An Improved Feature Set for Hyperspectral Image Classification: Harmonic Analysis Optimized by Multiscale Guided Filter

Wei Zhang, Peijun Du, *Senior Member, IEEE*, Cong Lin, Pingjie Fu, Xin Wang, Xuyu Bai, Hongrui Zheng, Junshi Xia, *Senior Member, IEEE*, and Alim Samat, *Member, IEEE*

***Abstract*—Effective features derived from an original hyperspectral image (HSI) are quite important to improve the classification performance. An improved feature set, namely HGFM, is constructed by integrating harmonic analysis (HA) optimized by a multiscale guided filter (GF) with morphological operation for HSI classification. To establish HGFM, HA is first adopted to convert the HSI from spectral space to the frequency domain represented by amplitude, phase, and residual. With the first component of minimum noise fraction obtained from the original HSI as the guidance image, the harmonic components are then processed by the multiscale GF. Finally, the obtained results are then operated via morphological opening by reconstruction and closing by reconstruction to generate an improved feature set for classification. The HGFM features are input to an ensemble learning (EL) based on classification framework, in which EL plays an auxiliary role to enhance the classification stability and reliability. Three commonly used HSIs are used for experiments, and different feature sets are evaluated by comparing EL and rotation forest, support vector machine optimized by particle swarm optimization, random forest, and others. Compared with benchmark feature sets, the proposed HGFM feature set can better depict the details of objects easily, and the experimental results confirm the effectiveness in terms of classification accuracy and generalization ability.**

***Index Terms*—Feature set, harmonic analysis (HA), hyperspectral image (HSI) classification, morphological operation, multiscale guided filter.**

I. Introduction

A HYPERSPECTRAL remote sensing technique is able to capture an efficient description of the materials observed by the sensor with fine spectral resolution [1], [2]. With the advancement of imaging spectrometer technologies, the hyperspectral image (HSI) collects the spectral information of ground objects by using hundreds or thousands of narrow spectral channels, which generally range from the visible through near- and mid-infrared to the thermal infrared portions [3], [4]. Owing to high spectral resolution and good discrimination capacity, the HSI has been benefitting various practical applications, e.g., classification, target detection, data fusion, mineral mapping, environmental management, and so forth [5]–[9]. Classification is one of the most important tasks in hyperspectral remote sensing analysis [10]. However, the HSI is usually provided with high-dimensional and vast data volume, which brings the challenges of high computational cost and Hughes phenomenon in classification [11]. As a consequence, several critical efforts, such as feature extraction and band selection algorithms, are taken into consideration for dimensionality reduction to address the aforementioned issues.

Suitable features are crucial to improve the classification performance of the HSI, and feature extraction aims to transform the original data into specific feature space by certain criteria [12]. Spectral–spatial features, which incorporate spatial context and spectral information simultaneously, are commonly accepted to deal with the ill-posed problems to improve the performance of classification [13]–[15]. Compared with the global optimization method, the spectral–spatial features based on local optimization are more popular because of lower computational cost.

Numerous features have been developed in the past few decades; some researchers focus on the development of spatial features to facilitate the combination with spectral features, including gray-level co-occurrence matrix (GLCM) [16], extended morphological profiles (EMPs) [17], extended attribute profiles (EAPs) [18], and Gabor filtering features [19]. These methods have achieved good classification accuracies; however, massive features with a certain degree of redundancy could be difficult to use by the classifiers due to the lack of selection mechanism. In addition, some studies mainly exploit the spectral and spatial information separately, and then, the features are superimposed in series. For example, principal component analysis (PCA) [20] and minimum noise fraction (MNF) [21] are used to obtain spectral features. Spatial information can be obtained through filtering [22], morphology [23], low-rank representation [24], and so on. However, the features generated

Manuscript received April 22, 2020; revised June 22, 2020; accepted June 29, 2020. Date of publication July 3, 2020; date of current version July 16, 2020. This work was supported in part by the Natural Science Foundation of China under Grant 41631176. *(Corresponding author: Peijun Du.)*

Wei Zhang, Peijun Du, Cong Lin, Xin Wang, Xuyu Bai, and Hongrui Zheng are with the Key Laboratory for Land Satellite Remote Sensing Applications of Ministry of Natural Resources, Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, and the School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China (e-mail: zhangwrs@163.com; dupjrs@gmail.com; lcnjucumt@126.com; wangxrs@126.com; dg1727001@smail.nju.edu.cn; myzhenghr@126.com).

Pingjie Fu is with the College of Surveying and Geo-Informatics, Shandong Jianzhu University, Jinan 250101, China (e-mail: 13918@sdjzu.edu.cn).

Junshi Xia is with the Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Tokyo 103-0027, Japan (e-mail: junshi.xia@riken.jp).

Alim Samat is with the Xinjiang Institute of Ecology and Geography and the CAS Research Center for Ecology and Environment of Central Asia, Chinese Academy of Sciences, Ürümqi 830011, China (e-mail: alim.smt@gmail.com).

Digital Object Identifier 10.1109/JSTARS.2020.3006772

by this way are prone to carry noise, and the interest of spectral signatures may not be emphasized, affecting the final output of classification. Furthermore, another family of methods to obtain spectral–spatial features is based on image segmentation [25]–[27]. Various segmentation techniques are adopted to segment an HSI into different homogeneous regions according to intensity or texture, and the strategies of multiple method combination and majority voting are usually applied. Nevertheless, the qualities of features acquired by segmentation are dependent on advanced segmentation techniques that are time-consuming. Moreover, deep features have received widespread attention in recent years, in which a deconvolutional layer generates enlarged and dense graphs to extract high-level features [28]–[30]. Although deep features have produced outstanding classification results, hardware requirements and high computational complexity still need to be effectively mitigated.

To sum up, it is necessary to enhance the classification performance of HSIs with limited training samples by exploring and mining multiple feature extraction methods. Meanwhile, we should strengthen the research of frequency-domain information for a remotely sensed image. Different from the conventional spectral-domain-based and relatively complex theoretical methods, recently, harmonic analysis (HA) [31], [32] exerts good hyperspectral classification performance by converting the spectral signatures into multiple frequency-domain components, yielding more discriminative feature sets. In the remote sensing community, HA is mainly applied in time-series analysis, phenology, and change detection [33]–[35]. As a whole, the energy information of objects in the HA feature set is significantly highlighted, which leads to better identification of object structures and more accurate classification. Nevertheless, the interferences of noise and rough object boundaries in the HA features have a certain impact on the classification. Specifically, the noise in spectral and spatial domains may result in the salt-and-pepper effects in the classification maps. How to reduce the noise and mine powerful information through HA to improve the classification performance is worthy of further attention. A very convenient way to alleviate noise is filtering; a wide range of filtering methods have been proposed by using appropriate algorithms and sufficient data, such as mean filtering [36], median filtering [37], Gaussian filtering [38], and edge preservation filtering [22]. Among these methods, particularly, the edge preservation filtering can not only decrease image noise, but also keep the edge structure of objects clear rather than blurred. Therefore, edge-preserving filtering (e.g., bilateral filtering [39] and guided filtering (GF) [40]) is a possible valuable choice to address aforementioned problems of HA. However, to the best of our knowledge, very few efforts have made to cope with such concerns in terms of hyperspectral classification. Hence, a joint HA and GF feature extractor would be potential to preserve edge and reduce the presence of noisy or redundant features that degrade the classification performance.

In this article, the objective is to further exploit the feature set of HSIs from the frequency perspective and solve the limitations of traditional HA methods by embedding spatial filtering and morphological operations. We proposed an ensemble classification scheme based on the improved feature extraction method of HA optimized by multiscale GF (HGFM). The main innovations and contributions are: an HGFM feature set is constructed by integrating HA and multiscale GF with morphological operations, which is expected to optimize the feature representation and enhance the classification performance. Through the ensemble learning (EL) method, the desired applicability of the HGFM is comprehensively verified in high-, medium-, and low-spatial-resolution hyperspectral scenes.

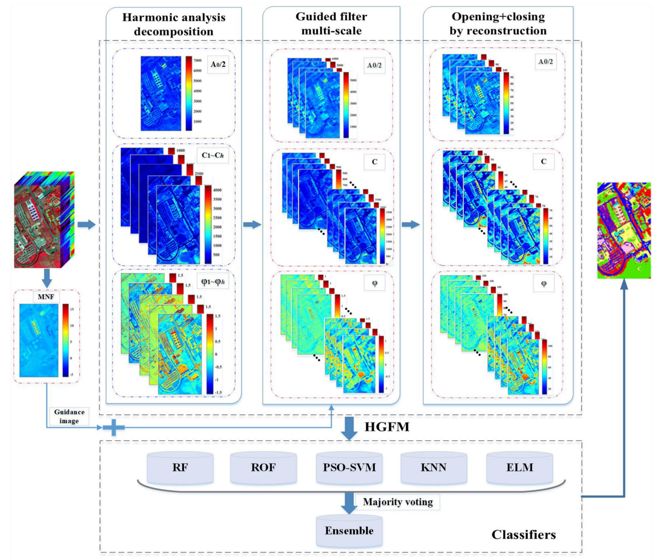


Fig. 1. Overall procedure of HSI classification by using the HGFM feature set based on EL.

The rest of this article is organized as follows. Section II reviews related approaches of HA and GF, and details the proposed method. The datasets, experimental setup, and discussion of results are presented in Section III. Finally, Section IV concludes this article.

II. Methods

Fig. 1 shows a schematic illustration of the classification method, consisting of four steps: HA for feature transform, GF for multiscale feature extraction, construction of HGFM, and classification based on EL.

A. HA for Feature Transform

Some studies have demonstrated HA's potential for HSI classification [31], [32]. It stands out in two aspects: 1) HA is designed to extract the features account for high-dimensional properties and information between different bands of HSI, through which more precise description can be acquired; and 2) the frequency-domain features produced by HA can reflect the target characteristics in multiple levels from the perspective of energy information, which is more distinguishable than the spectral-domain features.

The HA permits a spectral signature to be expressed as the sum of a series of overlapped sine and cosine waves (harmonics) [41], that is, energy components, such as amplitude, phase, and residual. The physical meaning of harmonics is to represent the

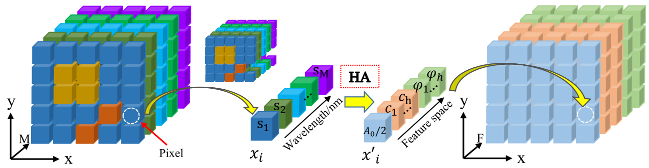


Fig. 2. Example of converting pixel's spectral signature of HSIs from the spectral domain to the frequency domain by using HA.

average energy spectrum of each pixel, express energy floating on different bands, and indicate the position where the amplitude occurs [30]. We define the notations that will be adopted throughout this part. Let $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_N] \in \mathbf{R}^{M\times N}$ be the HSI with an M-dimensional spectrum of each pixel $\boldsymbol{x}_i = [s_1, s_2, \ldots, s_M]^\top$, where N is the total number of pixels. The pixel-based spectral vector $\boldsymbol{x}_i$ can be expressed by harmonics through HA with formulations as

$$\hat{\boldsymbol{x}}_i = \frac{A_0}{2} + \sum_{h=1}^{h_{\max}} \left[A_h \cos\left(2\pi hi/M\right) + B_h \sin\left(2\pi hi/M\right)\right] \quad (1)$$

$$\hat{\boldsymbol{x}}_i = \frac{A_0}{2} + \sum_{h=1}^{h_{\max}} \left[C_h \sin\left(2\pi hi/M + \varphi_h\right)\right]. \quad (2)$$

In (1) and (2), the expressions of A_h, B_h, C_h, and φ_h are as follows:

$$A_h = \frac{M}{2}\sum_{i=1}^{M} x_i \cos\left(2\pi hi/M\right) \quad (3)$$

$$B_h = \frac{M}{2}\sum_{i=1}^{M} x_i \sin\left(2\pi hi/M\right) \quad (4)$$

$$C_h = \sqrt{{A_h}^2 + {B_h}^2} \quad (5)$$

$$\varphi_h = \arctan\left(-A_h/B_h\right) \quad (6)$$

where i denotes the band index, h and $h_{\max}$ refer to the index and number of harmonics, respectively, $C_h \sin(2\pi hi/M + \varphi_h)$ is the hth harmonic of x_i with M dimensions, and C_h and φ_h represent the amplitude and phase of the hth harmonic, respectively.

