Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Dynamic classifier selection using spectral-spatial information for hyperspectral image classification

Hongjun Su
Bin Yong
Peijun Du
Hao Liu
Chen Chen
Kui Liu

SPIE.

Dynamic classifier selection using spectral-spatial information for hyperspectral image classification

Hongjun Su,[a,b,*] Bin Yong,[a,*] Peijun Du,[b,*] Hao Liu,[c] Chen Chen,[d] and Kui Liu[d]

[a]Hohai University, State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing 210098, China

[b]Nanjing University, Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing 210046, China

[c]Wuhan University, State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan 430079, China

[d]University of Texas at Dallas, Department of Electrical Engineering, Richardson, Texas 75080-3021, United States

Abstract. This paper presents a dynamic classifier selection approach for hyperspectral image classification, in which both spatial and spectral information are used to determine a pixel's label once the remaining classified pixels' neighborhood meets the threshold. For volumetric texture feature extraction, a volumetric gray level co-occurrence matrix is used; for spectral feature extraction, a minimum estimated abundance covariance-based band selection is used. Two hyperspectral remote sensing datasets, HYDICE Washington DC Mall and AVIRIS Indian Pines, are employed to evaluate the performance of the developed method. The classification accuracies of the two datasets are improved by 1.13% and 4.47%, respectively, compared with the traditional algorithms using spectral information. The experimental results demonstrate that the integration of spectral information with volumetric textural features can improve the classification performance for hyperspectral images. © *2014 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JRS.8.085095]

Keywords: dynamic classifier selection; volumetric textural feature; spectral feature; hyperspectral image classification.

Paper 14137SS received Mar. 9, 2014; revised manuscript received Jul. 9, 2014; accepted for publication Jul. 15, 2014; published online Aug. 22, 2014.

1 Introduction

Hyperspectral remote sensing plays an important role in land use/cover classification and mapping. Spectral information has been widely utilized in hyperspectral image classification. However, related works that focused on geometric and textural features are still limited.[1,2] Many approaches have shown that the textural features combined with spectral information have great potential to improve the classification accuracy.[3] Texture analysis methods include structural methods,[4] statistical methods,[5] model-based methods,[6] and transform-based methods.[6] Statistical methods consist of a histogram, gray-level co-occurrence matrix (GLCM), and gray-level run-length. Although GLCM is a very popular technique, it is worth noting that this approach is only applied to single-band images.[5,7,8] Tsai et al.[9] extended the GLCM model to a three-dimensional (3-D) space, and proposed a volume GLCM (VGLCM) model to extract the texture features of hyperspectral data. The results obtained indicated that the VGLCM model outperformed the GLCM model, and the VGLCM model captures the relationship between neighboring spectral bands. Moreover, it has been verified that combining volumetric texture features computed by the VGLCM algorithm can improve the classification accuracy for hyperspectral image analysis.

*Address all correspondence to: Hongjun Su, E-mail: hjsurs@163.com; Bin Yong, E-mail: yongbin_hhu@126.com; Peijun Du, E-mail: dupjrs@gmail.com

Multiple classifier systems (MCSs), which employ different feature descriptors and classifiers to gain a more robust, reliable, and efficient recognition performance, have been verified to be an effective solution for improving classification performance.[10,11] Over the past few decades, MCS and related ensemble methods for classifiers have been developed based on machine learning, neural networks, pattern recognition,[12–14] and so on. Commonly, two strategies, i.e., fusion- and selection-based methods, are used to merge an ensemble of classifiers (EoC). Fusion-based methods apply individual classifiers in parallel and combine the outputs to achieve consensus.[15] To guarantee the improvement in classification, the individual error of each classifier is required to form an ensemble system,[16] which is difficult to achieve. In contrast, selection-based methods directly choose the classifier with the best performance from an EoC for a given pattern.[17] Moreover, selection-based methods tend to have superior performances over fusion-based methods under the condition that one classifier in an EoC strongly dominates others.[18] Selection-based methods include static classifier selection (SCS) and dynamic classifier selection (DCS). The difference between SCS and DCS is that the best classifier selected by SCS methods is for all test patterns, while DCS methods only choose the classifier that best suits the current pattern[19] to ensure a reasonable classification result for each pattern.

Previous research aimed at utilizing the spectral information of hyperspectral images for classifier combination and classification.[20,21] Recently, more focus has been placed on spatial information.[22–24] However, spatial and spectral information are treated as data sources of the classifier in most DCS-related works. In practice, pixels are spatially related. In other words, it is highly probable that two adjacent pixels belong to the same class. Therefore, spectral information with the support of spatial information can significantly improve the hyperspectral image classification accuracy.[25,26] In this paper, we propose a novel dynamic classifier ensemble method by combining spectral information and volumetric textural features. In the proposed algorithm, spatial and spectral information are used to determine the label with the condition that the classified pixels' percentage of the unlabeled pixels' neighborhood meets a specific threshold. Meanwhile, the minimum estimated abundance covariance (MEAC)-based band selection and the VGLCM model are used to extract the spatial and spectral features for dynamic classifier ensembles.

In the proposed algorithm, spatial and spectral information are not only used as features in the preprocessing step, but also used for assigning the label to a pattern. In this approach, the label information of the pixels, when an EoC agrees, can be also adopted as spatial information[22] When it comes to assigning labels to test patterns, the proposed DCS algorithm utilizes the spatial information to check whether the proportion of the labeled pixels around the pattern exceeds the threshold. If it does, spatial and spectral information around the pattern will be used to determine the label of the pattern; otherwise, spectral information is used. Furthermore, the VGLCM algorithm is applied to extract the spatial information from hyperspectral images. There are two differences between the proposed DCS algorithm and other DCS algorithms. (1) Preprocessing is required by the proposed algorithm. In this step, pixels on which all member classifiers agree are directly assigned using member classifiers' classification results. As more spatial information is provided, a more accurate prediction can be made. (2) The proposed DCS algorithm will accordingly choose different computing methods based on the statistics of the current pixel's neighborhood.

