# Intrinsic Image Decomposition for Feature Extraction of Hyperspectral Images

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

***Abstract*—In this paper, a novel feature extraction method based on intrinsic image decomposition (IID) is proposed for hyperspectral image classification. The proposed method consists of the following steps. First, the spectral dimension of the hyperspectral image is reduced with averaging-based image fusion. Then, the dimension reduced image is partitioned into several subsets of adjacent bands. Next, the reflectance and shading components of each subset are estimated with an optimization-based IID technique. Finally, pixel-wise classification is performed only on the reflectance components, which reflect the material-dependent properties of different objects. Experimental results show that, with the proposed feature extraction method, the support vector machine classifier is able to obtain much higher classification accuracy even when the number of training samples is quite small. This demonstrates that IID is indeed an effective way for feature extraction of hyperspectral images.**

***Index Terms*—Feature extraction, hyperspectral image, image fusion, intrinsic image decomposition (IID), support vector machines (SVMs).**

## I. Introduction

HYPERSPECTRAL images are images of high spectral dimensionality. Every pixel in a hyperspectral image contains hundreds of spectral channels that correspond to the detailed spectrum of reflected light. The rich spectral information of hyperspectral images can be used for accurate classification of different materials and thus has been widely used in many practical applications such as monitoring of the environment [1] and precision agriculture [2].

For hyperspectral image classification, supervised classifiers such as Bayesian estimation method [3], decision tree [4], neural networks [5], [6], support vector machines (SVMs) [7], [8], sparse representation [9], [10], genetic algorithm [11], and kernel-based techniques [12]–[14] have demonstrated very good performances in terms of high classification accuracies. However, when very limited labeled information is available, most of these spectral classifiers are still not able to obtain satisfactory classification accuracies [15]. The reason is that although the high spectral resolution of hyperspectral images gives the possibility to detect and distinguish various objects with an improved accuracy, the large dimensionality of the data in the spectral domain also leads to difficulties in computer processing. Specifically, as the dimension of the data space becomes higher, the number of training samples needed to depict the statistical behavior of the data also increases exponentially, which indicates the "Hughes" phenomenon [16]. Moreover, a high spectral dimensionality also means that the computational burden of image classification will be quite large. To solve these problems, many feature extraction methods have been proposed to reduce the spectral dimension of the hyperspectral data while preserving the class separability of different objects [17]–[20]. Typical feature extraction methods such as principal component analysis (PCA) [21], independent component analysis (ICA) [22], and linear discriminant analysis [23], [24] have been widely used in commercial software.

In addition, the spectral characteristics, which are the primary discrimination features in hyperspectral images, spatial attributes have also been demonstrated to be very useful in increasing the classification accuracy. In recent years, feature extraction and classification methods that make full use of the spatial contextual information of hyperpsectral imagery have been studied extensively [25]–[33]. Specifically, spectral–spatial feature extraction and classification techniques aim at assigning each image pixel to one class by considering both its own spectral values and the spatial information extracted from its neighborhood. For instance, Benediktsson *et al.* defined an adaptive neighborhood for each pixel by multiple morphological operations such that the closest neighborhood information could be preserved in the morphological features [18], [34]–[36]. Another type of spectral–spatial classification method is based on the integration of pixel-wise classification and segmentation of hyperspectral images [37]–[39]. Aside from the two types of spectral–spatial classification methods aforementioned, other powerful tools that make full use of the spatial information of the hyperspectral image have also been proposed for hyperspectral image classification. These approaches include minimum spanning forest [40], conditional random fields [41], edge-preserving filters [19], [42], hypergraph [43], tensor representation [17], empirical mode decomposition [44], [45], anisotropic diffusion [46], [47], and joint collaborative representation [48].

Intrinsic image decomposition (IID) is a challenging problem in computer vision that aims at modeling the perceiving

Manuscript received June 26, 2014; revised August 8, 2014; accepted September 4, 2014. This paper was supported in part by the National Natural Science Foundation for Distinguished Young Scholars of China under Grant 61325007, by the National Natural Science Foundation of China under Grant 61172161, by the Fundamental Research Funds for the Central Universities, Hunan University, and by the Chinese Scholarship Award for Excellent Doctoral Student.

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

J. A. Benediktsson is with the Faculty of Electrical and Computer Engineering, University of Iceland, 101 Reykjavik, Iceland (e-mail: benedikt@hi.is).

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

Digital Object Identifier 10.1109/TGRS.2014.2358615

function of human vision to distinguish the reflectance and shading of the objects from a single image [49]. Since the intrinsic components of an image reflect different physical characteristics of the scene, e.g., reflectance, illumination, and shading, many classic computer vision and pattern recognition problems such as image segmentation [50] and object detection [51] can benefit from the IID technique.

