Hyperspectral Image Classification for Mapping Agricultural Tillage Practices

Qiong Ran[1*], Wei Li[1] , Qian Du[2] , Chenghai Yang[3]

[1] *College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 100029 China;*

[2] *Department of Electrical and Computer Engineering, Mississippi State University, 39762 USA*

[3] *USDA Southern Plains Agricultural Research Center at College Station, Texas, 77845 USA*

* Corresponding author email: ranqiong@mail.buct.edu.cn (Qiong Ran)

Abstract: An efficient classification framework for mapping agricultural tillage practice using hyperspectral remote sensing imagery is proposed, which has the potential to be implemented practically to provide rapid, accurate, and objective surveying data for precision agricultural management and appraisal from large-scale remote sensing images. It includes a local region filter (i.e., Gaussian low-pass filter (GLF)) to extract spatial/spectral features, a dimensionality reduction process (i.e., local fisher's discriminate analysis (LFDA)), and the traditional k-nearest neighbor (KNN) classifier, which is denoted as GLF-LFDA-KNN. Compared to our previously used local average filter (LAF) and adaptive weighted filter (AWF), the GLF considers spatial features in a small neighborhood, but it emphasizes the central pixel itself and it is data-independent; therefore it can achieve the balance between classification accuracy and computational complexity. The KNN classifier has lower computational complexity, compared to the traditional support vector machine (SVM). After classification separability is enhanced by the GLF and LFDA, the less powerful KNN can outperform SVM and the overall computational cost stays lower. The proposed framework can also outperform the support vector machine with composite kernel (SVM-CK) that uses spatial-spectral features.

Keywords: Conservation tillage, hyperspectral data, feature extraction, spatial-spectral classification, agricultural remote sensing.

I. INTRODUCTION

Conservation tillage management has been advocated for the purpose of soil preservation and sustainable crop production [1, 2]. Conservational tillage practice induces less surface disturbance and

leaves more crop residues, which can decrease runoff rate, improve soil and water quality, and increase organic matter. The demand for mapping of crop tillage practices has been brought up for precision agricultural management and appraisal [3]. However, current methods to mapping massive crop tillage practices are mainly with field investigations, which are labor-costing, time-consuming, subjective, and difficult to generate widely distributed survey data. Remote sensing technology provides a more rapid, accurate, and objective solution [3]. Moreover, the vast data from remote sensing in agriculture requires more efficient approaches in data analytics, including tillage mapping, in support of management decisions.

Some methods have been explored for classification of tillage with multispectral remote sensing images with indices or classifiers [4, 5, 6]. However, these methods need data acquired at a specific time [4], or are based on multi-temporal data with revisits over the observed field [5, 6], making the task difficult and costly. Hyperspectral data provides more subtle spectral information about the imaged scene with hundreds of narrow contiguous bands, and is able to reveal spectral discrepancy among different tillage conditions, thereby providing a feasible solution for agricultural tillage practice mapping [7]. However, challenge still remains because of spectral similarities. For fields with different tillage conditions, a spectrum can be taken as mixture of soil and crop residues, and soil moisture. The spectral difference of the field with different tillage conditions mainly come from the proportion of soil and crop residue, and soil moisture. Thus, for better discrimination of different tillage practices, the shortwave infrared spectrum needs to be adopted to reveal more details about vegetation and moisture conditions. In this study, we choose Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data with shortwave infrared information in the experiments.

Numerous classification algorithms have been developed using hyperspectral data with the rich spectral information [8-11]. Many pixel-wise classifiers are based solely on spectral signatures. For these methods, pixels are classified independently without considering the relations between spatially adjacent pixels. However, recent investigations have demonstrated the importance of spatial information, and more classifiers have utilized spectral-spatial features. This is particularly true for tillage practice classification because different tillage systems have different spatial features. Classifiers with spatial information can be mainly divided into two categories, the one integrating texture-related features and the one integrating contextual relation. The texture information is often combined to generate a joint feature vector for pixel-wise classification [11-18]. For example, morphological profile (MP) has been introduced for classification [12]. Later, MP features were integrated with spectral features in hyperspectral image analysis [13]. Mean feature is combined with spectral features for support vector machine (SVM) based classification [14]. Texture features, such as digital wavelet transform (DWT) [15], 2-D Gabor features [16], 3-D Gabor features [17] and gray level

co-occurrence matrices (GLCM) [18], have also been investigated. The methods integrating contextual information include the relaxation labeling, and those with the random field models, such as conditional random field, Markov random field (MRF), hidden Markov model (HMD), etc. [19-22].

Applying a local-region filter (LRF) has been demonstrated as an efficient approach for combining spatial-spectral information [23-24], which has much lower complexity than the aforementioned methods such as MF and GLCM. It is particularly useful for images with large homogeneous areas, such as agricultural remote sensing images. An LRF filter an image by replacing the central pixel (vector) with the generated feature according to the weights assigned to its neighbors. The processed image may have improved discriminative ability by reducing within-class spectral variations and intrinsic noise. An LRF, such as local average filter (LAF), has been tested for classification of conservation tillage practices [25]. The LAF assigns identical weights to the surrounding neighboring pixels and may be suboptimal for its rough averaging procedure. In [24], the adaptive weighted filter (AWF) was proposed to improve the performance of LAF by considering pixel discrepancy with a small neighborhood.

In this paper, we will investigate the performance of LAF and AWF in conservation tillage mapping. Considering the fact that practical remote sensing data often have large scales, the computational cost of an LRF is a concern. Although the LAF basically does not introduce additional cost in computing, its performance may be poor in heterogeneous areas; on the other hand, the AWF outperforms in [24] but with much higher computational cost because pairwise pixel similarity has to be measured. In this paper, we propose to use the cost-effective Gaussian low-pass filter (GLF), which assigns larger weights to spatial (not feature) nearest neighbors. It is expected that the data-independent GLF can outperform the LAF, and is computationally much more efficient than the data-dependent AWF, making it more suitable to large-scale image data analysis.