The rest of the paper is organized as follows. In the next section, spatial and spectral information extraction methods are described. Section 3 presents the proposed DCS algorithm. A presentation of the experimental results and analysis for two hyperspectral datasets follow in Sec. 4. Finally, summarizing remarks and conclusions are stated in Sec. 5.

2 Spatial and Spectral Information Extraction

2.1 *Spectral Feature Extraction*

2.1.1 *Principal component analysis*

Principal component analysis (PCA) is a decorrelation technique that is widely used in dimensionality reduction and data compression for hyperspectral images.[27] PCA maps a data vector

from the original space of q variables to a new space of q variables, which are uncorrelated over the dataset. However, not all the principal components have to be kept. Keeping only the first L principal components that are produced by using only the first L loading vectors, yields the truncated transformation

$$\mathbf{T}_L = \mathbf{X}\mathbf{W}_L, \tag{1}$$

where $\mathbf{X}$ is the original data matrix, $\mathbf{W}_L$ is the projection matrix, and the dimension reduced matrix $\mathbf{T}_L$ has n rows but only L columns. Matrix $\mathbf{T}_L$ maximizes the variance in the original data while minimizing the total squared reconstruction error $\|\mathbf{X} - \mathbf{T}_L\|^2{}_2$. The selected number of principle components (PCs) is smaller than the dimension of the original data, and can be treated as the dimension reduced spectral features of the hyperspectral remote sensing data.

2.1.2 *Minimum estimated abundance covariance-based band selection*

MEAC is a supervised band selection algorithm proposed by Yang et al.[28] Let us assume there are p classes present in an image. Based on the linear mixture model, a pixel $\mathbf{r}$ is considered as the mixing result of the endmembers of the p classes. Let the endmember matrix be $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_p]$. The pixel $\mathbf{r}$ can be expressed as

$$\mathbf{r} = \mathbf{S}\alpha + \mathbf{n}, \tag{2}$$

where $\alpha = (\alpha_1\alpha_2\alpha_p)^T$ is the non-negative abundance vector and $\mathbf{n}$ the uncorrelated white noise with $E(\mathbf{n}) = \mathbf{0}$ and $\mathrm{Cov}(\mathbf{n}) = \sigma^2\mathbf{I}$ ($\mathbf{I}$ is an identity matrix). Intuitively, the selected bands should let the deviation of α from the actual α be as small as possible. If only parts of the classes are known, this is equivalent to

$$\arg\min_{\Phi^S}\{\mathrm{trace}[(\mathbf{S}^T\Sigma^{-1}\mathbf{S})^{-1}]\}, \tag{3}$$

where Φ^S is the selected band subset and Σ is the noise covariance matrix. The resulting band selection algorithm is referred to as the MEAC method.

The MEAC algorithm does not require training samples. It only needs the class signatures. In addition, it is not necessary to examine the entire original bands or band combinations. With a forwarding search and the initial band pair selection, this method quickly completes band selection.

2.2 *Spatial Feature Extraction by Volumetric Gray Level Co-Occurrence Matrix*

The GLCM model has been a widely adopted algorithm for image texture analysis since it was proposed by Haralick in the 1970s.[7] The model is a measurement of the second-order statistical changes in gray level as well as a function describing the structure of the texture. It reflects the information of direction, neighboring spacing, and magnitude of changes in the gray level. It is a common method for analyzing the local patterns and pixel arrangements of images.

VGLCM is another commonly used texture extraction method proposed in Ref. 9, and the procedures for texture computation used by VGLCM and GLCM are different. The GLCM model uses a two-dimensional (2-D) moving window in 2-D space. However, the VGLCM applies a moving box in 3-D space to calculate the texture. For a hyperspectral image cube with $n = 4$ gray levels (6 bits), the co-occurrence matrix, M, is an n by n matrix. The matrix elements within a moving box, W, at a given displacement $d = (dx, dy, dz)$, are defined as[9]

$$M(i,j) = \sum_{x=1}^{Wx-dx}\sum_{y=1}^{Wy-dy}\sum_{z=1}^{Wz-dz}\begin{cases}1, & W(x,y,z) = i \quad \text{and} \quad W(x+dx, y+dy, z+dz) = j \\ 0, & \text{otherwise}\end{cases}, \tag{4}$$

where i and j are the values of pairwise pixels, and x, y, and z are denoted as the positions in the moving box. $M(i,j)$ is the value of a 3-D GLCM element, which reflects how often the gray levels of two pixels, $G(x,y,z)$ and $G(x+dx, y+dy, z+dz)$, are equal to i and j, respectively,

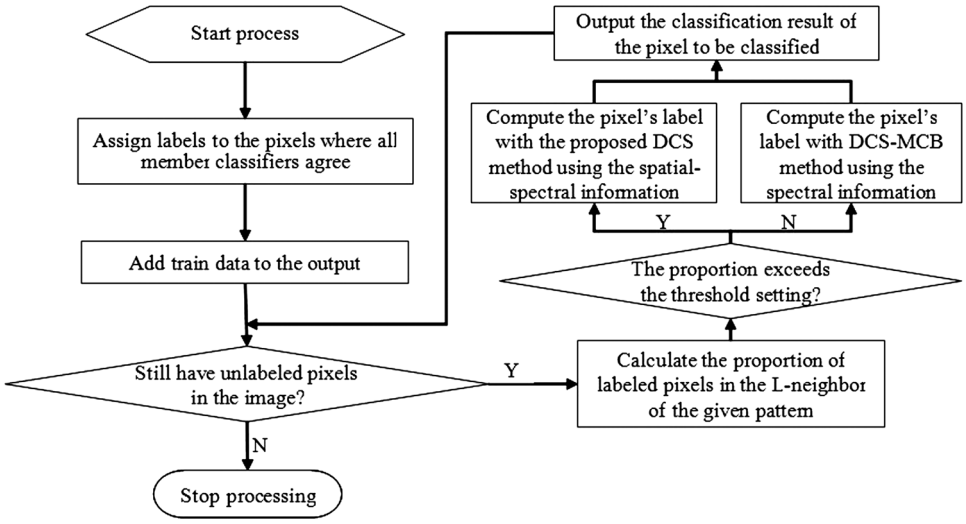


Fig. 1 Flow chart of the proposed DCS algorithm.