Based on the idea of IID, the pixel values of hyperspectral images are determined by two factors: The spectral reflectance, which is determined by the material of different objects, and the shading component, which consists of light and shape-dependent properties. Since the second factor is not directly related to the material of the object, IID is adopted in this paper to remove useless spatial information preserved in the shading component of the hyperspectral image. Specifically, the IID-based feature extraction method proposed in this paper consists of the following steps: First, averaging-based image fusion is adopted to reduce the spectral dimension of the hyperspectral data. Second, an optimization-based IID method [52] is applied to decompose the dimension reduced hyperspectral data into intrinsic components, and only the reflectance components of the image are preserved for classification.

Experiments performed on different images captured by different hyperspectral sensors demonstrate the high accuracies of the proposed method. More importantly, the experimental results show that the proposed method can still obtain very high classification accuracies when the number of training samples is relatively small. Taking one of the most widely used hyperspectral remote sensing data as an example, i.e., the University of Pavia ROSIS data, the proposed method can lead to an overall classification accuracy of 97.92% when only 0.4% of the scene are used as labeled training samples.

Compared with the existing spectral–spatial classification methods, the main novelties and contributions of this paper lie in the following two aspects.

1) IID is applied for feature extraction of hyperspectral images for the first time. The proposed method shows the advantage of using the perceiving function of human vision, i.e., the ability to separate reflectance and shading components from a single image, for feature extraction of hyperpsectral images.
2) Instead of making full use of the spatial information, the proposed method aims at removing useless spatial information in the hyperspectral image such as shading and image noise. Compared with some existing feature extraction methods [18], [19], this paper shows that IID is a better solution in terms of higher classification accuracies.

The remainder of this paper is organized as follows. In Section II, IID and related works are reviewed. Section III descries the proposed IID-based feature extraction method. The experimental results are presented in Section IV, and finally, conclusions are given in Section V.

## II. IID

In the tenth century, one perceiving function of human vision is noted by the scientist Alhazen as follows: "*Nothing of what is visible, apart from light and color, can be perceived by pure sensation, but only by discernment, inference, and recognition, in addition to sensation*" [53]. In 1978, the "intrinsic images" problem is defined based on Alhazen's observations [54], and this problem has recently gathered a great interest in the computer vision field. Specifically, the "intrinsic images" problem can be represented by the following equation:

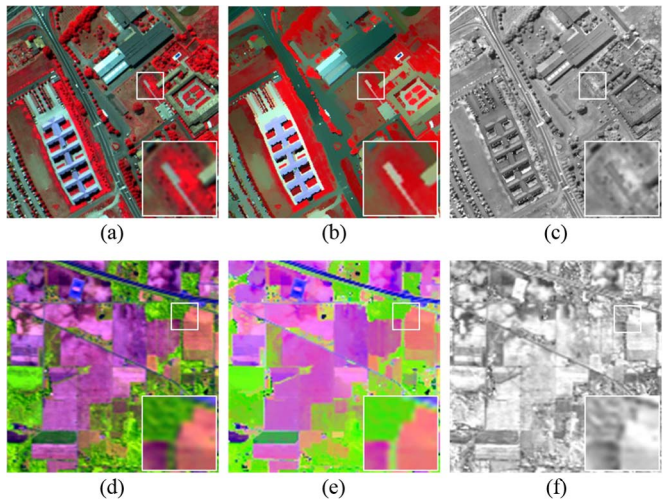



Fig. 1. Two intuitive examples of IID. From left to right: (a) and (d) two input color images, (b) and (e) reflectance components, and (c) and (f) shading components.

