

# Extended Random Walker-Based Classification of Hyperspectral Images

Xudong Kang, *Student Member, IEEE*, Shutao Li, *Member, IEEE*, Leyuan Fang, *Student Member, IEEE*, Meixiu Li, and Jón Atli Benediktsson, *Fellow, IEEE*

**Abstract**—This paper introduces a novel spectral–spatial classification method for hyperspectral images based on extended random walkers (ERWs), which consists of two main steps. First, a widely used pixelwise classifier, i.e., the support vector machine (SVM), is adopted to obtain classification probability maps for a hyperspectral image, which reflect the probabilities that each hyperspectral pixel belongs to different classes. Then, the obtained pixelwise probability maps are optimized with the ERW algorithm that encodes the spatial information of the hyperspectral image in a weighted graph. Specifically, the class of a test pixel is determined based on three factors, i.e., the pixelwise statistics information learned by a SVM classifier, the spatial correlation among adjacent pixels modeled by the weights of graph edges, and the connectedness between the training and test samples modeled by random walkers. Since the three factors are all well considered in the ERW-based global optimization framework, the proposed method shows very good classification performances for three widely used real hyperspectral data sets even when the number of training samples is relatively small.

**Index Terms**—Extended random walkers (ERWs), graph, hyperspectral image, optimization, spectral–spatial image classification.

## I. INTRODUCTION

**H**YPERSPECTRAL image classification gives a high-level understanding of remotely sensed scenes and is therefore now widely used in different application domains such as environment monitoring [1], precision agriculture [2], and national defense [3]. However, because of the special characteristics of hyperspectral data sets, image classification in the hyperspectral domain still has many unresolved problems [4].

For instance, the high dimensionality of hyperspectral data sets involves the “Hughes” phenomenon in classification [5]. The “Hughes” phenomenon refers to the fact that, if the number of training samples is fixed, the classification accuracy may de-

crease significantly for some supervised classification methods as the data dimensionality increases beyond a certain number of features. In order to deal with this difficulty, several solutions have been developed such as feature extraction [6]–[8] and discriminative learning [9], [10]. Feature extraction methods such as principal component analysis (PCA) [7], independent component analysis [6], and linear discriminant analysis [8], [11], [12] project the high-dimensional data into a low dimensional feature space while preserving the discriminative information of different classes. Furthermore, discriminative learning approaches such as support vector machines (SVMs) [9], multinomial logistic regression [10], and artificial immune networks [13] learn the class distributions in high-dimensional spaces by inferring the nonlinear boundaries between classes in feature space. These methods can effectively tackle the aforementioned difficulties caused by high dimensionality.

In addition to research on how to overcome the “Hughes” phenomenon, another active research topic for hyperspectral image classification is how to make full use of the spatial information of the data in order to further improve the classification accuracy [14]. To achieve this objective, intensive work has been performed in the last decade to develop spectral–spatial hyperspectral image classification methods. For example, spatial feature extraction methods [15]–[21] have been proposed to define an adaptive neighborhood for each pixel by local filtering operations so that the adaptive local neighborhood information could be preserved in the resulting features for classification. Furthermore, an optimal set of the resulting spatial features can be selected to further improve the performance of feature extraction [22].

In addition to spatial feature extraction, image segmentation is a widely used technique for spectral–spatial image classification. Specifically, segmentation-based methods perform the decision fusion of image segmentation and pixelwise classification to make full use of the spatial information of hyperspectral images. For this kind of methods, the automatic segmentation of hyperspectral images is a challenging task, and thus, many different hyperspectral image segmentation methods have been proposed such as watershed [23], partitional clustering [24], hierarchical segmentation [25], and stochastic minimum spanning forest [26], [27].

In recent years, other types of spectral–spatial methods such as nonlocal joint collaborative representation [28], spatial kernel-based methods [29], [30], and probabilistic modeling-based methods [10], [31]–[34] have also been successfully applied for hyperspectral image classification. For instance, probabilistic modeling-based methods first estimate the

Manuscript received January 28, 2014; revised March 24, 2014; accepted April 18, 2014. This paper was supported in part by the National Natural Science Foundation for Distinguished Young Scholars of China under Grant 61325007, by the National Natural Science Foundation of China under Grant 61172161, by the Fundamental Research Funds for the Central Universities, by Hunan University, and by the Chinese Scholarship Award for Excellent Doctoral Student.

X. Kang, S. Li, L. Fang, and M. Li are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: xudong\_kang@hnu.edu.cn; shutao\_li@hnu.edu.cn; fangleyuan@gmail.com).

J. A. Benediktsson is with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland (e-mail: benedikt@hi.is).

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/TGRS.2014.2319373

probability of each pixel belonging to a specified class using a pixelwise classifier such as SVM. Then, the pixelwise probabilities are optimized with local or global optimization algorithms which consider the spatial information of the image. Finally, the classification result is obtained by assigning each pixel with the label which gives the maximum probability.

In this paper, a novel probabilistic modeling-based spectral-spatial classification method is proposed. The extended random walker (ERW) [35] algorithm is used for the spectral-spatial classification of hyperspectral images, which consists of the following steps: First, pixelwise SVM classification is used to obtain a set of probability maps which measure the probability that a pixel belongs to a particular class. Then, the ERW algorithm is adopted to calculate a set of optimized probabilities in order for the class of each pixel to be determined based on the maximum probability. In the ERW-based optimization framework, the pixelwise spectral information, the spatial information between adjacent pixels, and the distance information between training and test pixels are combined. Therefore, the proposed method can produce classification results with much higher accuracies than those obtained by the SVM method, even when the number of training samples is relatively small.

The rest of this paper is organized as follows. The ERW algorithm is described in Section II. The proposed ERW-based classification method is introduced in Section III. Experimental results are given in Section IV. Finally, conclusions are given in Section V.

## II. ERWs

The random walker (RW) algorithm was originally developed for image segmentation and has been successfully applied in a wide variety of image processing applications [36], [37]. Specifically, the algorithm models an image as a weighted graph  $G = (V, E)$  with vertices  $v \in V$  and edges  $e \in E$ , in which vertices refer to the pixels in the image and edges refer to the links connecting two adjacent pixels. Each edge  $e_{ij}$  connecting the  $i$ th and  $j$ th pixels has a weight  $\omega(e_{ij}) = e^{-\beta(v_i - v_j)^2}$  which defines the intensity difference between the two pixels. Here,  $\beta$  is a free parameter. Based on the weighted graph representation, the RW segmentation algorithm consists of the following steps: First, a set of marked pixels  $V_M$  is given, in which each pixel  $v_i \in V_M$  has been assigned a label  $n$  from the set  $\mathcal{L} = \{1, \dots, N\}$ . Second, with the marked pixels, the RW algorithm computes the probability  $\mathbf{p}_{in}$  that an RW starting its walk from a unlabeled pixel  $i$  first reaches a pixel belonging to  $V_M$  with label  $n \in \mathcal{L}$ . Finally, the remaining unlabeled pixels  $V_U = V \cap \overline{V_M}$  are assigned labels selected from  $\mathcal{L}$  which gives the greatest probabilities.