Support vector machine (SVM) is a powerful classifier [12]. However, its complexity may be high since a large optimization problem has to be solved during its training process. The traditional k-nearest neighbor (KNN) classifier is much simpler, but less powerful. Here, we propose to use a dimensionality reduction process, called locality-preserving-based discriminate analysis (LFDA) [9, 11], to improve class separability. When the KNN is applied to the LFDA-transformed data, the final classification accuracy can be as high as that of SVM. Note that the major computational cost in LFDA is in an eigen-decomposition process (here only the first eigenvector is used), which is lower than that of SVM.

The proposed methods, including the step of spatial-spectral feature extraction with a LRF, are denoted as LAF-LFDA-KNN, AWF-LFDA-KNN, and GLF-LFDA-KNN, respectively. They are validated with two real hyperspectral datasets, and compared with SVM [12], SVM with composite

kernel (SVM-CK) [14], the traditional KNN, LAF followed by KNN (LAF-KNN), AWF followed by KNN (AWF-KNN), and GLF followed by KNN (GLF-KNN). The experimental results demonstrate that the proposed framework (i.e., an LRF followed by LFDA+KNN) can apparently improve the classification performance, and can be applied to agricultural tillage condition mapping. In particular, GLF-LFDA-KNN requires lowest computational cost but offers better classification performance than SVM-CK.

The remainder of this paper is organized as follows. Section II introduces the proposed methods. Section III provides experimental results. Section IV draws the conclusion.

II. PROPOSED METHOD

2.1. Local Region Filter

Since mapping of the conservation tillage practices involves several classes with very similar spectral features, extracting discriminant features to maximize class separability is critical. An LRF is adopted because its capability of efficiently extracting spectral-spatial features has been demonstrated in our previous work; in addition, such a low-pass filter can smooth out noise and trivial spectral variations. It is simply realized by spatial convolution with a sliding window, and then the central pixel is replaced with the filtering output, which can be expressed as

$$G'(s,t) = G \otimes W = \sum_{c=-\frac{m-1}{2}}^{\frac{m-1}{2}} \sum_{l=-\frac{m-1}{2}}^{\frac{m-1}{2}} G(s+c,t+l)W(c,l) \tag{1}$$

where G represents the original image of size $m' \times n' \times d$ ($m' \times n'$ is the spatial size and d is the number of spectral bands), G' represents the generated image after spatial filtering, and W is the filter with a $m \times m$ window (m is usually an odd number) with all the weights be positive and sum-to-one.

An LAF uses the same weights for all the neighboring pixels in the window. Obviously, pixels within a local neighborhood have heterogeneity or may belong to different classes. Thus, in AWF, the element $w_{i,j}$ is large only when an adjacent pixel is similar to the central pixel. The popular Gaussian kernel function is utilized to describe the divergence between two vectors, and the relationship $L_{i,j}$ between $\mathbf{x}_i$ and $\mathbf{x}_j$ is computed as

$$L_{i,j} = \exp(\frac{-\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2}{\sigma}) \tag{2}$$

where $\mathbf{x}_j$ represents the neighbor pixels around $\mathbf{x}_i$, and σ is a parameter of the Gaussian kernel function, which is set according to the following equation

$$\sigma = median(\left\| x_t - \bar{x} \right\|^2) \tag{3}$$

where $\bar{x}$ is the mean of all training samples $x_t, t = 1,\ldots\ldots,$ m*m [26]. The weights is normalized as

$$w_{i,j} = \frac{L_{i,j}}{\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{m} L_{i,j}} \tag{4}$$

The weights in AWF are obtained according to the similarity distances measured between the central pixel and its surrounding pixels, thus the surrounding samples that are dissimilar to the central sample to be classified provide less contribution to the convolution process, which are provided with smaller weights. Thus, it is robust to less homogeneous image scenes. However, the AWF is computationally expensive due to its data-dependent nature.

In the GLF, the equation similar to Eq. (2) is used for weight assignment, but the distance between two pixels is based on their spatial coordinates. Obviously, the central pixel itself is assigned the largest weight, and the four nearest neighbors the next, with enlargement of the distance the weight value follows the trends of a Gaussian curve. Such weight assignment results to a fixed filter, and is data-independent as in the LAF, resulting in low computational cost. Compared to the LAF, the GLF still provides smoothing effect but image details are less blurred.

2.2. Local Fisher's Discriminate Analysis

After spatial-spectral features being generated, LFDA [9, 11] is adopted to reduce feature dimensions and enhance class separability. LFDA is a supervised dimensionality-reduction technique which is designed to handle multimodal, non-Gaussian distributions. In essence, LFDA combines the properties of linear discriminate analysis (LDA) [27] and local preserving projection (LPP) [28]. LFDA preserves neighborhood relationships in the embedding by employing an "affinity" matrix $\mathbf{A}_{i,j}$, and define the "local" between-class $\mathbf{S}_b$ and within-class $\mathbf{S}_w$ scatter matrices as

$$\mathbf{S}_b = \frac{1}{2}\sum_{i,j}\mathbf{V}_{i,j}^{(b)}\left(s_i - s_j\right)\left(s_i - s_j\right)^T \tag{5}$$

$$S_w = \frac{1}{2}\sum_{i,j} V_{i,j}^{(w)}\left(s_i - s_j\right)\left(s_i - s_j\right)^T \tag{6}$$

where s_i and s_j are two training samples, and $\mathbf{V}^{(b)}$ and $\mathbf{V}^{(w)}$ are matrices defined as

$$V_{i,j}^{(b)} = \begin{cases} A_{i,j}(1/n - 1/n_l), & if \ y_i = y_j = l \\ 1/n, & if \ y_i \neq y_j \end{cases} \tag{7}$$

$$V_{i,j}^{(w)} = \begin{cases} A_{i,j}/n_l, & if \ y_i = y_j = l \\ 0, & if \ y_i \neq y_j \end{cases} \tag{8}$$

where y_i and y_j are the class labels for the two training samples class labels, and n_l is the number of training samples available for the l-th class.