within a moving box.[9] In a hyperspectral data cube, d is usually set as one pixel in distance. For each pixel that has 26 neighboring pixels, there are 26 or 13 combinations (ignoring the opposite directions) in the horizontal and vertical directions.

The texture features extracted by GLCM largely depend on the quantization level, the size of the moving window, the distance and angle between pairwise pixels, and the statistical measures. The size of the moving window accounts for 90.4% of the classification variability, 7.1% is explained by the statistics used as texture measures, and only a small portion is accounted for by the quantization level as well as the distance and the angle between pixel pairs.[29]

For the parameters of the VGLCM algorithm, the semivariogram function, an efficient tool for analyzing the spatial heterogeneity, is used for window size analysis. Since not all the statistical measures are suitable for describing the texture features,[7] we choose variance, contrast, dissimilarity, energy, entropy, and homogeneity as measures with which to extract the texture features.

2.3 *Spatial Pixel Neighborhood Information*

Using spatial information to compute the label of the current pixel can be explained with spatial correlation-based knowledge as follows: pixels tend to be more similar if they are closer in space. In classification applications, if most neighborhoods of the current pixel belong to a class, it is highly probable that the current pixel also belongs to the same class.[30] However, a prediction made based on the spatial correlation using only spatial information is not as reliable as that made using spectral information since spectral information-based methods compute the label of the current pixel by means of the spectral similarity, which has an exact physical meaning. When only a few labeled pixels are available in the current pixel's neighborhood, adding spatial information will not obviously improve the classification performance. In contrast, sometimes it has a negative impact. However, as the proportion of neighboring labeled pixels increases, spatial information will provide more reliable support in determining the labels.

3 Proposed Dynamic Classifier Selection Algorithm

3.1 *Dynamic Classifier Selection*

The DCS methodology in the proposed algorithm is dynamically used to select a classifier from an EoC that best suits the current pattern, aimed at utilizing the strengths of each individual classifier while avoiding their weaknesses.[17] Giacinto first proposed the idea of DCS to avoid the assumption of classifier independence. Woods developed the idea of DCS using a local accuracy estimation,[15] and a series of modifications were made.[20,21] A theoretical

framework for DCS was described in Ref. 31, and a new DCS algorithm based on priori selection and posteriori selection methods was proposed. Some other works have considered combining ensemble methods with DCS methods.[19,20] The diversity of DCS and its influence were explored in Ref. 32. In this work, our study focuses on improving the DCS method by using the spectral and spatial information together.

3.2 *Proposed Dynamic Classifier Selection with Spectral-Spatial Information*

Above-mentioned DCS algorithms only use spectral information to select the best classifier while ignoring spatial information. Specifically, they have the following disadvantages:

1. DCS algorithms using only spectral information of hyperspectral images reach a bottleneck if no other new information is added. In order to improve the classification performance, spatial information especially of the neighboring pixels' information for hyperspectral images should be taken into consideration.
2. DCS algorithms using the spectral-spatial information have more support for the label assigned to the pattern, although comparison only with spectral information is considered.
3. Spectral information-based classification methods assign a label to the current pixel by comparing the spectral similarity of the current pixel with the training data. If neighborhood information of the current pixel is considered, it will abolish the spectral similarity compared to all sample patterns.

The proposed DCS algorithm utilizes the spectral information as well as spatial information to determine the current pattern. The neighborhoods around the pattern provide spatial information, but not all neighborhoods provide meaningful spatial information. Only the neighborhoods that are known or already classified provide support information for classification, thus a preprocessing step is required. In this step, labels are directly assigned to the pixels where all member classifiers from an EoC agree. The labels of the training pixels are already known. However, pixels used as training data might not keep the original labels after the training process due to training errors, which might reduce the final classification rate. Therefore, we can directly assign labels to the pixels used as training data. For any remaining classified pixels, calculate the proportion of the labeled pixels in an $L \times L$ window ($L = 5$ in this paper) around the unlabeled pixels. If the proportion exceeds the threshold (set as 70%),[33] this means the spatial information of the current pixel is strong enough to assign a label to the current pixel.

3.3 *Framework of Proposed Multiple Classifier System with Spectral-Spatial Information*

Many applications have proven that informative texture features are able to enhance the discriminative power for hyperspectral images' classification.[1–4] In our previous research, dimensionality reduction algorithms have been proposed,[28,34] and the classification accuracy significantly improved. Therefore, it is possible to improve the classification performance by combining the volumetric texture features with dimension reduced spectral features.

In the proposed fusion schemes as shown in Table 1, the volumetric texture features are extracted using the VGLCM algorithm, and the dimension reduced spectral features are obtained by using the MEAC-based band selection algorithm. All the features are used as inputs to a support vector machine (SVM) classifier. The detailed schemes are: Scheme II–texture features fused with all the bands of the original data; Scheme III–texture features fused with the selected bands, PCs of the original data after PCA compression; Scheme IV–texture features fused with selected bands from the original data.