Based on the deep connection between RWs and electronic networks [37], instead of performing a random walk simulation for each pixel, the probability  $\mathbf{p}_{in}$  that an RW starting its walk from a pixel  $i$  first reaches a labeled pixel can be directly calculated by minimizing the following energy function:

$$E_{spatial}^n(\mathbf{p}_n) = \mathbf{p}_n^T L \mathbf{p}_n \quad (1)$$

where  $L$  represents the Laplacian matrix of the graph

$$L_{ij} = \begin{cases} d_i & \text{if } i = j \\ -\omega(e_{ij}) & \text{if } i \text{ and } j \text{ are adjacent pixels} \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

In this matrix,  $d_i = \sum \omega(e_{ij})$  refers to the degree of the  $i$ th pixel which is the sum of all weighed edges connecting to  $v_i$ . With labeled pixels  $V_M$ , the energy function shown in (1) has an analytical solution which can be calculated by solving a system of linear equations. The detailed solution can be found in [37]. However, one property of RWs is that this algorithm will not work if  $V_M$  is empty. More importantly, the spectral information of the hyperspectral image pixels cannot be appropriately considered in the RW-based segmentation framework. Therefore, the RW algorithm cannot be directly applied for probabilistic modeling-based classification. To solve this problem, this paper adopts an improved version of the original RW algorithm, i.e., the ERWs [35]. Specifically, aside from the spatial function shown in (1), another aspatial energy function is given as follows:

$$E_{aspatial}^n(\mathbf{p}_n) = \sum_{q=1, q \neq n}^N \mathbf{p}_q^T \Lambda_q \mathbf{p}_q + (\mathbf{p}_n - 1)^T \Lambda_n (\mathbf{p}_n - 1) \quad (3)$$

where  $\Lambda_n$  is a diagonal matrix where the values on the diagonal measure the initial probability (priors)  $\mathbf{r}_{in}$  for each pixel  $v_i$ . In [35], the initial probability is estimated according to the density distribution of pixel values (intensities) for image segmentation. Thus, the spatial and aspatial energy functions can be combined as follows:

$$E^n(\mathbf{p}_n) = E_{spatial}^n(\mathbf{p}_n) + \gamma E_{aspatial}^n(\mathbf{p}_n) \quad (4)$$

where  $\gamma$  is a free parameter. Similar to the solution of (1), the deep connection between RWs and electronic networks ensures that the probabilities  $\mathbf{p}_n$  can be estimated by simply solving a system of linear equations. The detailed solution can be found in [35]. Once the optimized probabilities are known, the last step of ERW-based segmentation is to assign each unlabeled pixel  $v_i \in V_U$  with the label  $n \in \mathcal{L}$  which gives the greatest probability.

## III. PROPOSED APPROACH

Fig. 1 shows the schematic of the proposed ERW-based spectral-spatial hyperspectral image classification method which consists of two main steps: First, SVM is adopted to estimate the initial probability maps which measure the probabilities that each pixel of a hyperspectral image belongs to different classes. Second, the probabilities are refined with the ERW to determine the class of each pixel.

### A. Initial Probability Estimation With SVM

Let  $\mathcal{L} = \{1, \dots, N\}$  be a set of labels, and let  $n \in \mathcal{L}$  refer to the  $n$ th class. Furthermore, let  $\mathcal{S} \equiv \{1, \dots, i\}$  be the set of the pixels of the hyperspectral image, and let  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_i) \in$

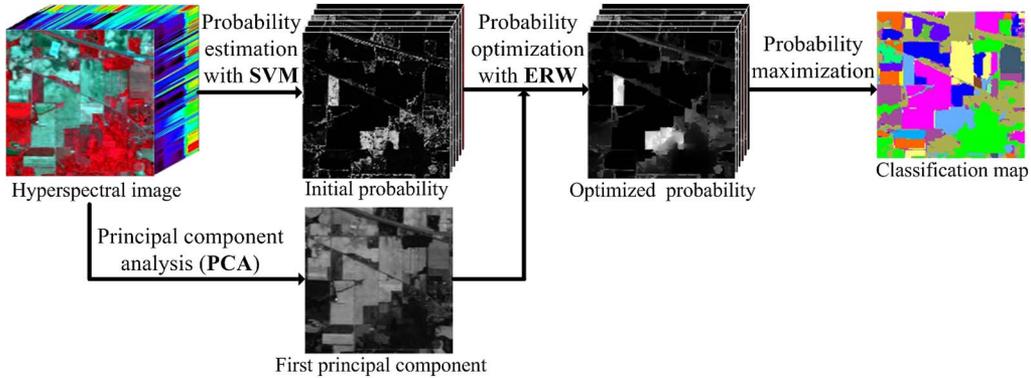


Fig. 1. Schematic of the proposed ERW-based spectral-spatial classification method.

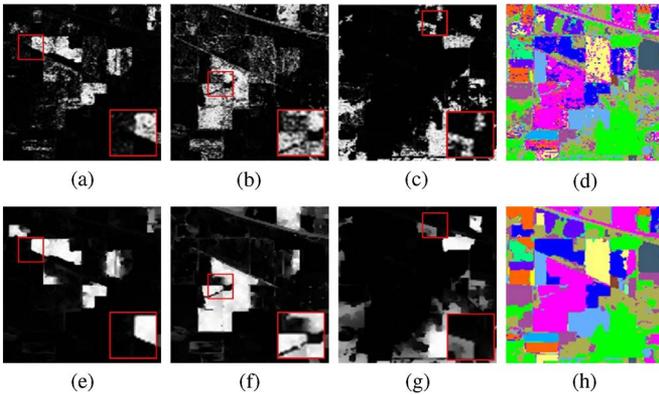


Fig. 2. (a)–(c) Three initial probability maps estimated by SVM corresponding to different types of crops in the scene. (d) Classification map obtained by SVM. (e)–(g) Optimized probability maps with the ERW. (h) Classification map obtained by the proposed ERW-based classification method. (a)  $r_1$ . (b)  $r_5$ . (c)  $r_{14}$ . (d) SVM. (e)  $p_1$ . (f)  $p_5$ . (g)  $p_{14}$ . (h) ERW.

$\mathbb{R}^{d \times i}$  represent the input hyperspectral image, in which each pixel is a  $d$ -dimensional pixel vector. Given  $\tau$  training samples  $T_\tau \equiv \{(\mathbf{x}_1, c_1), \dots, (\mathbf{x}_\tau, c_\tau)\} \in (\mathbb{R}^d \times \mathcal{L})^\tau$ , by using the probability estimation function of the SVM classifier [38], the probability maps  $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_n)$  can be obtained, in which  $r_{in} \in [0, 1]$  is the probability that a pixel  $i$  belongs to the  $n$ th class. From Fig. 2(a)–(c), it can be seen that the probability maps estimated by pixelwise SVM usually look noisy since spatial information is not considered. Thus, the corresponding classification map obtained by choosing the maximum of probability [see Fig. 2(d)] also contains serious noise.

### B. Probability Optimization With ERW

In order to make full use of the available spatial information, this paper introduces the ERW algorithm for probability optimization. The ERW-based optimization method consists of two main steps: First, the spectral dimension of the hyperspectral image is reduced with PCA in order to construct a weighted graph  $G = (V, E)$ , which can well preserve the spatial and structure information of the hyperspectral image in the mean squared sense. In this graph,  $V$  is a set of pixels in the first principal component.  $E$  is a set of edges which link the pixels of the first principal component (eight neighbors are considered for each pixel). A weight  $\omega(e_{ij}) = e^{-\beta(v_i - v_j)^2}$  is defined for each edge  $e_{ij}$  to model the intensity difference between the

adjacent pixels in the weighted graph. Then, the optimized probabilities are obtained by minimizing the energy function presented in (4), in which the first term  $E_{spatial}^n$  models the spatial correlation among adjacent pixels and the second term  $E_{aspatial}^n$  integrates the initial probability maps  $\mathbf{r}_n$  estimated by the pixelwise SVM.