Similar to LDA, LFDA seeks to find a linear transformation such that the within-class scatter is minimized and the between-class scatter is maximized. The LFDA solution is obtained by maximizing the Rayleigh quotient, and the solution can be simply obtained by solving a generalized eigen-decomposition problem [29]. It is expected that a less powerful classifier, such as KNN, can offer satisfactory performance on the LFDA-transformed data.

2.3. K-Nearest Neighbor Classifier

The typical KNN classifier follows the LFDA process to determine the final class label due to its simplicity. It attempts to find out which training sample is nearest to the testing sample according to a given distance measure, and then assign the class label of the training sample to the testing sample. Euclidean distance is commonly used to measure the similarity between a training sample and a testing sample.

III. EXPERIMENTAL ANALYSIS

The proposed methods are tested with two real datasets containing the tillage information to compare the classification performance.

3.1. Hyperspectral Data

The first HSI dataset was acquired using National Aeronautics and Space Administration (NASA)'s AVIRIS sensor and was collected over northwest Indiana's Indian Pines test site in June 1992. This scene represents a vegetation-classification scenario with 145×145 pixels in the 0.4- to 2.45 μm region of the visible and infrared spectrum with a spatial resolution of 20m. The scene contains two-thirds

agriculture, and one-third forest. In this research, a total number of 200 bands are used after removal of water-absorption bands.

The second dataset was also collected by the AVIRIS sensor, capturing an area over Salinas Valley, California, with a spatial resolution of 3.7m. The clipped image comprises of 236 × 217 pixels.



(a)   (b)
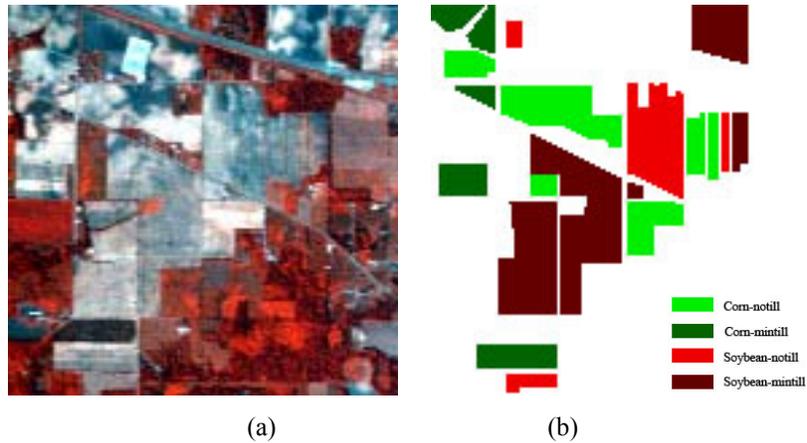
Fig. 1 Indian pines dataset: (a) false-color image generated by using bands 36, 15 and 10; (b) ground truth for the tillage-related classes.

TABLE I

PER-CLASS SAMPLES FOR THE INDIAN PINES DATASET.

No.	Name	Number of samples
1	Alfalfa	46
2	**Corn-notill**	**1428**
3	**Corn-mintill**	**830**
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	**Soybean-notill**	**972**
11	**Soybean-mintill**	**2455**
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Building-grass-trees-drives	386
16	Stone-steel-towers	93
Total		10249
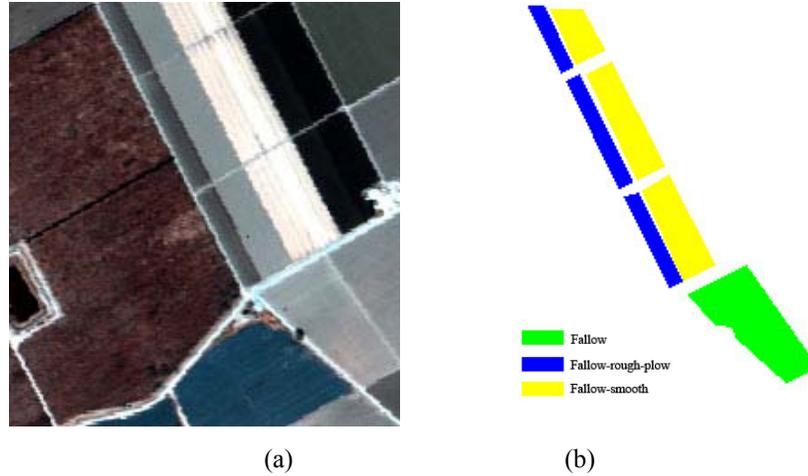
Fig. 2 Salinas dataset: (a) false-color image generated by using bands 36, 15 and 10; (b) ground truth for the tillage-related classes.

TABLE II

PER-CLASS SAMPLES FOR THE SALINAS DATASET.

No.	Name	Number of samples
— Class		
1	Brocoli-green-weeds-1	2009
2	Brocoli-green-weeds-2	3726
3	**Fallow**	**1976**
4	**Fallow-rough-plow**	**1394**
5	**Fallow-smooth**	**2678**
6	Stubble	3959
7	Celery	3579
8	Grapes-untrained	11271
9	Soil-vineyard-develop	6203
10	Corn-senesced-green-weeds	3278
11	Lettuce-romaine-4wk	1068
12	Lettuce-romaine-5wk	1927
13	Lettuce-romaine-6wk	916
14	Lettuce-romaine-7wk	1070
15	Vineyard-untrained	7268
16	Vinyard-vertical-trellis	1807
Total		54129

The two images and the chosen samples are shown in Figs. 1 and 2, respectively. The class description and the sample numbers for these two datasets are shown in Tables I and II, respectively.

