A Novel MRF-Based Multifeature Fusion for Classification of Remote Sensing Images

Qikai Lu, Xin Huang, *Senior Member, IEEE*, Jun Li, *Senior Member, IEEE*, and Liangpei Zhang, *Senior Member, IEEE*

***Abstract*—The spatial information has been proved to be effective in improving the performance of spectral-based classification. However, it is difficult to describe different image scenes by using monofeature owing to complexity of the geospatial scenes. In this letter, a novel framework is developed to combine the multiple spectral and spatial features based on the Markov random field (MRF). Specifically, the pixels in an image are separated into reliable and unreliable ones according to the decision of multifeature classifications. The labels of the reliable pixels can be conveniently determined, but the unreliable pixels are then classified by fusing the multifeature classification results and reducing the classification uncertainties based on the MRF optimization. Experiments are conducted on three multispectral high-resolution images to verify the effectiveness of the proposed method. Several state-of-the-art multifeature classification methods are also achieved for the purpose of comparison. Moreover, three classifiers (i.e., multinomial logistic regression, support vector machines, and random forest) are used to test the performance of the proposed framework. It is shown that the proposed method can effectively integrate multiple features, yield promising results, and outperform other approaches compared.**

***Index Terms*—Classification, data fusion, high resolution, Markov random field (MRF), multifeature.**

I. Introduction

In RECENT years, availability of high spatial resolution images has provided abundant detailed and structural information for the land surface. However the increase in spatial resolution can lead to increase in intraclass variation and decrease in interclass variation in an image, which essentially reduces the separability among different classes in the spectral domain [1]. Consequently, traditional spectral-based classification approaches are inadequate for the high spatial resolution-image interpretation.

In this context, the information that describes the spatial distribution of pixels is exploited to enhance the feature space for improving classification performances.

Although various spatial features have been developed to improve the classification, it is difficult to find a certain kind of feature that can effectively represent various kinds of scenes owing to the complexity of geospatial landscapes. A popular approach to address this problem is to stack the different features as a vector and input them into classifiers, i.e., vector stacking (VS) [2]. The VS approach has been widely used in the applications owing to its simple and convenient implementation; it, however, always leads to a high-dimensional feature space that needs large memory space and high computation workload. Moreover, direct concatenation of different spatial features does not necessarily result in higher classification accuracy. In order to adequately take advantage of multiple features, Huang and Zhang proposed several algorithms for combination of multiple features based on support vector machines (SVMs), e.g., certainty voting (C-voting) and probabilistic fusion (P-fusion) [3]. Specifically, C-voting is an improved majority voting (MV) algorithm by assigning different weights to different classifiers according to the classification uncertainties. P-fusion is actually a soft voting by considering the probabilistic outputs of different classifiers, weighted by their uncertainties. Li *et al.* proposed a multiple-feature learning approach that is able to adaptively exploit information from linearly and nonlinearly derived features [4]. Zhang *et al.* developed a patch alignment framework to obtain a unified low-dimensional representation of multiple features for exploring the specific properties of each feature [5]. Li *et al.* proposed a joint collaborative representation classification method with multitask learning, in which multiple complementary features from different perspectives are considered [6]. However, these existing algorithms are in a pixel-by-pixel manner, which ignores the spatial relationship between adjacent pixels.

In this context, we propose a novel multiple-feature fusion approach for the classification of remote sensing images. The spatial features employed in this letter include gray-level co-occurrence matrix (GLCM) [7], differential morphological profiles (DMPs) [8], and multi-index features (MIFs) [9]. By concatenating the spectral bands and a certain kind of spatial feature, a series of classifications are performed. Subsequently, an image can be separated into reliable and unreliable pixels, depending on the classification results. MV is employed to

Manuscript received December 2, 2015; revised January 12, 2016; accepted January 18, 2016. Date of publication February 29, 2016; date of current version March 23, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 91338111, by the China National Science Fund for Excellent Young Scholars under Grant 41522110, and by the Foundation for the Author of National Excellent Doctoral Dissertation of China under Grant 201348. *(Corresponding author: Xin Huang.)*

Q. Lu is with the School of Electronic Information, Wuhan University, Wuhan 430072, China.

X. Huang is with the School of Remote Sensing and Information Engineering and the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: huang_whu@163.com).

J. Li is with the School of Geography and Planning and Guangdong Key Laboratory for Urbanization and Geo-Simulation, Sun Yat-sen University, Guangzhou 510275, China.

