

Cascade Superpixel Regularized Gabor Feature Fusion for Hyperspectral Image Classification

Sen Jia[✉], Senior Member, IEEE, Zhijie Lin, Bin Deng, Jiasong Zhu[✉], and Qingquan Li

Abstract—A 3-D Gabor wavelet provides an effective way to obtain the spectral–spatial-fused features for hyperspectral image, which has shown advantageous performance for material classification and recognition. In this paper, instead of separately employing the Gabor magnitude and phase features, which, respectively, reflect the intensity and variation of surface materials in local area, a cascade superpixel regularized Gabor feature fusion (CSRGFF) approach has been proposed. First, the Gabor filters with particular orientation are utilized to obtain Gabor features (including magnitude and phase) from the original hyperspectral image. Second, a support vector machine (SVM)-based probability representation strategy is developed to fully exploit the decision information in SVM output, and the achieved confidence score can make the following fusion with Gabor phase more effective. Meanwhile, the quadrant bit coding and Hamming distance metric are applied to encode the Gabor phase features and measure sample similarity in sequence. Third, the carefully defined characteristics of two kinds of features are directly combined together without any weighting operation to describe the weight of samples belonging to each class. Finally, a series of superpixel graphs extracted from the raw hyperspectral image with different numbers of superpixels are employed to successively regularize the weighting cube from over-segmentation to under-segmentation, and the classification performance gradually improves with the decrease in the number of superpixels in the regularization procedure. Four widely used real hyperspectral images have been conducted, and the experimental results constantly demonstrate the superiority of our CSRGFF approach over several state-of-the-art methods.

Index Terms—Cascade superpixel regularization (SR), feature fusion, hyperspectral image classification, image segmentation.

I. INTRODUCTION

THROUGH imaging the same spatial area at different wavelengths by highly sensitive sensors, hundreds of spectral bands can be simultaneously acquired, and the so-called hyperspectral image provides the opportunity to

Manuscript received August 1, 2018; revised April 2, 2019 and June 3, 2019; accepted June 4, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61671307 and Grant 41871329, in part by the Guangdong Special Support Program of Top-notch Young Professionals under Grant 2015TQ01X238, and in part by the Shenzhen Scientific Research and Development Funding Program under Grant JCYJ20180305124802421 and Grant JCYJ20170818092931604. (Corresponding author: Jiasong Zhu.)

S. Jia, Z. Lin, and B. Deng are with the Guangdong Key Laboratory of Urban Informatics, Shenzhen University, Shenzhen, China, and also with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: senjia@szu.edu.cn; 18814103637@163.com; szubing@qq.com).

J. Zhu and Q. Li are with the Guangdong Key Laboratory of Urban Informatics, Shenzhen University, Shenzhen 518060, China (e-mail: zhujisong@gmail.com; liqq@szu.edu.cn).

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/TNNLS.2019.2921564

accurately classify the surface materials [1]–[4]. However, the number of labeled samples is usually limited in practice (caused by the difficulty of manually labeling), and thus, it is difficult to effectively train a supervised classifier. Furthermore, the high spectral dimension of hyperspectral image further exacerbates the small sample size problem, which heavily degrades the applicability of hyperspectral data. Therefore, the small sample size problem can be the most challenging issue for material identification of hyperspectral image [5], [6].

Specifically, it can be easily observed from hyperspectral image that the distribution of surface materials presents meaningful spatial structure, for example, the structure of man-made objects (such as road and building) is regular, while the structure of nature objects (such as grass and tree) is continuous, and thus, there is a popular belief in the literature that the spatial structural information is a very important complement to the spectral signatures for feature representation of different land-cover types [7].

In the last two decades, the spectral–spatial feature extraction techniques have been extensively studied in the field of hyperspectral image classification, in which one popular method is the morphological analysis [8]–[10]. Concretely, two mathematical operators, opening and closing, are applied to the selected representative bands or the principal component analysis (PCA) extracted features. Since the spatial structure of materials is different from each other, multi-level parameters of the two operators are used to obtain adequate morphological profiles. Furthermore, several extensions of morphological profiles, including morphological attribute profile [11], [12], multiple-structure-element morphological profiles [13], and multiple morphological component analysis [14], have been developed to characterize the spatial information more accurately. Likewise, the spatial texture information can also be exploited by other methods, such as subspace projection [15]–[17] and joint sparse representation [18]. After the spatial features have been obtained, they can be directly applied or combined with the spectral features to carry out the classification task. Alternatively, the spatial–contextual structure can also be characterized by conditional random field [19] and segmentation-based methods (including watershed transformation [20] and graph cut [21]), which can be employed as the postprocessing step to enhance the spatial homogeneity of material distribution. However, the main limitation with these two kinds of methods is that the spatial and spectral features are separately conducted, and thus, the joint spatial–spectral relationship in the 3-D hyperspectral data cannot be adequately investigated.

Due to the strong feature extraction ability and large feature representation capacity, deep learning technique has drawn increasing attention in hyperspectral image classification [22]–[24]. Concretely, Chen *et al.* [25] performed the classification task with stacked autoencoders that can extract high-level abstraction features. A 1-D convolutional neural network (CNN) [26] and a 2D CNN [27] were also employed. Furthermore, Li *et al.* [28] designed a 3D CNN model that can simultaneously extract the spatial–spectral features. Meanwhile, regarding the spectral signature as a long sequence, a recurrent neural network (RNN) has also been introduced in hyperspectral image classification [29]. Xu *et al.* [30] built a novel unified network that combines long short term memory (LSTM) and CNN to, respectively, obtain the spectral and spatial features. Nevertheless, the excellent performance of deep learning-based methods usually depends on a large amount of training samples that are used to optimize the huge number of parameters in the CNN model. Yu *et al.* [31] constructed a novel CNN model with data augmentation and 1×1 convolutional kernels. Besides, semi-supervised CNNs with limited training samples were utilized to improve the classification performance [32], [33].

Naturally, in order to effectively represent the internal 3-D spatial–spectral mechanism of hyperspectral image, a feasible way is to introduce a set of predefined 3-D filters and directly interact with the raw hyperspectral image, including a 3-D discrete wavelet transform [34], a 3-D morphological operator [35], a 3-D local binary pattern (LBP) operator [36], a 3-D Gabor transform [37], [38], and a 3-D scattering wavelet transform [39]. Due to the powerful representative ability and moderate computational requirement, the 3-D Gabor transform has attracted much attention in recent years and a series of 3-D Gabor transform-based methods has been developed [40], [41]. Specifically, through extending conventional 2-D Gabor transform into 3-D domains, dozens of 3-D Gabor filters with different scales and orientations have been defined, which can sufficiently characterize the signal variance in the local 3-D region. In the following, each 3-D Gabor filter is directly convolved with the raw hyperspectral image, and the extracted Gabor magnitude features can be utilized for hyperspectral region/pixelwise classification. Furthermore, in order to increase the discriminative ability of the extracted large amount of Gabor magnitude features, collaborative representation and Fisher discrimination criterion have been, respectively, adopted [42], [43]. Alternatively, 3-D Gabor phase features, which are sensitive to the changing positions and traditionally omitted, have been properly encoded and employed for hyperspectral image classification, and remarkable results (both classification accuracy and computational efficiency) have been achieved [44]. Therefore, due to the successful application of Gabor magnitude and phase, which reflect the properties of surface materials from different aspects (intensity and variation), it is desirable to consider the two kinds of powerful features in a unified framework for hyperspectral image analysis.

Moreover, since the spatial distribution of surface objects is usually regular and continuous, the structural information

can provide important contextual description of materials for hyperspectral image analysis. Fortunately, along with the advance of image segmentation in the computer vision field, superpixel extraction could obtain spatially homogeneous representation of objects [45], [46], which is an invaluable tool and has been recently introduced for hyperspectral image classification. Specifically, several works focus on superpixel-based dictionary learning, in which the superpixel map is used to facilitate the dictionary learning process. More precisely, Fang *et al.* [47] proposed a superpixel-based sparse representation model, which learns a discriminative dictionary within each superpixel via joint sparse coding strategy. Li *et al.* [48] proposed a superpixel-level collaborative representation model for dictionary learning of multiple modalities of hyperspectral image, and multitask learning framework is subsequently adopted to combine the reconstruction error. Similarly, He *et al.* [49] employed the superpixel segmentation method to group spatial pixels, which was then utilized to assist the following sparse and low-rank regularized dictionary learning process. Besides, the superpixel methods have also been used for feature extraction and dimensionality reduction of hyperspectral image. Fang *et al.* [50] developed a superpixel-based multiple kernel learning methods, in which an oversegmented superpixel map is generated and used for kernel-based spatial feature extraction. Meanwhile, Jia *et al.* [51] proposed a superpixel-based dimensionality reduction technique to reduce the dimensions of the extracted large number of 2-D Gabor features. In fact, it can be easily observed from the aforementioned methods that the obtained superpixel segmentation map is used as a preprocessing step to enhance the relationship among pixels within the superpixel, which greatly facilitates the following data processing (both dictionary learning and feature extraction). However, it is not a trivial task to precisely estimate the number of extracted superpixels, which controls the segmentation scale of the scene and has a crucial impact on hyperspectral image analysis. Unfortunately, the estimation problem of the number of extracted superpixels remains unsolved. It is also reasonable to take the superpixel map as a postprocessing step to improve the spatial regularity and preserve the spatial continuity of hyperspectral image [52], [53], but the superpixel regularization (SR) procedure should be carried out in a more elegant way.

In this paper, in order to integrate the advantages of 3-D Gabor features (both magnitude and phase) and SR, a cascade superpixel regularized Gabor feature fusion (CSRGFF) approach, named CSRGFF, has been proposed. First, instead of considering Gabor filters with multiple directions that could exploit the whole 3-D domains, only a single Gabor filter that is parallel to the spectral orientation is picked out. Just as pointed out in [44], since the chosen filter is consistent with the internal structure of hyperspectral image, the extracted Gabor features can be more discriminative than the other ones. Evidently, the reliability and efficiency of the Gabor feature extraction process have been substantially improved. Furthermore, in order to sufficiently represent the spatial–spectral correlation of material, the Gabor filters with four different scales are utilized and directly convolved with

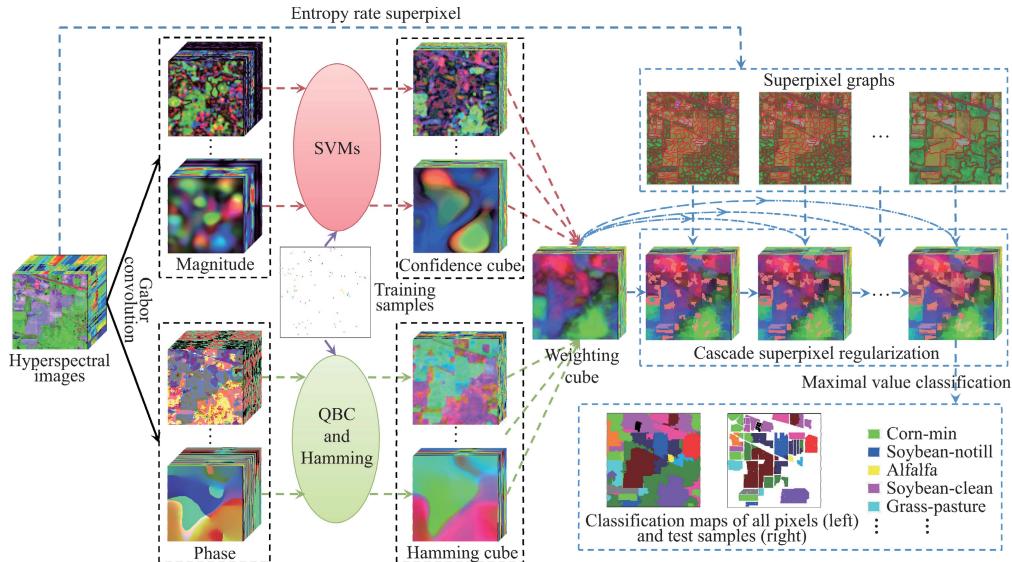


Fig. 1. Schematic of the proposed CSRGFF approach for hyperspectral image classification.

the raw hyperspectral image, resulting in eight Gabor feature cubes (including four Gabor magnitudes and four Gabor phases).