Through HA decomposition, feature sets $\boldsymbol{X}' = [\boldsymbol{x}_1^{'}, \boldsymbol{x}_2^{'}, \ldots, \boldsymbol{x}_N^{'}] \in \mathbf{R}^{(2\times h_{\max}+1)\times N}$ can be obtained for each pixel $\boldsymbol{x}' = [A_0/2, C_1, C_2, \ldots, C_{h_{\max}}, \varphi_1, \varphi_2, \ldots, \varphi_{h_{\max}}]$, $i \in \{1, 2, \ldots, N\}$. The final dimension of the HA feature set is $F = 2 \times h_{\max} + 1$ (see Fig. 2).

B. Guided Filter for Multiscale Feature Generation

To weaken interference of noise and rough boundary, the GF should be introduced. Beyond smoothing and denoising function, the GF conduces to transfer the spatial edge information of the guidance image[1] to the output image accurately and makes the filtered image more valuable [42]. The applications of GF mainly focus on image enhancement [43], target recognition [44], anomaly detection [45], etc. In essence, GF is based on the local linear relationship model between the guidance image and the input image, and its calculation time is not related to the size of the filter [42].

Specifically, assuming that the guidance image is $\boldsymbol{I}$ and the filtered output image $\boldsymbol{O}$ is obtained by the window ω_k with filter radius r centered at pixel k, O_i can be formulated as

$$O_i = a_k I_i + b_k \quad \forall i \in \omega_k \quad (7)$$

where $\boldsymbol{\omega}_k$ is a square window with the size of $(2r+1)\times(2r+1)$, and i represents pixel indexes, and a_k and b_k are the coefficients, which remain constant in the window $\boldsymbol{\omega}_k$.

To evaluate a_k and b_k, the minimization cost function in the window $\boldsymbol{\omega}_k$ can be defined as follows:

$$E(a_k, b_k) = \sum_{i\in\omega_k}\left((a_k I_i + b_k - p_i)^2 + \varepsilon {a_k}^2\right) \quad (8)$$

where ε denotes the regularization parameter to penalize large a_k. Furthermore, adopting the *linear ridge regression* [46], the key of a_k and b_k can be represented by

$$a_k = \frac{\frac{1}{|\omega|}\sum I_i p_i - \mu_k \bar{p}_k}{{\sigma_k}^2 + \varepsilon} \quad (9)$$

$$b_k = p_k - a_k \mu_k \quad (10)$$

$$\bar{p}_k = \frac{1}{|\omega|}\sum_{i\in\omega_k} p_i \quad (11)$$

where μ_k and ${\sigma_k}^2$ represent the mean and variance of $\boldsymbol{I}$ in $\boldsymbol{\omega}_k$, respectively. $|\omega|$ denotes the number of pixels in $bm\omega_k$, $\boldsymbol{p}$ is the filtering input image, and $\bar{p}_k$ refers to the mean of $\boldsymbol{p}$ in $\boldsymbol{\omega}_k$.

Nonetheless, the pixel i can be located in multiple different windows $\boldsymbol{\omega}_k$, which contribute to the values of $\{a_k, b_k\}$ and the change of O_i. Therefore, it is necessary to determine the mean of $\{a_k, b_k\}$ in $\boldsymbol{\omega}_k$ with pixel i as the center, and then, O_i can be given by [47]

$$O_i = \frac{1}{|\omega|}\sum_{k\in\omega_i}(a_k I_i + b_k) = \bar{a}_i I_i + \bar{b}_i. \quad (12)$$

Through (12), it can be deduced that $\boldsymbol{O}$ and $\boldsymbol{I}$ are linear in the window $\boldsymbol{\omega}_k$. Hence, when the guidance image $\boldsymbol{O}$ contains edge information, the output image $\boldsymbol{O}$ can retain the edge information at the corresponding position. Significantly, the two adjustable parameters involved in the calculation of GF are r and ε, which control the filter window size and blur degree, respectively. In this article, we use $G_r^{\varepsilon}(\boldsymbol{p}, \boldsymbol{I})$ to describe GF operations.

Especially, it should be highlighted that the single-scale feature generated by GF may induce the phenomenon of attenuating objects, causing difficultly to express the multiscale structure information of the objects in the HSI. Consequently, the tactics of multiscale factors should be considered for the generation of the GF feature set.

[1]The guided filter performs edge-preserving smoothing on an image by employing the content of other image, called a guidance image, to affect the filtering.

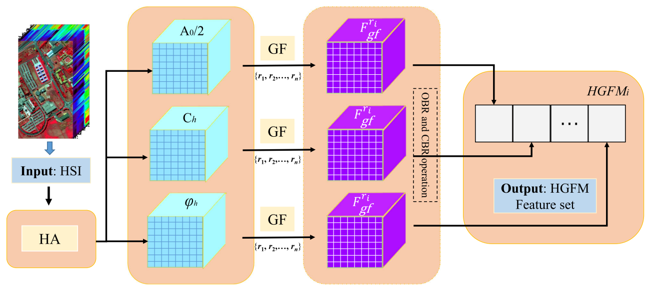


Fig. 3. Construction of the proposed HGFM feature set.

C. HGFM Feature Set

Integrating the different algorithms is one of the effective ways to improve the classification performance [48]. To significantly enhance the classification effect, the new feature set is expected to inherit the merits of HA and GF. The main procedures of the proposed HGFM feature set are depicted in Fig. 3. First, the HSI is carried out by HA to achieve the destination of dimensionality reduction and feature transform in the frequency domain. Then, the first component of MNF is used as the guidance image that determines the gradient information of the output image, and the extracted HA feature sets are filtered through multiscale GF. Simultaneously, the morphological opening by reconstruction (OBR) and closing by reconstruction (CBR) with the average operation are introduced to optimize acquired features. Finally, by concatenating all information into a single stacked vector, the new feature set HGFM can be derived.

The HA for HSIs is used to obtain the collection of harmonic features, which can be represented as

$$H_{f,h}(\boldsymbol{X}) = [\boldsymbol{A}_0/2, \boldsymbol{C}_i, \boldsymbol{\varphi}_i] \in \mathbf{R}^{(2\times h_{\max}+1)\times N} \tag{13}$$

where $H_{f,h}$ refer to the HA operation, X is the HSI, and $A_0/2$, C_i, and φ_i denote the remainder of HA, amplitude, and phase of the hth harmonic through HA function f along with h index of harmonics, respectively, $\boldsymbol{i} = \{1, 2, \ldots, h_{\max}\}$.

The objects are provided with scale attributes, and the window sequences of different sizes are helpful to obtain the multiscale detailed information of images [49], [50]. Inspired by this idea, the multiscale GF operation for the obtained HA feature set is carried out by using a series of window sequences, which can be expressed by

$$G_r^{\varepsilon}(\boldsymbol{p}, \boldsymbol{I}) = \left[\boldsymbol{F}_{gf}^{r_1}, \boldsymbol{F}_{gf}^{r_2}, \ldots, \boldsymbol{F}_{gf}^{r_n}\right] \tag{14}$$

where $\boldsymbol{p}$ and $\boldsymbol{I}$ represent HA feature set and guidance image, respectively, ε is a regularization parameter, and $\boldsymbol{F}_{gf}$ refers to the feature set acquired by GF with filtering radius $\boldsymbol{r} = \{r_1, r_2, \ldots, r_n\}$. Since ε has little contribution to the filtering output [51], it can be set to a fixed value of 10^{-4}. Concerning the choice of guidance image, MNF is utilized because it renders an optimal representation of the image in the signal-to-noise ratio sense.

Morphological operators are the collection of filters based on set theory, in which the two fundamental operations are erosion and dilation [52]. They are usually used to integrate contextual information of images based on a fixed-shape structural element (SE), which determines the neighborhood boundary of the pixels [53]. The dilation of the eroded image and the erosion of the dilated image are known as morphological opening and closing, respectively. Furthermore, morphological OBR and CBR are more effective than opening and closing operations, which easily lead to the breakage of the outline shape and the mismatch of the target in the image [54]. Thus, OBR and CBR are applied to optimize the features acquired by multiscale GF. Morphological reconstruction includes two images: the former is a marker, which belongs to the starting point of transformation, and the latter is a mask that is used to constrain the transformation process. Let δ and ρ denote erosion and dilation, respectively, and let $\boldsymbol{g}$ be the grayscale image, $\boldsymbol{b}$ be the SE. Mathematically, OBR and CBR can be given by

$$RO_g^{\rho}(\boldsymbol{\mu}) = \min\left\{x_{\boldsymbol{g}}, \rho_b^k(\delta_b(\boldsymbol{g}))\right\} |\rho_b^k(\delta_b(\boldsymbol{g})) \tag{15}$$

$$RC_g^{\delta}(\boldsymbol{\mu}) = \max\left\{x_{\boldsymbol{g}}, \delta_b^k(\rho_b(\boldsymbol{g}))\right\} |\delta_b^k(\rho_b(\boldsymbol{g})) \tag{16}$$

where $RO_g^{\rho}(\boldsymbol{\mu})$ means OBR from marker bmu to $\boldsymbol{g}$ in (15), $\boldsymbol{\mu} = \delta_b(\boldsymbol{g})$, and $RC_g^{\delta}(\boldsymbol{\mu})$ means CBR from marker $\boldsymbol{g}$ to $\boldsymbol{g}$ in (16), $\boldsymbol{\mu} = \rho_b(\boldsymbol{g})$. The process of morphological reconstruction is iterated until the reconstructed image at iteration k is the same as image obtained at iteration $k-1$.

To make the approach more effective, two crucial strategies are enabled in this study: the first one is to adopt the SE with flat circle, and the latter is to embrace the average results of successive OBR and CBR operations. Both strategies can maintain the rich texture information, as well as decrease the dimensions of the feature.