L. Zhang is with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China.

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

Digital Object Identifier 10.1109/LGRS.2016.2521418

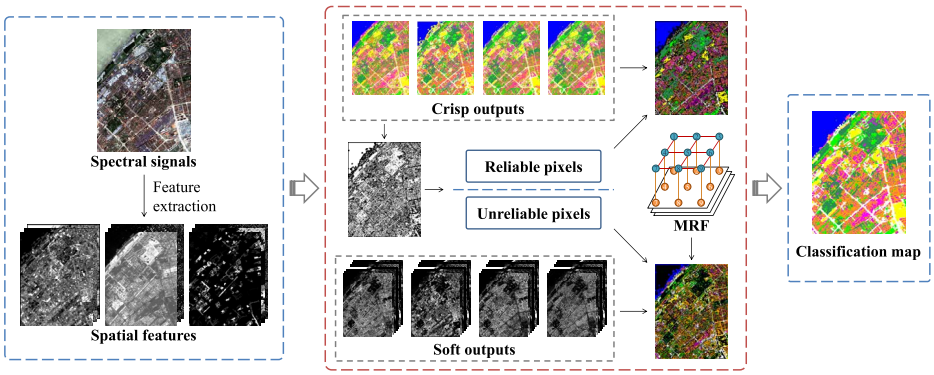

Fig. 1. Flowchart of the proposed MRF-based multifeature fusion framework.

determine the labels of reliable pixels, where the most frequent class achieved by multiple classifications is assigned to the pixel [10]. Then, the neighborhood information is taken into account for identification of the unreliable pixels since the information from an individual pixel is insufficient to classify them. Herein, we propose to use Markov random field (MRF) for determining the labels of the unreliable pixels since MRF is able to make an iteratively optimized decision to the unreliable pixels by considering the probabilistic outputs from adjacent pixels.

Three data sets acquired by different sensors are used in the experiments for algorithm validation, and the proposed approach is compared with the state-of-the-art multifeature classification methods.

II. Multifeature Extraction

A. Multifeature Extraction

GLCM describes how often pairs of specific values with a specific spatial relationship occur in a window and can be expressed as $f_{\mathrm{GLCM}}(w, m)$ with w and m indicating the window size and texture measures, respectively. Owing to the multiscale characteristic inherent in the high spatial resolution image, a single window size is insufficient to represent the spatial information at different scales. Therefore, a series of window sizes are considered in our work for the generation of GLCM. Furthermore, based on the obtained GLCM, several statistical measures can be derived to describe the texture information in an image.

Morphological profiles (MPs) are generated based on the opening and closing by reconstruction with structural elements (SEs) of increasing size. As the derivative of MPs, DMPs can be regarded as a measurement of the gray value variation of the MPs for every step of an increasing SE series [8]. In particular, to exploit the multiscale characteristics of an image, a series of SEs with different sizes are considered in this letter.

MIFs aim at using low-dimensional information indices to replace the high-dimensional and low-level features for image description [9]. In this letter, MIFs are constructed by the normalized difference vegetation index, morphological building index, and morphological shadow index. Readers can refer to [9] for more details.

B. Proposed MRF-Based Multifeature Fusion Scheme

To effectively take advantage of the multiple features and describe complex image scenes, a novel framework is developed for integrating the multiple-feature sources (see Fig. 1), which is introduced in the following steps.

1) *Single-feature classification*: Spectral bands and a certain kind of spatial feature are concatenated and used for generating both crisp (class label) and soft (probabilistic) outputs for each pixel.
2) *Determining reliable and unreliable pixels*: The pixels in an image are divided into reliable and unreliable ones according to the results of the single-feature classifications and the full-spectral classification. In this letter, the pixels that are assigned to the same labels by all classifiers are defined as the reliable ones, whereas the rest are regarded as unreliable.
3) *Evaluation of classification certainty*: Based on the soft output of the classification, the certainty of a pixel can be defined by considering the multiclass probabilistic outputs. Specifically, the specificity measure [11] is employed to describe the certainty of classification

$$S(x) = \sum_{k=1}^{K-1} \frac{1}{k} \cdot \left(\widehat{p}_k(x) - \widehat{p}_{k+1}(x) \right) \tag{1}$$