Second, since the characteristic of the two kinds of features is totally different, the combination process should be specifically designed. Concretely, due to the stability of Gabor magnitude features, a support vector machine (SVM)-based probability representation strategy is developed to fully exploit the decision information in SVM output, and the achieved confidence score can better comply with the coding scheme of Gabor phase. Alternatively, for each Gabor phase feature cube, which is unstable compared to the magnitude feature, the quadrant bit coding (QBC) is used to encode the phase feature and Hamming distance metric is then adopted to measure sample similarity. Third, the well-defined similarity measures of two kinds of features are directly combined together without any weighting operation to describe the weight of samples belonging to each class (called weighting cube). Finally, a series of superpixel graphs is extracted from the raw hyperspectral image with different numbers of superpixels, which are employed to regularize the weighting cube in the cascade mode from over-segmentation (i.e., the number of superpixels is large) to under-segmentation (i.e., the number of superpixels is small), and the classification performance gradually improves with the decrease of the number of superpixels in the regularization procedure. For illustrative purposes, Fig. 1 shows the system block diagram of the proposed framework, and the Indian Pines hyperspectral image data are used here. Compared with other related state-of-the-art methods, our CSRGFF approach has four main contributions, which are listed in the following.

- 1) First of all, we propose to jointly exploit the two kinds of powerful Gabor features, i.e., magnitude and phase, to accomplish hyperspectral image classification. Since both the Gabor magnitude and phase features have shown favorable discriminative ability on hyperspectral image analysis, and they, respectively,

reflect the intensity and variation of surface materials in local area, it is beneficial to combine the two kinds of features together to carry out the classification task.

- 2) Second, a new strategy that systematically investigates the inter-class relationship and transforms the output of SVMs' classifier into a reasonable confidence representation has been specifically designed. (Clearly, other probability-transformed classifiers can also be used. The reason that SVMs' classifier is employed here is due to its high effectiveness) Here, the voting scheme used in the one-against-one strategy of SVMs and summation of the predicted values of each class are unified in a single framework, and thus, the combined result, called confidence score, is more effective to characterize the inter-class relationship. Moreover, since the magnitude-based confidence score and phase coding-based Hamming distance have been transformed into the same range, they can be directly fused for hyperspectral image analysis without any weighting operation, and hence the practicability of the proposed method can be ensured.
- 3) Another distinct advantage of CSRGFF is that the parameter estimation of the number of extracted superpixels is unnecessary. Especially, the multi-scale superpixel graphs are used to successively regularize the weighting cube from over-segmentation to under-segmentation, and the classification performance gradually improves with the decrease of the number of superpixels in the regularization procedure. More importantly, since the numbers of superpixel of both over-segmentation and under-segmentation are easy to choose and robust to the classification performance, the parameters included in the proposed CSRGFF approach can be properly set and kept unchanged for hyperspectral images with different spatial resolutions.
- 4) In CSRGFF, only four 3-D Gabor filters are utilized for feature extraction, and the obtained eight Gabor

TABLE I

DEFINITIONS OF ALL MATHEMATICAL VARIABLES USED IN THIS PAPER

Abbreviation	Meaning
R	Hyperspectral image
A, L	Training set and corresponding labels
n, C	Number of training samples and classes
K, S	Number of superpixels and ERS-based superpixel graph
B, E	Initial and terminal number of superpixels
X, Y, Z	Three dimensions of R
Ψ, f, θ, φ	The 3D Gabor filters and related parameters
G, M, P	The extracted 3D Gabor, magnitude and phase features
D, C	The decision matrix and confidence cube
Q, H	The QBC cube and Hamming cube
W, Z	The weighting cube and summed regularization cube

magnitude and phase feature cubes are applied for classification without dimensionality reduction. Meanwhile, the time complexity of the proposed method is basically linear to the spatial size of the hyperspectral image, which is applicable for hyperspectral image with large coverage.

A number of experiments have been conducted with several state-of-the-art methods on four real hyperspectral image data sets, and the results consistently demonstrate that our CSRGFF approach could achieve higher classification performance than other ones. Table I summarizes the definitions of all mathematical variables used throughout this paper.

The rest of this paper is structured as follows. Some closely related works are reviewed in Section II, including 3-D Gabor wavelets and entropy rate superpixel (ERS) segmentation. In Section III, our proposed CSRGFF method is presented in detail. Afterward, the experimental results on four real hyperspectral remote sensing data are reported in Section IV. Finally, in Section V, we provide some concluding remarks.

II. RELATED WORKS

A. 3-D Gabor Wavelets

Due to the powerful signal representative ability of Gabor wavelets (it has been found that the optimal time and frequency resolution could be achieved by the Gabor wavelet transform, and then both the 1-D signal and the 2-D image can be effectively analyzed in the time–frequency domains) [54], it has been recently extended into 3-D formation to comply with the joint spectral–spatial structure of the hyperspectral image [38]. Concretely, for a 3-D Gabor filter in the 3-D frequency domain (u, v, w) , three parameters should be clearly defined, including the central frequency (denoted as f) and two orientation parameters (denoted as θ and φ). As shown in Fig. 2, θ and φ , respectively, depict the angles of vector f with the w -axis and projection of f onto the uv plane with the u -axis. A 3-D Gabor filter can be formulated as

$$\begin{aligned} \Psi_{\{u, v, w\}}(x, y, z) &= \frac{1}{(2\pi)^{2/3}\sigma^3} \exp(j2\pi(xu + yv + zw)) \\ &\quad \times \exp\left(-\frac{x^2 + y^2 + z^2}{2\sigma^2}\right) \\ u &= f \sin \theta \cos \varphi, \quad v = f \sin \theta \sin \varphi, \quad w = f \cos \theta \end{aligned} \quad (1)$$

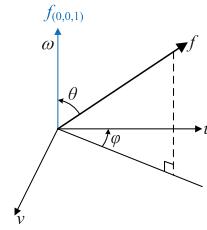


Fig. 2. Gabor filters in the 3-D domains. Blue line: chosen direction of the Gabor filter used in this paper.

where (x, y, z) represent the sampled variables in the 3-D domains. Since the internal structure of the hyperspectral image is unknown, it is necessary to apply Gabor filters with different scales and orientations on hyperspectral image to obtain sufficient features for classification. Generally, the orientation parameters θ and φ vary in the steps of 45° , resulting in 13 Gabor filters with different directions for each scale in the 3-D domains, and considerable redundancy exists in the obtained large number of 3-D Gabor features. Recent advance in the literature [44] has revealed that the Gabor features extracted by the Gabor filter parallel to the spectral direction, that is, orientation $\theta = 0$ (the direction of the blue line in Fig. 2), contain more discriminative information than the ones with other orientations, since the chosen Gabor filter is in accord with the spectral–spatial structure of hyperspectral image. Therefore, only one Gabor filter in each scale is applied for Gabor feature extraction, which can be derived from (1)

$$\begin{aligned} \Psi_{\{0, 0, f\}}(x, y, z) &= \frac{1}{(2\pi)^{2/3}\sigma^3} \exp(j2\pi fz) \\ &\quad \times \exp\left(-\frac{x^2 + y^2 + z^2}{2\sigma^2}\right). \end{aligned} \quad (2)$$

In this paper, four different scales have been considered, i.e., f takes the values of $[0.5, 0.25, 0.125, 0.0625]$, and the corresponding four 3-D Gabor filters are denoted as Ψ_i , ($i = 1, \dots, 4$) for simplicity. Evidently, the four retained Gabor feature cubes can significantly improve the efficiency of the feature extraction procedure.

Furthermore, for each Gabor filter Ψ_i , ($i = 1, \dots, 4$), the coefficient $\mathbf{G}_i(x, y, z) = (\mathbf{R} \otimes \Psi_i)(x, y, z)$ is used to characterize the change information of local area around location (x, y, z) (we suppose that $\mathbf{R} \in \mathbb{R}^{X \times Y \times Z}$ denotes the hyperspectral image cube, in which $X \times Y$ represent the imaged scene, and Z is the number of spectral bands), where \otimes is the convolution operation. It can be easily observed from (2) that the Gabor wavelet Ψ_i is complex; therefore, the convolutional coefficient \mathbf{G}_i can be divided into two parts, i.e., the magnitude \mathbf{M}_i and phase \mathbf{P}_i , which can be computed by

$$\mathbf{M}_i(x, y, z) = \sqrt{\operatorname{Re}(\mathbf{G}_i(x, y, z))^2 + \operatorname{Im}(\mathbf{G}_i(x, y, z))^2} \quad (3)$$

$$\mathbf{P}_i(x, y, z) = \arctan(\operatorname{Im}(\mathbf{G}_i(x, y, z)), \operatorname{Re}(\mathbf{G}_i(x, y, z))) \quad (4)$$

where Im and Re operators obtain the imaginary and real units of the complex number $\mathbf{G}_i(x, y, z)$, respectively. Clearly, the size (both spectral and spatial) of \mathbf{M}_i , ($i = 1, \dots, 4$) and \mathbf{P}_i , ($i = 1, \dots, 4$) is the same as the original hyperspectral image \mathbf{R} .

B. Entropy Rate Superpixel Segmentation

A perceptually uniform region, called superpixel, can extract useful spatial structure information of surface materials in a hyperspectral image. In order to acquire homogeneous superpixels with controllable sizes, ERS [45] segmentation is used in this paper. ERS is a graph-based clustering method, which has demonstrated excellent superpixel extraction performance in the field of computer vision. The method consists of two components: entropy rate of random walk on a graph and a balancing term. The aim of the entropy rate function is to make each superpixel more compact and homogeneous, whereas the balancing term is to reduce the number of unbalanced superpixels. The procedure of generating superpixels in an image can be briefly described as the following four steps.