Note that these two images contain many classes, but only the tillage-related classes are concerned in this research, which are highlighted in Tables I and II. In the experimental setting, 10 training samples for each class are randomly selected from the training sets, and all the remaining samples are used to test the classification accuracy. To avoid the bias, the experiments are repeated 20 times and all the classification accuracies are averaged.

3.2. Parameter Tuning

First of all, the classification performance of the LRF filters with various window sizes m is discussed. The window size m is investigated with a step interval of 2, starting from 1, which is actually with no prior filtering process. Fig. 3 illustrates the changing trends of the overall classification accuracy with different parameter settings with 10 trials on the Indian Pines dataset over the class of *corn-notill* and *corn-mintill*.

(a) LAF-LFDA-KNN

(b) GLF-LFDA-KNN

(c) AWF-LFDA-KNN
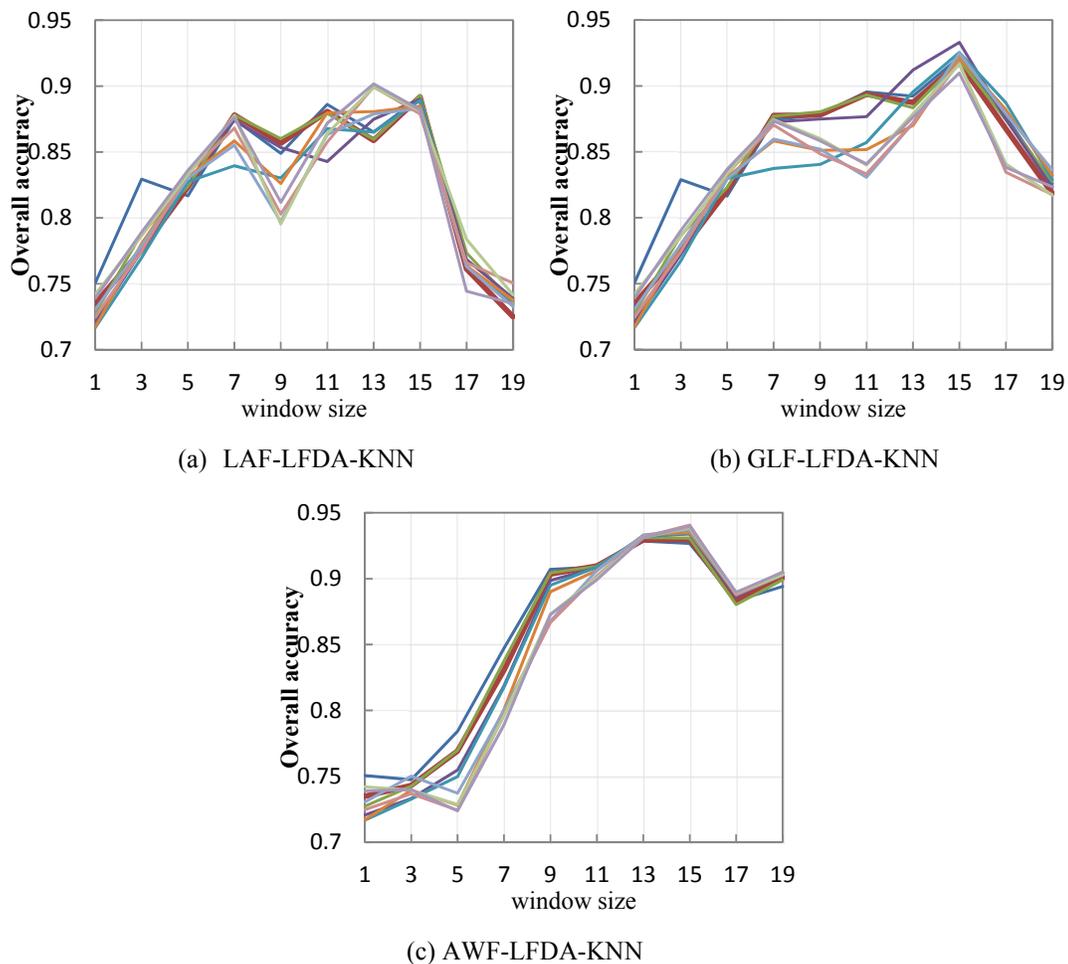
Fig. 3. Overall classification accuracy in the Indian Pines experiment with various window sizes (different curves indicate different trials).

The change of the overall accuracy with enlargement of window sizes reveals the characteristics of the LRF, which is sensitive to kernel size. Rapid falls can be seen for LAF-LFDA-KNN with window sizes exceeding the optimum value and the curve trends for repeated trials vary. The overall classification accuracy may even be lower than the original value, indicating possible distortion to the central pixel information. While for GLF-LFDA-KNN, the stability is better, and the accuracy at the optimum window size is higher. For AWF-LFDA-KNN, the curves are more stable, and the trends for repeated trials are consistent, implying more steady performance. This makes sense because AWF adaptively assigns weights according to the spectral similarity to the central pixel. From Fig. 3, the optimum window size for classifying *corn-notill* and *corn-mintill* is 15.