In the proposed framework, four types of features from the designed schemes are used for the classification steps, namely the MCS system (Fig. 1). The process of computing the labels of the remaining unclassified pixels is divided into three steps. At first, a preprocessing step is required. In this step, pixels where all member classifiers agree are directly assigned labels according to the classification result of each member classifier. The computing method is chosen based on the spatial information provided by the current pattern. The proportion of labeled pixels in the current pixel's L-neighbors is obtained. If the proportion of already labeled pixels exceeds the

Table 1 Feature fusion schemes for hyperspectral image classification.

	Classification with spectral and texture
Scheme I	All bands
Scheme II	All bands +6 texture features
Scheme III	PCA +6 texture features + selected bands
Scheme IV	Selected bands +6 texture features

Input:

X – The hyperspectral image

Tr – Train data of the hyperspectral image

$<R_1,R_2,\ldots,R_n>$ - Classification results of each member classifier

Output:

Classification result of the hyperspectral image using SSI-Vote algorithm

Method:

1.Assign neighbor edge $l=L$

2.Assign labeled pixels proportion threshold $\delta=P$

3.**for** i = 1 to M

4. **for** j = 1 to N

5. **if** all member classifier agree on X_{ij}

6. assign $C_i(X_{ij})$ to X_{ij}

7. **end if**

8. **end for**

9.**end for**

10.Add train data Tr into X

11.**while** there still remains unlabeled X_{ij} Do

12. Analyze the proportion of labeled pixels of the unlabeled pixel's L-neighbor

13. **if** the proportion exceeds δ

14. Compute the label of X_{ij} with SSI-Vote algorithm using spatial and spectral information

15. **else**

16. Compute the label of X_{ij} with DCS-MCB algorithm using spectral information

17. **end if**

18. output the label of X_{ij}

19.**end while**

Fig. 2 Pseudocode of the proposed DCS algorithm.

threshold, compute the label of the current pattern with spatial information; otherwise, compute the label of the current pattern with the DCS-MCB algorithm using the spectral information until every pixel in the hyperspectral image is assigned a label. The pseudocode of the proposed DCS-SSI algorithm is shown in Fig. 2.

4 Experiment Results and Analysis

4.1 *Comparison Method and Classifier used in Multiple Classifier System*

For comparison purpose, the DCS-LCA method is used in the experiments. The basic idea of this scheme is to estimate each individual classifier's accuracy in local regions surrounding a test sample, and use the decision of the most locally accurate classifier. Local accuracy is estimated using output classes. More details can be found in Ref. 15.

The classifiers used in the DCS involve SVM with poly, Gaussian, RBF kernel, KNN, and diagquadratic classifiers. All these classifiers are implemented in MATLAB®.

4.2 *Experiment I: HYDICE Washington DC Mall*

The first experimental image is the HYDICE subimage scene with 304×301 pixels over the Washington DC Mall area as used in Ref. 35. After bad band removal, 191 bands are used in the experiments. There are six classes including roof, tree, grass, water, road, and trail in the dataset. Six class centers are used for band selection. Class center means the center of the class distribution, i.e., a mean value. The image in pseudocolor is shown in Fig. 3. As seen in the figure, the roof areas exhibit obvious different spectral signatures.

First, the box size for the VGLCM is analyzed, and $7 \times 7 \times 7$ is chosen as the best box size and with angles (135 and 135 deg). It is a direction in 3-D space. Figure 4 shows the texture features including variance, contrast, difference, energy, entropy, and homogeneity extracted using the VGLCM algorithm.

To evaluate the performance of the proposed method, the fused features from three schemes (i.e., 191 bands of the original data, the first five PCs after PCA compression, and five selected bands from the original data) are then used for classification with different classifiers, respectively. The classification results of Washington DC Mall are listed in Table 2. Specifically, for the case of six VGLCM features fused with all bands, the overall classification accuracy is improved and even reaches 96.08%. As shown in Table 2, the classification accuracies are improved after fusion of texture features and spectral information.

Figure 5 presents the classification maps generated from using different classification methods. The major misclassification occurs between trail (in yellow) and roof (in orange). It can be seen that with six texture features, the proposed method can slightly reduce the yellow (trail) areas that are supposed to be in orange for roof (highlighted in two circles). Tables 3–4 tabulate the confusion matrices for SVM and the proposed method, and illustrate the improvement in class separation, particularly for the roof class.

4.3 *Experiment II: AVIRIS Indian Pines*

The second experimental image is an AVIRIS subimage scene taken over northwest Indiana's Indian Pines with 145×145 pixels and 202 bands. 16 different land-cover classes are presented based on the ground truth.[36] In this experiment, nine classes are used. The experimental results

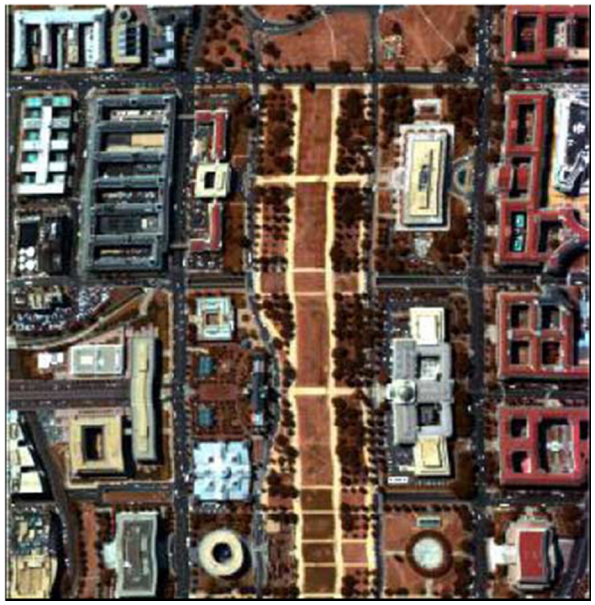


Fig. 3 Image used in HYDICE experiment.