Ultimately, on the basis of integrating HA and GF, the proposed HGFM feature set can be represented by

$$\mathrm{HGFM}(\boldsymbol{F}_{gf}^{r_i}) = \mathrm{RC}_{\mathrm{g}}^{\delta}(\rho_b(\boldsymbol{F}_{gf}^{r_i}))|\mathrm{RO}_{\mathrm{g}}^{\rho}(\delta_b(\boldsymbol{F}_{gf}^{r_i})) = [\boldsymbol{HGFM}_i] \tag{17}$$

where r_i is the filtering radius of the window, in which $\boldsymbol{i} = \{1, 2, \ldots, n\}$, and $\boldsymbol{HGFM}_i$ represents the acquired HGFM feature set.

D. EL for Classification

There is no single classifier suitable for all classification tasks, and the generalization ability of the ensemble method could be better than that of an individual classifier [55], [56]. EL is a suitable alternative approach to deal with challenges of multisource data and multidisciplinary application [57]. By combining the outputs of multiple single classifiers with some approaches (e.g., majority voting, Bayesian average, etc.), EL overcomes the dilemma of no free lunch [56] and is able to produce more accurate classification results, and its effectiveness has been demonstrated from statistical, expressional, and computational perspectives [57]–[59]. Thus, EL is conducive to enhancing the adaptability of the HGFM feature set, and it is also a powerful way to construct the framework for HSI classification.

The suitable algorithms, including random forest (RF), rotation forest (RoF), support vector machine optimized by particle swarm optimization (PSO-SVM), K-nearest neighbor (KNN),

TABLE I
DIMENSIONS OF FEATURE SETS OBTAINED BY DIFFERENT EXTRACTION METHODS

HSI[1]	RAW	HA	GF	HGFM	EAP_a	EAP_d	EAP_i	EAP_s	EMP	GLCM	IFRF
A	200	35	36	34	36	36	36	36	36	32	20
B	103	35	36	34	36	36	36	36	36	32	21
C	204	35	36	34	36	36	36	36	36	32	21

Note: A—Indian Pines, B—Pavia University, C—Salinas.

TABLE II
CLASSIFICATION ACCURACIES OF UNIVERSITY OF PAVIA ROSIS USING DIFFERENT FEATURE SETS BASED ON EL

Class	Samples		RAW	HA	GF	HGFM	EAP_a	EAP_d	EAP_i	EAP_s	EMP	GLCM	IFRF
	Train	Test											
Asphalt	50	6581	72.53±0.03	50.41±0.03	77.85±0.03	95.64±0.01	96.55±0.02	92.38±0.01	95.02±0.01	73.09±0.02	92.35±0.02	77.18±0.03	80.65±0.06
Meadows	50	18599	78.51±0.04	74.06±0.05	86.35±0.02	96.18±0.02	92.02±0.03	83.84±0.03	94.47±0.02	95±0.01	90.33±0.04	81.38±0.03	95.31±0.01
Gravel	50	2049	78.77±0.04	72.15±0.03	85.49±0.03	93.35±0.02	95.7±0.02	86.56±0.04	97.62±0.01	95.9±0.02	93.62±0.03	60.67±0.04	94.08±0.02
Trees	50	3014	94.22±0.02	93.45±0.02	96.07±0.01	96.07±0.02	96.67±0.01	94.25±0.02	96.41±0.01	94.89±0.02	99.03±0	92.74±0.02	90.82±0.02
Painted metal sheets	50	1295	99.23±0	97.41±0.01	99.95±0	99.88±0	99.34±0	99.19±0	99.59±0	99.69±0	99.83±0	99.64±0	99.51±0
Bare Soil	50	4979	72.21±0.03	71.18±0.06	86.92±0.04	98.47±0.01	92.46±0.02	88.04±0.04	98.76±0.01	98.32±0.01	92.73±0.02	70.45±0.06	98.77±0.01
Bitumen	50	1280	89.83±0.01	88.59±0.02	93.58±0.01	97.55±0.01	99.83±0	96.61±0.01	99.84±0	93.59±0.02	96.3±0.01	85.86±0.03	98.09±0.01
Self-Blocking Bricks	50	3632	71.64±0.05	65.15±0.04	90.95±0.02	95.44±0.01	95.63±0.01	92.22±0.02	96.4±0.02	87.42±0.03	97.5±0.01	62.27±0.04	82.43±0.04
Shadows	50	897	98.8±0	97.5±0.01	99.86±0	98.53±0.01	99.84±0	99.78±0	99.96±0	100±0	99.65±0	98.91±0.01	80.84±0.06
OA (%)	-	-	78.79±0.02	72.22±0.02	87.06±0.01	**96.36±0.01**	94.22±0.01	88.45±0.01	95.95±0.01	91.58±0.01	92.99±0.02	78.67±0.01	91.86±0.01
AA (%)	-	-	83.97±0.01	78.88±0.01	90.78±0.01	96.79±0	96.45±0	92.54±0	**97.56±0**	93.10±0	95.7±0.01	81.01±0.01	91.17±0.01
κ	-	-	0.73±0.02	0.65±0.02	0.83±0.01	**0.95±0.01**	0.92±0.02	0.85±0.01	0.95±0.01	0.89±0.01	0.91±0.02	0.72±0.01	0.89±0.02

and extreme learning machine (ELM), are selected as the base classifiers for multiple EL in this article. The RF is a statistical learning method by using the bootstrap resampling theory, which has been proven to have a high predictability accuracy and fine tolerance for outliers and noise [60], [61]. The RoF focuses on generating the rotation feature space by using PCA to improve feature diversity, and it performs a good performance on improving the classification results [62], [63]. The PSO-SVM is a powerful classifier in dealing with small-size training samples and nonlinear and high-dimensional problems [64]. KNN is a theoretically mature method, in which the category of unlabeled data is determined by the samples in setting distance range. The simplicity and small-size training samples of KNN are in favor of classification [65]. ELM is a commonly used algorithm to solve single-hidden-layer feedforward networks, which is widely used in various fields due to the ability of fast learning, good versatility, and simple parameter setting [66]. Moreover, the majority voting is adopted as a combination strategy of EL.

Finally, the proposed classification framework is constructed based on the improved HGFM feature set and the EL (see Fig. 1).

III. EXPERIMENTS AND ANALYSIS

A. Hyperspectral Datasets

To investigate the classification performance of proposed the approach, three public hyperspectral datasets are adopted in our experiments, which are available online. They are gathered by airborne visible infrared imaging spectrometer (AVIRIS) and reflective optics spectrographic imaging system (ROSIS) sensors.

1) University of Pavia ROSIS: This image was acquired by the ROSIS over University of Pavia, Italy, on July 8, 2002. The image size is 610 × 340 pixels, with 103 spectral bands after 12 noisy bands removal (wavelength range 0.43–0.86 m) and 1.3-m/pixel geometric resolution. The ground truth map consists of nine classes and 42 776 labeled pixels, as shown in Table II. The three-band color composite image and the ground truth map are depicted in Fig. 4(a).

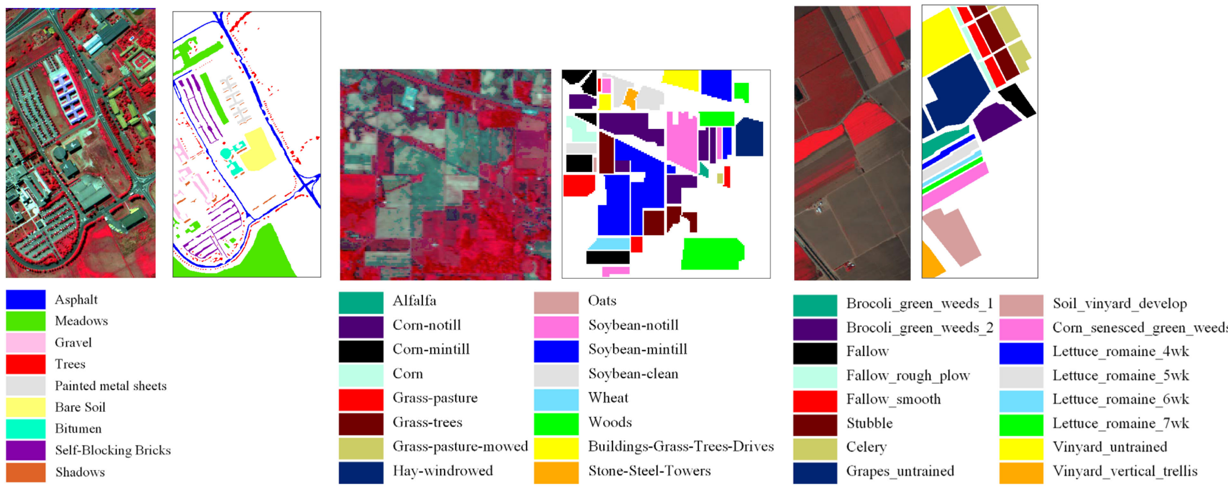
2) Indian Pines AVIRIS: It was captured by the AVIRIS sensor over the region of northwestern Indiana, USA, on June 12, 1992. The image is composed of 145×145 pixels and 200 spectral channels (wavelength range of 0.4–2.5 m) with the spatial resolution of 20 m/pixel. The ground truth of the scene contains 16 classes with 10 249 samples in total, which are detailed in Table III. Meanwhile, the false color image of three bands and ground truth map are described in Fig. 4(b).

3) Salinas AVIRIS: This scene was collected by AVIRIS, which covers the area of Salinas Valley, California. It has a size of 512×217 pixels and is characterized by high spatial resolution of 3.7 m/pixel and 204 bands (wavelength range of 0.4–2.5 m) after discarding several water absorption bands. The scene reference data comprise a total of 54 129 labeled pixels distributed in 16 classes (see Table IV); false-color composite image and the ground truth map are shown in Fig. 4(c).

B. Experimental Settings

In order to fully investigate the effectiveness of HGFM, we compare it with other benchmark features obtained from the full bands (RAW), HA, GF, and EAPs, where EAPs specifically include the area of the region (EAP_a), diagonal of the box bounding the regions (EAP_d), the moment of inertia (EAP_i), and standard deviation (EAP_s) [49]. Besides, three more feature extraction methods are added for comparison, which are GLCM [67], EMPs [68], and image fusion and recursive filtering (IFRF) [69]. In all experiments, it should be noted that the notations RAW, HA, GF, EAP_a, EAP_d, EAP_i, and EAP_s represent the feature set acquired by corresponding methods. The dimensions of all feature sets are described in Table I. In addition, the accuracy measures of individual class accuracy [%], overall accuracy (OA) [%], average accuracy (AA) [%], and kappa coefficient (κ) along with standard deviation are adopted to assess the classification performance. The classification results are obtained by the average of ten independent Monte Carlo runs.

In the feature extraction procedure, the h of HA is set to 8, the guidance image I used in GF is the first principal component of



(a) (b) (c)

Fig. 4. False-color image of HSI and the corresponding ground truth map of (a) The University of Pavia ROSIS, (b) Indian Pines AVIRIS, and (c) Salinas AVIRIS.