3.3. Classification Performance

The overall classification accuracies for the two datasets with different methods are listed in Table III. As can be seen from this table, the results of LAF-KNN, GLF-KNN and AWF-KNN are much better than KNN and SVM, and the result of AWF-KNN is comparable to that of SVM-CK, revealing the distinguishing performance of the LRF filtering process. In addition, the overall accuracies for LAF-LFDA-KNN, GLF-LFDA-KNN and AWF-LFDA-KNN are much better than other methods such as KNN, LAF-KNN, GLF-KNN, AWF-KNN, SVM and SVM-CK, implying the important role LFDA plays. Moreover, the results with AWF-LFDA-KNN are consistently higher than LAF-LFDA-KNN by 1-3%, and GLF-LFDA-KNN follows with an improvement of 1-2%, showing the improved performance for adaptive combination of the local region information.

TABLE III

OVERALL CLASSIFICATION ACCURACY FOR DIFFERENT DATASETS.

	Indian pines-Corn	Indian pines-Soybean	Salinas
KNN	60.15%	66.65%	92.31%
SVM	65.69%	69.23%	93.62%
SVM-CK	78.00%	76.68%	95.10%
LAF-KNN	70.76%	70.85%	93.81%
GLF-KNN	72.51%	73.42%	94.52%
AWF-KNN	77.26%	78.06%	96.26%
LAF-LFDA-KNN	89.46%	90.09%	99.10%
GLF-LFDA-KNN	91.20%	92.12%	99.41%
AWF-LFDA-KNN	92.02%	93.87%	99.60%

The classification maps are shown in Figs. 4-6, which can be compared with the ground truth maps in Figs. 1-2. From the results, LAF-LFDA-KNN, GLF-LFDA-KNN and AWF-LFDA-KNN, with the local region filtering and dimension reduction process, yield evident improvements in producing smoother classification maps. With more adaptive combination of the information from the local region, AWF-LFDA-KNN provides the best performance, followed by GLF-LFDA-KNN.


(a) KNN (b) SVM (c) SVM-CK

(d) LAF-KNN (e) GLF-KNN (f) AWF-KNN

(g) LAF-LFDA-KNN (h) GLF-LFDA-KNN (i) AWF-LFDA-KNN
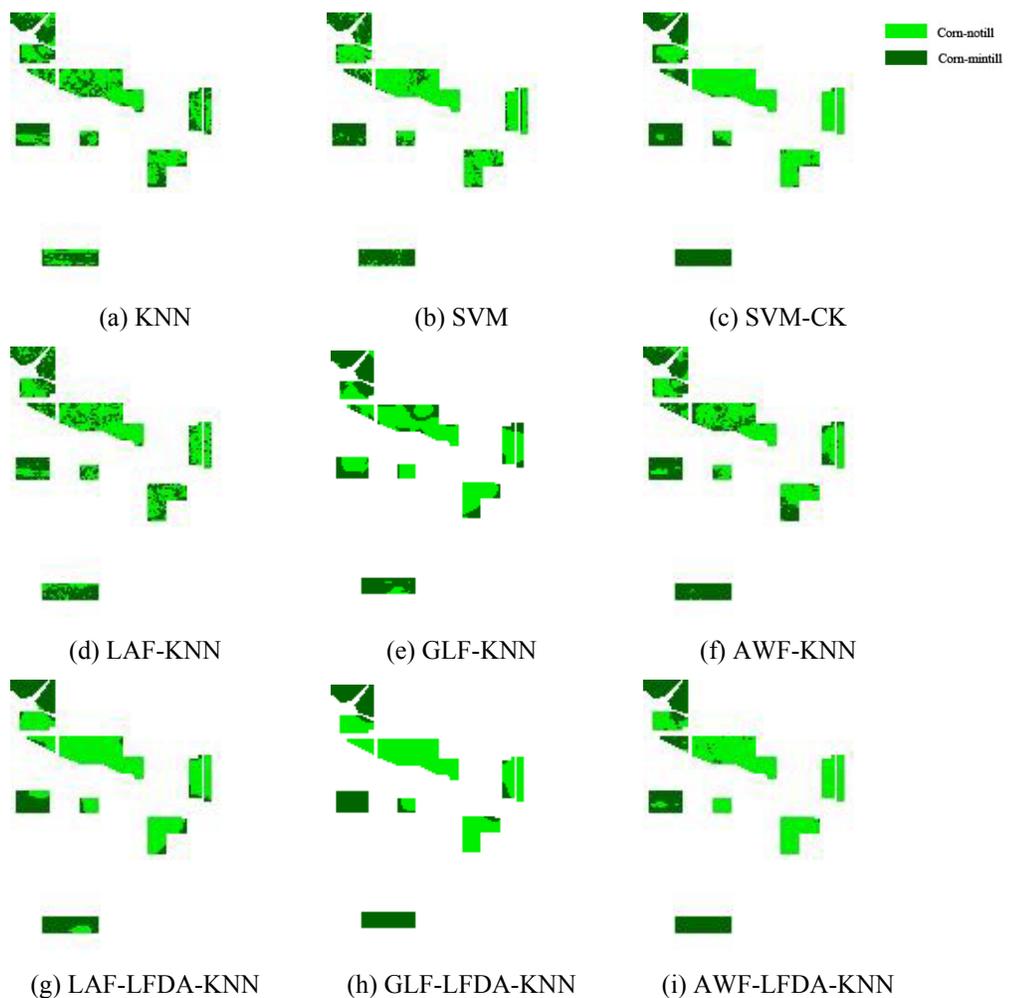
Fig. 4. Classification map with Indian pines dataset for corn tillage conditions.