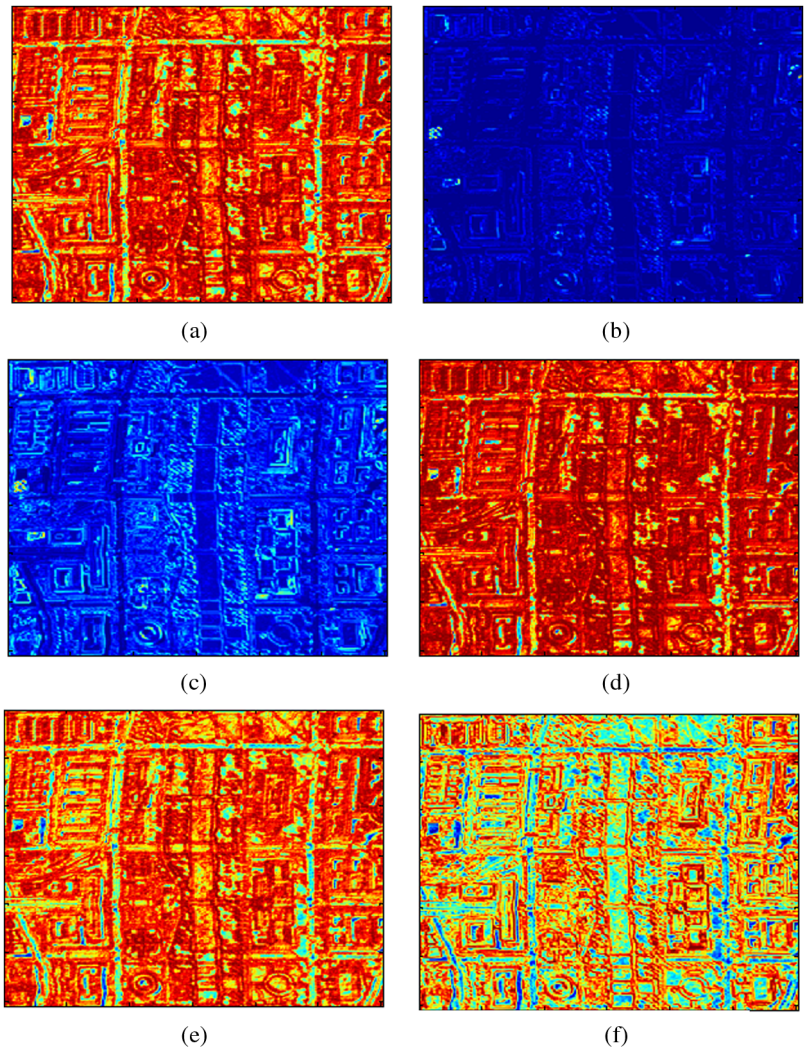
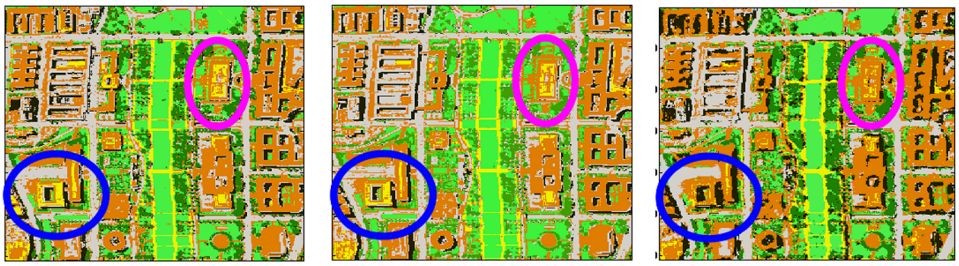


Fig. 4 Extracted textures using VGLCM. (a) Variance, (b) contrast, (c) dissimilarity, (d) energy, (e) entropy, (f) homogeneity.

Table 2 Classification results using different method for HYDICE DC Mall data.

		SVM	DCS-LCA	The proposed DCS method
All bands	OA	0.9495	0.9442	**0.9539**
	Kappa	0.9388	0.9322	**0.9440**
VGLCM 6+all bands	OA	0.9522	0.9535	**0.9608**
	Kappa	0.9422	0.9454	**0.9524**
VGLCM 6+5 PCs+5 selected bands	OA	0.9512	0.9524	**0.9525**
	Kappa	0.9408	0.9411	**0.9416**
VGLCM 6+5 selected bands	OA	0.9503	0.9512	**0.9526**
	Kappa	0.9397	0.9401	**0.9425**



(a) (b) (c)

Fig. 5 Classification maps for HYDICE DC Mall data. (a) SVM, (b) DCS-LCA, (c) the proposed DCS method.

Table 3 Confusion matrix of HYDICE experiment (SVM with all bands).

		Ground truth							
		Road	Grass	Trail	Tree	Shadow	Roof	Number of classified pixels	Users accuracy (%)
Classified	Road	878	0	20	0	0	42	892	98.43
	Grass	0	893	0	3	3	0	899	99.33
	Trail	2	0	547	0	2	0	551	99.27
	Tree	0	0	0	611	0	131	742	82.35
	Shadow	0	17	0	0	651	0	668	97.46
	Roof	12	0	0	9	0	950	971	97.84
Number of ground truth pixels		892	910	567	623	656	1123	OA = 94.95	Kappa = 0.9388
Producers accuracy (%)		98.43	98.13	96.47	98.07	99.24	84.59		

Table 4 Confusion matrix of HYDICE experiment (the proposed DCS with VGLCM 6+all bands).