1) *Graph Construction:* Let V represent the set of vertices v_i (i.e., the pixels) and F denote the set of edges $e_{i,j}$, then a graph $G = (V, F)$ is constructed by using eight-connected grid graph structure. Meanwhile, the weight of each edge $e_{i,j}$ is defined as

$$w_{i,j} = \exp\left(-\frac{d(v_i, v_j)^2}{2\sigma^2}\right), \quad e_{i,j} \in F \quad (5)$$

where $d(v_i, v_j)$ is defined as the intensity difference multiplied by the spatial distance.

2) *Entropy Rate:* Let A be the selected edge set, then the entropy rate of the random walk on the graph $G = (V, A)$ is used as a criterion to obtain compact and homogeneous clusters and it is written as a set function

$$\mathcal{H}(A) = - \sum_i u_i \sum_j p_{i,j}(A) \log(p_{i,j}(A)) \quad (6)$$

where the transition probabilities $p_{i,j}$ in the set function are given in the following:

$$p_{i,j}(A) = \begin{cases} \frac{w_{i,j}}{w_i} & \text{if } i \neq j \text{ and } e_{i,j} \in A, \\ 0 & \text{if } i \neq j \text{ and } e_{i,j} \notin A, \\ 1 - \frac{\sum_{j:e_{i,j} \in A} w_{i,j}}{w_i} & \text{if } i = j. \end{cases}$$

$$w_i = \sum_{k:e_{i,k} \in F} w_{i,k}, \quad w_T = \sum_{i=1}^{|V|} w_i, \quad u_i = \frac{w_i}{w_T}. \quad (7)$$

The entropy rate is a monotonically increasing function since the inclusion of any edges increases the uncertainty of a jump of the random walk. However, the increase can be large when the edges are selected to form the compact and homogeneous clusters.

3) *Balancing Function:* A balancing function is utilized to encourage clusters with similar sizes and it is defined as

$$\mathcal{B}(A) = H(Z_A) - N_A = - \sum_i PZ_A(i) \log(PZ_A(i)) - N_A \quad (8)$$

where A is the selected edge set, Z_A is the distribution of the cluster membership, and N_A is the number of connected components in the graph. For example, let the graph partitioning for the edge set A be $S_A = \{S_1, S_2, \dots, S_{N_A}\}$, then the distribution of Z_A is equal to

$$PZ_A(i) = \frac{|S_i|}{|V|}, \quad i = \{1, \dots, N_A\}. \quad (9)$$

Similar to the entropy rate, the balancing function is also a monotonically increasing function and becomes large when selecting edges from similarly sized clusters.

4) *Optimization:* The objective function combines the entropy rate and the balancing function, thus favoring the compact, homogenous, and balanced clusters. Clustering is achieved by optimizing the objective function on the edge set as

$$\max_A \mathcal{H}(A) + \lambda \mathcal{B}(A) \quad (10)$$

$$\text{subject to } A \subseteq F \text{ and } N_A \geq K \quad (11)$$

where $\lambda \geq 0$ is the weight of the balancing term and it can automatically adjust based on: 1) the number of superpixels K and 2) a data-dependent dynamic parameter β that is computed from the input image [45]. Before presenting the proposed CSRGFF approach, it is worth to point out that since the superpixel map extracted by the ERS method often exhibits a balanced distribution (i.e., the superpixel size is similar with each other), the spatial locality of surface materials can be well characterized, and thus, the efficiency of the well-designed cascade SR procedure can be ensured.

III. PROPOSED APPROACH

The proposed CSRGFF algorithm involves the following three steps: confidence cube computation, Hamming cube computation, and cascade SR.

A. Confidence Cube Computation

The confidence cubes are calculated based on the Gabor magnitude feature cubes \mathbf{M}_i , ($i = 1, \dots, 4$), which are acquired by (3). In particular, since the characteristic of Gabor magnitude and phase features is totally different, the computed confidence cubes should be specifically designed to comply with the coding scheme of Gabor phase. Concretely, for every pixel (or sample) in each Gabor magnitude feature cube \mathbf{M}_i , our goal is to get the pixel's confidence scores that belong to each class and then the corresponding confidence cube \mathbf{C}_i can be achieved. For a certain pixel in \mathbf{M}_i , we denote its confidence score belonging to class c as C_i^c . In order to better represent the confidence score C_i^c , SVMs' classifier is used, in which radial basic function (RBF) kernel and one-against-one scheme are employed for multi-class classification [55]. Suppose there are C significant materials existing in the scene, $C \times (C - 1)$ classifiers can thus be constructed by SVMs with training samples. For each classifier corresponding to any two classes c_1 and c_2 , a decision value δ for the certain pixel is obtained by a voting process. The decision value δ can be considered as the distance from the pixel to the optimal hyperplane [56]; therefore, this value can be used to construct confidence score to different classes. Mathematically, in our proposed method, a decision matrix $\mathbf{D} \in \mathbb{R}^{C \times C}$ is built for the certain pixel, in which $\mathbf{D}_{c_1 c_2} = \delta$ if $\delta > 0$, else $\mathbf{D}_{c_2 c_1} = -\delta$ and the other entries in \mathbf{D} are set as zero. After the decision matrix \mathbf{D} is built, the confidence score of certain pixel belonging to class c in \mathbf{M}_i is computed by

$$C_i^c = \frac{\sum_{j=1}^C \mathbf{D}_{cj}}{2 \times n_c} + \frac{\sqrt{n_c}}{2 \times \sqrt{C}}, \quad c = 1, \dots, C \quad (12)$$

where n_c is the number of nonzero entries in the c th row of the matrix \mathbf{D} . It can be evidently observed from (12) that the range of the obtained SVM-based probability output is $[0, 1]$. Meanwhile, the larger the confidence score C_i^c , the higher probability the certain pixel belonging to the c th class. After all the pixels in the Gabor Magnitude feature cube \mathbf{M}_i have been processed by the carefully designed confidence score representation, the confidence cube $\mathbf{C}_i \in \mathbb{R}^{X \times Y \times C}$ can thus be achieved. So finally, we can gain four different confidence cubes $\mathbf{C}_i, i = 1, \dots, 4$ based on four different Gabor magnitude feature cubes $\mathbf{M}_i, i = 1, \dots, 4$.

B. Hamming Cube Computation

As an important feature representation, Gabor phase has been successfully applied in the hyperspectral image for feature extraction in our previous work [44]. In this paper, in order to further improve classification performance as well as effectively fuse the Gabor magnitude and phase features, QBC [57] and Hamming distance-based matching are conducted to generate the Hamming cubes. QBC is a coding method that encodes phase information according to the quadrant it belongs to. Concretely, for each value $\mathbf{P}_i(x, y, z)$ in the i th Gabor phase feature cube, it is encoded into two bits, $(\mathbf{Q}_i^{Re}(x, y, z), \mathbf{Q}_i^{Im}(x, y, z))$, by the following rules:

$$\mathbf{Q}_i^{Re}(x, y, z) = \begin{cases} 1 & \text{if } \mathbf{P}_i(x, y, z) \in \{\text{I}, \text{IV}\} \\ 0 & \text{if } \mathbf{P}_i(x, y, z) \in \{\text{II}, \text{III}\} \end{cases} \quad (13)$$

$$\mathbf{Q}_i^{Im}(x, y, z) = \begin{cases} 1 & \text{if } \mathbf{P}_i(x, y, z) \in \{\text{I}, \text{II}\} \\ 0 & \text{if } \mathbf{P}_i(x, y, z) \in \{\text{III}, \text{IV}\}. \end{cases} \quad (14)$$

From (13) and (14), it can be easily seen that the QBC codes of the phase value are encoded based on quadrant that it lies. That is, any two phase values will be encoded into two same bits both in \mathbf{Q}^{Re} and \mathbf{Q}^{Im} if they lie in the same quadrant, while the coding bits will be completely different in the opposite quadrant (i.e., quadrants I and III or quadrants II and IV). After four Gabor phase feature cubes have been encoded from the QBC coding method, eight binary feature cubes, \mathbf{Q}_i^{Re} and $\mathbf{Q}_i^{Im}, i = 1, \dots, 4$, have been obtained. For any two certain pixels \mathbf{p} and \mathbf{q} with spatial positions (x_p, y_p) and (x_q, y_q) , in order to measure their similarity based on the QBC codes, the Hamming distance [57] is adopted. Concretely, the Hamming distance between the pixels \mathbf{p} and \mathbf{q} in the i th QBC feature cube is computed by

$$h_i(\mathbf{p}, \mathbf{q}) = \frac{\sum_{z=1}^Z (\text{bit}_i^{Re} + \text{bit}_i^{Im})}{2 \times Z} \quad (15)$$

$$\begin{aligned} \text{bit}_i^{Re} &= \mathbf{Q}_i^{Re}(x_p, y_p, z) \oplus \mathbf{Q}_i^{Re}(x_q, y_q, z) \\ \text{bit}_i^{Im} &= \mathbf{Q}_i^{Im}(x_p, y_p, z) \oplus \mathbf{Q}_i^{Im}(x_q, y_q, z) \end{aligned} \quad (16)$$

where \oplus is the bit XOR operator, i.e., the result is equal to zero if and only if two bits, $\mathbf{Q}_i(x_p, y_p, z)$ and $\mathbf{Q}_i(x_q, y_q, z)$, are equal. From that, it can be easily found that the value of Hamming distance is between 0 and 1 and the smaller the value, the more similar the pixels \mathbf{p} and \mathbf{q} . Based on the Hamming distance calculation between two pixels, we define

a pixel-to-class Hamming distance. Precisely, the Hamming distance between the pixel \mathbf{p} and the c th class is computed by

$$\mathcal{H}_i^c = \min_{\text{label}(\mathbf{A}_j)=c} h_i(\mathbf{p}, \mathbf{A}_j) \quad (17)$$

where \mathbf{A} is the training set composed of some labeled pixels and \mathbf{A}_j is the j th sample (or pixel) in the training set \mathbf{A} . The number i represents that such calculation is computed based on the i th QBC feature cube. The same as the confidence score achieved from the Gabor magnitude features, the range of Hamming distance is also between 0 and 1. The pixel-to-category Hamming distance of all the pixels can be calculated in this way, and finally, four Hamming cubes, $\mathbf{H}_i \in \mathbb{R}^{X \times Y \times C}, i = 1, \dots, 4$, can be generated.