(a) KNN (b) SVM (c) SVM-CK

(d) LAF-KNN   (e) GLF-KNN   (f) AWF-KNN

(g) LAF-LFDA-KNN   (h) GLF-LFDA-KNN   (i) AWF-LFDA-KNN
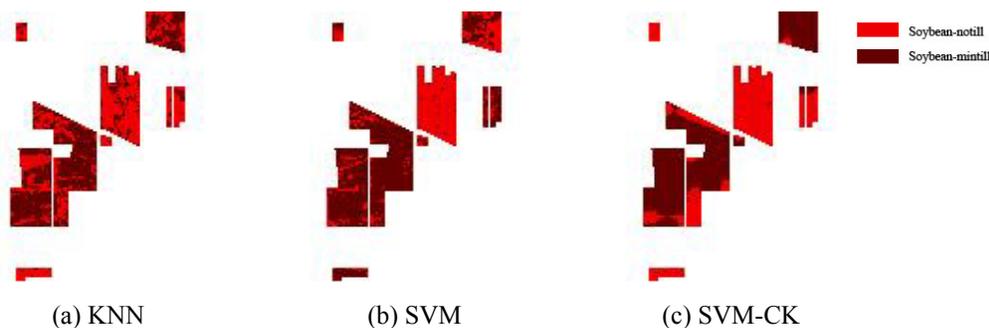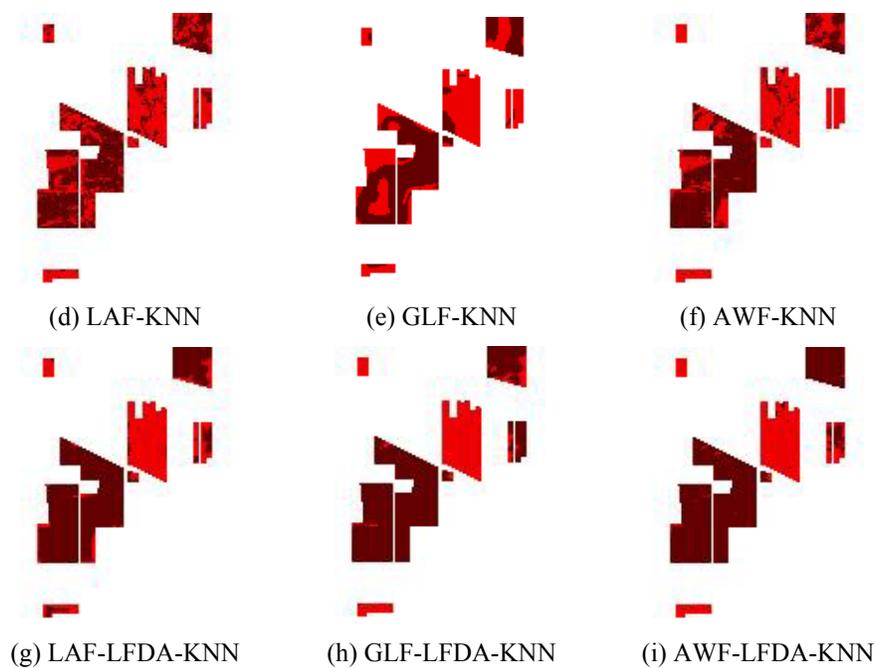
Fig. 5. Classification map with Indian pines dataset for soybean tillage conditions.

■ Fallow
■ Fallow-rough-plow
■ Fallow-smooth

(a) KNN   (b) SVM   (c) SVM-CK

(d) LAF-KNN   (e) GLF-KNN   (f) AWF-KNN

(g) LAF-LFDA-KNN (h) GLF-LFDA-KNN (i) AWF-LFDA-KNN
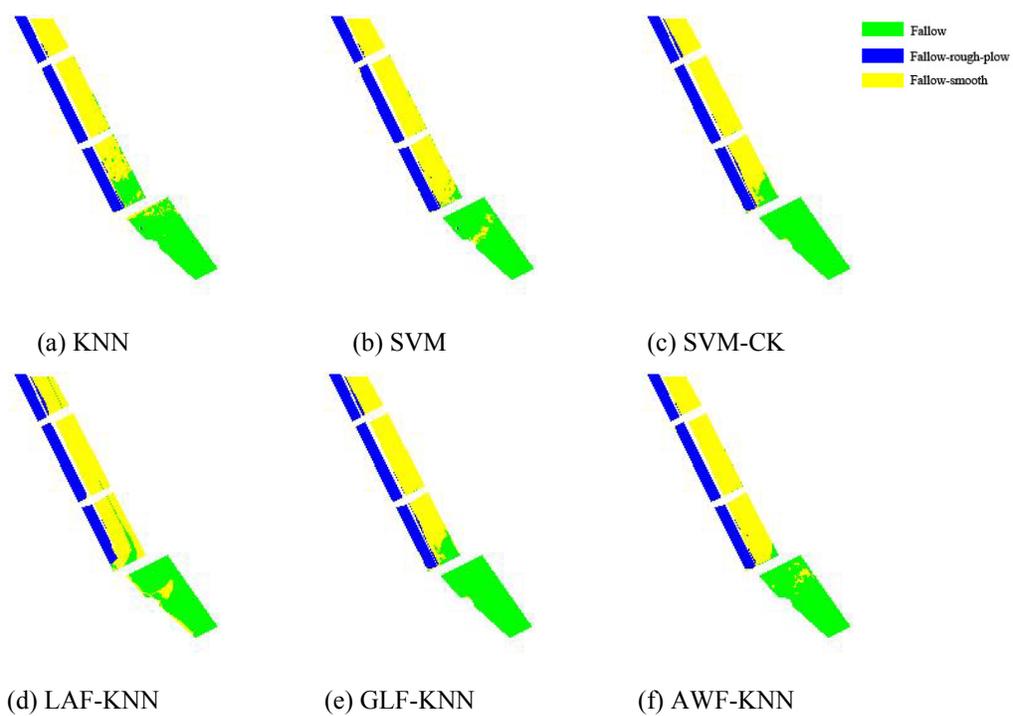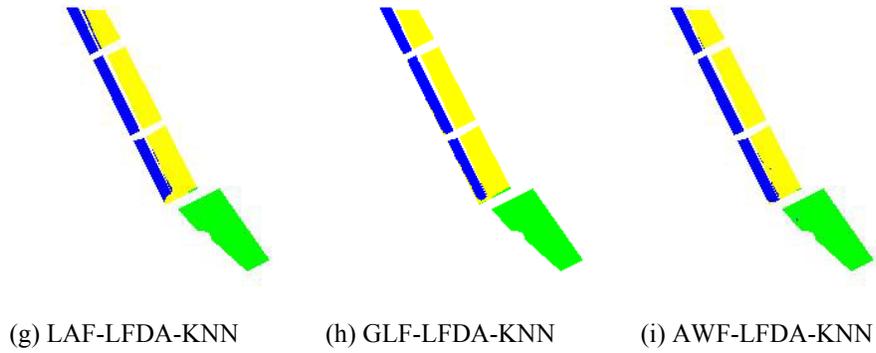
Fig. 6. Classification map with Salinas dataset for different tillage conditions.