TABLE III
CLASSIFICATION ACCURACIES OF INDIAN PINES AVIRIS USING DIFFERENT FEATURE SETS BASED ON EL

Class	Samples		RAW	HA	GF	HGFM	EAP$_a$	EAP$_d$	EAP$_i$	EAP$_s$	EMP	GLCM	IFRF
	Train	Test											
Alfalfa	5	41	44.63±0.14	41.71±0.17	80.49±0.1	95.12±0.04	95.37±0.01	95.12±0	89.27±0.06	90.24±0.05	92.2±0.03	20±0.09	93.66±0.05
Corn-no till	143	1285	70.57±0.02	65.97±0.03	87.95±0.01	92.96±0.01	87.34±0.02	86.93±0.02	85.04±0.02	80.64±0.03	91.97±0.02	55.31±0.02	94.37±0.02
Corn-min till	83	747	57.68±0.04	56.17±0.05	76.96±0.05	96.77±0.02	86.53±0.02	84.78±0.02	87.52±0.02	57.10±0.03	93.82±0.02	37.62±0.04	96.87±0.01
Corn	24	213	45.31±0.08	43.05±0.06	56.76±0.07	94.84±0.03	81.55±0.05	82.58±0.04	68.83±0.05	59.30±0.07	90.33±0.06	36.01±0.05	87.89±0.06
Grass-pasture	48	435	87.45±0.02	88.14±0.02	79.72±0.05	95.75±0.02	91.31±0.02	90.37±0.02	89.31±0.02	87.89±0.03	91.82±0.03	48.57±0.04	94.09±0.02
Grass-trees	73	657	97.64±0.01	96.29±0.02	95.69±0.02	99±0.01	98.26±0.01	97.50±0.02	98.68±0.01	96.85±0.01	99.3±0	89.27±0.03	99.21±0.01
Grass-pasture-mowed	3	25	72.8±0.08	70.4±0.13	86.8±0.08	92.8±0.04	86.80±0.06	81.60±0.13	67.20±0.13	62.80±0.11	92.8±0.06	6.4±0.07	99.2±0.03
Hay-windrowed	48	430	98.98±0.01	99±0.01	99.07±0.01	99.88±0	99.91±0	99.91±0	100±0	99.88±0	99.91±0	96.98±0.02	99.93±0
Oats	2	18	16.11±0.08	14.44±0.09	43.89±0.2	88.33±0.13	19.44±0.11	20.56±0.12	30.56±0.14	34.44±0.11	81.11±0.24	5.56±0.05	62.22±0.21
Soybean-no till	97	875	72.53±0.02	69.29±0.04	88.46±0.02	91.31±0.03	87.35±0.02	88.53±0.02	87.05±0.02	78.67±0.02	90.71±0.01	49.62±0.03	97.67±0.01
Soybean-min till	246	2209	86.79±0.02	85.23±0.02	93.82±0.02	97.09±0.01	94.99±0.01	95.47±0.01	96.12±0.01	90.73±0.01	96.32±0.01	81.2±0.02	98.82±0
Soybean-clean	59	534	60.51±0.03	48.86±0.04	90.3±0.01	89.34±0.03	82.49±0.02	83.61±0.03	81.14±0.02	84.61±0.03	89.55±0.02	31.97±0.04	94.57±0.02
Wheat	21	184	97.28±0.01	95.92±0.02	98.32±0.01	99.51±0	99.08±0	99.08±0	97.12±0.02	97.83±0.02	99.08±0.01	73.26±0.1	99.62±0
Woods	127	1138	97.22±0.01	97.22±0.01	99.33±0	99.67±0	98.16±0.01	99.59±0	99.76±0	99.26±0	99.13±0	90.95±0.02	99.49±0
Buildings-Grass-Trees-Drives	39	347	46.48±0.05	46.54±0.03	84.93±0.06	98.76±0.01	92.97±0.03	94.67±0.02	86.60±0.04	96.71±0.02	98.39±0.01	49.51±0.04	97.2±0.01
Stone-Steel-Towers	9	84	90.24±0.06	89.05±0.03	98.81±0.01	100±0	95.95±0.03	97.02±0.02	94.88±0.03	98.21±0.01	98.33±0.02	43.21±0.08	94.64±0.05
OA (%)	-	-	79.36±0.01	77.07±0.01	90.09±0.01	96.06±0	92.01±0	92.26±0	91.43±0	86.40±0.01	95.06±0	65.83±0.01	**97.20±0**
AA (%)	-	-	71.39±0.02	69.21±0.02	85.08±0.02	**95.7±0.01**	87.34±0.01	87.33±0.01	84.94±0.02	82.20±0.01	94.05±0.02	50.96±0.01	94.34±0.02
κ	-	-	0.76±0.01	0.74±0.01	0.89±0.01	**0.96±0**	0.91±0	0.91±0	0.90±0	0.84±0.01	0.94±0	0.60±0.01	0.96±0

TABLE IV
CLASSIFICATION ACCURACIES OF SALINAS AVIRIS USING DIFFERENT FEATURE SETS BASED ON EL

Class	Samples		RAW	HA	GF	HGFM	EAP$_a$	EAP$_d$	EAP$_i$	EAP$_s$	EMP	GLCM	IFRF
	Train	Test											
Brocoli_green_weeds_1	50	1959	99.42±0	98.9±0.01	99.87±0	99.93±0	99.39±0	99.11±0	100±0	98.43±0.01	98.75±0.01	17.88±0.03	100±0
Brocoli_green_weeds_2	50	3676	98.36±0.01	98.53±0	99.46±0	99.53±0	99.57±0	98.05±0.02	99.82±0	98.70±0.01	99.66±0	55.18±0.04	99.39±0.01
Fallow	50	1926	95.05±0.02	97.95±0.01	95.64±0.03	99.83±0	99.48±0.01	98.95±0.01	99.91±0	98.96±0.01	99.06±0.01	90.94±0.02	100±0
Fallow_rough_plow	50	1344	99.45±0	99.52±0	98.84±0.01	99.82±0	99.46±0	99.35±0	99.32±0.01	99.45±0	99.88±0	97.95±0.01	99.23±0
Fallow_smooth	50	2628	97.04±0.02	97.56±0.01	97.02±0.01	97.64±0.01	98.53±0.01	97.88±0.01	98.50±0	98.57±0	96.41±0.01	90.32±0.03	99.14±0
Stubble	50	3909	99.51±0	99.42±0	99.64±0	99.92±0	99.72±0	98.68±0.01	99.72±0	99.65±0	99.83±0	99.44±0.01	99.86±0
Celery	50	3529	99.30±0	99.25±0	99.31±0	99.31±0.01	99.53±0	99.20±0	99.60±0	99.76±0	98.01±0.01	91.55±0.01	99.71±0
Grapes_untrained	50	11221	76.80±0.04	73.78±0.02	73.75±0.04	91.86±0.02	83.52±0.03	70.22±0.02	82.52±0.04	70.32±0.06	83.08±0.03	57.31±0.09	88.43±0.03
Soil_vinyard_develop	50	6153	97.30±0.01	97.27±0.01	99.06±0	99.81±0	98.09±0.01	93.10±0.02	99.27±0	98.90±0.01	99.19±0.01	93.17±0.02	99.98±0
Corn_senesced_green_weeds	50	3228	88.38±0.02	89.85±0.02	89.75±0.02	97.38±0.02	96.63±0.01	92.94±0.02	95.60±0.01	91.95±0.01	96.61±0.01	84.61±0.02	99.47±0
Lettuce_romaine_4wk	50	1018	96.22±0.01	97.79±0.02	97.06±0.01	99.97±0	96.72±0.01	94.61±0.02	97.07±0.01	95.21±0.01	97.56±0.01	96.55±0.02	99.42±0.01
Lettuce_romaine_5wk	50	1877	99.93±0	99.91±0	99.57±0	100±0	99.91±0	94.42±0.02	99.99±0	99.86±0	99.93±0	95.9±0.02	99.49±0.01
Lettuce_romaine_6wk	50	866	98.04±0.01	97.81±0.01	97.41±0.02	98.04±0.01	98.26±0.01	97.34±0.01	98.23±0.01	98.57±0.01	97.79±0.01	78.6±0.04	98.41±0.01
Lettuce_romaine_7wk	50	1020	94.34±0.01	95.96±0.01	95.85±0.01	97.66±0.01	97.33±0.02	93.98±0.03	95.52±0.02	94.26±0.02	98.6±0.01	72±0.04	99.01±0.01
Vinyard_untrained	50	7218	66.97±0.04	67.65±0.02	82.52±0.03	95.08±0.01	86.85±0.03	76.90±0.04	86.22±0.03	68.66±0.05	87.19±0.03	57.76±0.08	93.94±0.03
Vinyard_vertical_trellis	50	1757	97.72±0.01	97.00±0.02	97.35±0.02	98.97±0.01	100±0	99.77±0	99.13±0	99.99±0	98.95±0.01	94.57±0.02	99.28±0
OA (%)	-	-	88.79±0	88.48±0	90.65±0.01	**97.12±0.01**	93.97±0.01	88.44±0.01	93.75±0.01	88.36±0.01	93.83±0.01	74.75±0.01	96.70±0.01
AA (%)	-	-	93.99±0	94.26±0	95.13±0	**98.42±0**	97.06±0	94.03±0	96.90±0	94.45±0	96.91±0	79.61±0.01	98.42±0
κ	-	-	0.88±0	0.87±0	0.90±0.01	**0.97±0.01**	0.93±0.01	0.87±0.01	0.93±0.01	0.87±0.01	0.93±0.01	0.72±0.01	0.96±0.01

PCA, and input images are first four components of PCA; the filter radius r ranges from 1 to 9. To be specific, the guidance image of HGFM is the first component of MNF, and the radius r is 1–2. In the combination operation of OBR and CBR, the SE scales are $[3 \times 3, 7 \times 7, 11 \times 11]$, $[7 \times 7, 9 \times 9, 11 \times 11]$, and $[13 \times 13, 17 \times 17, 21 \times 21]$, respectively. The corresponding parameters of EAPs are assigned values by default according to [18]. The GLCM feature set originates from the first four principal components of image, including mean, variance, contrast, homogeneity, entropy, dissimilarity, second moment, and correlation. Based on the first four principal components of the image, the EMP feature set is constructed by using circular structure elements with a step size increment of 2, and four opening and closing are operated for each principal component, respectively. We adopt the method and default parameters provided by Kang *et al.* [69] to produce an IFRF feature set.