C. Cascade Superpixel Regularization

The confidence score reflects the similarity between the pixel and the class and this similarity will be higher if the score is larger. By contrast, the Hamming distance measures the distance between the pixel and the class; therefore, the larger the value, the smaller the similarity between them. Based on the analysis, the confidence cubes $\mathbf{C}_i, i = 1, \dots, 4$ and Hamming cubes $\mathbf{H}_i, i = 1, \dots, 4$, respectively, obtained from the Gabor magnitude and phase features are combined together to construct the weighting cube

$$\mathbf{W} = \sum_{i=1}^4 (\mathbf{C}_i - \mathbf{H}_i). \quad (18)$$

In practice, it is commonly accepted that introducing more parameters can increase the flexibility of the algorithm, but the complexity and usability of the algorithm can also be degraded. Since the ranges of $\mathbf{C}_i, i = 1, \dots, 4$ and $\mathbf{H}_i, i = 1, \dots, 4$ are the same, i.e., they both take values from 0 to 1, they can be directly combined without balancing the two terms, and the practicability of the proposed method can be ensured. Through constructing such a weighting cube \mathbf{W} , the well-defined similarity measures of two different kinds of features are integrated, and the decision accuracy can thus be increased. Clearly, the weighting cube can be used to classify each pixel directly by assigning to the class with the maximal weight, which can actually present a good classification result. However, the spatial information is not fully utilized in this way. In this paper, in order to well reflect the spatial regularity of surface materials in the scene and further boost the classification performance, we utilize a simple yet effective method to regularize the weighting cube, i.e., the weight of each pixel in the same superpixel is regularized to be equal. Algorithm 1 describes the process of SR for hyperspectral image. In the algorithm, the ERS is applied to obtain the superpixel graph (it is worth pointing out that the first three PCA bands are used for the ERS method), and then the regularization will be carried out in each superpixel one by one. If a superpixel contains only one training sample, it is reasonable to consider that all pixels in this superpixel belong to the label of the single training sample. Since each superpixel is homogenous as much as possible, the weight of this class will be set to 1 and others to 0 for every pixel in this superpixel. Otherwise, all

Algorithm 1 SR for Hyperspectral Image

```

1: INPUT: raw HSI data  $\mathbf{R}^{X \times Y \times Z}$ , the weighting cube  $\mathbf{W} \in \mathbb{R}^{X \times Y \times C}$ , the number of superpixel  $K$ ;
2: OUTPUT: the regularized weighting cube  $\mathbf{U} \in \mathbb{R}^{X \times Y \times C}$ ;
3: BEGIN
4: using ERS to obtain the superpixel graph  $\mathbf{S}$  with  $K$  superpixels;
5:  $\mathbf{U} = \text{zeros}(X, Y, C)$ ;
6: for  $k = 1$  to  $K$  do
7:   if  $\mathbf{S}_k$  contains only one training sample then
8:      $\mathbf{U}_k = \mathbf{W}(\mathbf{I}_k, :)$ ; %  $\mathbf{I}_k$  is the spatial position of the
      kth superpixel;
9:      $\mathbf{u} = \text{zeros}(C, 1)$ ;
10:     $\mathbf{u}_c = 1$ ; % suppose the training sample belongs to the
      c-th class;
11:    each position in  $\mathbf{U}_k$  is equal to  $\mathbf{u}$ ;
12:   else
13:      $\mathbf{U}_k = \mathbf{W}(\mathbf{I}_k, :)$ ;
14:      $\mathbf{u} = \text{mean}(\mathbf{U}_k, 2)$ ;
15:     each position in  $\mathbf{U}_k$  is equal to  $\mathbf{u}$  divided by the number
      of pixels in the superpixel;
16:   end if
17: end for
18: END

```

these pixels' weights of each class are replaced by the mean value of them. Through regularizing the weighting cube in this way, the spatial information of surface materials contained in superpixels can be integrated into the weighting cube.

Nevertheless, the number of superpixels is not easy to be precisely estimated in the step of superpixel generation. Specifically, if the number of superpixels is too large (corresponding to over-segmentation), the number of pixels in each superpixel is relatively small and the same material may be partitioned into several superpixels, which could decrease the classification performance with small training sample scenario. Alternatively, if the number of superpixels is too small (corresponding to under-segmentation), the number of pixels in each superpixel is relatively large and several materials may be contained in a single superpixel, and thus, the classification process can be misled.

To tackle the problem, a cascade SR procedure has been specifically designed. In the proposed model, instead of directly estimating the number of superpixels, a series of superpixel graphs is extracted from the raw hyperspectral image with different numbers of superpixels, which are employed to regularize the weighting cube in a cascade mode from over-segmentation to under-segmentation. The regularized results are summed together to get the final weighting cube \mathbf{Z} , and the classification map of the hyperspectral image can thus be achieved by taking the maximal weight of each pixel. The total procedure of our CSRGFF approach is presented in Algorithm 2 (it is worth to mention that in order to guarantee the efficiency of the regularization procedure, a step size should be taken in superpixel extraction from over-segmentation to under-segmentation, which is simply set

Algorithm 2 CSRGFF for Hyperspectral Image Classification

```

1: INPUT: raw HSI data  $\mathbf{R}^{X \times Y \times Z}$ , training set  $\mathbf{A} \in \mathbb{R}^{Z \times n}$ 
   (with  $n$  samples) from  $C$  classes, the labels of the training
   set  $\mathcal{L} \in \mathbb{R}^{1 \times n}$ ; the initial ( $B$ ) and terminal ( $E$ ) number of
   superpixels;
2: OUTPUT: the predicted labels of the whole data set  $\mathcal{C} \in \mathbb{R}^{X \times Y}$ ;
3: BEGIN
4: using equation (2) to create four 3D Gabor filters, denoted
   as  $\Psi_i$ , ( $i = 1, \dots, 4$ );
5: for  $i = 1$  to  $4$  do
6:    $\mathbf{G}_i = \mathbf{R} \otimes \Psi_i$ ;
7:   using equations (3) and (4) to obtain the magnitude  $\mathbf{M}_i$ 
   and phase  $\mathbf{P}_i$ ;
8:   for  $\mathbf{M}_i$ , using LibSVM software package to obtain the
   confidence cube  $\mathbf{C}_i \in \mathbb{R}^{X \times Y \times C}$ ;
9:   for  $\mathbf{P}_i$ , using QBC and Hamming distance measure to
   obtain the Hamming cube  $\mathbf{H}_i \in \mathbb{R}^{X \times Y \times C}$ ;
10: end for
11: the weighting cube  $\mathbf{W} = \sum_{i=1}^4 (\mathbf{C}_i - \mathbf{H}_i)$ ;
12:  $\mathbf{Z} = \text{zeros}(X, Y, C)$ ;
13: for  $j = B$  to  $E$  do
14:    $\mathbf{Z} = \mathbf{Z} + \text{SR}(\mathbf{R}, \mathbf{W}, j)$  [SR refers to Algorithm 1];
15: end for
16:  $[\mathbf{T}, \mathcal{C}] = \max(\mathbf{Z}, [ ], 3)$ ;
17: END

```

as 50). In Section IV, it can be found that the classification performance is not only insensitive to the chosen initial (B) and terminal (E) number of superpixels but also continuously increases along with the SR procedure; therefore, the estimation problem of the superpixel number can be avoided, which is considered as an important advantage of the proposed CSRGFF approach.

At last, the time complexity of our proposed CSRGFF approach is mathematically analyzed, which can be roughly divided into three parts, i.e., the SVM-based confidence cube, QBC-based Hamming cube, and cascade SR procedure. Specifically, the complexity of SVMs' classifier and each Gabor filter is, respectively, $O(n^3 Z)$ [58] and $O(3XYZ \log(XYZ)^1/2)$, while that of the QBC-based Hamming cube is approximately $O(8XYZ + nZ)$. The complexity of cascade SR is mainly related to that of the ERS algorithm, which approximates $O(K \log K)$. Since the Gabor features (both magnitude and phase) are just needed to be computed once, the complexity of the proposed approach is basically linear to the spatial size of the hyperspectral image (n and K are relatively small compared with the spatial size $X \times Y$), which is applicable for hyperspectral image with large coverage.

IV. EXPERIMENTS

This section presents the experimental results of four real hyperspectral images to demonstrate the superiority of our proposed CSRGFF approach. First, we show the robustness



Fig. 3. Ground-truth map of the Indian Pines data set (16 land cover classes).

TABLE II
LAND COVER CLASSES WITH THE NUMBER OF SAMPLES
FOR THE INDIAN PINES DATA

Class	Land Cover Type	No. of Samples
C1	Stone-steel-towers	95
C2	Hay-windrowed	489
C3	Corn-min Till	834
C4	Soybean-no Till	968
C5	Alfalfa	54
C6	Soybean-clean Till	614
C7	Grass/Pasture	497
C8	Woods	1294
C9	Bldg-Grass-Tree-Drives	380
C10	Grass/Pasture-mowed	26
C11	Corn	234
C12	Oats	20
C13	Corn-no Till	1434
C14	Soybean-min Till	2468
C15	Grass/Trees	747
C16	Wheat	212
	Total	10366

of parameter setting and continuously increased classification accuracy achieved by our method. Then, the proposed CSRGFF approach is compared with other state-of-the-art methods in different small training sample scenarios.

A. Data Sets and Experimental Setup

Four real hyperspectral image data sets are used in the experiments and they are introduced as follows.

- 1) The first data set is the classic Indian Pines data set, which was acquired by the AVIRIS sensor over the area of Northwestern Indiana, and it contains 10 366 labeled pixels with 16 ground-truth classes. Since the data were collected over 20 years ago (1992), the spatial resolution is as low as 20 m/pixel, and the spatial dimension of the data is 145 × 145. Within the original 224 bands, four zero bands and 35 lower SNR bands affected by atmospheric absorption have been discarded in the experiments, and thus, the rest of the 185 bands are preserved. Its ground truth and detail information about each class are shown in Fig. 3 and Table II.
- 2) The second hyperspectral data set is the Salinas scene, which was also collected by the AVIRIS instrument but acquired in the place of the Salinas Valley, CA, USA. This image data set has a high spatial resolution of 3.7 m/pixel. The size of the Salinas scene is 512 × 217 with 224 spectral bands. Similarly, we discard 20 water absorption bands and the remainder 204 channels are used in our experiments. The Salinas data set is composed of 54 129 labeled pixels and 16 land-cover classes.

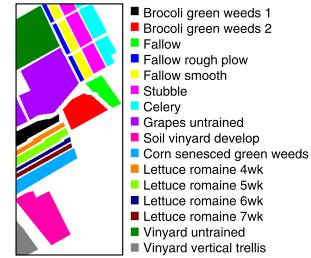


Fig. 4. Ground-truth map of the Salinas data set (16 land cover classes).

TABLE III
LAND COVER CLASSES WITH THE NUMBER OF SAMPLES FOR THE SALINAS DATA

Class	Land Cover Type	No. of Samples
C1	Brocoli-green-weeds-1	2009
C2	Brocoli-green-weeds-2	3726
C3	Fallow	1976
C4	Fallow-rough-plow	1394
C5	Fallow-smooth	2678
C6	Stubble	3959
C7	Celery	3579
C8	Grapes-untrained	11271
C9	Soil-vinyard-develop	6203
C10	Corn-senesced-green-weeds	3278
C11	Lettuce-romaine-4wk	1068
C12	Lettuce-romaine-5wk	1927
C13	Lettuce-romaine-6wk	916
C14	Lettuce-romaine-7wk	1070
C15	Vinyard-untrained	7268
C16	Vinyard-vertical-trellis	1807
	Total	54129

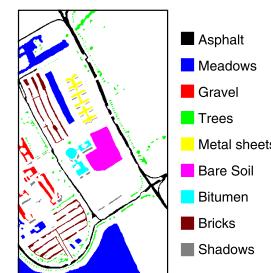


Fig. 5. Ground-truth map of the Pavia University data set (9 land cover classes).

Detailed information about its ground truth can be found in Fig. 4 and Table III.