where K indicates the number of classes, and $\widehat{p}_1(x), \ldots, \widehat{p}_k(x), \ldots, \widehat{p}_K(x)$ is the probabilistic output for pixel x in a descending order. A high value of $S(x)$, ranging from 0 to 1, signifies a large classification certainty.
4) *Identification of the reliable pixels*: The labels of the reliable pixels can be conveniently determined by MV,

$$C(x) = \underset{k \in [1,K]}{\arg\max} \left(\sum_{f=0}^{F} I\left(C^f(x) = k \right) \right) \tag{2}$$

where $C(x)$ denotes the label of pixel x, and $I(\cdot)$ represents the indicator function. F is the number of spatial features, and $C^f(x)$ is the label of pixel x with the feature f. In particular, $f = 0$ represents the spectral-based classification.

5) *Multifeature P-fusion*: For an unreliable pixel, we compute its multifeature weighted probabilistic function for each class and the classification certainty is used as the weight

$$\tilde{p}_k(x_{\text{un}}) = \frac{\sum_{f=0}^{F} S^f(x_{\text{un}}) \cdot p_k^f(x_{\text{un}})}{\sum_{f=0}^{F} S^f(x_{\text{un}})} \tag{3}$$

where $\tilde{p}_k(x_{\text{un}}) \in [0,1]$, which satisfies $\sum_{k=1}^{K} \tilde{p}_k(x_{\text{un}}) = 1$, indicate the fused probability of the unreliable pixel x_{un} belonging to the class k, and $p_k^f(x_{\text{un}})$ and $S^f(x_{\text{un}})$ represent the probability that a pixel x_{un} belongs to class k and its classification certainty with the feature f, respectively.

6) *Classification of the unreliable pixels*: Different from the reliable pixels, it is inappropriate to employ MV to identify the unreliable ones as the multifeature classifiers do not achieve an agreement. To better interpret the label of the unreliable pixels, the spatial relationship between the pixels in an image should be explored. In this letter, a novel algorithm based on MRF is developed to use the neighboring information for identification of the unreliable pixels.

The MRF model can smooth the classification results in a homogeneous area and at the same time preserve the boundary and details of objects. A detailed analysis of the MRF-based classification approach has been introduced in [12]. It can be formulated as

$$U = \sum U_{\text{data}}(x) + \beta \sum_{u \in N_x} U_{\text{spa}}(x,u) \tag{4}$$

where $U_{\text{data}}(x)$ is the data term, and $U_{\text{spa}}(x,u)$ is the spatial term. N_x indicates the eight-neighborhood of pixel x, and u is the pixel located in N_x. Parameter β is a nonnegative constant that controls the contribution of the spatial term. The data term indicates how likely a pixel belongs to a certain class with the observed data. In the proposed scheme, the multifeature weighted probability $\tilde{p}_k(x_{\text{un}})$ is used as the data term in order to integrate the multiple-feature sources in the MRF optimization. Thus, the data term can be written as

$$U_{\text{data}}(x_{\text{un}}) = -\text{In}(\tilde{p}_k(x_{\text{un}})). \tag{5}$$

On the other hand, the spatial term models the relationship between a pixel and its neighborhood. Depending on the assumption that the adjacent pixels are likely to belong to the same class [13], the spatial term is defined as follows:

$$U_{\text{spa}}(x_{\text{un}}, u) = 1 - \delta\left(C(x_{\text{un}}), C(u)\right) \tag{6}$$

$$\delta\left(C(x), C(u)\right) = \begin{cases} 1, & \text{if } C(x) = C(u) \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

where $\delta(\cdot,\cdot)$ is the Kronecker delta function. Therefore, the energy function of the proposed multifeature MRF model can be written as

$$U = -\sum \text{In}\left(\tilde{p}_k(x_{\text{un}})\right) + \beta \sum_{u \in N_{x_{\text{un}}}} \left(1 - \delta\left(C(x_{\text{un}}), C(u)\right)\right). \tag{8}$$