At the classification stage, the training sample sets are randomly selected from the ground truth. In this research, the

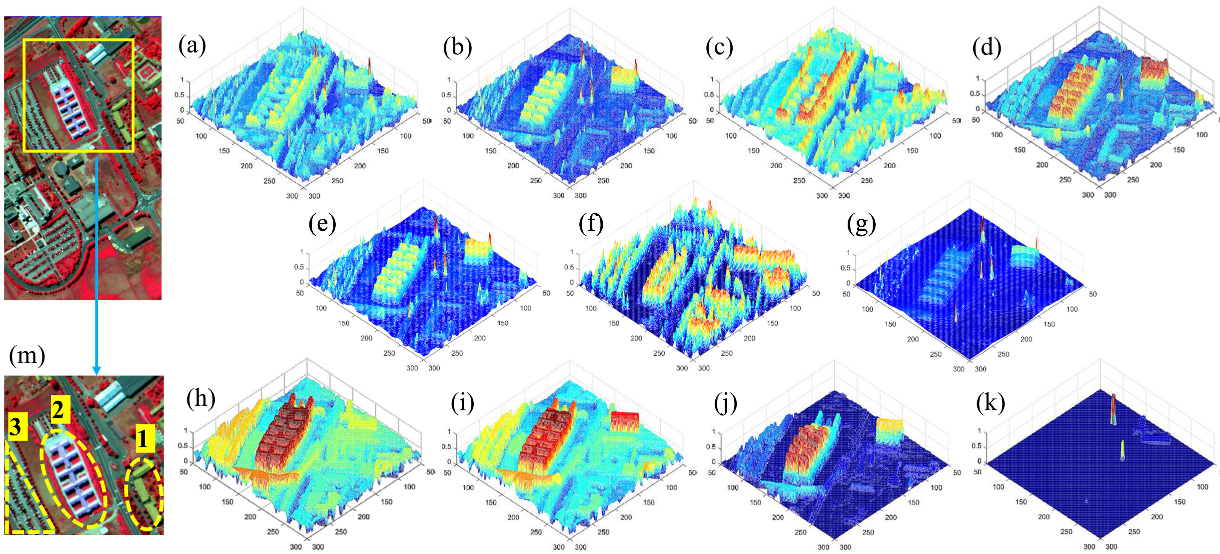


Fig. 5. Visual comparison of feature set attributes on (a) RAW, (b) HA, (c) GF, (d) HGFM, (e) EMP, (f) GLCM, (g) IFRF, (h) EAP_a, (i) EAP_d, (j) EAP_i, and (k) EAP_s from (m) local original image.

number of decision trees in RF and RoF is set to 10. The number of features in the subset of the RF adopts the default value (the maximum integer not greater than the square root of the number of features used). PSO-SVM is carried out with the support of the LIBSVM [70]. The KNN uses Euclidean distance, and the k value is determined by the minimum error rate through the training data. In ELM with the activation function of Sigmoid, the number of hidden neurons is assigned to 256. All the experiments are implemented in MATLAB R2017a on Intel Core i7-6700 Desktop PC with 3.4-GHz CPU and 32 GB of RAM.

C. Visual Comparison on Feature Sets

Fig. 5 illustrates the attributes of different feature sets (only a part of the feature set is shown to emphasize respective salient characteristics). To compare the effect of feature extraction, three regions are marked in the original image (see Fig. 5). The proposed HGFM feature set differs from RAW, HA, GF, EMP, and GLCM: the target is raised and the noise is effectively smoothed simultaneously, and the edge information of object is well preserved. Some small objects are ignored in the IFRF feature set. EAP_s fails to extract image features effectively, and other feature sets of EAPs have distinct sensitivity to different ground object types. Therefore, EAPs require multifeature stacking for optimal performance. It can be seen that HGFM manifests its excellent discriminative performance in fewer dimensions.

D. Experimental Results

1) Experiment on University of Pavia ROSIS Image: Table II illustrates the classification results of different feature sets based on EL. The optimal OA, AA, and κ of feature sets are marked in bold, and the classification maps of University of Pavia ROSIS are shown in Fig. 6(a). To test the performance of the feature sets, we randomly select 50 samples per class of ground truth to train classifiers. From the results of classification mapping, the feature sets of HGFM and IFRF have less noise estimations and smooth boundary. However, due to excessive smoothing, the mapping effect of the IFRF feature set is relatively poor, and the boundary of some ground objects is not accurate. In general, OA, AA, and κ obtained by the proposed HGFM feature set are evidently competitive with other feature sets based on EL from Table II. Furthermore, based on the analysis of optimal accuracy based on individual classifiers, it can be found that the best OA of the RAW is 79.98%, the classification accuracy of the HA feature set is not ideal through PSO-SVM classifier, and optimum OA is merely 72.58%. The classification accuracies of the GF feature set are substantially improved; the optimum OA obtained by EL is 87.06%. Among the four diverse EAPs feature sets, the maximal OA obtained by EAP_i with the PSO-SVM classifier is up to 97.12%. In particular, the proposed HGFM feature set is significant in enhancing classification performance, and OA and AA are up to 96.36% and 96.79% respectively; κ reached 0.95. Despite that the optimal accuracies of HGFM are slightly lower than EAP_i, the acquired classification results basically reached the ideal situation in the condition of small-size training samples and fewer feature dimension. Moreover, in classification maps, the HGFM feature set can improve the classification effect of Asphalt, Meadows, and Bare Soil relied on the proposed classification framework. Hence, the HGFM feature set has better classification performance under limited training samples because of adequately exploiting the spectral and spatial information.

To intuitively express and compare the classification accuracies of different feature sets assisted by the individual classifiers and EL, the radar graphs of optimum classification accuracy (OA, AA, and κ) are illustrated in Fig. 7. It can be concluded that the EL plays a key role, which makes HGFM, GF, EAP_a, and EAP_d always exert outstanding classification capability. In addition, PSO-SVM and RoF perform great performance.

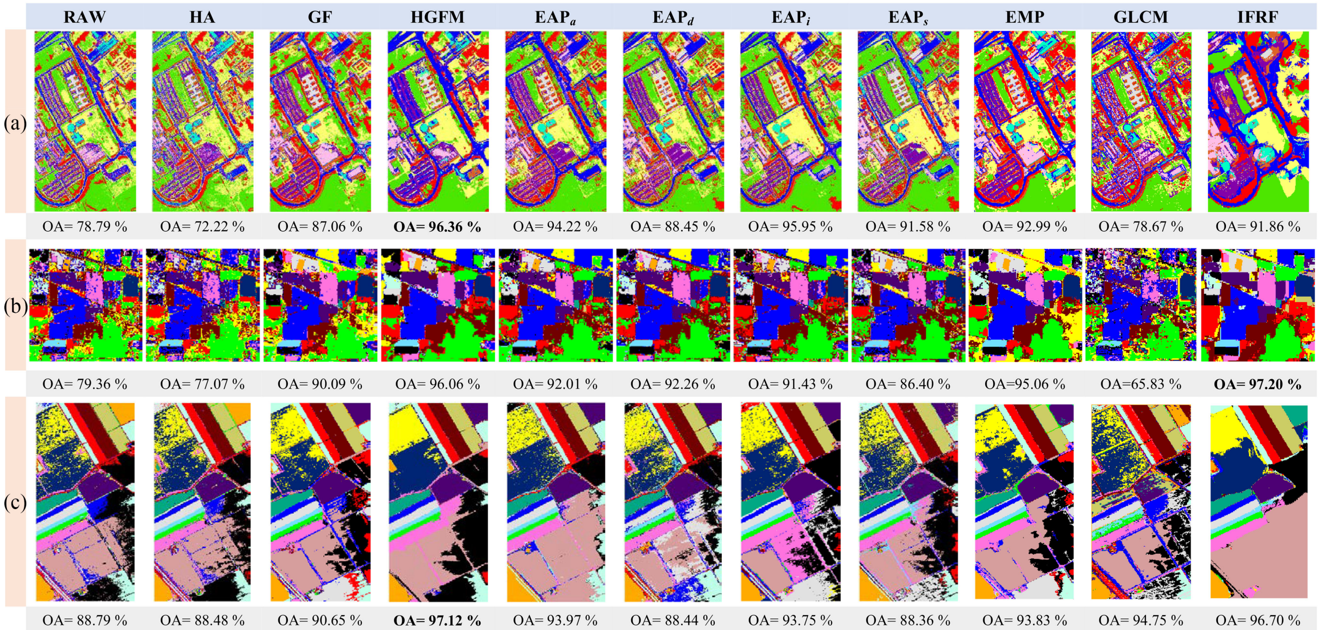


Fig. 6. Classification maps and OA (%) of (a) The University of Pavia ROSIS, (b) Indian Pines AVIRIS, and (c) Salinas AVIRIS based on EL using different feature sets.

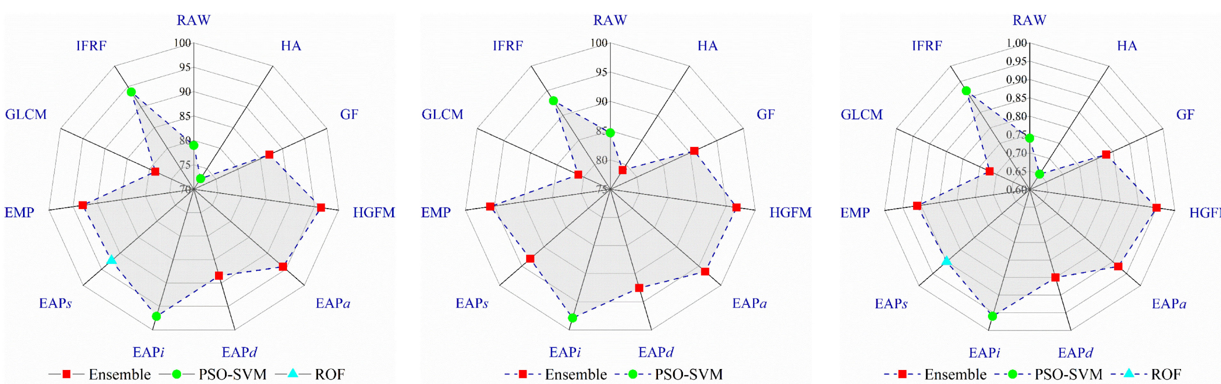


Fig. 7. Radar graphs of optimal OA, AA, and κ obtained by different feature sets based on the proposed classification framework of University of Pavia ROSIS image.

2) Experiment on Indian Pines AVIRIS Image: The existence of low spatial resolution and the mixed pixels make the classification task of this scene challengeable. For this reason, 10% training samples per class are selected from the ground truth to train the classifiers. All classification accuracies based on the proposed classification framework are shown in Table III. Classification maps are presented in Fig. 6(b). Based on the analysis of Table III, HGFM achieves the wonderful classification performance with respect to OA, AA, and κ. Considering the individual classifiers, the classification accuracies of the HGFM feature set are remarkable, in which the optimal values of OA and AA obtained by PSO-SVM are all greater than 96%. Compared with RAW and HA feature sets, the optimum classification accuracy of HGFM is increased by about 15% and 5%, respectively. By comparing HGFM with EAP_s, it can be seen that the best OA (92.58%) obtained by EAP_d is lower than that (96.08%) of HGFM. In terms of classification results, the classification performance of the ELM classifier is unstable. For instance, when ELM is used for the classification of the RAW feature set, the individual class accuracy of Alfalfa and Oats is 0. However, it is gratifying that HGFM has fine classification accuracy based on individual classifiers (e.g., ELM), which fully illustrates the adaptability of the HGFM feature set. When there are only two training samples of Oats, the HGFM feature set achieves the best classification accuracy of 88.33%, which is 7% higher than EMP. Although the OA of the IFRF feature set is slightly higher than that of HGFM in the classification of Indian Pines AVIRIS, the AA and κ of the IFRF feature set are relatively poor, and the estimations of the edge area is imprecise. In general, the HGFM feature set performs the

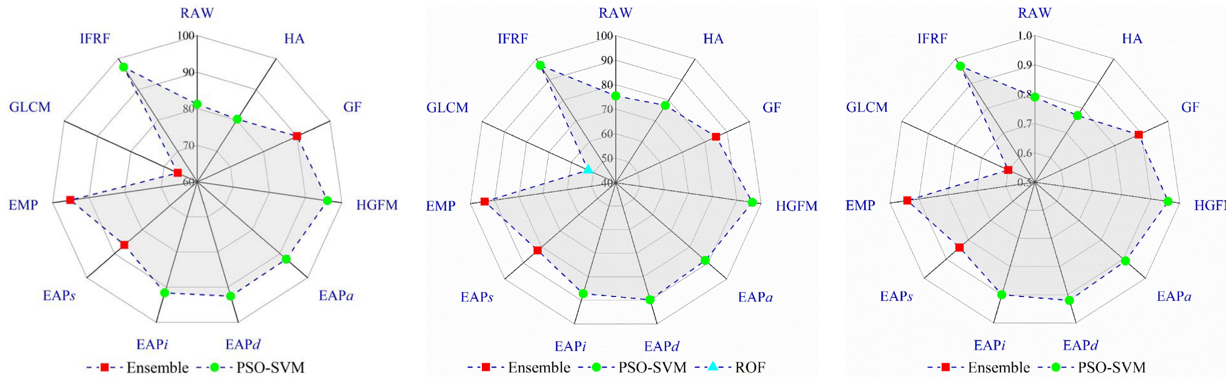


Fig. 8. Radar graphs of optimal OA, AA, and κ obtained by different feature sets based on the proposed classification framework of University of Indian Pines AVIRIS image.