The combination of spatial information with a pixelwise classification approach is not new in the literature. Energy-based spectral-spatial classification methods model the spatial information by defining a spatial energy term and perform minimization on the total energy, e.g., using the widely used Markov random field (MRF) method. However, the conventional MRF is based on the assumption that the adjacent pixels (four or eight neighbors) should have the same class labels as the central pixel. Furthermore, in the optimization process, each pixel uses the same weighting coefficient for its spatial term regardless if it is within a homogeneous region or on an object boundary. As a result, although classification accuracy is improved in homogeneous regions, pixels at class boundaries may be wrongly classified [39]. In contrast to the MRF method, the ERW method assumes that only these adjacent pixels that have similar intensity values should have the same class labels as the central pixel. In other words, pixels on the same side of an edge will tend to have similar labels because these pixels tend to have similar pixel intensities. Therefore, the ERW method not only improves the accuracy in homogeneous regions but also ensures that the optimized probabilities align well with real object boundaries.

For example, Fig. 2(e)–(g) shows three optimized probability maps corresponding to three different landscapes, respectively. From the figures, it can be seen that the optimized probabilities have two properties: First, errors in homogeneous regions can be effectively removed. Second, optimized probabilities are usually aligned well with real object boundaries. Fig. 2(h) shows the corresponding classification map obtained by labeling each pixel with the class which gives the highest probability. It can be seen that the accuracy of SVM can be effectively improved for those pixels in homogeneous regions and those near class boundaries. Furthermore, the comparison between Fig. 2(d) and (h) shows that some objects can be classified more accurately after the optimization. The reason is that, aside from the spatial correlation among adjacent pixels, the connectedness between training and test samples is also considered in the ERW-based method.

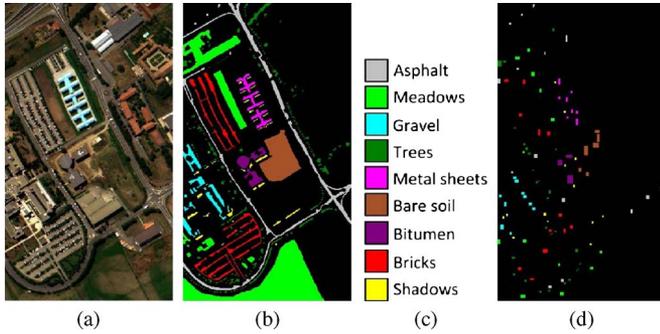


Fig. 3. University of Pavia data set. (a) Three-band color composite. (b) Reference data. (c) Color code. (d) Publicly available training set which consists of 3921 training samples.

#### IV. EXPERIMENTS

##### A. Data Sets

In this paper, the performance of different classification methods is evaluated using three hyperspectral data sets. The first hyperspectral image was acquired with the Reflective Optics System Imaging Spectrometer (ROSIS-03) optical sensor which provides 115 bands with a spectral coverage ranging from 0.43 to 0.86  $\mu\text{m}$ . The other two images were acquired by the Airborne Visible/Infrared Imaging Spectrometer sensor with a spectral coverage ranging from 0.4 to 2.5  $\mu\text{m}$ . The three data sets are detailed as follows.

1) *University of Pavia Data Set*: The University of Pavia test set is of the campus at the University of Pavia, Pavia, Italy. It is of size  $610 \times 340$  pixels. The spatial resolution is 1.3 m per pixel. Before classification, 12 bands were removed due to noise. Fig. 3 shows the three-band false color image and the reference data of the University of Pavia data set, in which nine classes of interest are considered: trees, asphalt, bitumen, gravel, metal sheets, shadow, bricks, meadows, and bare soil. Furthermore, Fig. 3 shows a manually selected training set which is publicly available and widely used in related publications [10], [14], [39], [40].

2) *Salinas Data Set*: The second test set is the Salinas image which was captured at Salinas Valley, California. This image contains 224 bands of size  $512 \times 217$ . Twenty water absorption bands were discarded before classification. Fig. 4 shows the three-band false color image and the reference data of the Salinas data set, in which the reference data contains 16 classes of interest, which represent mostly different types of crops.

3) *Indian Pines Data Set*: The third test set took place over the agricultural Indian Pine test site in Northwestern Indiana. It was  $145 \times 145$  pixels. The spatial resolution is 20 m per pixel. Twelve channels were removed due to noise. The remaining 200 spectral channels were processed. A three-band false color image and the reference data which contains 16 classes of interest are presented in Fig. 5.

##### B. Parameter Settings

In the following experiments, the proposed method is compared to other widely used spectral-spatial classification methods, including the original SVM [9], extended morphological profiles (EMPs) [15], logistic regression via variable splitting

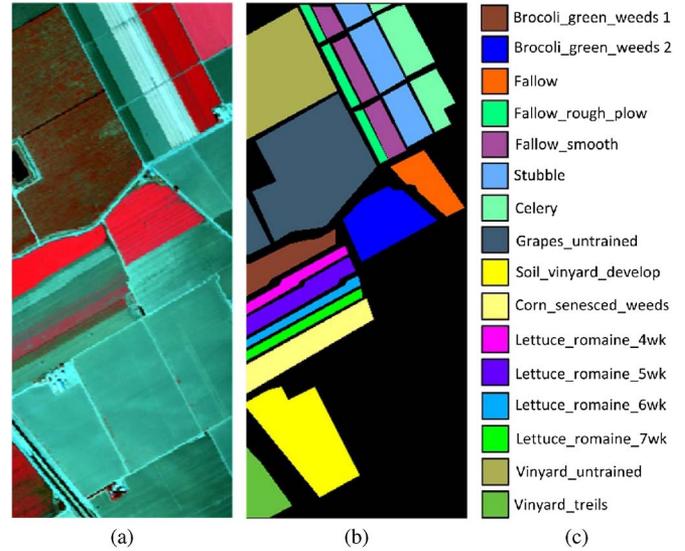


Fig. 4. Salinas data set. (a) Three-band color composite. (b) Reference data. (c) Color code.

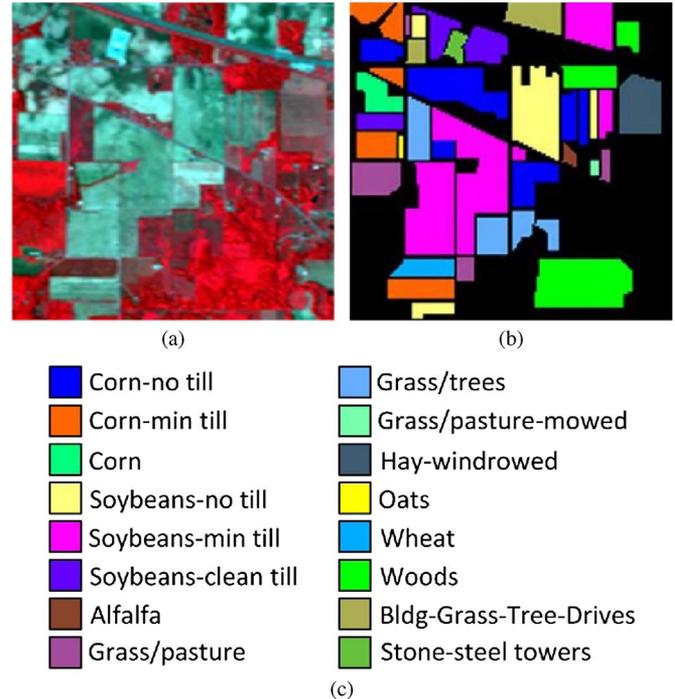


Fig. 5. Indian Pines data set. (a) Three-band color composite. (b) Reference data. (c) Color code.

and augmented Lagrangian-multilevel logistic (LORSAL-MLL) [10], maximizer of the posterior marginal by loopy belief propagation (MPM-LBP) in [40], edge-preserving filtering (EPF) [31], and image fusion and recursive filtering (IFRF) [20]. The SVM method is implemented using the functions provided by the library for support vector machines [38]. The parameters for SVM were selected using a fivefold cross-validation. The EMPs are constructed using the first four principal components and a circular structural element with a step size increment of two. Four openings and closings were computed for each principal component [14]. Furthermore, the LORSAL-MLL, MPM-LBP, EPF, and IFRF methods are respectively implemented using the default parameters given in [10], [20], [31], and [40].