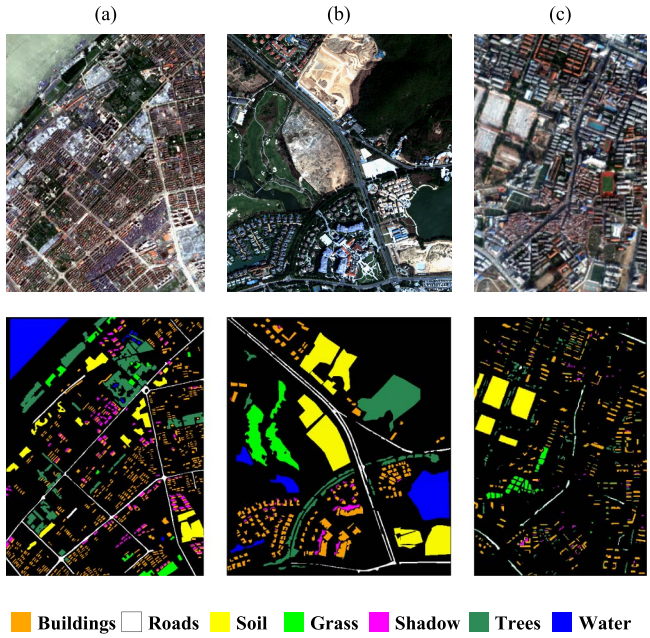

Fig. 2. Test images and their reference samples. (a) ZY-3, (b) WV-2, and (c) GeoEye-1 data sets.

Please note that the proposed MRF classification optimization only focuses on the unreliable pixels. In this way, the multifeature probabilistic outputs in a neighborhood of each unreliable pixel are utilized for the classification.

III. Experiments

A. Data Sets' Description

To validate the performance of the proposed method, the experiments are conducted on three multispectral remote sensing data sets. The first one was collected by the ZY-3 Satellite over an urban area from Wuhan City in central China. It contains 651 pixels × 499 pixels with a spatial resolution of 5.8 m and four bands. The second data set is an image with eight multispectral channels, collected by the WorldView-2 (WV-2) Satellite in Hainan Province. The image shows a rural area with 600 pixels × 520 pixels at a spatial resolution of 2 m. The third one is also from Wuhan City, obtained by GeoEye-1 with four bands. The image is composed of 908 pixels × 607 pixels with a spatial resolution of 2 m/pixel. The images and reference data are presented in Fig. 2(a)–(c) for the ZY-3, WV-2, and GeoEye-1 data sets, respectively.

B. Experimental Setup

In this letter, principal component (PC) analysis is used to extract the base images for GLCM and DMPs. The first and second PCs are selected since 99% of information contained in the image can be represented by them. The GLCM measures considered include contrast and homogeneity, and their window sizes are set to 3, 5, 7, 9, and 11. The radius of the disk SE for DMPs' feature is set to 1, 2, 3, 4, and 5.

Three classifiers are considered in the experiments, including SVMs, multinomial logistic regression via variable splitting

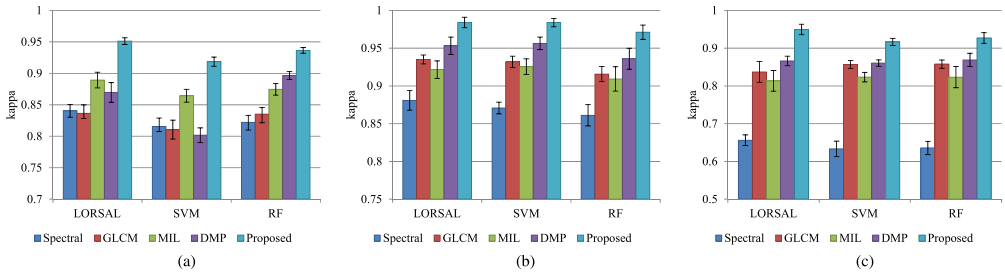

Fig. 3. Kappa coefficients and standard deviations obtained by different methods for (a) ZY-3, (b) WV-2, and (c) GeoEye-1 data sets, respectively.

and augmented Lagrangian (LORSAL), and random forest (RF). The parameters for SVM are set as kernel = radial basis function (RBF), penalty coefficient = 100, and RBF bandwidth = $1/d$ (d is the dimension of the input feature). For RF, the number of trees is set to 500.

For each data set, 50 samples per class are randomly selected for training the classification model. Moreover, all of the experiments are repeated ten times, and the mean and standard deviation of the kappa coefficients are calculated for the accuracy assessment. Parameter β is set to 1 in the experiments. Additionally, the value of β can be determined by using the particle swarm optimization algorithm [14].