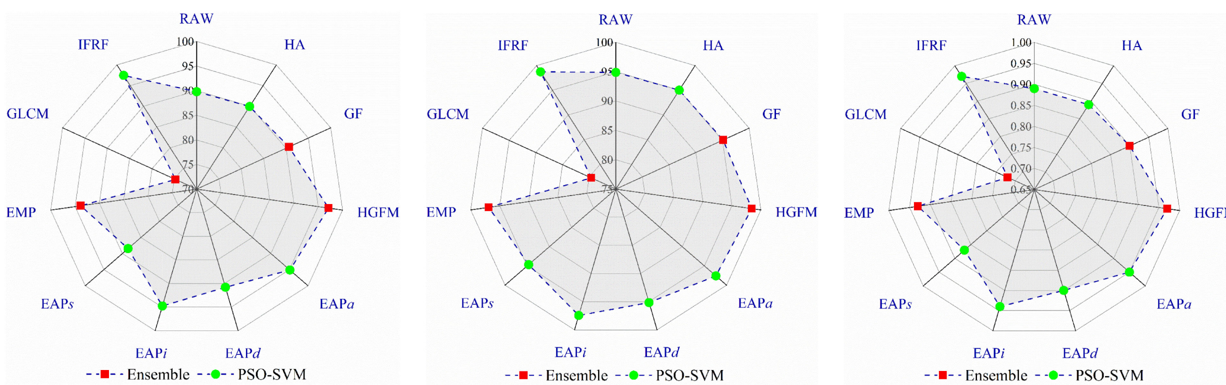


Fig. 9. Radar graphs of optimal OA, AA, and κ obtained by different feature sets based on the proposed classification framework of the University of Salinas AVIRIS image.

better classification performance in terms of visual quality and objective metrics (AA and κ) compared with the other feature sets.

Fig. 8 shows radar graphs of the supreme accuracy obtained by different feature sets with the proposed classification framework. According to radar graphs, the combination of the HGFM feature set and PSO-SVM achieves the optimal OA, AA, and κ. What is more, the classification effect of GF and EAP feature sets relied on EL is noteworthy.

3) Experiment on Salinas AVIRIS Image: Table IV illustrates the average classification results over ten independent Monte Carlo runs using different feature sets based on EL. Classification maps are shown in Fig. 6(c). Concretely, solely 50 training samples per class are randomly selected for the training of classifiers. In light of the classification results of Table IV, the OA (97.12%), AA (98.42%), and κ (0.97) of the HGFM feature set based on EL are optimal. Compared with the best classification accuracy of RAW, HA, and GF feature sets, OA of HGFM is improved by more than six percentage points. The first-best OA obtained by EAP_a relied on PSO-SVM in EAPs is approximately two percentage points lower than that of HGFM. Simultaneously, it has been found that the joint of HGFM and EL significantly improved classification accuracies of grapes_untrained, corn_senesced_green_weeds, and vinyard_untrained.

Obviously, the classification framework of HGFM combined with EL performs outstanding classification performance, and the individual classifier PSO-SVM integrated with feature sets demonstrates fine capability in the classification task, as shown in Fig. 9.

E. Effects on Parameter Selection

1) Effect of HA Dimensionality: In the construction of the HGFM feature set, HA is a critical part of data dimensionality reduction and frequency-domain information extraction. Therefore, it is necessary to evaluate the impact of the h-index of HA on classification accuracy. In the process of investigation, the sample selection and parameter settings are consistent with the experimental settings in Section III-B. The relevant results are depicted in Fig. 10, and the following three main characteristics are found.

1) The OA is gradually improved along with the increase of HA dimensionality. The classification results of the ELM classifier are undesirable, and the supreme OA of Indian Pines and Pavia University is merely about 55%.

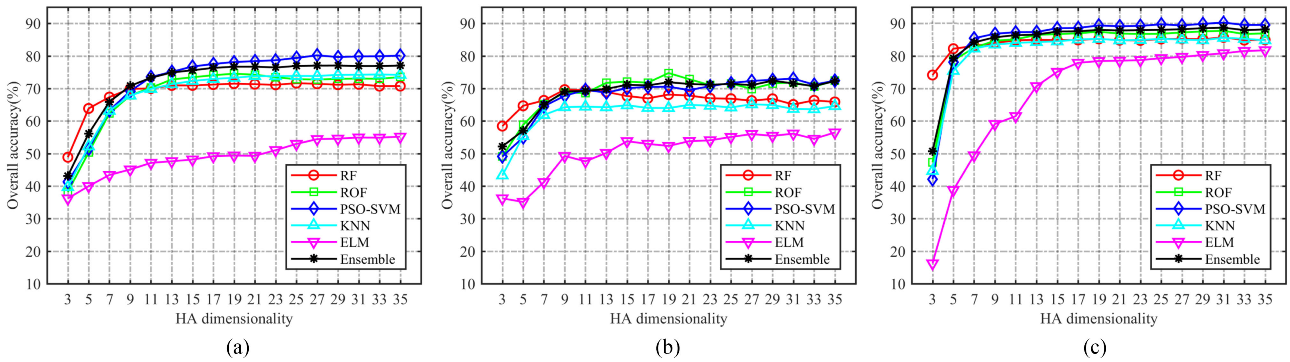


Fig. 10. OA (%) obtained by HA feature set with different dimensionality based on classifiers (RF, RoF, PSO-SVM, KNN, ELM, and Ensemble) in (a) Indian Pines AVIRIS, (b) The University of Pavia ROSIS, and (c) Salinas AVIRIS.

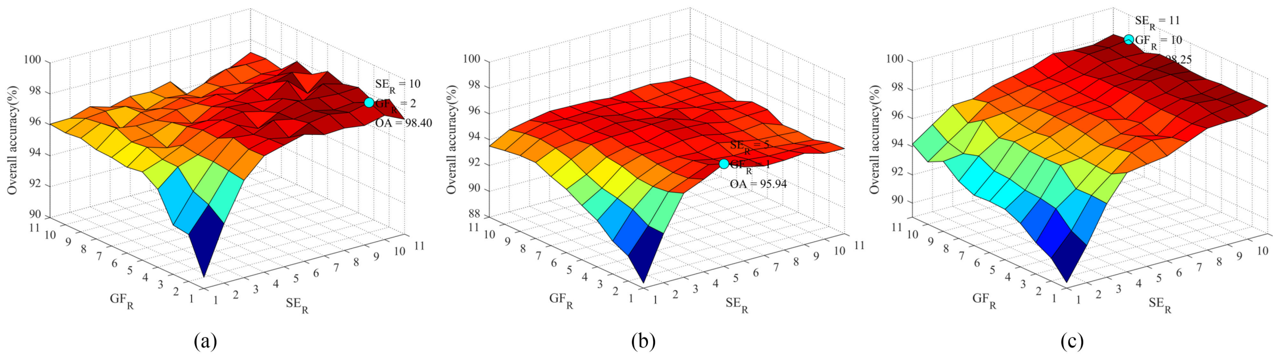


Fig. 11. OA of the HGFM feature set based on a combination of different SE and GF window radii using EL in (a) The University of Pavia ROSIS, (b) Indian Pines AVIRIS, and (c) Salinas AVIRIS.

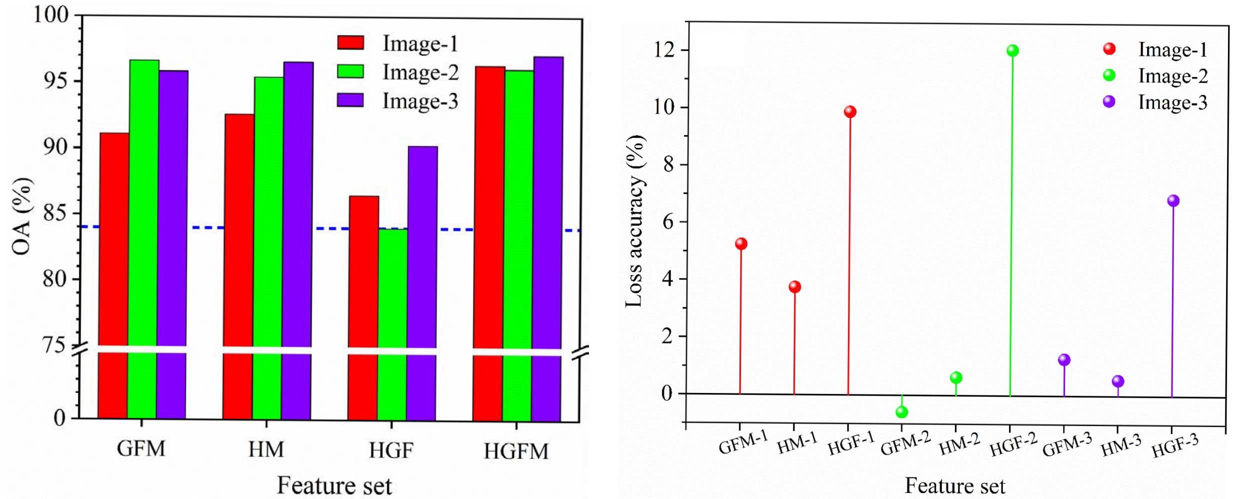
2) The dimensionality of HA is different when the optimal OA is achieved in different datasets, and the base classifiers of PSO-SVM and RoF are very helpful to improve the classification performance of the HA feature set. For instance, the optimum OA of Indian Pines AVIRIS is 80.18% when the dimensionality of HA is 27, while the optimal OA is only 74.74% based on RoF when the dimensions of HA are 19 in the University of Pavia ROSIS. At the state of 31 dimensions of HA, the first-best OA of Salinas AVIRIS is 90.23% based on PSO-SVM.
3) When the dimensionality of the HA feature set is greater than or equal to 17, the increasing trend of OA of three datasets based on individual classifiers or EL is not obvious. Therefore, it is reasonable to assign h-index of HA to 8 (dimensions of HA are 17) as performing HGFM feature set extraction.