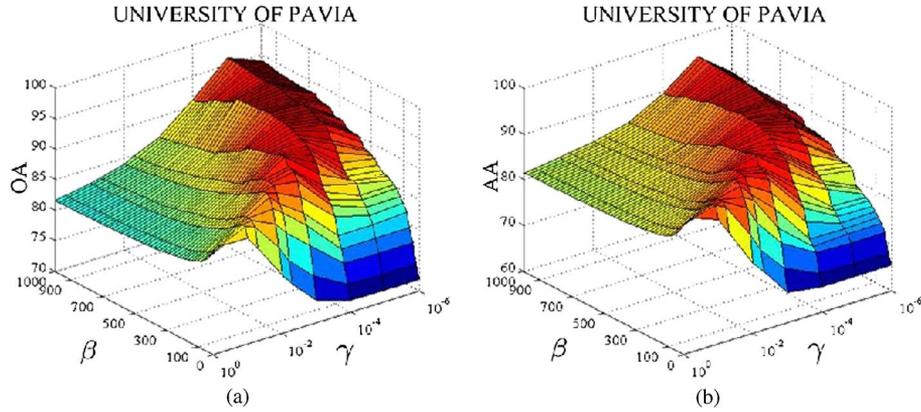


Fig. 6. Influence of the parameters on classification accuracies with different values of  $\beta$  and  $\gamma$  (University of Pavia data set). (a) OA and (b) AA.

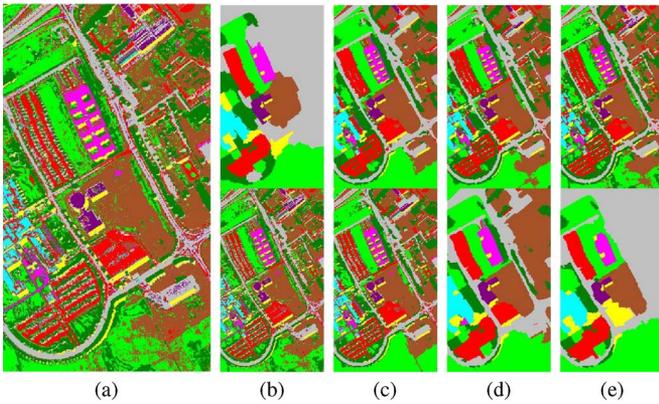


Fig. 7. Sensitivity analysis for the free parameters  $\beta$  and  $\gamma$  for the University of Pavia image. (a) is the classification map obtained by the SVM method. In the top row of (b)–(e),  $\gamma$  is fixed as  $10^{-5}$ . (b)–(e) respectively show the classification maps obtained by the proposed ERW method with (b)  $\beta = 100$ , (c)  $\beta = 700$ , (d)  $\beta = 10^3$ , and (e)  $\beta = 10^4$ . In the bottom row of (b)–(e),  $\beta$  is fixed as 700. (b)–(e) respectively show the classification maps obtained with (b)  $\gamma = 10^{-1}$ , (c)  $\gamma = 10^{-4}$ , (d)  $\gamma = 10^{-6}$ , and (e)  $\gamma = 10^{-7}$ .

For the proposed ERW method, the influence of parameters  $\beta$  and  $\gamma$  is evaluated by objective and visual analysis. A publicly available training and test set which consists of 3921 training and 40 002 test samples is adopted for the University of Pavia image (see Fig. 3). Fig. 6 shows the overall accuracy (OA) and the average accuracy (AA) of the proposed method when the two parameters  $\beta$  and  $\gamma$  vary. Furthermore, Fig. 7 shows the classification maps obtained by the SVM method and the proposed method with different values of  $\beta$  and  $\gamma$ . From Fig. 6, it can be first seen that, if  $\beta$  is smaller than 30, the OA and AA of the ERW method may decrease significantly. Moreover, the first row of Fig. 7(b) also shows that the classification map becomes oversmoothed as  $\beta$  becomes too small. The reason is that the value of the graph weight  $\omega(e_{ij}) = e^{-\beta(v_i - v_j)^2}$  for adjacent pixels tends to be one when  $\beta$  is quite small. In this situation, the spatial correlation between adjacent pixels cannot be well modeled by the edge weights. Similarly, if  $\beta$  is quite large, the weight  $\omega(e)$  will be close to zero. In this situation, the parameter  $\gamma$  which controls the dynamic range of the aspatial function is also suggested to be relatively small in order to ensure the balance of the spatial and aspatial energy functions shown in (4). Otherwise, when  $\gamma$  is quite large, the aspatial function will play a major role in classification, and thus, the

TABLE I  
GLOBAL CLASSIFICATION ACCURACIES IN PERCENTAGE FOR THE UNIVERSITY OF PAVIA IMAGE. THE RESULTS OF THE LORSAL-MLL [10], MPM-LBP [40], AND SVM-MRF METHODS (SVM-MRF-NE AND SVM-MRF-E) [39] REPORTED IN THE TABLE ARE RESPECTIVELY TAKEN FROM [10], [40], AND [39]. THESE METHODS USED THE SAME TRAINING AND TEST SET, THUS ALLOWING A FAIR COMPARISON

Methods	OA	AA	Kappa
SVM	80.99	88.28	76.16
LORSAL	80.11	87.70	75.09
SVM-MRF-NE	86.89	92.12	83.14
SVM-MRF-E	87.63	93.41	84.07
LORSAL-MLL	85.57	92.54	81.80
MPM-LBP	85.78	92.20	82.05
EPF	87.83	89.24	84.26
ERW	95.97	95.06	94.56
LORSAL-ERW	96.60	96.28	95.40

classification map looks quite similar to the classification map obtained by SVM [see the bottom row of Fig. 7(b), and compare it to Fig. 7(a)]. However, if  $\gamma$  is too small, the spatial effect may lead to the overcorrected classification maps shown in the bottom row of Fig. 7(d) and (e). Based on the aforementioned analysis in this paper,  $\beta = 710$  and  $\gamma = 10^{-5}$  are set as the default parameters for the following experiments in order to ensure the balance of the spatial and aspatial energy functions shown in (4).

### C. Experiments With the University of Pavia Data Set

In the first experiment, the proposed ERW algorithm is compared with other probabilistic modeling-based spectral-spatial classification methods, including the EPF [31], LORSAL-MLL [10], MPM-LBP [40], and SVM-MRF [39] methods. These methods are all based on the optimization of pixelwise classification results in order to improve the classification accuracy. Table I shows the OA, AA, and Kappa coefficient (Kappa) of the different methods, which are three widely used quality metrics for hyperspectral image classification [4], [14]. In this table, the results of the LORSAL-MLL, MPM-LBP, and SVM-MRF methods reported in the table are taken from [10], [40], and [39], respectively, where exactly the same training and test sets were used to produce the results and, thus, a fair comparison of these methods can be done. From Table I, it can be observed that