- 3) Third, the Pavia University hyperspectral image data were acquired by another sensor ROSIS-03 and it has the highest spatial resolution of 1.3 m/pixel among these four data sets. After 12 noisy bands are omitted from the original 115 spectral bands, 103 bands are kept for processing. The spatial size of the image is 610 × 340 and it covers nine classes with 42 776 labeled pixels in total. Fig. 5 and Table IV display its detailed information.
- 4) Finally, the University of Houston hyperspectral data set distributed in the 2013 IEEE GRSS data fusion contest is considered [59]. The Houston University data set contains 144 spectral bands ranging between

TABLE IV
LAND COVER CLASSES WITH THE NUMBER OF SAMPLES
FOR THE PAVIA UNIVERSITY DATA

Class	Land Cover Type	No. of Samples
C1	Asphalt	6631
C2	Meadow	18649
C3	Gravel	2099
C4	Trees	3064
C5	Metal sheets	1345
C6	Bare Soil	5029
C7	Bitumen	1330
C8	Bricks	3682
C9	Shadows	947
	Total	42776



Fig. 6. Ground-truth map of the Houston University data set (15 land cover classes).

TABLE V
LAND COVER CLASSES WITH THE NUMBER OF SAMPLES
FOR THE HOUSTON UNIVERSITY DATA

Class	Land Cover Type	No. of Samples
C1	Healthy grass	1251
C2	Stressed grass	1254
C3	Synthetic grass	697
C4	Tree	1244
C5	Soil	1242
C6	Water	325
C7	Residential	1268
C8	Commercial	1244
C9	Road	1252
C10	Highway	1227
C11	Railway	1235
C12	Parking lot 1	1233
C13	Parking lot 2	469
C14	Tennis court	428
C15	Running track	660
	Total	15029

380 and 1050 nm and has a large spatial dimension of 1905×349 . Its spatial resolution is 2.5 m/pixel. There exist 15 surface materials and 15 029 labeled samples in total, and detailed information can be seen in Fig. 6 and Table V.

1) Algorithms and Parameter Setting: We compare our method with seven algorithms, including two deep learning-based methods (3D-CNN [28] and LSTM-CNN [30]), two state-of-the-art spatial-spectral-integrated methods (generalized composite kernel (GCK) machines [60] and local binary pattern-based extreme learning machine (LBP-ELM) [61]), and three 3-D Gabor-based methods. The first Gabor-based method combines the four 3-D Gabor Magnitude cubes \mathbf{M}_i , ($i = 1, \dots, 4$) together, and then SVMs with RBF kernel are used for classification, which is named 3DGM-SVM. The second Gabor-based method (3DGPC-HDM) was proposed in our previous work [44]. In this method, only Gabor phase features are used and coded for classification. Moreover, in order to illustrate the effectiveness

of the designed cascade SR, the last Gabor-based comparing method is implemented by conducting the weighting cube \mathbf{W} , which is called 3DG-MP. With respect to the training sample selection, here a fixed number of labeled samples per class are randomly picked out from the labeled set to constitute the training set, while the rest of the labeled samples are used for testing. Each experiment is carried out ten times to decrease the sampling effect. Moreover, two classic metrics, overall accuracy (OA) and Kappa coefficient (κ), are adopted to evaluate the classification results, in which OA reflects the whole classification accuracy, while κ also considers the impact of classes [62].

Since the balancing parameter λ in the ERS method can be adjusted automatically by a default value, our CSRGFF approach has only two important parameters now, i.e., the initial (B) and terminal (E) number of superpixels, which, respectively, correspond to the over-segmentation and under-segmentation circumstance. Actually, over-segmenting and under-segmenting hyperspectral images are easy to implement, as long as the number of superpixels is set to be large (hundreds) and small (tens), respectively. In Section IV-A2, the insensitivity of the two parameters to the classification performance has been investigated.

2) Parameter Analysis of the Initial (B) and Terminal (E) Number of Superpixels: In the ERS method, it can be easily found that the number of superpixels can greatly influence the quality of the generating superpixels, which will, therefore, affect the performance of the corresponding SR. Fortunately, our proposed cascade SR has cleverly solved the problem. More precisely, since the SR procedure is carried out from over-segmentation to under-segmentation, the parameters of the initial (B) and terminal (E) number of superpixels are easy to choose and the classification accuracy has been found to be insensitive and robust to the variation of the two parameters.

Specifically, concerning different initial numbers of superpixels B , which varies from 500 to 1000 with a step size of 100 (here the terminal number of superpixels E is set as 50), the OA with different sizes of training set on the four hyperspectral images is shown in Fig. 7(a)–(d). It can be directly seen that the curves with different initial numbers of superpixels B nearly overlap together. Furthermore, when ten labeled samples per class are used for training, Fig. 7(e) shows the OA with different initial numbers of superpixels B , which almost remains consistent across all over-segmentation sizes for the four data sets. All these observation implies that the classification performance is insensitive and robust to the variation of parameter B .

Alternatively, the terminal number of superpixels E is analyzed in the same way. Fig. 8(a)–(e), respectively, show the OA as the functions of training set and parameter E on the four hyperspectral images, where E varies from 10 to 50, and B is set as 500. Similarly, the variation of parameter E hardly affects the overall accuracy, implying the robustness of classification performance to E . Although the initial (B) and terminal (E) number of superpixels (corresponding to over-segmentation and under-segmentation) can take large and small values, they are, respectively, set as 500 and 50 for all four hyperspectral images to increase

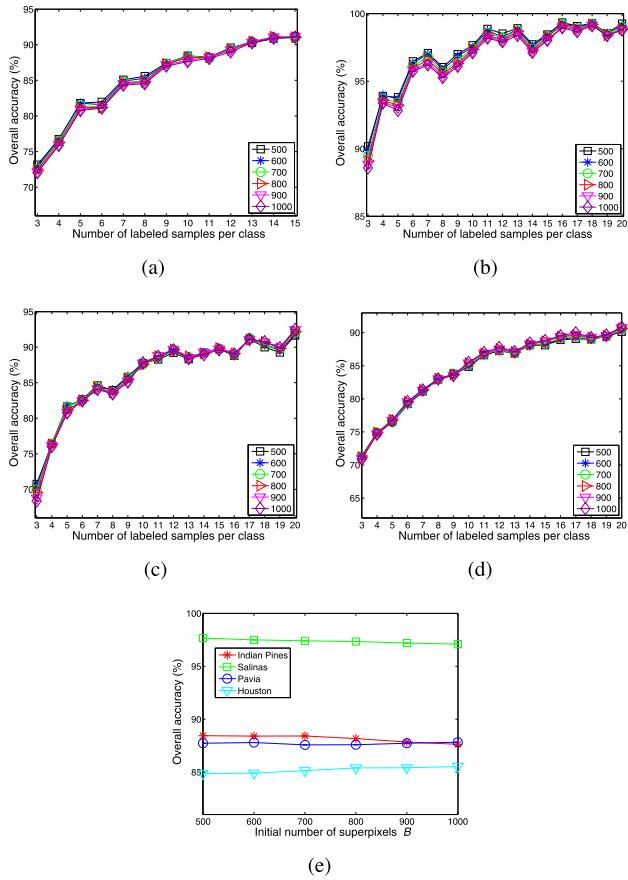


Fig. 7. Parameter analysis of the initial number of superpixels (B) on the four hyperspectral images. (a) Indian Pines, (b) Salinas, (c) Pavia, (d) Houston, and (e) overall accuracy versus the initial number of superpixels (B) on the four hyperspectral images with ten labeled samples per class as the training set.

the computational efficiency. Since a series of superpixel size from over-segmentation to under-segmentation in our cascade SR procedure has been fixed in advance (as aforementioned, the step size is set as 50) and can adapt to all data sets, the estimation of the number of extracted superpixels is avoided.

3) *Cascade Superpixel Regularization*: In order to demonstrate the superiority of the proposed cascade SR procedure, our CSRGFF approach is compared with single SR on the Gabor fused weighting cube \mathbf{W} , abbreviated as SSRGFF, which regularizes the weighting cube only one time. Fig. 9 shows the compared results of both methods on four hyperspectral images in the case of ten training samples per class, where the number of superpixels ranges from 500 to 50. It should be noted that for CSRGFF, the initial number of superpixels (B) is set as 500, while the values on the X-axis in Fig. 9 represent the terminal number of superpixels (E). For example, when E equals to 500, which is equal to the value of B , CSRGFF is just the same as SSRGFF. In addition, if E equals to 300, the SR process of CSRGFF is cascaded from 500 to 300. As expected, the performance of SSRGFF relies heavily on the number of superpixels, and it is unrealistic to accurately estimate the number of superpixels in practice. Alternatively, the classification accuracy of our

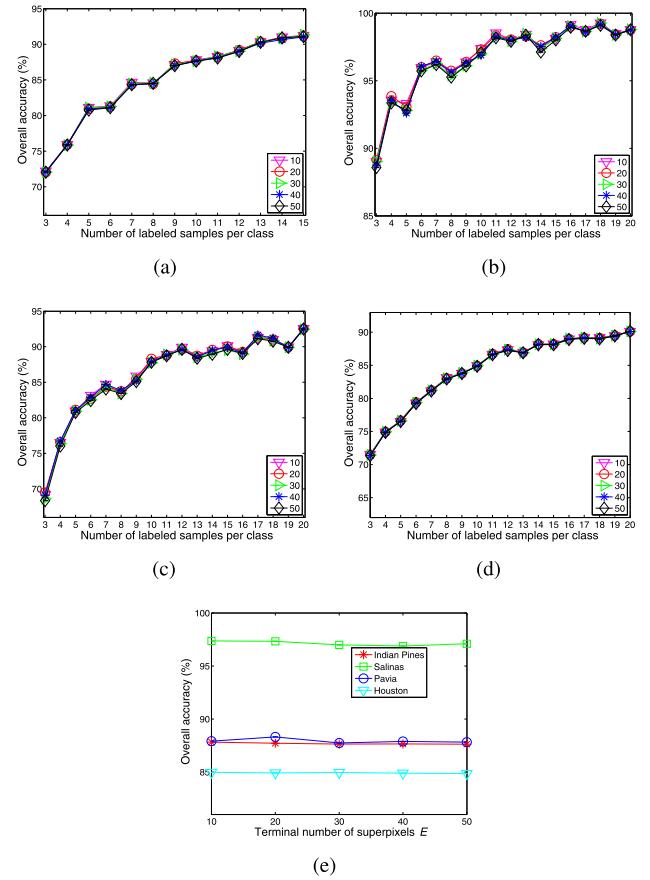


Fig. 8. Parameter analysis of the terminal number of superpixels (E) on the four hyperspectral images. (a) Indian Pines, (b) Salinas, (c) Pavia, (d) Houston, and (e) overall accuracy versus the terminal number of superpixels (E) on the four hyperspectral images with ten labeled samples per class as training set.