2) Effect of Windows Radius Combination on HGFM: To evaluate the effect of window radius of GF (GFR) and SE (SER) attached to OBR and CBR in the HGFM feature set, the window radius is taken separately as [1,2,...,11] during the process, and related parameters are all in accordance with the experimental settings in Section III-B. As shown in Fig. 11, HSI classification results are produced on the basis of the proposed classification framework.

When $\{\mathrm{GF_R} = 2, \mathrm{SE_R} = 10\}$, the optimal OA of the University of Pavia ROSIS is 98.40%, and the obtained best OA of Salinas AVIRIS is 98.25% when $\{\mathrm{GF_R} = 10, \mathrm{SE_R} = 11\}$. Above all, the optimum OA of these two datasets is over 98%, and the classification effect is remarkable. Although for Indian Pines AVIRIS, when $\{\mathrm{GF_R} = 1, \mathrm{SE_R} = 5\}$, the OA reaches a maximum (95.94%) and the classification performance is slightly poor. Therefore, different window radius combinations have certain effect on classification accuracy, and appropriate adjustments need to be considered in classification tasks. Besides, a single window scale may lead to the missing of the acquired feature details, and the multiscale combination can better weaken the influence of such cases. Thus, the multiscale combination and the average strategy adopted in this article are acceptable to a certain extent.

F. Contribution of Different Ingredients on HGFM

To analyze the contribution of different ingredients of the HGFM feature set, including HGFM without HA (GFM), HGFM without GF (HM), and HGFM without OBR/CBR (HGF). Meanwhile, corresponding results of the HGFM feature set are listed for comparison. We analyzed OA of HGFM feature sets and defined the loss accuracy to quantitatively describe the



(a) (b)

Fig. 12. (a) OA (%) and (b) loss accuracy of GFM, HM, HGF, and HGFM on the HSI, including Image-1 (the University of Pavia ROSIS), Image-2 (Indian Pines AVIRIS), and Image-3 (Salinas AVIRIS).

TABLE V
RUNNING TIME (SECONDS) FOR FEATURE EXTRACTION AND CLASSIFICATION OF DIFFERENT FEATURE SETS ON THE UNIVERSITY OF PAVIA ROSIS, INDIAN PINES, AND SALINAS

Feature set	Step	University of Pavia ROSIS	Indian Pines	Salinas
HGFM	Feature extraction	87.44 (82.17)	13.65 (12.38)	73.66 (69.87)
	Classification	8.14	11.65	6.79
HA	Feature extraction	325.48	50.61	280.59
	Classification	11.03	8.61	6.9
GF	Feature extraction	1.7	0.24	0.61
	Classification	8.14	12.14	9.19
EAP_a	Feature extraction	2.53	0.3	1.33
	Classification	7.67	13.09	7.63
EAP_d	Feature extraction	2.85	0.33	1.51
	Classification	8.63	14.12	8.41
EAP_i	Feature extraction	14.21	2.55	4.44
	Classification	6.23	9.08	6.21
EAP_s	Feature extraction	15.57	2.14	4.4
	Classification	6.18	9.08	6.33
EMP	Feature extraction	1.43	0.44	0.74
	Classification	8.58	14.09	9.19
GLCM	Feature extraction	28.09	4.76	15.76
	Classification	12.29	20.37	54.96
IFRF	Feature extraction	2.27	0.31	1.13
	Classification	6.13	5.25	4.13

contribution of each part. Loss accuracy can be expressed as OA of HGFM feature set minus OA of GFM, HM, or HGF. For example, the contribution of HA is measured by analyzing the loss accuracy of GFM, that is, the OA of HGFM is subtracted from that of GFM. If the obtained value is positive, it means that removing HA from HGFM will reduce the classification accuracy of the HGFM feature set. The larger the value, and the greater the contribution of HA. If the value is negative, it means that removing HA will improve the classification accuracy of the HGFM feature set, and the contribution of HA is negligible.

As shown in Fig. 12, the OA of the HGFM feature set in three common hyperspectral scenes is over 96%, and the classification effect is better than others. The classification accuracy (OA) of HGF is the lowest, which indicates that OBR/CBR is an important part in the construction of HGFM feature set. If OBR/CBR is removed, the classification performance of the HGFM feature set will decline. In general, the order of contribution is OBR/CBR $>$ HA $>$ GF during the construction of the HGFM feature set.

G. Complexity Analysis

Table V reports the running time of different feature sets on the three commonly used HSIs. All the experiments are implemented in MATLAB R2017a on Intel Core i7-6700 Desktop PC with 3.4-GHz CPU and 32 GB of RAM. It should be noted that the computational time is mainly composed of two parts: feature extraction time and classification time. From Table V, the classification time of the HGFM feature set is acceptable compared with GF, EAP_d, EMP, and GLCM feature sets. However, it is not computationally efficient in comparison to other feature sets such as IFRF. It can be found that the main time consumption of HGFM is concentrated on the feature extraction.

The time in brackets represents the cost of building HA, which is 82.17, 12.38, and 69.87 s for University of Pavia ROSIS, Indian Pines, and Salinas, respectively. The reason is that the HA requires multiple iterations to decompose components. One of our ongoing efforts is to design graphics processing unit to speed up this process.

IV. Conclusion

To improve the discrimination capacity from feature space, the HGFM feature set is established and used for ensemble classification of HSIs. The main advantages of the HGFM feature set lie in optimizing the feature representation and enhancing the classification performance simultaneously.

The proposed methodology is investigated by three publicly available hyperspectral scenes: The University of Pavia ROSIS, Indian Pines AVIRIS, and Salinas AVIRIS images. The classification results of HGFM are compared with some other benchmark feature sets. It reveals that the proposed HGFM can improve classification accuracies with small-size training samples, and the optimal OA obtained by the new feature set is all higher than 96%. In addition, the proposed classification framework shows a certain generalization ability and fine classification performance of HSIs, which is helpful to provide a powerful alternative approach. Furthermore, the sensitivity of the parameters in the HGFM feature set is also investigated, and the base classifiers PSO-SVM and RoF exhibit a good classification effect.

In future research, we would like to extend the proposed classification framework to the classification of multisource data (e.g., Lidar, SAR, etc.) and the exploration of HA with other spatial algorithms. Moreover, the connection to deep learning can be explored and bring about more novel way to exploit and choose the features.

References

[1] 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.

[2] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, 2009.

[3] Q. X. Tong, Y. Q. Xue, and L. F. 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.

[4] L. Bruzzone, M. M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3363–3373, Nov. 2006.

[5] P. J. Du, Z. H. Xue, J. Li, and A. Plaza, "Learning discriminative sparse representations for hyperspectral image classification," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1089–1104, Sep. 2015.

[6] S. Mahesh, D. S. Jayas, J. Paliwal, and N. D. G. White, "Hyperspectral imaging to classify and monitor quality of agricultural materials," *J. Stored Products Res.*, vol. 61, pp. 17–26, 2015.

[7] L. R. Gao, D. Yao, Q. T. Li, L. N. Zhuang, B. Zhang, and J. M. Bioucas-Dias, "A new low-rank representation based hyperspectral image denoising method for mineral mapping," *Remote Sens.*, vol. 9, no. 11, 2017, Art. no. 1145.

[8] S. Bernabe, C. Garcia, F. D. Igual, G. Botella, M. Prieto-Matias, and A. Plaza, "Portability study of an openCL algorithm for automatic target detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9499–9511, Nov. 2019.

[9] N. Kayet *et al.*, "Forest health assessment for geo-environmental planning and management in hilltop mining areas using Hyperion and Landsat data," *Ecol. Indicators*, vol. 106, 2019, Art. no. 105471.

[10] P. Du *et al.*, "Advances of four machine learning methods for spatial data handling: A review," *J. Geovisualization Spatial Anal.*, vol. 4, no. 1, 2020, Art. no. 13.

[11] G. F. Hughes, "On mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. 14, no. 1, pp. 55–63, Jan. 1968.

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

[13] 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.

[14] X. Guo, X. Huang, L. F. Zhang, L. P. Zhang, A. Plaza, and J. A. Benediktsson, "Support tensor machines for classification of hyperspectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3248–3264, Jun. 2016.

[15] M. Imani and H. Ghassemian, "Edge-preserving-based collaborative representation for spectral-spatial classification," *Int. J. Remote Sens.*, vol. 38, no. 20, pp. 5524–5545, 2017.

[16] D. J. Marceau, P. J. Howarth, J. M. M. Dubois, and D. J. Gratton, "Evaluation of the grey-level co-occurrence matrix method for land-cover classification using spot imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 4, pp. 513–519, Jul. 1990.

[17] J. A. Benediktsson, M. Pesaresi, and K. Arnason, "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.

[18] M. D. Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 542–546, May 2011.

[19] S. Jia, L. L. Shen, and Q. 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.

[20] M. D. Farrell and R. M. Mersereau, "On the impact of PCA dimension reduction for hyperspectral detection of difficult targets," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 192–195, Apr. 2005.

[21] L. Ballanti, L. Blesius, E. Hines, and B. Kruse, "Tree species classification using hyperspectral imagery: A comparison of two classifiers," *Remote Sens.*, vol. 8, no. 6, 2016, Art. no. 445.

[22] Z. Li *et al.*, "Feature extraction method based on spectral dimensional edge preservation filtering for hyperspectral image classification," *Int. J. Remote Sens.*, vol. 41, no. 1, pp. 90–113, 2019.

[23] P. Ghamisi *et al.*, "New frontiers in spectral-spatial hyperspectral image classification the latest advances based on mathematical morphology, Markov random fields, segmentation, sparse representation, and deep learning," *IEEE Geosci. Remote Sens. Mag.*, vol. 6, no. 3, pp. 10–43, Sep. 2018.

[24] X. J. Zhang, C. Xu, M. Li, and X. L. Sun, "Sparse and low-rank coupling image segmentation model via nonconvex regularization," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 29, no. 2, 2015, Art. no. 1555004.

[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] 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, Cybern.*, vol. 40, no. 5, pp. 1267–1279, Oct. 2010.

[27] J. K. Chen, J. S. Xia, P. J. Du, and J. Chanussot, "Combining rotation forest and multiscale segmentation for the classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4060–4072, Sep. 2016.

[28] A. Romero, C. Gatta, and G. Camps-Valls, "Unsupervised deep feature extraction for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1349–1362, Mar. 2016.

[29] W. Zhao 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, Aug. 2016.

[30] J. Li, X. Zhao, Y. Li, Q. Du, B. Xi, and J. Hu, "Classification of hyperspectral imagery using a new fully convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 292–296, Feb. 2018.

[31] Z. H. Xue, P. J. Du, and H. J. Su, "Harmonic analysis for hyperspectral image classification integrated with PSO optimized SVM," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2131–2146, Jun. 2014.

[32] K. Yang, F. Liu, Y. Sun, H. Wei, and G. Shi, "Classification algorithm of hyperspectral imagery by harmonic analysis and spectral angle mapping," *J. Image Graph.*, vol. 20, no. 6, pp. 836–844, 2015.

[33] S. Schlaffer, P. Matgen, M. Hollaus, and W. Wagner, "Flood detection from multi-temporal SAR data using harmonic analysis and change detection," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 38, pp. 15–24, 2015.