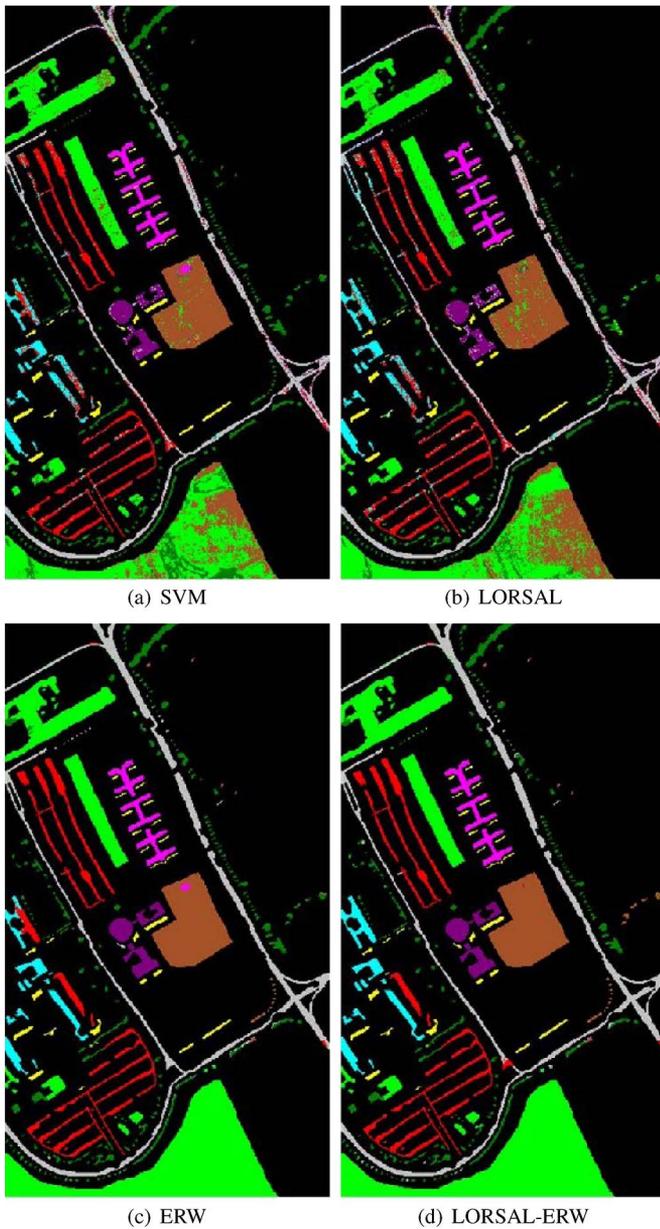


Fig. 8. Classification results (University of Pavia image) obtained by (a) SVM method, (b) LORSAL method, (c) ERW method, and (d) LORSAL-ERW method.

all spectral–spatial based algorithms exhibit better performance than SVM in terms of accuracies, indicating the importance of including spatial information in the probabilistic framework. Furthermore, the proposed ERW method provides accuracies which are comparable to those obtained by other probabilistic modeling-based spectral–spatial methods.

In the second experiment, the proposed method is used to improve the accuracies of other types of pixelwise classifiers such as the logistic regression via variable splitting and augmented Lagrangian (LORSAL) classifier [10]. Specifically, instead of estimating the initial probability with the SVM, the LORSAL method can also be adopted for the estimation of initial probability  $r$  in Section III-A. Fig. 8 shows the classification results obtained by the SVM, LORSAL, ERW, and LORSAL-ERW methods, respectively. It can be seen that both the ERW and LORSAL-ERW methods are able to improve the performance

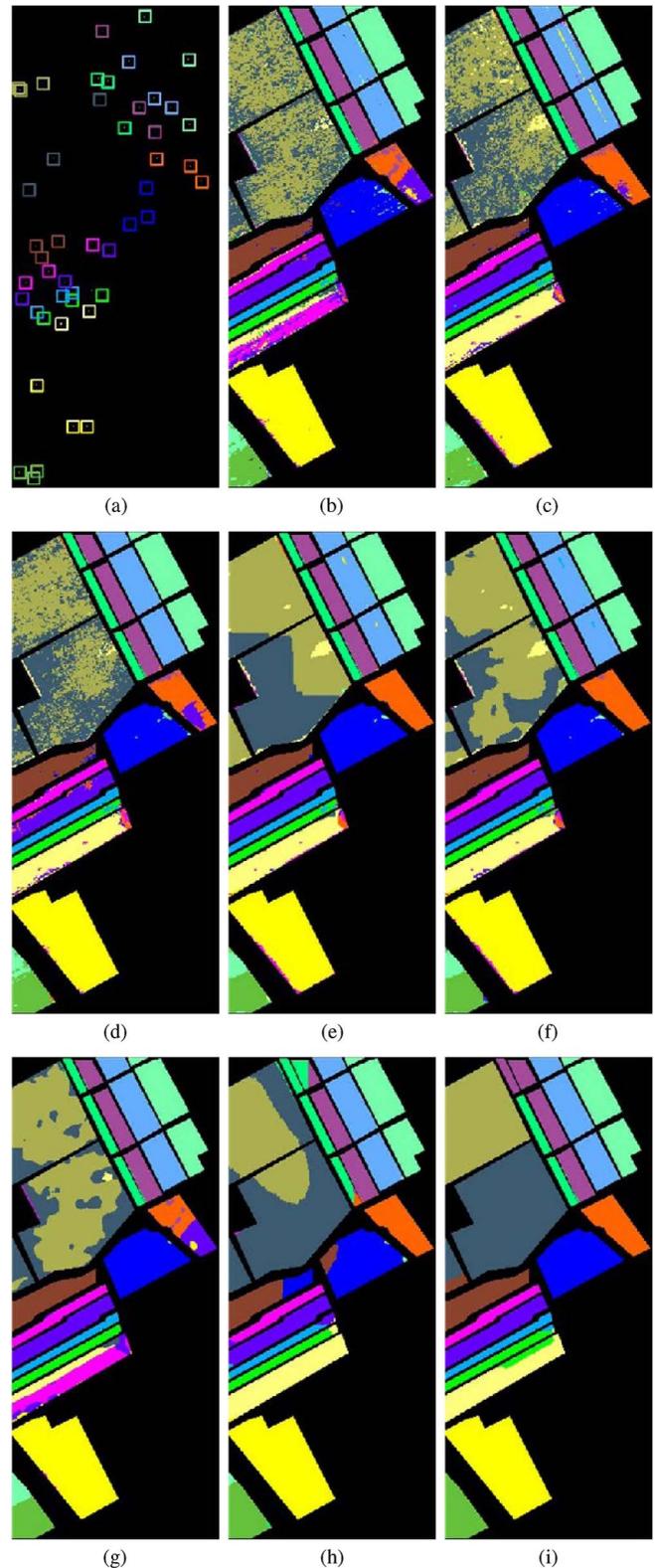


Fig. 9. (a) Training set which has three training pixels for each class, and the classification results (Salinas image) obtained by (b) SVM method, (c) LORSAL method, (d) EMP method, (e) LORSAL-MLL method, (f) MPM-LBP method, (g) EPF method, (h) IFRF method, and (i) ERW method. Numbers in the parentheses refer to the overall classification accuracies given in percent. (a) Training set. (b) SVM (74.1). (c) LORSAL (81.9). (d) EMP (85.1). (e) LORSAL-MLL (88.0). (f) MPM-LBP (82.6). (g) EPF (76.9). (h) IFRF (89.1). (i) ERW (98.0).