TABLE IV

COMPUTATIONAL COST FOR DIFFERENT METHODS AND DATASETS.

	Indian pines-Corn	Indian pines-Soybean	Salinas
KNN	0.26s	0.38s	0.68s
SVM	0.43s	0.54s	0.73s
SVM-CK	1.1s	1.7s	3.37s
LAF-KNN	0.67s	0.71s	1.33s
GLF-KNN	0.75s	0.93s	1.54s
AWF-KNN	1.04s	1.27s	2.99s
LAF-LFDA-KNN	0.68s	0.76s	1.39s
GLF-LFDA-KNN	0.79s	1.14s	1.86s
AWF-LFDA-KNN	1.12s	1.54s	3.11s

The computational cost is listed in Table IV. All the experiments are carried out on a computer with an Intel i7-4770 3.4GHZ CPU, 4GB of RAM and 64bit operating system. The two images to be processed are with 145×145 pixels and 236×217 pixels, and the number of samples for the datasets of Indian pines-Corn, Indian pines-Soybean and Salinas are 2238, 3407, 6018, respectively. Note that all classifiers related with KNN are implemented purely in MATLAB, and SVM is implemented in the popular *libsvm* package which uses MEX function to call C program in MATLAB. As is shown in Table III and IV, KNN can have competitive classification performance as SVM while with lower computation costs. When combining with some spatial filters, such as LAF, GLF and AWF, the precision is even greatly improved. It is worth mentioning that GLF-LFDA-KNN requires much less computational cost as compared to AWF-LFDA-KNN and SVM-CK, while providing comparable classification accuracy to AWF-LFDA-KNN and much better performance than SVM-CK as shown in

Table III. Thus, to be concluded, the proposed GLF-LFDA-KNN is promising for application in data mining of the big data, besides the fact that it has better classification performance with acceptable computational cost.

IV. CONCLUSION

In this paper, we proposed an efficient spatial-spectral classification framework for mapping agricultural tillage practices using hyperspectral remote sensing imagery. Considering the large scale of a remote sensing image data, we advocate the cost-effective GLF for spatial-spectral feature extraction and the simple KNN for classification. To improve the classification power for KNN, the LFDA is applied to improve class separability. The final GLF-LFDA-KNN can offer comparable classification accuracy as AWF-LFDA-KNN with much lower computational cost. Compared to SVM-CK using spatial-spectral features as well, GLF-LFDA-KNN outperforms in terms of both classification accuracy and computational efficiency. Since the LRF filtering process for each pixel is independent, it is potential to further reduce the computation burden with parallel computing when dealing with a large remote sensing scene.

References

[1]   A. Bannari, K. Staenz, and K. S. Khurshid, "Remote sensing of crop residue using hyperion (EO-1) data," in *Proceedings of the International Geoscience and Remote Sensing Symposium*, Barcelona, Spain, July, 2795–2799 (2007).

[2]   E. Choe, S. Hong, and Y. Kim, "Down-scaling of satellite hyperspectral images for monitoring croplands," in *Proceedings of the International Geoscience and Remote Sensing Symposium*, Honolulu, Hawaii, July, 3873–3874 (2010).

[3]   B. Zheng, J. B. Campbell, G. Serbin, J. M. Galbraith, "Remote sensing of crop residue and tillage practice: Present capabilities and future prospects," Soil and Tillage Research, 138, 26-34(2014).

[4]   S. South, J. Qi, D. P. Lusch, "Optimal classification methods for mapping agricultural tillage practices," Remote Sensing of Environment, 91(1), 90-97(2004).

[5] B. Zheng, J. B. Campbell, K. M. Beurs, "Remote sensing of crop residue cover using multi-temporal Landsat imagery," Remote Sensing of Environment, 117, 177-183(2012).

[6] R. L. L. J. D. Watts, S. L. Powell and T. Hilker, "Improved classification of conservation tillage adoption using high temporal and synthetic satellite imagery," *Remote Sensing of Environment*, 115, 66–75 (2011).

[7] C. S. T Daughtry, P. C. Doraiswamy, E. R. Hunt Jr., A. J. Stern, J. E. McMurtrey III, J. H. Prueger, "Remote sensing of crop residue cover and soil tillage intensity," Soil and Tillage Research, 91(1-2), 101-34(2014).

[8] Q. Du, "Modified Fisher's linear discriminant analysis for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, 4(4), 503-507 (2007).

[9] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving dimensionality reduction and classification for hyperspectral image analysis," *IEEE Trans. Geosci. Remote Sens.*, 50(4), 1185–1198 (2012).

[10] W. Li, E.W.Tramel, and S.Prasad, "Nearest regularized subspace for hyperspectral classification," *IEEE Trans. on Geosci. and Remote Sens.*, 52, 477-489 (2014).

