	6.36	3.11	0	99.1	0.2	0.5
10	9.55	0	55.84	33.42	0.5	0	95.2	4.5
11	9.06	0	12.37	27.0	6.6	0.9	3.4	89.9
Average recognition rates [%]	63.17				94.3			

Table V shows the confusion matrixes for the sub-image in amplitude spectrum DLDA subspace and phase spectrum DLDA subspace. The amplitude spectrum DLDA subspace and phase spectrum DLDA subspace are achieved by performing DLDA in the amplitude spectrum and phase spectrum space, respectively.

TABLE V.
CONFUSION MATRIXES FOR SUB-IMAGE IN AMPLITUDE SPECTRUM DLDA SUBSPACE AND PHASE SPECTRUM DLDA SUBSPACE

	Amplitude spectrum DLDA subspace (3 dimensions)				Phase spectrum DLDA subspace (3 dimensions)			
Class labels	2	6	10	11	2	6	10	11
2	91.4	0	0.5	6.7	84.7	0.3	2.4	9.2
6	0.1	99.7	0.2	0.8	0	98.7	0.2	0.9
10	1.5	0	95.5	6	3.3	0.5	95.4	9.5
11	7	0.3	3.8	86.6	12	0.5	2.0	80.3
Average recognition rates [%]	93.3				89.8			

It can be seen from Table V that, the average recognition rates in amplitude spectrum DLDA subspace and phase spectrum DLDA subspace are both very high. Comparing Table V with Table III, we can see that, the average recognition rate in the amplitude spectrum DLDA subspace is 0.82 percentage point larger than that in spectral DLDA subspace. Table V indicates that, the amplitude spectrum and phase spectrum of a spectral curve both contain important separability information. Comparing Table V with Table IV, it can be seen that, combination spectrum contains more separability information than the individual amplitude spectrum and phase spectrum.

IV. CONCLUSION

A novel hyperspectral images terrain classification method is presented in this paper. Firstly, the combination spectrum of each pixel spectral curve is achieved by combining the amplitude spectrum and corresponding phase spectrum; secondly, the combination spectrum DLDA subspace is obtained by performing DLDA in the combination spectrum space; finally, minimum distance classifier is designed in the achieved combination spectrum DLDA subspace to evaluate the feature extraction performance. The experimental results for AVIRIS hyperspectral image show that, comparing with the spectral DLDA subspace method, the present method can improve the terrain classification accuracy. The experimental results also verify that, the Fourier spectrum of pixel spectral curve contains important separability information for hyperspectral images terrain classification.

ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China (No. 61003199), the Special Foundation for Natural Science of the Education Department of Shaanxi Province of China (No. 2013JK1129 and No.2013JK1075), and the Young Scholars Plan Project of Xi'an University of Posts and Telecommunications (No. ZL2012-11).

REFERENCES

[1] S. Ding, "Spectral and wavelet-based feature selection with particle swarm optimization for hyperspectral classification," *Journal of Software*, vol. 6, no. 7, pp. 1248–1256, 2011.

[2] G. Hughes, "On the mean accuracy of statistical pattern recognition," *IEEE Transactions on Information Theory*, vol. IT-14, no. 1, pp. 55–63, 1968.

[3] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Human Genetics*, vol. 7, no. 2, pp. 179-188, 1936.

[4] C. R. Rao, "The utilization of multiple measurements in problems of biological classification," *Journal of the Royal Statistical Society, Series B (Methodological)*, vol. 10, no. 2, pp. 159-203, 1948.

[5] Z. Li, G. Tan, Y. Li, "Fault diagnosis based on improved kernel Fisher discriminant analysis," *Journal of Software*, vol. 7, no. 12, pp. 2657–2662, 2012.

[6] G. Wang, X. Li, K. He, "Kernel local fuzzy clustering margin Fisher discriminant method faced on fault diagnosis," *Journal of Software*, vol. 6, no. 10, pp. 1993–2000, 2011.

[7] T. V. Bandos, L. Bruzzone, G. Camps-Valls, et al, "Classification of hyperspectral images with regularized linear discriminant analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 3, pp. 862-873, 2009.

[8] W. Li, S. Prasad, J. E. Fowler, et al, "Locality-preserving discriminant analysis in kernel-induced feature spaces for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 5, pp. 894 – 898, 2011.

[9] W. Liao, A. Pizurica, P. Scheunders, et al, "Semisupervised local discriminant analysis for feature extraction in hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 1, pp. 184-198, 2013.

[10] L. Wei, S. Prasad, J. E. Fowler, et al, "Locality-preserving dimensionality reduction and classification for hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1185-1198, 2012.

[11] H. Yu, J. Yang, “A direct LDA algorithm for high-dimensional data-with application to face recognition,” *Pattern Recognition*, vol. 34, no. 10, pp. 2067-2070, 2001.
[12] Z. Wang, X. Sun, “An efficient discriminant analysis algorithm for document classification,” *Journal of Software*, vol. 6, no. 7, pp. 1265–1272, 2011.
[13] K. Fukunaga, *Introduction to statistical pattern recognition*. USA: Academic Press, 1990.
[14] A. K. Jain, R. P. W. Duin, J. C. Mao, “Statistical pattern recognition: a review,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, 2000.

Jing Liu received the B.Eng. degree in electronic engineering from Xi’an Technological University, Xi’an, China, in 1995; she received the M.Eng. degree and Ph.D. degree in computer science and engineering from Xidian University, Xi’an, China, in 2004 and 2009, respectively.

She is currently an instructor with School of Electronic Engineering, Xi’an University of Posts and Telecommunications. Her research interests are in the fields of remote sensing image automatic target recognition, artificial intelligence, and machine learning.

Yi Liu received the B.Eng. degree from Xidian University in 1999; he received the M.Eng. degree and Ph.D. degree from Xidian University in 2002 and 2013 respectively.

He is currently an instructor with School of Electronic Engineering, Xidian University. His research interests are in the fields of pattern recognition and machine learning.

Jin Wu received the B. Eng. degree from Xi’an Jiaotong university in 1998; she received the M.Eng. degree from Xi’an Jiaotong university in 2001.

She is currently an associate professor with School of Electronic Engineering, Xi’an University of Posts and Telecommunications. Her research interests are in the fields of intelligent information processing.