TABLE II  
INDIVIDUAL CLASS ACCURACIES, OA, AA, AND KAPPA COEFFICIENT (IN PERCENT) OF THE SVM [9], LORSAL [10], EMP [15], LORSAL-MLL [10], MPM-LBP [40], EPF [31], IFRF [20], AND ERW METHODS FOR THE SALINAS DATA SET

Class	Methods									
	Training	Test	SVM	LORSAL	EMP	LORSAL-MLL	MPM-LBP	EPF	IFRF	ERW
Weeds_1	3	2006	98.20	97.46	<b>100</b>	99.10	99.35	<b>100</b>	83.31	88.60
Weeds_2	3	3723	99.01	98.47	98.33	99.22	98.74	<b>100</b>	87.09	<b>100</b>
Fallow	3	1973	91.03	85.71	82.81	99.80	97.87	99.53	95.68	<b>99.95</b>
Fallow_r	3	1391	95.91	<b>99.57</b>	97.67	99.71	99.64	96.00	76.79	97.51
Fallow_s	3	2675	90.12	98.47	94.96	98.62	98.62	95.99	99.27	88.54
Stubble	3	3956	<b>100</b>	94.59	<b>100</b>	98.63	97.83	<b>100</b>	<b>100</b>	<b>100</b>
Celery	3	3576	84.68	99.92	88.90	99.94	<b>99.97</b>	89.11	97.17	<b>99.97</b>
Grapes_u	3	11268	63.68	55.80	78.15	60.57	39.53	73.54	81.19	<b>99.96</b>
Soil_v	3	6200	95.06	96.90	99.32	97.24	97.55	98.24	99.31	<b>100</b>
Corn_s	3	3275	72.10	82.56	91.90	89.62	87.57	92.42	97.64	<b>99.76</b>
Lettuce_4wk	3	1065	35.92	90.05	80.45	91.36	91.27	34.80	99.02	<b>100</b>
Lettuce_5wk	3	1924	58.07	99.64	75.83	<b>100</b>	<b>100</b>	61.83	91.24	<b>100</b>
Lettuce_6wk	3	913	89.92	98.58	93.71	98.69	98.58	<b>100</b>	97.43	94.68
Lettuce_7wk	3	1067	84.67	89.97	97.70	92.78	91.75	<b>98.60</b>	98.23	72.12
Vinyard_u	3	7265	43.23	63.68	57.43	89.04	83.55	45.41	71.73	<b>100</b>
Vinyard_v	3	1804	94.02	72.28	98.39	76.77	76.94	<b>100</b>	<b>100</b>	<b>100</b>
OA			74.16	81.93	85.10	87.98	82.60	76.94	89.08	<b>98.00</b>
AA			80.98	88.98	88.82	93.19	91.17	86.59	92.19	<b>96.32</b>
Kappa			71.44	80.00	83.47	86.71	80.82	74.58	87.84	<b>97.77</b>

of pixelwise classification. Table I shows that the LORSAL-ERW is able to obtain OA = 96.60%, AA = 96.28%, and Kappa = 95.40% which are even higher than those obtained by the ERW for this example.

Finally, the computing time of the ERW-based optimization is also evaluated using MATLAB on a computer with 3.5-GHz CPU and 8-GB memory. It is found that the ERW-based optimization method is relatively fast, taking about 1.9 s for processing the University of Pavia image. The reason is that, although a global energy function should be solved for the proposed method, the function presented in (4) actually has a closed form analytic solution which can be easily calculated. Therefore, it is possible to conclude that the proposed ERW method is indeed a good candidate for spectral-spatial hyper-spectral image classification due to its advantage in improving classification accuracy and its computing efficiency.

#### D. Experiments With the Salinas Data Set

An experiment is performed on the Salinas image which contains 16 different classes. Fig. 9(a) shows the labeled 48 samples used for training (each class is trained with only three labeled samples). Fig. 9(b)–(i) shows the classification results obtained by different methods, including the SVM, LORSAL, EMP, LORSAL-MLL, MPM-LBP, EPF, and IFRF methods. It can be seen that, compared with other widely used spectral-spatial methods, the classification map obtained by the ERW method looks more similar to the reference data shown in Fig. 4(b). Furthermore, Table II records the individual classification accuracies, OA, AA, and Kappa of different methods. It can be seen that the proposed method gives the highest classification accuracies for most classes. Moreover, the proposed method also shows the best performance in terms of OA = 98.00%, AA = 96.32%, and Kappa = 97.77%. This experiment demonstrates that the ERW method works much better than traditional spectral-spatial methods in terms of accuracies, particularly when the number of training samples is limited.

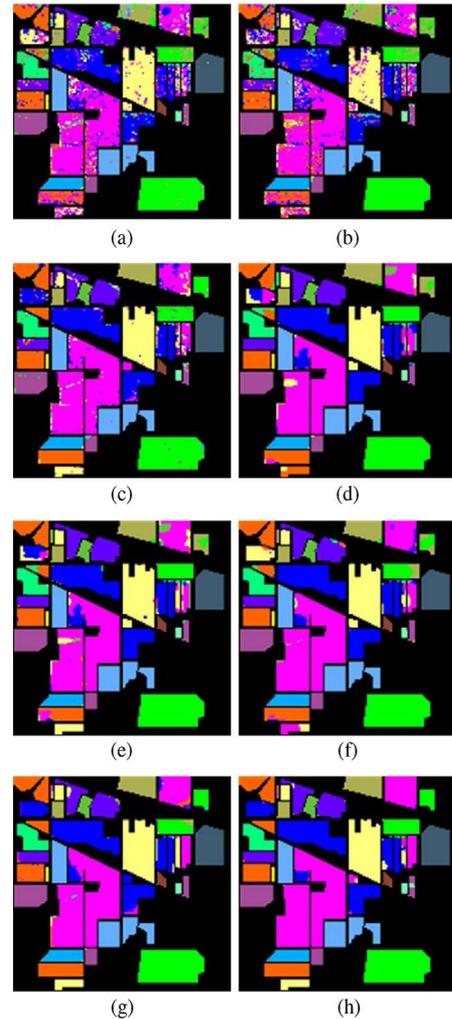


Fig. 10. Classification results (Indian Pines image) obtained by (a) SVM method, (b) LORSAL method, (c) EMP method, (d) LORSAL-MLL method, (e) MPM-LBP method, (f) EPF method, (g) IFRF method, and (h) ERW method. Numbers in the parentheses refer to the overall classification accuracies given in percent. (a) SVM (81.3). (b) LORSAL (81.7). (c) EMP (93.6). (d) LORSAL-MLL (92.1). (e) MPM-LBP (93.1). (f) EPF (93.2). (g) IFRF (96.4). (h) ERW (98.7).

TABLE III  
CLASSIFICATION ACCURACIES (IN PERCENT) FOR THE SVM [9], LORSAL [10], EMP [15], LORSAL-MLL [10], MPM-LBP [40], EPF [31], IFRF [20], AND ERW METHODS AS AN AVERAGE AFTER TEN REPEATED EXPERIMENTS. NUMBERS IN PARENTHESES REFER TO THE VARIANCE OF THE REPEATED EXPERIMENTS