[11] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving discriminant analysis in kernel-induced feature spaces for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, 8(5), 894-898 (2011).

[12] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson, "Spectraland spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, 46(11), 3804–3814 (2008).

[13] 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.*, 43(3), 480-491 (2005).

[14] 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.*, 3(1), 93–97 (2006).

[15] Y. Qian, M. Ye, and J. Zhou, "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features," *IEEE Trans. Geosci. Remote Sens.*, 51, 2276-2291 (2013).

[16] X. Zhang, Q. Song, R. Liu, W. Wang, and L. Jiao, "Modified co-training with spectral and spatial views for semisupervised hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, 7(6), 2044-2055 (2014).

[17] L. Shen and S. Jia, "Three-dimensional Gabor wavelets for pixel-based hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, 49(12), 5039-5046 (2011).

[18] F. Tsai and Jhe-Syuan Lai, "Feature extraction of hyperspectral image cubes using three dimensional gray level co-occurrence," *IEEE Trans. Geosci. Remote Sens.*, 51, 3504-3513 (2013).

[19] W. Li, S. Prasad and J. E. Fowler, "Hyperspectral image classification using Gaussian mixture models and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, 11(1), 153-157 (2014).

[20] Y. Tarabalka, M. Fauvel and J. Chanussot, "SVM- and MRF-based method for accurate classification of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, 7(4), 736 − 740 (2010).

[21] B. Zhang, Sh. Li and X. Jia, "Adaptive Markov random field approach for classification of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, 8(5), 973-977 (2011).

[22] P. Zhong, R. Wang, "Modeling and classifying hyperspectral imagery by CRFs with sparse higher order potentials," *IEEE Trans. Geosci. Remote Sens.*, 49(2), 688-705 (2011).

[23] W. Li and Q. Du, "An efficient spatial-spectral classification method for hyperspectral classification method for hyperspectral imagery," Proc. SPIE, Beijing, China, DOI: 9124/912410 2014.

[24] Q. Ran, W. Li, Q. Du, M. Xiong, "Hyperspectral image classification with improved local-region filter," *Journal of Applied Remote Sensing*, 8(1), article ID: 085088, (2014).

[25] W. Li, Q. Ran, Q. Du, C. Yang, "Improved classification of conservation tillage practices using hyperspectral imagery with spatial-spectral features," in *Agro-geoinformatic (Agro-geoinformatics 2014),Third International Conference on*, Beijing, China, DOI: 10.1109/6910589 (2014).

[26] L. Zhang, W. D. Zhou, P. C. Chang, J. Liu, Z. Yan, T. Wang, F. Z. Li, "Kernel sparse representation-based classifier," IEEE Trans. on Signal Processing, 60 (4) (2012).

[27] S. Mika, G. Ratsch, J. Weston, B, Scholkopf, and K.-R. Muller, "Fisher discriminant analysis with kernels," in *Proc. IEEE Neural Netw. Signal Process. Workshop*, 41-48 (1999).

[28] M. Sugiyama, "Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis," *J. of Mach. Learn. Res.*, 8, 1027-1061 (2007).

[29] R. O. Duda, P. E. Hart, and D. G. Stocrk, *Pattern Classification*. John-Wiley and Sons, New York, NY (2001).

Qiong Ran received her Ph.D. degrees from the Institute of Remote Sensing Applications, Chinese Academy of Sciences (CAS), Beijing, China, in 2009. She has published over 10 papers in China and abroad. Her research interests include image acquisition, image processing, hyperspectral image analysis and applications.

Wei Li received his Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, in 2012. Subsequently, he spent one year as a postdoctoral researcher at the University of California, Davis. He is currently with the College of Information Science and Technology at Beijing University of Chemical Technology, Beijing, China. His research interests include statistical pattern recognition, hyperspectral image analysis, and data compression.

Qian Du received her Ph.D. degree in electrical engineering from University of Maryland Baltimore County in 2000. Currently, she is Bobby Shackouls Professor in the Department of Electrical and Computer Engineering at Mississippi State University. Her research interests include hyperspectral remote sensing image analysis, pattern classification, data compression, and neural networks. Dr. Du is an associate editor for Journal of Applied Remote Sensing. She is a Fellow of SPIE.

Chenghai Yang is an agricultural engineer with the U.S. Department of Agriculture-Agricultural Research Service's Aerial Application Technology Research Unit at College Station, Texas. He received his Ph.D. degree in agricultural engineering from the University of Idaho in 1994. His current research is focused on the development and evaluation of remote sensing technologies for monitoring crop growth and pest conditions and assessing the performance of site-specific chemical applications.

List of Figures:

List of Tables:

	Corn-notill
	Corn-mintill
	Soybean-notill
	Soybean-mintill

(a)   (b)

(a)

(b)

(a) LAF-LFDA-KNN

(b) GLF-LFDA-KNN

(c) AWF-LFDA-KNN

Corn-notill
Corn-mintill

(a) KNN   (b) SVM   (c) SVM-CK

(d) LAF-KNN   (e) GLF-KNN   (f) AWF-KNN

(g) LAF-LFDA-KNN   (h) GLF-LFDA-KNN   (i) AWF-LFDA-KNN

Soybean-notill
Soybean-mintill

(a) KNN

(b) SVM

(c) SVM-CK

(d) LAF-KNN

(e) GLF-KNN

(f) AWF-KNN

(g) LAF-LFDA-KNN

(h) GLF-LFDA-KNN

(i) AWF-LFDA-KNN

	Fallow
	Fallow-rough-plow
	Fallow-smooth

(a) KNN

(b) SVM

(c) SVM-CK

(d) LAF-KNN

(e) GLF-KNN

(f) AWF-KNN

(g) LAF-LFDA-KNN

(h) GLF-LFDA-KNN

(i) AWF-LFDA-KNN