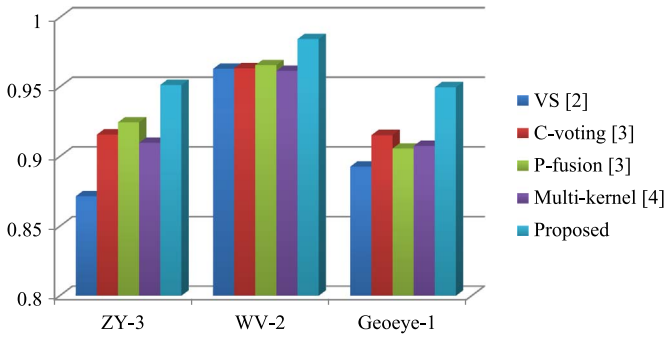

Fig. 4. Comparison between different multifeature fusion approaches, where the vertical axis indicates the kappa coefficient.

C. Experimental Results

The classification accuracy values obtained by different algorithms are provided in Fig. 3 for the ZY-3, WV-2, and GeoEye-1 data sets, respectively. As seen from the results, the classification performances in the three test data sets exhibit similar trends. For the spectral-based classification, the corresponding results show that the spectral information is difficult to accurately classify the high-resolution urban areas. Consideration of spatial features enhances discriminability between the similar information classes, resulting in higher accuracy values. However, we can also notice that, although the spatial features are able to complement the spectral information, the performances and effectiveness of different features are quite different in the test data sets. It implies that an inappropriate use of spatial information may reduce the image classification accuracy, and it is impossible to pick out a certain feature that is suitable for all scenes. From these figures, it is clear that the proposed multiple-feature fusion methods are superior to the monofeature classifications in terms of the higher classification accuracy values and lower standard deviations.

IV. Comparisons and Discussions

The proposed MRF-based multifeature fusion scheme is compared with the state-of-the-art methods, e.g., VS [2], C-voting, P-fusion [3], and the multikernel learning framework proposed by Li *et al.* [4]. Fig. 4 presents the accuracy values of different multifeature classification approaches for the ZY-3, WV-2, and GeoEye-1 data sets, respectively. In this test, LORSAL is used as the baseline classifier. It can be observed that the existing state-of-the-art multifeature classification approaches achieve similar performances, but the proposed method outperforms them in all the cases. The accuracy increments achieved can be attributed to the MRF-based classification optimization, which is able to simultaneously exploit the multifeature-integrated probability and the neighborhood response obtained from the spatial term. In this way, the proposed multifeature strategy can classify the unreliable pixels more accurately.

Table I demonstrates the classification accuracy values for reliable and unreliable pixels in the ZY-3, WV-2, and GeoEye-1 data sets, respectively. The first noteworthy observation from the results is that the accuracies for the unreliable pixels are much lower than the reliable ones. This can be understandable since the reliable pixels are more likely to be assigned to the correct classes since all the classifiers point to the same prediction. The classification accuracy values are highly dependent on the precision of interpreting the unreliable pixels. By assessing the accuracy on the unreliable samples only, it can be more clearly seen that the proposed method can achieve significantly better results compared with other multifeature fusion strategies.Furthermore, Fig. 5 reports the percentage of the unreliable pixels for each class, where it can be found that buildings, roads, and soil, which show similar spectral properties, are the most difficult to identify.

V. Conclusion and Further Work

In this letter, we have proposed a new framework for the classification of remote sensing image by fusing multiple-feature sources. Based on the outputs of the full-spectral and single-feature classifications, the proposed method divides the pixels in the image into reliable and unreliable ones. The reliable

TABLE I
CLASSIFICATION ACCURACY VALUES FOR THE RELIABLE AND UNRELIABLE PIXELS OBTAINED BY DIFFERENT APPROACHES IN THE TEST DATA SETS

Datasets	Classifier	Reliable pixels	Unreliable pixels							
			Spectral	GLCM	MIL	DMP	MV	C-voting	P-fusion	Proposed
ZY-3	LORSAL	0.992	0.420	0.405	0.601	0.530	0.643	0.698	0.733	0.833
	SVM	0.970	0.465	0.447	0.619	0.411	0.629	0.717	0.737	0.795
	RF	0.966	0.287	0.342	0.525	0.629	0.560	0.613	0.641	0.815
WV-2	LORSAL	0.996	0.219	0.578	0.475	0.702	0.693	0.759	0.778	0.910
	SVM	0.992	0.161	0.491	0.450	0.693	0.674	0.748	0.750	0.917
	RF	0.980	0.094	0.465	0.425	0.627	0.560	0.547	0.578	0.903
Geoeye-1	LORSAL	0.986	0.051	0.531	0.458	0.620	0.566	0.754	0.725	0.863
	SVM	0.959	0.029	0.550	0.434	0.592	0.463	0.685	0.669	0.782
	RF	0.966	-0.039	0.583	0.479	0.637	0.513	0.621	0.641	0.823