CSRGFF approach improves continuously as the decrease of parameter E , verifying the effectiveness of the proposed technique. Since the superpixel maps with the different numbers of superpixels can reflect the spatial structure of materials in different scales, that is, the homogeneity of each superpixel can be largely guaranteed in over-segmentation, while the shape and boundary attributes of surface objects can be well characterized in under-segmentation (shown in Fig. 1), it is rational to regularize the weighting cube from over-segmentation to under-segmentation in a cascade mode, and our CSRGFF approach provides a simple yet effective way to fuse the superpixel maps.

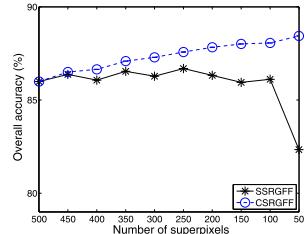
B. Experimental Results

In this section, the proposed CSRGFF approach has been quantitatively and visually compared with seven alternatives, including LSTM-CNN, 3D-CNN, GCK, LBP-ELM, 3DGM-SVM, 3DGPC-HDM, and 3DG-MP. The classification performance of the eight compared methods is exploited with different numbers of training sizes, as shown in Figs. 10–13. Here, each experiment is repeated ten times, and both the mean and standard variation are reported. Since the smallest set in Indian Pines hyperspectral data set only contains 20 labeled samples, i.e., the 12th class (Oats), the training set in Fig. 10

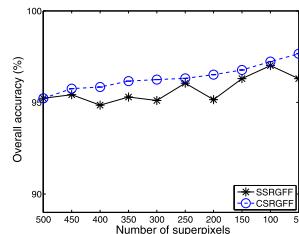
TABLE VI

CLASSIFICATION PERFORMANCE USING LSTM-CNN, 3D-CNN, GCK, LBP-ELM, 3DGM-SVM, 3DGPC-HDM, 3DG-MP, AND CSRGFF ON THE INDIAN PINES DATA SET WITH TEN LABELED SAMPLES PER CLASS AS TRAINING SET

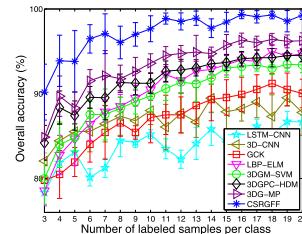
Class	LSTM-CNN		3D-CNN		GCK		LBP-ELM		3DGM-SVM		3DGPC-HDM		3DG-MP		CSRGFF	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
C1	97.26	1.66	95.47	5.41	94.94	3.04	98.00	2.22	97.18	3.05	98.59	2.13	96.24	3.83	97.29	3.92
C2	87.30	7.60	93.11	7.97	93.70	3.82	99.77	0.46	97.79	1.49	99.98	0.07	99.83	0.22	99.52	0.58
C3	48.32	7.64	57.90	13.69	61.75	9.94	70.12	9.09	77.89	6.82	71.53	9.57	76.58	9.23	83.45	11.32
C4	71.85	7.16	67.08	9.02	63.57	7.70	86.68	3.78	72.84	9.14	82.18	7.00	77.01	9.06	88.41	5.10
C5	82.04	5.86	96.85	5.31	84.55	8.95	99.77	0.72	97.27	3.68	98.64	2.87	97.27	3.68	100.00	0.00
C6	54.89	9.34	53.71	7.65	52.45	11.62	71.27	5.03	73.44	5.35	77.50	7.04	83.03	6.07	91.57	6.33
C7	69.42	5.34	83.32	7.21	78.99	7.63	83.72	9.42	82.48	9.24	81.40	7.70	84.21	9.63	85.26	9.03
C8	78.10	10.68	85.16	6.59	79.40	11.27	91.39	8.17	85.78	9.70	89.45	9.29	90.48	10.74	91.29	10.71
C9	50.84	14.92	74.29	6.83	74.19	9.61	93.57	7.77	92.70	7.05	94.08	7.05	94.24	7.03	90.78	7.57
C10	95.77	2.18	99.62	1.22	94.38	5.47	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00	96.25	5.27
C11	79.66	6.94	71.54	7.15	64.46	6.63	96.52	3.51	96.16	3.72	96.74	3.66	97.72	2.69	97.99	3.24
C12	98.00	4.83	100.00	0.00	98.00	4.22	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00	100.00	0.00
C13	55.17	10.78	52.54	8.51	47.21	9.08	69.92	7.46	68.16	9.71	74.00	10.04	73.92	8.62	78.55	8.83
C14	55.07	8.92	55.75	8.99	65.50	8.15	71.59	7.48	78.22	5.19	80.05	8.89	80.95	6.15	87.26	8.46
C15	84.08	6.81	91.95	3.74	90.00	3.80	88.49	9.36	89.08	5.12	88.78	5.34	92.40	3.76	95.14	4.17
C16	97.55	1.27	98.73	1.44	98.66	0.57	97.43	5.43	97.92	4.59	97.72	3.45	97.43	3.87	100.00	0.00
OA	65.25	2.26	68.47	2.74	68.53	2.11	80.52	2.35	80.46	1.57	82.89	2.33	83.86	2.16	88.44	2.57
Kappa	0.61	0.02	0.65	0.03	0.65	0.02	0.78	0.03	0.78	0.02	0.81	0.03	0.82	0.02	0.87	0.03



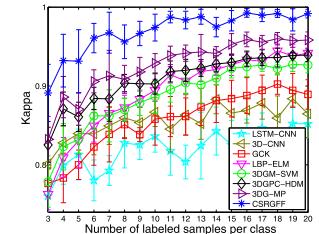
(a)



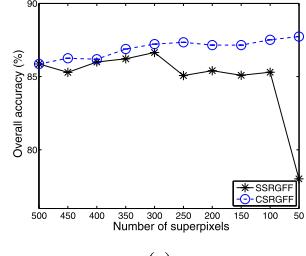
(b)



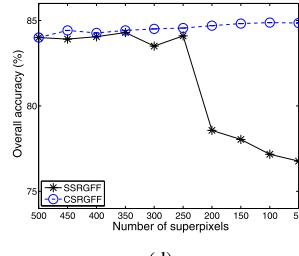
(a)



(b)

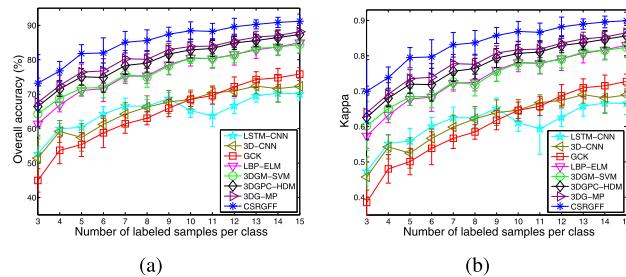


(c)

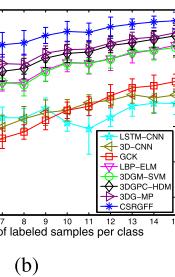


(d)

Fig. 9. Comparison of SSRGFF and CSRGFF with different numbers of superpixels on the four hyperspectral images when ten labeled samples per class are used for training. (a) Indian Pines. (b) Pavia. (c) Salinas. (d) Houston.

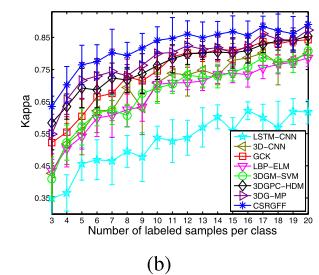


(a)

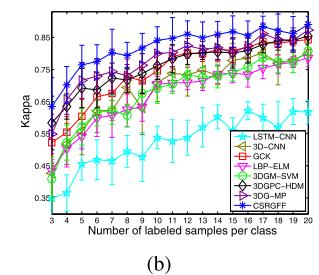


(b)

Fig. 10. Indian Pines hyperspectral data set. (a) OA and (b) Kappa as functions of the number of labeled samples per class.



(a)



(b)

Fig. 11. Salinas hyperspectral data set. (a) OA and (b) Kappa as functions of the number of labeled samples per class.

labeled samples. From the four figures, it can be clearly noted that our CSRGFF approach always outperforms the compared ones. It is worth to point out that the excellent performance of deep learning-based methods usually depends on a large amount of training samples that are used to optimize the huge number of parameters in the CNN model. For hyperspectral image classification, small sample size is a general problem in practice (which is the main concern of this paper), so the deep learning-based methods can be less efficient. When only ten labeled samples per class are used for training, the detailed classification performance on the four real hyperspectral images, including the accuracies with each class, OA, and Kappa, is listed in Tables VI–IX, respectively. Likewise,

TABLE VII

CLASSIFICATION PERFORMANCE USING LSTM-CNN, 3D-CNN, GCK, LBP-ELM, 3DGM-SVM, 3DGPC-HDM, 3DG-MP, AND CSRGFF ON THE SALINAS DATA SET WITH TEN LABELED SAMPLES PER CLASS AS TRAINING SET

Class	LSTM-CNN		3D-CNN		GCK		LBP-ELM		3DGM-SVM		3DGPC-HDM		3DG-MP		CSRGFF	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
C1	96.43	3.92	99.60	0.78	97.62	2.21	97.94	1.66	95.37	5.33	99.95	0.11	96.44	4.23	100.00	0.00
C2	98.66	1.58	99.92	0.13	98.83	0.66	95.34	2.33	97.10	5.25	99.88	0.17	99.07	1.89	99.52	1.51
C3	94.12	5.05	98.65	0.98	97.67	3.43	98.61	2.69	98.78	1.63	99.69	0.35	99.95	0.10	100.00	0.00
C4	99.43	0.62	98.43	1.53	98.98	0.82	96.98	3.16	95.17	2.05	98.02	1.02	96.14	2.02	98.43	4.91
C5	97.03	0.65	96.96	1.41	97.36	0.47	93.18	2.54	88.23	4.85	90.61	4.55	94.71	3.03	99.01	0.18
C6	99.67	0.24	98.85	1.43	98.12	1.65	89.00	6.75	98.26	2.07	99.82	0.39	99.24	1.16	99.85	0.00
C7	98.97	0.86	98.97	1.13	99.16	0.34	95.03	4.05	95.03	3.05	99.53	0.45	97.98	1.63	100.00	0.00
C8	57.67	15.69	68.37	10.84	63.34	12.82	81.03	7.76	71.51	11.14	74.64	8.98	84.22	10.18	93.37	8.85
C9	97.60	1.05	99.65	0.37	99.51	0.14	93.37	4.88	99.64	0.12	99.99	0.03	99.91	0.14	100.00	0.00
C10	86.01	4.35	89.49	3.79	92.03	3.34	91.30	3.74	95.64	1.85	94.91	3.18	96.69	2.23	95.93	3.73
C11	94.57	2.70	98.24	1.14	95.21	2.00	98.91	1.00	98.10	1.26	98.88	1.06	99.52	0.42	100.00	0.00
C12	99.34	0.43	99.94	0.12	99.70	0.18	90.53	3.72	98.52	1.31	98.93	0.91	98.48	1.66	100.00	0.00
C13	98.11	1.05	99.37	0.92	98.85	0.52	91.78	6.98	97.15	3.64	96.37	3.63	98.39	3.00	98.58	0.51
C14	91.88	2.65	97.17	1.80	92.12	5.06	92.61	4.12	94.62	5.60	95.32	3.22	96.03	4.16	94.90	4.55
C15	72.87	11.40	69.63	14.95	73.06	9.81	88.61	4.89	84.67	8.25	82.41	7.19	85.51	7.66	96.82	5.61
C16	94.37	3.43	96.02	1.79	96.47	2.76	96.85	7.74	96.94	1.67	99.72	0.44	98.91	1.00	99.15	0.30
OA	85.22	1.89	88.01	1.64	87.19	2.99	90.48	1.27	89.76	1.90	91.23	1.18	93.60	1.11	97.66	1.53
Kappa	0.84	0.02	0.87	0.02	0.86	0.03	0.89	0.01	0.89	0.02	0.90	0.01	0.93	0.01	0.97	0.02