		Ground truth							
		Road	Grass	Trail	Tree	Shadow	Roof	Number of classified pixels	Users accuracy (%)
Classified	Road	889	0	1	0	0	28	918	96.84
	Grass	0	888	0	39	7	0	934	95.07
	Trail	2	0	556	6	0	10	574	96.86
	Tree	0	0	0	574	0	55	629	91.26
	Shadow	0	22	0	0	647	0	669	96.71
	Roof	1	0	10	4	2	1030	1047	98.38
Number of ground truth pixels		892	910	567	623	656	1123	OA = 96.08	Kappa = 0.9524
Producers accuracy (%)		99.66	97.58	98.06	92.13	98.63	91.72		

have indicated that the best box size to describe the Indian Pines dataset is $9 \times 9 \times 9$ for VGLCM. The extracted texture features are then combined with the original dataset, 5 PCs, and 15 selected bands using the MEAC algorithm. The classification results of the Indian Pines dataset are reported in Table 5. We can see that the classification results of the proposed method outperformed other methods.

For visual comparison, the classification maps are presented in Fig. 6. By fusing texture information, the classification accuracy is considerably improved.

Tables 6–7 tabulate the confusion matrices of SVM and the proposed method for AVIRIS Pines data, which show a classification improvement in class separation, particularly for eight out of nine classes.

Table 5 Classification results using different methods for the Indian Pines dataset.

		SVM	DCS-LCA	The proposed DCS method
All bands	OA	0.8985	0.9250	**0.9432**
	Kappa	0.8809	0.9143	**0.9333**
VGLCM 6+all bands	OA	0.9112	0.9268	**0.9355**
	Kappa	0.8959	0.9163	**0.9239**
VGLCM 6+ 5 PCs+all bands	OA	0.9157	0.9277	**0.9421**
	Kappa	0.9011	0.9140	**0.9318**
VGLCM 6+5PCs+5 selected bands	OA	0.7921	0.9098	**0.9428**
	Kappa	0.7540	0.8944	**0.9326**

Table 6 Confusion matrix of AVIRIS experiment (SVM with all bands).

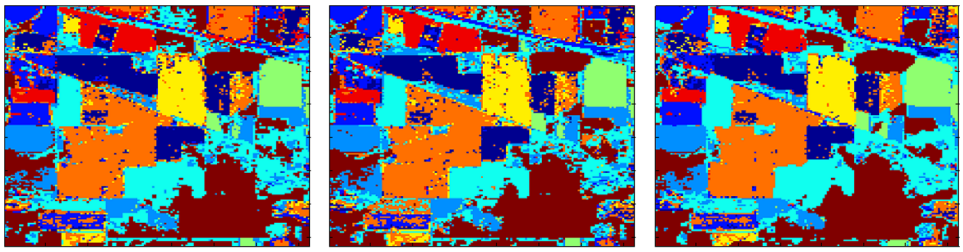
		Ground truth									Number of classified pixels	Users accuracy (%)
		1	2	3	4	5	6	7	8	9		
Classified	1	958	26	3	0	0	21	85	2	0	1095	87.49
	2	17	496	0	0	0	2	23	11	0	549	90.35
	3	0	0	362	4	3	4	5	2	2	382	94.76
	4	5	1	2	561	1	1	5	3	5	584	96.06
	5	0	0	0	0	354	0	1	2	0	357	99.16
	6	25	3	3	0	0	580	67	5	0	683	84.92
	7	85	88	2	0	0	119	1609	9	4	1916	83.98
	8	5	20	5	0	0	3	25	426	0	484	88.02
	9	0	0	5	2	0	0	0	0	951	958	99.27
Number of ground truth pixels		1095	634	382	567	358	730	1820	460	962	OA = 89.85 Kappa = 0.8809	
Producers accuracy (%)		87.49	78.23	94.76	98.94	98.88	79.45	88.41	92.61	98.86		

Note: 1-corn-no till, 2-corn-min till, 3-grass/pasture, 4-grass/trees, 5-hay-windrowed, 6-soybeans-no till, 7-soybeans-min till, 8-soybeans-clean till, 9-woods.

Table 7 Confusion matrix of AVIRIS experiment (the proposed DCS with all bands).

		Ground truth									Number of classified pixels	Users accuracy (%)
		1	2	3	4	5	6	7	8	9		
Classified	1	988	14	1	0	0	8	11	0	0	1022	96.67
	2	15	541	2	0	0	5	13	4	0	580	93.28
	3	1	0	364	0	0	0	2	0	0	367	99.18
	4	3	1	4	566	0	5	12	13	8	612	92.48
	5	0	0	0	0	358	0	0	2	0	360	99.44
	6	20	2	1	0	0	649	17	4	0	693	93.65
	7	64	72	6	0	0	62	1757	4	0	1965	89.41
	8	4	4	3	0	0	0	6	433	0	450	96.22
	9	0	0	1	1	0	1	2	0	954	959	99.48
Number of ground truth pixels		1095	634	382	567	358	730	1820	460	962	OA = 94.32 Kappa = 0.9333	
Producers accuracy (%)		90.23	85.33	95.29	99.82	100	88.90	96.54	94.13	99.17		

Note: 1-corn-no till, 2-corn-min till, 3-grass/pasture, 4-grass/trees, 5-hay-windrowed, 6-soybeans-no till, 7-soybeans-min till, 8-soybeans-clean till, 9-woods.



(a) (b) (c)

Fig. 6 Classification maps for the Indian Pines dataset. (a) SVM, (b) DCS-LCA, (c) the proposed DCS method.