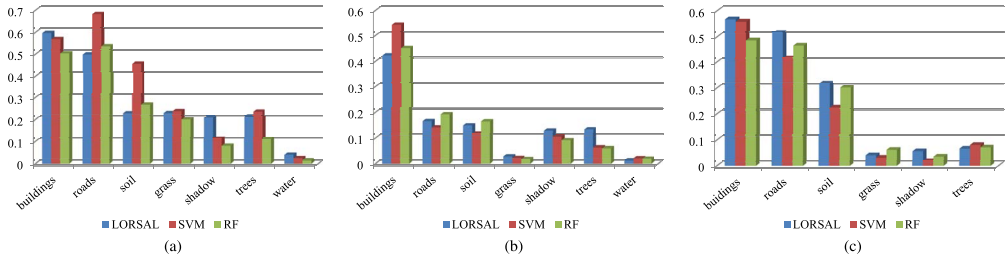

Fig. 5. Percentage of the unreliable pixels in the test set for the (a) ZY-3, (b) WV-2, and (c) GeoEye-1 data sets.

pixels can be simply classified by MV, and the unreliable pixels are identified in an MRF framework by taking the multifeature probabilistic outputs and their neighboring response into consideration. The experimental results demonstrate that performances of the proposed method surpass those of state-of-the-art multifeature fusion approaches. As shown in the experiments, the difference of classification performances for different multifeature methods is mainly attributed to the misclassification of the unreliable pixels. It can be stated that the proposed method finds an effective solution to the issue and hence significantly improves the accuracy for the unreliable pixels.

ACKNOWLEDGMENT

The authors would like to thank Prof. D. Landgrebe (Purdue University, West Lafayette, IN, USA) for providing the free downloads of the AVIRIS data set over Indian Pines. The authors also gratefully appreciate the insightful comments and suggestions from the editors and the anonymous reviewers, which greatly helped them improve the quality of this letter.

REFERENCES

[1] L. Bruzzone and L. Carlin, "A multilevel context-based system for classification of very high spatial resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2587–2600, Sep. 2006.

[2] X. Huang and L. Zhang, "Comparison of vector stacking, multi-SVMs fuzzy output, and multi-SVMs voting methods for multiscale VHR urban mapping," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 261–265, Apr. 2010.

[3] X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 257–272, Jan. 2013.

[4] J. Li *et al.*, "Multiple feature learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1592–1606, Mar. 2015.

[5] L. Zhang, L. Zhang, D. Tao, and X. Huang, "On combining multiple features for hyperspectral remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 879–893, Mar. 2012.

[6] J. Li, H. Zhang, L. Zhang, X. Huang, and L. Zhang, "Joint collaborative representation with multitask learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5923–5936, Sep. 2014.

[7] F. R. D. Siqueira, W. R. Schwartz, and H. Pedrini, "Multi-scale gray level co-occurrence matrices for texture description," *Neurocomputing*, vol. 120, pp. 336–345, Nov. 2013.

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

[9] X. Huang, Q. Lu, and L. Zhang, "A multi-index learning approach for classification of high-resolution remotely sensed images over urban areas," *ISPRS J. Photogramm. Remote Sens.*, vol. 90, pp. 36–48, Apr. 2014.

[10] A. B. Santos, A. D. A. Araújo, and D. Menotti, "Combining multiple classification methods for hyperspectral data interpretation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1450–1459, Jun. 2013.

[11] R. R. Yager, "On the specificity of a possibility distribution," *Fuzzy Sets Syst.*, vol. 50, no. 3, pp. 279–292, Sep. 1992.

[12] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM- and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.

[13] X. Huang, Q. Lu, L. Zhang, and A. Plaza, "New postprocessing methods for remote sensing image classification: A systematic study," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7140–7159, Nov. 2014.

[14] G. Liu, A. Wang, and Y. Zhao, "An efficient image segmentation method based on fuzzy particle swarm optimization and Markov random field model," in *Proc. 7th Int. Conf. WiCOM*, 2011, pp. 1–4.