TABLE VIII

CLASSIFICATION PERFORMANCE USING LSTM-CNN, 3D-CNN, GCK, LBP-ELM, 3DGM-SVM, 3DGPC-HDM, 3DG-MP, AND CSRGFF ON THE PAVIA UNIVERSITY DATA SET WITH TEN LABELED SAMPLES PER CLASS AS TRAINING SET

Class	LSTM-CNN		3D-CNN		GCK		LBP-ELM		3DGM-SVM		3DGPC-HDM		3DG-MP		CSRGFF	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
C1	55.32	9.20	70.38	9.52	83.47	6.03	62.00	4.90	70.21	7.33	72.48	7.28	76.35	6.94	83.33	7.41
C2	55.25	9.93	78.22	8.72	75.13	11.31	78.00	6.28	73.67	8.27	86.21	6.07	86.38	6.31	88.00	9.30
C3	41.55	24.79	82.40	6.12	79.94	7.76	83.61	8.05	82.92	7.61	89.12	6.23	85.95	5.89	97.10	5.29
C4	91.22	5.98	83.37	4.63	86.89	5.61	61.68	7.49	69.30	6.33	75.59	6.69	74.75	8.09	65.02	11.39
C5	99.45	0.36	99.69	0.46	89.33	11.08	91.60	5.99	99.64	0.38	98.26	1.19	99.95	0.14	99.33	0.34
C6	61.80	12.57	64.09	8.88	78.05	8.46	87.59	8.17	90.53	7.74	87.38	7.76	96.91	2.86	99.88	0.35
C7	92.26	5.86	88.60	3.68	94.75	1.19	93.30	4.87	89.78	4.89	97.99	1.28	93.36	3.92	100.00	0.00
C8	65.88	18.26	76.28	7.30	81.20	5.28	83.76	5.39	60.58	8.66	57.22	9.29	70.86	8.06	91.96	9.30
C9	99.92	0.12	97.55	1.07	97.95	2.36	60.11	4.57	85.71	4.30	71.48	7.81	85.80	5.83	51.30	10.23
OA	62.38	3.49	77.18	4.98	79.92	4.82	76.75	2.13	75.71	4.13	81.53	3.61	84.50	2.89	87.73	3.49
Kappa	0.54	0.03	0.71	0.06	0.75	0.06	0.70	0.02	0.69	0.05	0.76	0.04	0.80	0.03	0.84	0.04

TABLE IX

CLASSIFICATION PERFORMANCE USING LSTM-CNN, 3D-CNN, GCK, LBP-ELM, 3DGM-SVM, 3DGPC-HDM, 3DG-MP, AND CSRGFF ON THE HOUSTON UNIVERSITY DATA SET WITH TEN LABELED SAMPLES PER CLASS AS TRAINING SET

Class	LSTM-CNN		3D-CNN		GCK		LBP-ELM		3DGM-SVM		3DGPC-HDM		3DG-MP		CSRGFF	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
C1	91.11	5.68	92.29	5.23	85.29	8.09	62.95	5.83	72.88	7.16	86.08	4.24	91.72	6.14	83.97	8.45
C2	92.79	5.80	89.00	10.25	89.00	7.17	51.78	9.72	68.89	13.32	82.03	6.94	91.47	5.55	87.89	8.49
C3	96.07	6.74	95.64	1.99	99.53	0.64	93.45	4.09	96.51	2.23	99.53	0.37	94.70	2.60	100.00	0.00
C4	92.72	2.16	90.42	4.53	87.82	4.98	54.12	7.95	71.81	6.76	74.76	7.08	89.26	4.31	84.63	7.50
C5	94.91	6.12	91.86	5.97	96.85	1.79	82.91	5.10	81.60	10.23	95.74	4.07	93.56	7.99	98.34	1.26
C6	91.60	8.77	79.51	7.87	89.30	5.18	92.10	5.11	83.68	6.67	89.21	4.81	84.70	4.57	94.22	5.15
C7	69.20	10.36	64.60	12.51	77.62	7.66	60.76	7.97	66.44	9.87	62.80	5.33	71.03	11.17	74.07	9.29
C8	49.17	9.75	54.96	11.25	40.24	6.65	50.49	6.83	64.15	19.27	58.77	11.78	56.38	9.98	62.43	10.43
C9	68.64	12.66	53.28	9.08	73.06	8.14	48.73	9.67	71.46	10.81	50.31	5.24	76.78	8.79	80.17	6.63
C10	74.27	13.65	74.56	14.10	92.37	7.95	92.85	7.30	78.82	8.78	84.73	6.70	80.57	8.75	88.27	8.86
C11	71.35	5.04	71.88	5.86	84.16	7.80	92.16	4.13	69.94	8.30	77.16	6.66	76.14	9.11	94.14	6.14
C12	56.38	9.85	63.90	10.74	74.28	5.64	65.10	9.02	61.13	7.18	66.88	10.39	63.23	11.58	69.67	11.33
C13	56.99	11.26	35.54	5.09	54.64	9.58	87.08	4.94	82.77	5.99	83.05	5.07	75.71	8.91	87.87	4.86
C14	98.71	1.18	91.21	8.05	98.73	1.08	100.00	0.00	96.72	2.66	96.94	2.50	96.00	4.22	100.00	0.00
C15	97.71	2.92	93.50	3.52	97.88	1.27	90.45	5.87	89.15	5.15	95.49	6.00	97.08	1.27	100.00	0.00
OA	78.34	1.99	75.83	3.04	81.67	1.39	70.57	1.27	74.08	1.29	77.27	2.00	81.03	2.04	84.85	1.68
Kappa	0.77	0.02	0.74	0.03	0.80	0.02	0.68	0.01	0.72	0.01	0.75	0.02	0.80	0.02	0.84	0.02

the classification performance of CSRGFF is higher than the rest of the methods in most cases. Furthermore, Fig. 14 shows the training set (here ten labeled samples per class are used for training), the test set, and the classification maps obtained by various methods in one experiment. Obviously, the classification map achieved by CSRGFF [see Fig. 14(j)]

is more consistent with the ground truth [see Fig. 14(b)]. All these observations have demonstrated the superiority of the proposed CSRGFF approach over the other state-of-the-art ones. Due to the page limitation, the classification maps with the other three hyperspectral image data sets are omitted.

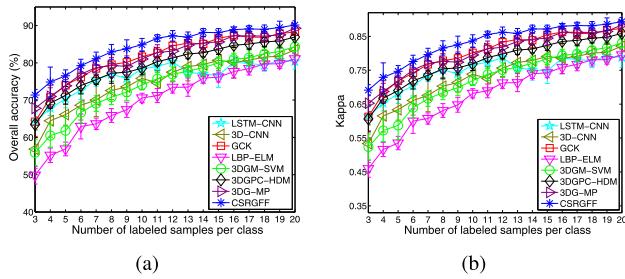


Fig. 13. Houston University hyperspectral data set. (a) OA and (b) Kappa as functions of the number of labeled samples per class.

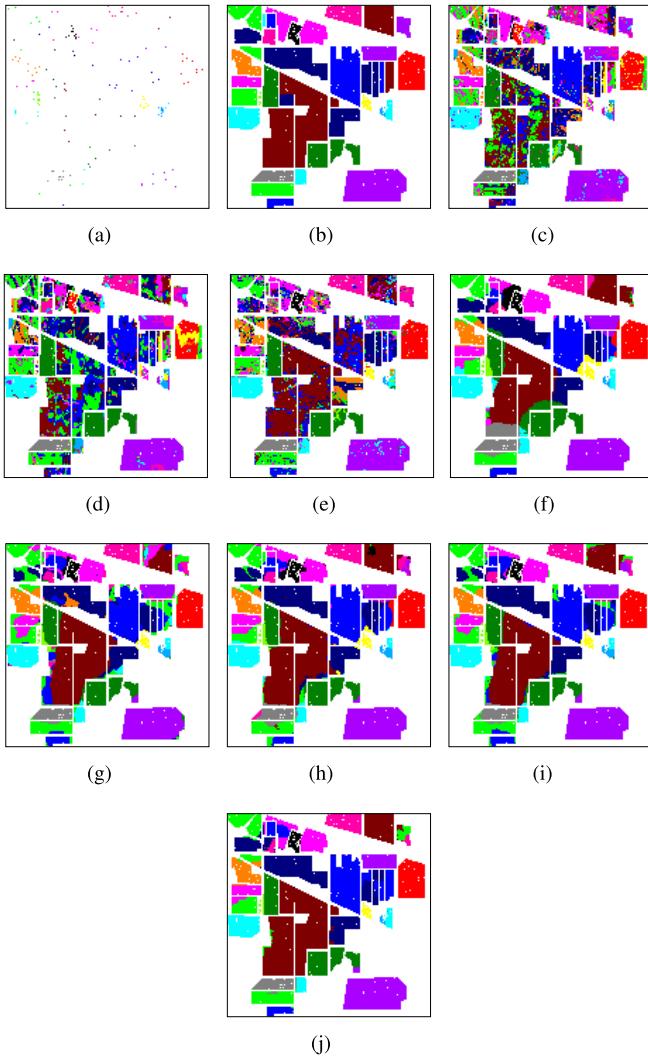


Fig. 14. Indian Pines data set. (a) Training set (ten labeled samples per class). (b) Test set (the rest of the labeled samples). Classification maps obtained by (c) LSTM-CNN (66.45%), (d) 3D-CNN (62.97%), (e) GCK (71.39%), (f) LBP-ELM (83.90%), (g) 3DGM-SVM (81.87%), (h) 3DGPC-HDM (88.11%), (i) 3DG-MP (87.68%), and (j) CSRGFF (92.55%).

V. CONCLUSION

In this paper, we have proposed a novel CSRGFF approach for hyperspectral image classification. On the one hand, the two kinds of powerful and complementary Gabor features, including magnitude and phase, have been jointly conducted,

which adequately utilized the discriminative information contained in the Gabor features. Meanwhile, since the characteristic of the two kinds of features is totally different, an SVM-based probability representation method has been proposed to make the results obtained from Gabor magnitude comply with the coding scheme of Gabor phase, and both kinds of features can thus be jointly investigated for classification without any weighting parameter. On the other hand, a cascade SR procedure has been introduced, which can effectively take advantage of the superpixel maps from over-segmentation to under-segmentation, and thus, the estimation of the number of extracted superpixels is unnecessary.