4.4 *Statistical Significance Evaluation*

The nonparametric McNemar's test is applied to evaluate the statistical significance in accuracy improvement with the proposed methods.[37] It is based on the standardized normal test statistic. For the two methods to be compared, let f_{11} denote the number of samples that both methods can correctly classify, f_{22} is the number of samples that both cannot, f_{12} is the number of samples misclassified by method 1 but not method 2, and f_{21} is the number of samples misclassified by method 2 but not method 1. Thus, the McNemar's test statistic for these two methods can be defined as

$$z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}}. \tag{5}$$

For a 5% level of significance, the corresponding $|z|$ value is 1.96; a $|z|$ value greater than this quantity means two methods have a significant performance discrepancy. Moreover, the sign of z indicates whether method 1 outperforms method 2 ($Z > 0$ or vice versa. Table 8 tabulates the z

Table 8 Z values in the McNemar's Test.

	The proposed DCS method			
	HYDICE DC MALL		AVIRIS Pine	
	z	Significant?(5%)	z	Significant?(5%)
SVM	2.57	Yes	6.21	Yes
DCS-LCA	1.78	Yes	5.18	Yes

values between the proposed method (method 1) and the competing methods (method 2). It can be seen that the proposed DCS method statistically outperforms other competing methods.

5 Conclusions

In this paper, a DCS approach that integrates both spectral and spatial information (especially a pixel's neighborhood) for classifying the hyperspectral images was proposed. In this method, both spatial and spectral information have been used to calculate the label on the condition that the classified pixels' percentage of the unlabeled pixels' neighbor meets the threshold. Additionally, the classification Additionally, the spectral features were generated by MEAC-based band selection, and the volumetric texture features was extracted using volumetric GLCM. It has shown that the use of volumetric texture features exhibited a better classification accuracy than the traditional spectral information algorithms. Thus, it can be concluded that the proposed DCS method leads to an improved classification performance for hyperspectral imagery.

Acknowledgments

This paper was partially supported by National Natural Science Foundation of China (Grant Nos. 41201341, 51379056), the Open Research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (Grant No. 12R02), Key Laboratory of Satellite Mapping Technology and Application, National Administration of Surveying, Mapping and Geoinformation (KLSMTA-201301), and Key Laboratory of Advanced Engineering Surveying of National Administration of Surveying, Mapping and Geoinformation (Grant No. TJES1301).

References

1. M. D. Mura et al., "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geosci. Remote Sens. Lett.* **8**(3), 542–546 (2011).
2. P. Gamba et al., "Improved VHR urban area mapping exploiting object boundaries," *IEEE Trans. Geosci. Remote Sens.* **45**(8), 2676–2682 (2007).
3. H. Yang et al., "Improving urban land use and land cover classification from high-spatial-resolution hyperspectral imagery using contextual information," *J. Appl. Remote Sens.* **4**(1), 041890 (2010).
4. 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).
5. J. Qiong and D. A. L andgrebe, "Adaptive bayesian contextual classification based on markov random fields," *IEEE Trans. Geosci. Remote Sens.* **40**(11), 2454–2463 (2002).
6. N. P. Angelo and V. Haertel, "On the application of Gabor filtering in supervised image classification," *Int. J. Remote Sens.* **24**(10), 2167–2189 (2003).

7. R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Texture features for image classification," *IEEE Trans. Sys. Man Cybern.* **3**(6), 610–621 (1973).
8. A. N. Nyoungui, E. Tonye, and A. Akono, "Evaluation of speckle filtering and texture analysis methods for land cover classification from SAR images," *Int. J. Remote Sens.* **23**(9), 1895–1925 (2002).
9. F. Tsai et al., "3-D computation of gray level co-occurrence in hyperspectral image cubes," *Lec. Notes Comput. Sci.* **4679**, 429–440 (2007).
10. M. Fauvel et al., "Advances in spectral–spatial classification of hyperspectral images," *Proc. IEEE* **101**(3), 652–675 (2013).
11. P. Du et al., "Multiple cassifier system for remote sensing image classification: a review," *Sensors* **12**(4), 4764–4792 (2012).
12. T. G. Dietterich, "Ensemble methods in machine learning," *Lec. Notes Comput. Sci.* **1857**, 1–15 (2000).
13. Z.-H. Zhou, J. Wu, and W. Tang, "Ensembling neural networks: many could be better than all," *Artif. Intell.* **137**(122), 239–263 (2002).
14. L. I. Kuncheva, J. C. Bezdek, and R. P. Duin, "Decision templates for multiple classifier fusion: an experimental comparison," *Pattern Recogn.* **34**(2), 299–314 (2001).
15. K. Woods, W. P. Kegelmeyer, Jr., and K. Bowyer, "Combination of multiple classifiers using local accuracy estimates," *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(4), 405–410 (1997).
16. K. Tumer and J. Ghosh, "Error correlation and error reduction in ensemble classifiers," *Connection Sci.* **8**(3), 385–404 (1996).
17. G. Giacinto and F. Roli, "Adaptive selection of image classifiers," *Lec. Notes Comput. Sci.* **1310**, 38–45 (1997).
18. L. I. Kuncheva, "Switching between selection and fusion in combining classifiers: an experiment," *IEEE Trans. Syst. Man. Cybern.* **32**(2), 146–156 (2002).
19. A. H. R. Ko, R. Sabourin, and J. A. S. Britto, "From dynamic classifier selection to dynamic ensemble selection," *Pattern Recogn.* **41**(5), 1718–1731 (2008).
20. P. C. Smits, "Multiple classifier systems for supervised remote sensing image classification based on dynamic classifier selection," *IEEE Trans. Geosci. Remote Sens.* **40**(4), 801–813 (2002).
21. L. Didaci et al., "A study on the performances of dynamic classifier selection based on local accuracy estimation," *Pattern Recogn.* **38**(11), 2188–2191 (2005).
22. Y. Tarabalka et al., "Multiple spectral-spatial classification approach for hyperspectral data," *IEEE Trans. Geosci. Remote Sens.* **48**(11), 4122–4132 (2010).
23. J. A. Benediktsson, J. Chanussot, and M. Fauvel, "Multiple classifier systems in remote sensing: from basics to recent developments," *Lec. Notes Comput. Sci.* **4472**, 501–512 (2007).
24. C. Chen et al., "Spectral-spatial classification of hyperspectral image based on kernel extreme learning machine," *Remote Sens.* **6**(6), 5795–5814 (2014).
25. T. M. Nguyen and Q. M. J. Wu, "Gaussian-mixture-model-based spatial neighborhood relationships for pixel labeling problem," *IEEE Trans. Syst. Man. Cybern.* **42**(1), 193–202 (2012).
26. C. Chen et al., "Spectral-spatial preprocessing using multihypothesis prediction for noise-robust hyperspectral image classification," *IEEE J. Sel. Top. Appl. Earth Observ.* **7**(4), 1047–1059 (2014).
27. Q. Du et al., "Noise-adjusted principal component analysis for buried radioactive target detection and classification," *IEEE Trans. Nucl. Sci.* **57**(6), 3760–3767 (2010).
28. H. Yang et al., "An efficient method for supervised hyperspectral band selection," *IEEE Geosci. Remote Sens. Lett.* **8**(1), 138–142 (2011).
29. D. J. Marceau et al., "Evaluation of the grey-level co-occurrence matrix method, for land-cover classification using SPOT imagery," *IEEE Trans. Geosci. Remote Sens.* **28**(4), 513–519 (1990).
30. W. R. Tobler, "A computer movie simulating urban growth in the Detroit region," *Econ. Geogr.* **46**(2), 234–240 (1970).
31. G. Giacinto and F. Roli, "Methods for dynamic classifier selection," in *Proc. 10th Int. Conf. on Image Analysis and Processing*, pp. 659–664, IEEE, Venice, Italy (1999).