Methods	Metrics	Number of Training Samples								
		1%	2%	5%	7%	10%	15%	20%	25%	30%
SVM	OA	52.76(3.7)	61.80(3.2)	74.38(1.3)	77.80(1.5)	80.86(1.4)	83.42(0.8)	85.95(0.8)	87.42(0.4)	88.36(0.6)
	AA	52.27(2.4)	58.88(2.2)	70.93(1.4)	74.65(1.8)	78.68(1.8)	81.20(1.1)	84.63(1.0)	85.59(0.8)	86.98(1.2)
	Kappa	47.10(3.9)	57.05(3.5)	70.01(1.3)	74.74(1.7)	78.18(1.5)	80.98(0.9)	83.76(0.9)	85.42(0.5)	86.42(0.7)
LORSAL	OA	58.44(3.6)	66.34(2.3)	76.67(1.1)	79.45(1.2)	81.83(0.5)	83.97(0.4)	86.05(0.5)	87.17(0.4)	87.80(0.4)
	AA	69.33(2.4)	77.04(1.3)	85.38(1.0)	87.14(0.7)	87.17(1.4)	89.06(0.7)	90.07(0.7)	91.21(1.0)	90.06(1.2)
	Kappa	53.60(3.8)	62.10(2.4)	73.60(1.2)	76.60(1.3)	79.26(0.5)	81.70(0.4)	83.88(0.5)	85.12(0.4)	85.77(0.4)
EMP	OA	63.97(4.2)	76.11(3.2)	87.25(1.5)	90.62(0.8)	92.35(0.9)	93.85(0.5)	95.00(0.4)	95.62(0.3)	96.08(0.2)
	AA	66.79(3.1)	74.54(2.4)	86.24(1.5)	87.57(1.5)	90.66(1.3)	92.69(1.0)	93.51(0.8)	94.34(1.1)	95.32(0.9)
	Kappa	59.65(4.5)	73.08(3.6)	85.49(1.7)	89.27(0.9)	91.21(1.0)	92.89(0.6)	94.19(0.4)	94.89(0.3)	95.40(0.2)
LORSAL-MLL	OA	66.08(5.1)	76.47(3.4)	88.13(2.3)	91.13(1.8)	93.10(1.3)	94.66(0.7)	96.19(0.8)	96.53(0.6)	97.24(0.6)
	AA	76.06(3.5)	84.94(2.0)	93.53(0.9)	94.91(1.0)	95.88(0.8)	97.02(0.6)	97.43(0.5)	97.98(0.6)	98.19(0.6)
	Kappa	62.11(5.5)	73.42(3.7)	86.52(2.5)	89.85(2.1)	92.10(1.4)	93.85(0.8)	95.59(1.0)	95.97(0.7)	96.77(0.6)
MPM-LBP	OA	60.16(5.5)	72.27(2.9)	86.67(1.7)	90.75(1.2)	93.12(1.0)	94.99(0.5)	96.63(0.4)	97.32(0.4)	97.73(0.4)
	AA	72.18(3.3)	81.45(1.3)	92.60(0.8)	94.86(0.5)	96.08(0.5)	97.09(0.5)	97.74(0.4)	98.36(0.2)	98.37(0.7)
	Kappa	56.11(5.7)	68.98(3.0)	84.89(1.9)	89.43(1.4)	92.12(1.1)	94.22(0.5)	96.09(0.4)	96.88(0.5)	97.34(0.5)
EPF	OA	65.10(5.2)	76.38(3.7)	89.24(0.7)	92.58(1.4)	94.46(1.0)	95.23(0.8)	96.38(1.2)	97.25(0.7)	98.00(0.8)
	AA	67.12(4.2)	79.36(3.3)	90.65(1.1)	93.34(1.0)	94.11(0.9)	94.79(0.6)	95.98(0.8)	96.69(0.8)	97.87(0.5)
	Kappa	60.69(5.6)	73.28(4.1)	87.75(0.8)	91.49(1.6)	93.64(1.1)	94.49(0.9)	95.80(1.3)	96.79(0.8)	97.66(0.9)
IFRF	OA	72.45(3.1)	84.75(2.2)	93.34(1.4)	96.19(0.5)	97.35(0.8)	98.46(0.3)	98.83(0.2)	98.82(0.3)	99.14(0.3)
	AA	69.81(2.6)	79.97(3.9)	93.03(1.4)	94.94(1.4)	96.58(1.2)	98.13(0.6)	98.46(0.4)	98.54(0.5)	98.96(0.5)
	Kappa	69.12(3.3)	82.73(2.5)	92.40(1.6)	95.62(0.5)	96.95(0.9)	98.22(0.4)	98.63(0.2)	98.62(0.3)	99.00(0.3)
ERW	OA	<b>79.33(3.9)</b>	<b>88.82(2.2)</b>	<b>95.95(1.3)</b>	<b>97.54(0.6)</b>	<b>98.44(0.4)</b>	<b>98.92(0.2)</b>	<b>99.15(0.2)</b>	<b>99.21(0.2)</b>	<b>99.40(0.1)</b>
	AA	<b>80.63(3.5)</b>	<b>86.02(3.5)</b>	<b>95.15(1.7)</b>	<b>96.26(1.7)</b>	<b>96.64(1.1)</b>	<b>98.41(0.8)</b>	<b>98.62(1.0)</b>	<b>98.73(1.0)</b>	<b>99.18(0.6)</b>
	Kappa	<b>76.65(4.2)</b>	<b>87.26(2.5)</b>	<b>95.37(1.5)</b>	<b>97.17(0.7)</b>	<b>98.20(0.5)</b>	<b>98.74(0.2)</b>	<b>99.01(0.2)</b>	<b>99.07(0.2)</b>	<b>99.30(0.2)</b>

### E. Experiments With the Indian Pines Data Set

In this section, an experiment is performed on the Indian Pines image. Fig. 10 shows the classification results obtained by different methods when the number of training samples is 10% of 10 249 reference samples (see Fig. 5). From this figure, it can be seen that the proposed method can effectively remove the noise in the SVM classification result. Furthermore, the ERW method also gives the highest OA = 98.7%. In order to objectively evaluate the performance of different methods, the experiment presented in Fig. 10 has been repeated ten times to estimate the mean and standard variance of OA, AA, and Kappa. Furthermore, experiments are also performed with different numbers of training samples (1%–30% of the reference data). As shown in Table III, the ERW method always outperforms other methods and even more significantly when there is a smaller number of labeled samples used for training. For example, when the number of training samples is 1% of the reference, the overall classification accuracy obtained by SVM is 52.76%. In this situation, the proposed method can still dramatically increase the accuracies of SVM. Specifically, the OA, AA, and Kappa obtained by ERW are 79.33%, 80.63%, and 76.65% percentages which are very close to those obtained by SVM (OA = 80.86%, AA = 78.68%, and Kappa = 78.18%) when the number of training samples is 10% of the reference samples.

## V. CONCLUSION

In this paper, a novel spectral–spatial image classification method based on ERWs is introduced. Through integrating spatial information in the ERW-based probabilistic optimization framework, the accuracy of SVM can be improved dramatically. Experiments performed on three widely used real

hyperspectral data sets demonstrate the outstanding performance of the proposed ERW method. Compared with other spectral–spatial classification methods, the major advantage of the proposed ERW method is that it is able to obtain high classification accuracies when the number of training samples is very small. Furthermore, since the energy function defined by ERW has a closed form analytic solution which can easily be determined, the method is relatively efficient computationally. Thus, the proposed method will be quite useful in real applications due to its high accuracy and low computational burden. Since the ERW method works well for classification with a small number of training samples, a topic of future research is to investigate whether the ERW method can be integrated into a semisupervised learning-based classification framework. Furthermore, the use of the ERW method in other remote sensing applications will also be further investigated.

## ACKNOWLEDGMENT

The authors would like to thank the Editor-in-Chief, the anonymous Associate Editor, and the reviewers for their insightful comments and suggestions which have greatly improved this paper.

## REFERENCES

- [1] J. Pontius, M. Martin, L. Plourde, and R. Hallett, "Ash decline assessment in emerald ash borer-infested regions: A test of tree-level, hyperspectral technologies," *Remote Sens. Environ.*, vol. 112, no. 5, pp. 2665–2676, May 2008.
- [2] J. L. Boggs, T. D. Tsegaye, T. L. Coleman, K. C. Reddy, and A. Fahsi, "Relationship between hyperspectral reflectance, soil nitrate-nitrogen, cotton leaf chlorophyll, and cotton yield: A step toward precision agriculture," *Remote Sens. Environ.*, vol. 22, no. 3, pp. 5–16, 2003.
- [3] G. Lampropoulos, T. Liu, S. Qian, and C. Fei, "Hyperspectral classification fusion for classifying different military targets," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2008, vol. 3, pp. 262–265.