[34] G. Yang, H. Shen, L. Zhang, Z. He, and X. Li, "A moving weighted harmonic analysis method for reconstructing high-quality spot vegetation NDVI time-series data," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 6008–6021, Nov. 2015.

[35] J. Zhou, L. Jia, and M. Menenti, "Reconstruction of global MODIS NDVI time series: Performance of harmonic analysis of time series (HANTS)," *Remote Sens. Environ.*, vol. 163, pp. 217–228, 2015.

[36] Z. Lu and J. He, "Spectral-spatial hyperspectral image classification with adaptive mean filter and jump regression detection," *Electron. Lett.*, vol. 51, no. 21, pp. 1658–1660, 2015.

[37] R. D. Phillips, C. E. Blinn, L. T. Watson, and R. H. Wynne, "An adaptive noise-filtering algorithm for AVIRIS data with implications for classification accuracy," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 9, pp. 3168–3179, Sep. 2009.

[38] F. Mirzapour and H. Ghassemian, "Multiscale gaussian derivative functions for hyperspectral image feature extraction," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 4, pp. 525–529, Apr. 2016.

[39] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 839–846.

[40] X. D. Kang, S. T. 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.

[41] J. Barros and R. I. Diego, "Analysis of harmonics in power systems using the wavelet-packet transform," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 1, pp. 63–69, Jan. 2008.

[42] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–409, Jun. 2013.

[43] N. H. Kaplan and I. Erer, "Remote sensing image enhancement via robust guided filtering," in *Proc. 9th Int. Conf. Recent Adv. Space Technol.*, 2019, pp. 447–450.

[44] Y. Xu, L. Wu, Z. Xie, and Z. Chen, "Building extraction in very high resolution remote sensing imagery using deep learning and guided filters," *Remote Sens.*, vol. 10, no. 1, 2018, Art. no. 144.

[45] W. Xie, T. Jiang, Y. Li, X. Jia, and J. Lei, "Structure tensor and guided filtering-based algorithm for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4218–4230, Jul. 2019.

[46] J. M. Hahne *et al.*, "Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 2, pp. 269–279, Mar. 2014.

[47] K. M. He, J. A. Sun, and X. O. Tang, *Guided Image Filtering* (ser. Lecture Notes in Computer Science), vol. 6311. Berlin, Germany: Springer-Verlag, 2010.

[48] E. L. Bullock, C. E. Woodcock, and C. E. Holden, "Improved change monitoring using an ensemble of time series algorithms," *Remote Sens. Environ.*, vol. 238, 2020, Art. no. 111165.

[49] P. Ghamisi, M. Dalla Mura, and J. A. Benediktsson, "A survey on spectral-spatial classification techniques based on attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2335–2353, May 2015.

[50] X. Huang, L. Zhang, and P. Li, "An adaptive multiscale information fusion approach for feature extraction and classification of IKONOS multispectral imagery over urban areas," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 654–658, Oct. 2007.

[51] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, Jul. 2013.

[52] V. Mesev, "Morphological image analysis: Principles and applications," *Environ. Planning B-Planning Des.*, vol. 28, no. 5, pp. 800–801, 2001.

[53] D. Tuia, F. Pacifici, M. Kanevski, and W. J. Emery, "Classification of very high spatial resolution imagery using mathematical morphology and support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3866–3879, Nov. 2009.

[54] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3747–3762, Oct. 2010.

[55] B. Krawczyk, L. L. Minku, J. Gama, J. Stefanowski, and M. Wozniak, "Ensemble learning for data stream analysis: A survey," *Inf. Fusion*, vol. 37, pp. 132–156, 2017.

[56] D. H. Wolpert, *The Supervised Learning No-Free-Lunch Theorems* (ser. Soft Computing and Industry: Recent Applications). London, U.K.: Springer-Verlag, 2002.

[57] P. Du, J. Xia, W. Zhang, K. Tan, Y. Liu, and S. Liu, "Multiple classifier system for remote sensing image classification: A review," *Sensors (Basel)*, vol. 12, no. 4, pp. 4764–92, 2012.

[58] Y. Ren, L. Zhang, and P. N. Suganthan, "Ensemble classification and regression-recent developments, applications and future directions," *IEEE Comput. Intell. Mag.*, vol. 11, no. 1, pp. 41–53, Feb. 2016.

[59] A. Samat *et al.*, "Direct, ECOC, ND and END frameworks—which one is the best? An empirical study of Sentinel-2A MSIL1C image classification for arid-land vegetation mapping in the Ili river delta, Kazakhstan," *Remote Sens.*, vol. 11, no. 16, 2019, Art. no. 1953.

[60] M. Belgiu and L. Dragut, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogrammetry Remote Sens.*, vol. 114, pp. 24–31, 2016.

[61] J. Xia, P. Ghamisi, N. Yokoya, and A. Iwasaki, "Random forest ensembles and extended multiextinction profiles for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 202–216, Jan. 2018.

[62] J. S. Xia, P. J. Du, X. Y. He, and J. Chanussot, "Hyperspectral remote sensing image classification based on rotation forest," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 239–243, Jan. 2014.

[63] J. S. Xia, J. Chanussot, P. J. Du, and X. Y. He, "Spectral-spatial classification for hyperspectral data using rotation forests with local feature extraction and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2532–2546, May 2015.

[64] A. Subasi, "Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders," *Comput. Biol. Med.*, vol. 43, no. 5, pp. 576–586, 2013.

[65] Z. Hanifelou, P. Adibi, S. A. Monadjemi, and H. Karshenas, "KNN-based multi-label twin support vector machine with priority of labels," *Neurocomputing*, vol. 322, pp. 177–186, 2018.

[66] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2004, vol. 2, pp. 985–990.

[67] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, May 1979.

[68] 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.

[69] 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.

[70] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. no. 27.

Wei Zhang received the M.S. degree in geodesy and survey engineering from the China University of Mining and Technology, Beijing, China, in 2019. He is currently working toward the Ph.D. degree in cartography and geographic information system with the School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing, China.

His research interests include image classification, machine learning, and heavy metal pollution monitoring with hyperspectral remote sensing.

Peijun Du (Senior Member, IEEE) received the Ph.D. degree in geodesy and survey engineering from the China University of Mining and Technology, Xuzhou, China, in 2001.

He is currently a Professor of Remote Sensing and Geographical Information Science with Nanjing University, Nanjing, China. He has authored or coauthored more than 70 articles in international peer-reviewed journals, and more than 100 papers in international conferences and Chinese journals. His research interests include remote sensing image processing and pattern recognition, hyperspectral remote sensing, and applications of geospatial information technologies.

Dr. Du is an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS. He also served as the Co-Chair of the Technical Committee of the 2009 URBAN Conference, the 2012 IAPR Workshop on Pattern Recognition in Remote Sensing, the 2014 International Workshop on Earth Observation and Remote Sensing Applications (EORSA), 2015 International Conference on Intelligent Earth Observing and Applications, and the 2015 Conference on Global Climate Change, the Co-Chair of the Local Organizing Committee of the 2009 Proceedings of Joint Urban Remote Sensing Event, the 2012 Workshop on Hyperspectral Image and Signal Processing (WHISPERS), and EORSA 2012, and a member of Scientific Committee or Technical Committee of other international conferences, including WHISPERS (2010–2016), the URBAN Conference (2011, 2013, and 2015), the International Workshop on the Analysis of Multitemporal Remote Sensing Images (2011, 2013, and 2015), the 2011 International Symposium on Image and Data Fusion, and the SPIE European Conference on Image and Signal Processing for Remote Sensing (2012–2016).

Cong Lin received the B.S. degree in surveying and mapping engineering from the School of Environment Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, China, in 2015. He is currently working toward the Ph.D. degree with the Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing, China.

His research interests include image processing, transfer learning, and scene classification.

Pingjie Fu received the B.S. degree in geological engineering and the M.S. degree in mineralogy, petrology, and mineral geology from the Shandong University of Science and Technology, Qingdao, China, in 2012 and 2015, respectively, and the Ph.D. degree in geodesy and survey engineering from the China University of Mining and Technology, Beijing, China, in 2019.

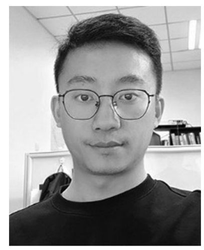
She is currently a Lecturer with the College of Surveying and Geo-Informatics, Shandong Jianzhu University, Jinan, China. Her research interests include environment remote sensing, signal processing, and soil heavy metal pollution monitoring with hyperspectral remote sensing.

Xin Wang received the B.S. degree in geographic information system from Lanzhou University, Lanzhou, China, in 2016. He is currently working toward the Ph.D. degree in cartography and geographic information systems with Nanjing University, Nanjing, China.

He is currently a Visiting Student with the Department of Geography and Planning, Queen's University, Kingston, ON, Canada. His research interests include multitemporal image processing, change detection, signal processing, and pattern recognition.

Xuyu Bai received the B.S. degree in resource environment and urban planning management from Beijing Normal University, Zhuhai, China, in 2013, and the M.A. degree in environmental management and consultancy from Lancaster University, Lancaster, U.K., in 2015. He is working toward the Ph.D. degree in environment and resource remote sensing with Nanjing University, Nanjing, China.

He is a Visiting Researcher with Lancaster University. His research interests include environment assessment, environment remote sensing, and urban remote sensing.



Hongrui Zheng received the M.S. degree in cartography and geographical information systems from Capital Normal University, Beijing, China, in 2014. He is currently working toward the Ph.D. degree in geographic information sciences with Nanjing University, Nanjing, China.

His research interests include time-series remote sensing imagery processing and applications.

Junshi Xia (Senior Member, IEEE) received the B.S. degree in geographic information systems and the Ph.D. degree in photogrammetry and remote sensing from the China University of Mining and Technology, Xuzhou, China, in 2008 and 2013, respectively, and the Ph.D. degree in image processing from the Grenoble Images Speech Signals and Automatics Laboratory, Grenoble Institute of Technology, Grenoble, France, in 2014.

From 2014 to 2015, he was a Visiting Scientist with the Department of Geographic Information Sciences, Nanjing University, Nanjing, China. From 2015 to 2016, he was a Postdoctoral Research Fellow with the University of Bordeaux, Bordeaux, France. Since 2016, he has been the Japan Society for the Promotion of Science Postdoctoral Overseas Research Fellow with the University of Tokyo, Tokyo, Japan. His research interests include multiple classifier systems in remote sensing, hyperspectral remote sensing image processing, and urban remote sensing.

Dr. Xia received the first place prize in the IEEE Geoscience and Remote Sensing Society Data Fusion Contest organized by the Image Analysis and Data Fusion Technical Committee, in 2017.

Alim Samat (Member, IEEE) received the B.S. degree in geographic information systems from Nanjing University, Nanjing, China, in 2009, the M.S. degree in photogrammetry and remote sensing from the China University of Mining and Technology, Xuzhou, China, in 2012, and the Ph.D. degree in cartography and geographic information systems from Nanjing University, in 2015.

He is currently an Associate Researcher with the State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Ürümqi, China. His current research interests include PolSAR and optical remote sensing for land applications, image processing and pattern recognition, and machine learning.

Dr. Samat is a Reviewer for several international journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and the *Pattern Recognition*.