32. A. M. P. Canuto et al., "Investigating the influence of the choice of the ensemble members in accuracy and diversity of selection-based and fusion-based methods for ensembles," *Pattern Recogn. Lett.* **28**(4), 472–486 (2007).
33. K. H. Riitters, J. D. Wickham, and T. G. Wade, "An indicator of forest dynamics using a shifting landscape mosaic," *Ecol. Indicators* **9**(1), 107–117 (2009).
34. K. Liu et al., "Optical flow and principal component analysis-based motion detection in outdoor videos," *EURASIP J Adv. Signal Process.* **2010**, 680623 (2010).
35. H. Su et al., "Adaptive affinity propagation with spectral angle mapper for semi-supervised hyperspectral band selection," *Appl. Opt.* **51**(14), 2656–2663 (2012).
36. C. Chen et al., "Reconstruction of hyperspectral imagery from random projections using multihypothesis prediction," *IEEE Trans. Geosci. Remote Sens.* **52**(1), 365–374 (2014).
37. G. M. Foody, "Thematic map comparison: evaluating the statistical significance of differences in classification accuracy," *Photogramm. Eng. Remote Sens.* **70**(5), 627–633 (2004).

Hongjun Su received his PhD degree in cartography and geography information system from Nanjing Normal University in 2011. He was a visiting student at Mississippi State University from 2009 to 2010. He is now an associate professor at the School of Earth Sciences and Engineering, Hohai University, Nanjing, China, and also with the Department of Geographical Information Science, Nanjing University, as a postdoctoral fellow. His main research interests include hyperspectral remote sensing dimensionality reduction, classification, and spectral unmixing.

Bin Yong received his PhD degree in cartography and geography information system, Nanjing University, Nanjing, China. He is currently a professor with State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, China. His research areas mainly include remote sensing retrievals and applications; remote sensing precipitation (radar, satellite, multisensor, multiplatform); hyperspectral image classification; and satellite-based surface hydrology. He is currently an American Geophysical Union member.

Peijun Du received his PhD degrees in cartography and geographic information system from China University of Mining and Technology in 2001. Currently, he is a professor of photogrammetry, remote sensing, and geographical information science with the Department of Geographical Information Science, Nanjing University, Nanjing, China. His research interests include remote sensing image processing and pattern recognition. He is the author of more than 100 research articles about remote sensing and geospatial information processing and applications.

Hao Liu received his BS degree in surveying and mapping engineering from Hohai University, Nanjing, China, in 2013. He is currently working toward a master's degree in the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China. His research interests include pattern recognition, image processing and computer vision.

Chen Chen received his BE degree in automation from Beijing Forestry University, Beijing, China, in 2009 and the MS degree in electrical engineering from Mississippi State University, Starkville, Mississippi, in 2012. He is currently working toward the PhD degree in the Department of Electrical Engineering, University of Texas at Dallas, Richardson, Texas. His research interests include compressed sensing, signal and image processing, pattern recognition, computer vision, and hyperspectral image analysis.

Kui Liu received his BE degree in electrical engineering from Nanchang University, Nanchang, China in 2005, and received the MS degree in electrical engineering from Mississippi State University, Starkville, Mississippi, in 2011. He is currently a graduate research assistant in the Department of Electrical Engineering at the University of Texas at Dallas as a member of the Signal and Image processing Laboratory. His research interests include real-time image processing 3-D computer vision, and machine learning.

2015-8

Dynamic Classifier Selection Using Spectral-Spatial Information for Hyperspectral Image Classification

UTD AUTHOR(S): Chen Chen and Kui Liu

Su, Hongjun, Bin Yong, Peijun Du, Hao Liu, et al. 2014. "Dynamic classifier selection using spectral-spatial information for hyperspectral image classification." Journal of Applied Remote Sensing 8(085095): doi:10.1117/1.JRS.8.085095.