- [4] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. Supplement 1, pp. S110–S122, Sep. 2009.
- [5] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 1, pp. 55–63, Jan. 1968.
- [6] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, "Hyperspectral image classification with independent component discriminant analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 4865–4876, Dec. 2011.
- [7] S. Prasad and L. M. Bruce, "Limitations of principal components analysis for hyperspectral target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 625–629, Oct. 2008.
- [8] 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.*, vol. 50, no. 4, pp. 1185–1198, Apr. 2012.
- [9] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.
- [10] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new Bayesian approach with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3947–3960, Oct. 2011.
- [11] W. Liao, A. Pizurica, P. Scheunders, W. Philips, and Y. Pi, "Semisupervised local discriminant analysis for feature extraction in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 184–198, Jan. 2013.
- [12] 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.*, vol. 8, no. 5, pp. 894–898, Sep. 2011.
- [13] Y. Zhong and L. Zhang, "An adaptive artificial immune network for supervised classification of multi-/hyperspectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 894–909, Mar. 2012.
- [14] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.
- [15] J. A. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.
- [16] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 466–479, Mar. 2005.
- [17] 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.
- [18] A. Erturk, M. Gullu, and S. Erturk, "Hyperspectral image classification using empirical mode decomposition with spectral gradient enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2787–2798, May 2013.
- [19] B. Demir and S. Erturk, "Empirical mode decomposition of hyperspectral images for support vector machine classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4071–4084, Nov. 2010.
- [20] X. Kang, S. Li, and J. A. Benediktsson, "Feature extraction of hyperspectral images with image fusion and recursive filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3742–3752, Jun. 2014.
- [21] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Tensor discriminative locality alignment for hyperspectral image spectral-spatial feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 242–256, Jan. 2013.
- [22] M. Pedergnana, P. Marpu, M. Mura, J. A. Benediktsson, and L. Bruzzone, "A novel technique for optimal feature selection in attribute profiles based on genetic algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3514–3528, Jun. 2013.
- [23] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recogn.*, vol. 43, no. 7, pp. 2367–2379, Jul. 2010.
- [24] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.
- [25] Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Multiple spectral-spatial classification approach for hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4122–4132, Nov. 2010.
- [26] K. Bernard, Y. Tarabalka, J. Angulo, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2008–2021, Apr. 2012.
- [27] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers," *IEEE Trans. Syst. Man Cybern. B*, vol. 40, no. 5, pp. 1267–1279, Oct. 2010.
- [28] J. Li, H. Zhang, Y. Huang, and L. Zhang, "Hyperspectral image classification by nonlocal joint collaborative representation with a locally adaptive dictionary," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3707–3719, Jun. 2014.
- [29] M. Fauvel, J. Chanussot, J. A. Benediktsson, and A. Villa, "Parsimonious Mahalanobis kernel for the classification of high dimensional data," *Pattern Recogn.*, vol. 46, no. 3, pp. 845–854, Mar. 2013.
- [30] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial-spectral kernel-based approach for the classification of remote-sensing images," *Pattern Recogn.*, vol. 45, no. 1, pp. 381–392, Jan. 2012.
- [31] X. Kang, S. Li, and J. A. Benediktsson, "Spectral-spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.
- [32] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, Mar. 2012.
- [33] Q. Jackson and D. Landgrebe, "Adaptive Bayesian contextual classification based on Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2454–2463, Nov. 2002.
- [34] G. Moser and S. B. Serpico, "Combining support vector machines and Markov random fields in an integrated framework for contextual image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2734–2752, May 2013.
- [35] L. Grady, "Multilabel random walker image segmentation using prior models," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recogn.*, Jun. 2005, vol. 1, pp. 763–770.
- [36] R. Shen, I. Cheng, J. Shi, and A. Basu, "Generalized random walks for fusion of multi-exposure images," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3634–3646, Dec. 2011.
- [37] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006.
- [38] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, Apr. 2011.
- [39] 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.
- [40] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial classification of hyperspectral data using loopy belief propagation and active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 844–856, Feb. 2013.



**Xudong Kang** (S'13) received the B.Sc. degree from Northeast University, Shenyang, China, in 2007. He is currently working toward the Ph.D. degree in electrical engineering of Hunan University, Changsha, China. In 2012–2013, he was working as a visiting Ph.D. student in electrical engineering at the University of Iceland, Reykjavik, Iceland.

He is engaged in image fusion, image super-resolution, pansharpening, and hyperspectral image classification.

Mr. Kang was the recipient of the Scholarship Award for Excellent Doctoral Student granted by the Chinese Ministry of Education, the Presidential Scholarship granted by Hunan University, and the National Graduate Scholarship.

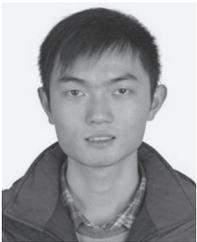


**Shutao Li** (M'07) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the Hunan University, Changsha, China, in 1995, 1997, and 2001, respectively.

He joined the College of Electrical and Information Engineering, Hunan University, in 2001. He was a Research Associate in the Department of Computer Science, Hong Kong University of Science and Technology, Kowloon, Hong Kong, from May 2001 to October 2001. From November 2002 to November 2003, he was a Postdoctoral Fellow at the Royal

Holloway College, University of London, London, U.K., working with Prof. John Shawe-Taylor. During April 2005 to June 2005, he visited the Department of Computer Science, Hong Kong University of Science and Technology, as a Visiting Professor. He is currently a Full Professor with the College of Electrical and Information Engineering, Hunan University. He has authored or coauthored more than 160 refereed papers. His professional interests are compressive sensing, sparse representation, image processing, and pattern recognition.

Dr. Li is on the Editorial Board of the journal *Information Fusion and the Sensing and Imaging*. He has won two 2nd-Grade National Awards at the Science and Technology Progress of China in 2004 and 2006.



**Leyuan Fang** (S'10) received the B.S. degree in electrical engineering from the Hunan University of Science and Technology, Changsha, China, in 2008. He joined the College of Electrical and Information Engineering, Hunan University, in 2008 for the Ph.D. degree program. In 2011–2013, he was working as a visiting Ph.D. student in the Department of Ophthalmology, Duke University, Durham, NC, USA, supported by the China Scholarship Council.

His research interests include sparse representation and multiresolution analysis applied to image

processing.

Mr. Fang as the recipient of the Scholarship Award for Excellent Doctoral Student granted by the Chinese Ministry of Education in 2011.



**Meixiu Li** received the B.Sc degree from the Hunan Institute of Engineering, Xiangtan, China, in 2012. She is currently working toward the M.Sc. degree in electrical engineering at Hunan University, Changsha, China.

Her research interest is spectral–spatial hyperspectral image classification.



**Jón Atli Benediktsson** (S'84–M'90–SM'99–F'04) received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, Iceland, in 1984 and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1987 and 1990, respectively.

He is currently the Prorector of Academic Affairs and a Professor of Electrical and Computer Engineering at the University of Iceland. His research interests are in remote sensing, image analysis, pattern

recognition, biomedical analysis of signals, and signal processing, and he has published extensively in those fields. He is a cofounder of the biomedical start-up company Oxymap.

Prof. Benediktsson is the 2011–2012 President of the IEEE Geoscience and Remote Sensing Society (GRSS) and has been on the GRSS Administrative Committee since 2000. He was the Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS) from 2003 to 2008 and has served as an Associate Editor of TGRS since 1999 and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS since 2003. He received the Stevan J. Kristof Award from Purdue University in 1991 as an outstanding graduate student in remote sensing. In 1997, he was the recipient of the Icelandic Research Council's Outstanding Young Researcher Award; in 2000, he was granted the IEEE Third Millennium Medal; in 2004, he was a corecipient of the University of Iceland's Technology Innovation Award; in 2006, he received the yearly research award from the Engineering Research Institute of the University of Iceland; and in 2007, he received the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society. He is a corecipient of the 2012 IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Best Paper Award and a corecipient of the 2013 IEEE GRSS High Impact Paper Award. In 2013, he received the Joint Icelandic Society of Chartered Engineers (VFI) and IEEE 2013 Electrical Engineer of the Year Award. He is a Fellow of the International Society for Optics and Photonics (SPIE). He is a member of Societas Scientiarum Islandica and Tau Beta Pi.