We have experimentally validated the insensitivity and robustness of the cascade SR process to the two important parameters, i.e., the initial (B) and terminal (E) number of superpixels, and hence, the values of all parameters in CSRGFF can be preset and remain unchanged for different hyperspectral image data sets. Meanwhile, the classification performance of CSRGFF is compared with several state-of-the-art methods, including LSTM-CNN, 3D-CNN, GCK, LBP-ELM, 3DGM-SVM, 3DGPC-HDM, and 3DG-MP. The experimental results always exhibit the superiority of the proposed CSRGFF approach. Since the cascade SR procedure is very simple and effective in practice, it can also be combined with other kinds of features to boost the classification performance.

REFERENCES

- [1] A. Plaza *et al.*, “Recent advances in techniques for hyperspectral image processing,” *Remote Sens. Environ.*, vol. 113, no. 1, pp. 110–122, Sep. 2009.
- [2] Q. Tong, Y. Xue, and L. Zhang, “Progress in hyperspectral remote sensing science and technology in China over the past three decades,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 1, pp. 70–91, Jan. 2014.
- [3] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [4] P. Ghamisi *et al.*, “Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art,” *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, Dec. 2017.
- [5] M. Imani and H. Ghassemian, “Feature extraction using attraction points for classification of hyperspectral images in a small sample size situation,” *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1986–1990, Nov. 2014.
- [6] N. Akhtar and A. Mian, “Nonparametric coupled Bayesian dictionary and classifier learning for hyperspectral classification,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 4038–4050, Aug. 2018.
- [7] 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.
- [8] J. A. Benediktsson, J. A. Palmason, and J. R. Steinsson, “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] W. Liao *et al.*, “Taking optimal advantage of fine spatial resolution: Promoting partial image reconstruction for the morphological analysis of very-high-resolution images,” *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 2, pp. 8–28, Jun. 2017.
- [10] P. Ghamisi, R. Souza, J. A. Benediktsson, X. X. Zhu, L. Rittner, and R. A. Lotufo, “Extinction profiles for the classification of remote sensing data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5631–5645, Oct. 2016.
- [11] M. D. Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, “Extended profiles with morphological attribute filters for the analysis of hyperspectral data,” *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5975–5991, Dec. 2010.

- [12] N. Falco, J. A. Benediktsson, and L. Bruzzone, "Spectral and spatial classification of hyperspectral images based on ICA and reduced morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 6223–6240, Nov. 2015.
- [13] Y. Gu, T. Liu, X. Jia, and J. A. Benediktsson, "Nonlinear multiple kernel learning with multiple-structure-element extended morphological profiles for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3235–3247, Jun. 2016.
- [14] X. Xu, J. Li, X. Huang, M. D. Mura, and A. Plaza, "Multiple morphological component analysis based decomposition for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 3083–3102, May 2016.
- [15] Q. Wang, J. Lin, and Y. Yuan, "Salient band selection for hyperspectral image classification via manifold ranking," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1279–1289, Jun. 2016.
- [16] S. Jia, G. Tang, J. Zhu, and Q. Li, "A novel ranking-based clustering approach for hyperspectral band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 88–102, Jan. 2016.
- [17] S. Yang, Z. Feng, M. Wang, and K. Zhang, "Self-paced learning-based probability subspace projection for hyperspectral image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 2, pp. 630–635, Feb. 2019.
- [18] J. Peng, L. Li, and Y. Y. Tang, "Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 6, pp. 1790–1802, Jun. 2019.
- [19] P. Zhong and R. Wang, "Jointly learning the hybrid CRF and MLR model for simultaneous denoising and classification of hyperspectral imagery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 7, pp. 1319–1334, Jul. 2014.
- [20] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognit.*, vol. 43, no. 7, pp. 2367–2379, 2010.
- [21] Y. Gao, X. Wang, Y. Cheng, and Z. Wang, "Dimensionality reduction for hyperspectral data based on class-aware tensor neighborhood graph and patch alignment," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 8, pp. 1582–1593, Aug. 2015.
- [22] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, Dec. 2017.
- [23] W. Shao and S. Du, "Spectral-spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544–4554, Oct. 2016.
- [24] F. Xing, Y. Xie, H. Su, F. Liu, and L. Yang, "Deep learning in microscopy image analysis: A survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 10, pp. 4550–4568, Oct. 2018.
- [25] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.
- [26] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 4959–4962.
- [27] M. Zhang, W. Li, and Q. Du, "Diverse region-based CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2623–2634, Jun. 2018.
- [28] Y. Li, H. Zhang, and Q. Shen, "Spectral-spatial classification of hyperspectral imagery with 3D convolutional neural network," *Remote Sens.*, vol. 9, no. 1, p. 67, 2017.
- [29] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639–3655, Jul. 2017.
- [30] Y. Xu, L. Zhang, B. Du, and F. Zhang, "Spectral-spatial unified networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5893–5909, Oct. 2018.
- [31] S. Yu, S. Jia, and C. Xu, "Convolutional neural networks for hyperspectral image classification," *Neurocomputing*, vol. 219, pp. 88–98, Jan. 2017.
- [32] B. Liu, X. Yu, P. Zhang, X. Tan, A. Yu, and Z. Xue, "A semi-supervised convolutional neural network for hyperspectral image classification," *Remote Sens. Lett.*, vol. 8, no. 9, pp. 839–848, Sep. 2017.
- [33] H. Wu and S. Prasad, "Semi-supervised deep learning using pseudo labels for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1259–1270, Mar. 2018.
- [34] Y. Qian, M. Ye, and J. Zhou, "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2276–2291, Apr. 2013.
- [35] B. Hou, T. Huang, and L. Jiao, "Spectral-spatial classification of hyperspectral data using 3-D morphological profile," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2364–2368, Dec. 2015.
- [36] S. Jia, J. Hu, J. Zhu, X. Jia, and Q. Li, "Three-dimensional local binary patterns for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 4, pp. 2399–2413, Apr. 2017.
- [37] T. C. Bau, S. Sarkar, and G. Healey, "Hyperspectral region classification using a three-dimensional Gabor filterbank," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3457–3464, Sep. 2010.
- [38] L. Shen and S. Jia, "Three-dimensional Gabor wavelets for pixel-based hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 5039–5046, Dec. 2011.
- [39] Y. Y. Tang, Y. Lu, and H. Yuan, "Hyperspectral image classification based on three-dimensional scattering wavelet transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2467–2480, May 2015.
- [40] J. Zhu, J. Hu, S. Jia, X. Jia, and Q. Li, "Multiple 3-D feature fusion framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 1873–1886, Apr. 2018.
- [41] S. Jia, K. Wu, J. Zhu, and X. Jia, "Spectral-spatial gabor surface feature fusion approach for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1142–1154, Feb. 2019.
- [42] S. Jia, L. Shen, and Q. Li, "Gabor feature-based collaborative representation for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 1118–1129, Feb. 2015.
- [43] S. Jia, J. Hu, Y. Xie, L. Shen, X. Jia, and Q. Li, "Gabor cube selection based multitask joint sparse representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3174–3187, Jun. 2016.
- [44] S. Jia, L. Shen, J. Zhu, and Q. Li, "A 3-D Gabor phase-based coding and matching framework for hyperspectral imagery classification," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1176–1188, Apr. 2018.
- [45] M.-Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 2097–2104.
- [46] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
- [47] L. Y. Fang, S. T. Li, X. D. Kang, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral images with a superpixel-based discriminative sparse model," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4186–4201, Aug. 2015.
- [48] J. Li, H. Zhang, and L. Zhang, "Efficient superpixel-level multitask joint sparse representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5338–5351, Oct. 2015.
- [49] Z. He, L. Liu, S. Zhou, and Y. Shen, "Learning group-based sparse and low-rank representation for hyperspectral image classification," *Pattern Recognit.*, vol. 60, pp. 1041–1056, Dec. 2016.
- [50] L. Fang, S. Li, W. Duan, and J. Ren, "Classification of hyperspectral images by exploiting spectral-spatial information of superpixel via multiple kernels," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6663–6674, Dec. 2015.
- [51] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, "Superpixel-based multitask learning framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2575–2588, May 2017.
- [52] T. Liu, Y. Gu, J. Chanussot, and M. Dalla Mura, "Multimorphological superpixel model for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 6950–6963, Dec. 2017.
- [53] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, "Local binary pattern-based hyperspectral image classification with superpixel guidance," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 749–759, Feb. 2018.
- [54] D. Gabor, "Theory of communication. Part 1: The analysis of information," *J. Inst. Elect. Eng. III, Radio Commun. Eng.*, vol. 93, no., pp. 429–457, Nov. 1946.
- [55] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, 2011.
- [56] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [57] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1148–1161, Nov. 1993.

- [58] A. Abdiansah and R. Wardoyo, "Time complexity analysis of support vector machines (SVM) in LibSVM," *Int. J. Comput. Appl.*, vol. 128, no. 3, pp. 28–34, Oct. 2015.
- [59] C. Debes *et al.*, "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2405–2418, Jun. 2014.
- [60] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4816–4829, Sep. 2013.
- [61] W. Li, C. Chen, H. Su, and Q. Du, "Local binary patterns and extreme learning machine for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3681–3693, Jul. 2015.
- [62] J. A. Richards, *Remote Sensing Digital Image Analysis—An Introduction*. New York, NY, USA: Springer, 2013.



Sen Jia (M'13–SM'17) received the B.E. and Ph.D. degrees from the College of Computer Science, Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively.

He is currently an Associate Professor with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China. His current research interests include hyperspectral image processing, signal and image processing, pattern recognition, and machine learning.



Zhijie Lin received the B.E. degree from the Guangzhou Medical University of Information Management and Information System, Guangzhou, China, in 2017. He is currently pursuing the master's degree in computer technology with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China.

His research interests include hyperspectral image classification, machine learning, and pattern recognition.



Bin Deng received the B.E. degree from the University of South China Agricultural, Guangzhou, China, in 2015. He is currently pursuing the master's degree in pattern recognition and intelligent system with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China.

His research interests include hyperspectral image classification, machine learning, and pattern recognition.



Jiasong Zhu received the B.E. degree from the Huazhong University of Science and Technology, Wuhan, China, in 1997, the M.E. degree from Wuhan University, Wuhan, in 2003, and the Ph.D. degree from The University of Hong Kong, Hong Kong, in 2008.

He is currently an Associate Professor with the College of Civil Engineering, Shenzhen University, Shenzhen, China. His current research interests include multi-sensor integration and data fusion, high-resolution image processing, and GIS applications in urban planning and transportation.



Qingquan Li received the M.S. degree in engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1988 and 1998, respectively.

From 1988 to 1996, he was an Assistant Professor with Wuhan University, where he became an Associate Professor from 1996 to 1998 and has been a Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing since 1998. He is currently the President of Shenzhen University, Shenzhen, China, where he is the Director of the Shenzhen Key Laboratory of Spatial Information Smart Sensing and Services. His current research interests include photogrammetry, remote sensing, and intelligent transportation systems.

Dr. Li is an Expert in modern traffic with the National 863 Plan and an Editorial Board Member of the *Surveying and Mapping Journal* and the *Wuhan University Journal—Information Science Edition*.