$$\mathbf{I}_i = \mathbf{R}_i \mathbf{S}_i \tag{1}$$

where $\mathbf{R}_i\mathbf{S}_i$ calculates the multiplication of the two metrics $\mathbf{R}_i$ and $\mathbf{S}_i$ element by element; $i$ refers to the $i$th pixel, $\mathbf{I}_i = (\mathbf{I}_{ir}, \mathbf{I}_{ig}, \mathbf{I}_{ib})$, $\mathbf{R}_i = (\mathbf{R}_{ir}, \mathbf{R}_{ig}, \mathbf{R}_{ib})$, $r$, $g$, and $b$ refer to the red, green, and blue channels of a color image, respectively. The aforementioned equation shows that a color image $\mathbf{I}$ [see Fig. 1(a) and (d)] can be decomposed into its spectral properties, known as reflectance $\mathbf{R}$ [see Fig. 1(b) and (e)], and its illumination and shape-dependent properties $\mathbf{S}$ [see Fig. 1(c) and (f)], such as shadows, specular highlights, and inter-reflectance.[1] Specifically, from Fig. 1(c) and (f), it can be shown that the pixel values in the shading component are usually determined by the shape and texture of the objects, and the illumination of the scene, which is not directly related to the materials of different objects. By contrast, as shown in the close-up view of Fig. 1(b), the discriminative colors of Meadows (red) and soil (hazel) actually represent the spectral signatures of different materials, which directly refer to different materials. A similar observation can be found for Fig. 1(e). For hyperspectral images, which have more than 100 data channels, it is easy to imagine that a much more informative spectral representation of different materials will be obtained and thus more classes are able to be discriminated with the reflectance $\mathbf{R}$. Motivated by this observation, this paper aims at extracting the spectral reflectance $\mathbf{R}$ of a hyperspectral image, in which the influence of the shading component $\mathbf{S}$, i.e., illumination of the scene and texture of the objects are removed. Based on

[1]The intrinsic components presented in this paper are produced using the method of Shen *et al.* [52].

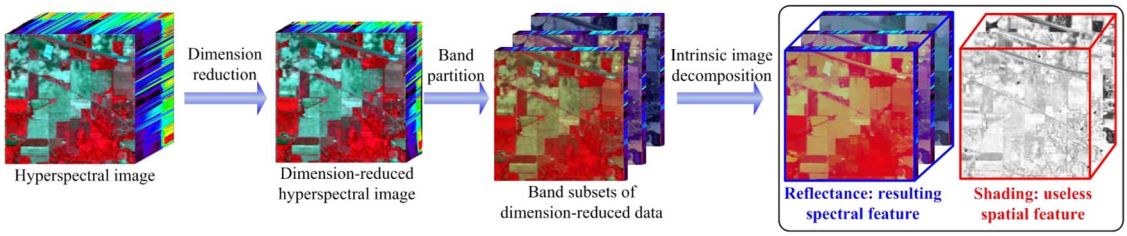



Fig. 2. Schematic of the proposed feature extraction method.

our knowledge, although IID has been successfully applied in computer vision [50], [55], our work is the first attempt to use the approach for feature extraction of hyperspectral images.

From (1), it can be seen that estimating $\mathbf{R}$ and $\mathbf{S}$ based on $\mathbf{I}$ is an unconstrained problem. Although many solutions have recently been proposed for this problem [52], [55], it is still considered as a challenge in computer vision research. Generally speaking, heuristic cues, extra constraints [49] or user intervention [52], [55] have to be exploited in order to estimate the intrinsic components. For example, one of the earliest image IID methods is based on the retinex algorithm, which assumes that large derivatives in the image are always attributed to reflectance changes, whereas smaller derivatives are usually due to shading changes in a scene [56]. In addition to the retinex-based methods, many other powerful tools have been proposed in recent years such as sparse representation [57] and the nonlocal method [58] to recover intrinsic components from a single image.

In this paper, an optimization-based IID method [52] is adopted for intrinsic decomposition of hyperspectral images because of the following reasons: First, it has been demonstrated in [52] that the optimization-based method is robust for images captured under different conditions, both indoor or outdoor. Second, no classification and complex mathematical models are required for the optimization-based method, and thus, the algorithm used in [52] can be easily extended for processing hyperspectral images that usually have more than three data channels. Here, the optimization-based IID method is briefly reviewed. Specifically, the method is based on the following weight function that models the reflectance value of one pixel as a weighted average of its adjacent reflectance values, i.e.,

$$\mathbf{R}_i = \sum_{j\in\omega_i} \alpha_{ij}\mathbf{R}_j,\ \alpha_{ij} = e^{-\left[(\mathbf{Y}_i-\mathbf{Y}_j)^2/\sigma^2_{i\mathbf{Y}}+\mathcal{A}(\mathbf{I}_i,\mathbf{I}_j)^2/\sigma^2_{i\mathbf{A}}\right]} \quad (2)$$

where $\omega_i$ is a local window of $(2r+1)\times(2r+1)$ pixels around pixel $i$, in which $r$ is the radius of the window; $\alpha_{ij}$ measures the similarity of the intensity value and the spectral angle between pixel $i$ and $j$; for images with multiple channels, $\mathbf{Y}$ represents the intensity image, which is calculated by averaging all the bands of the image; $\mathbf{R}_i = (\mathbf{R}_{ir}, \mathbf{R}_{ig}, \mathbf{R}_{ib})$, $\mathbf{R}_j = (\mathbf{R}_{jr}, \mathbf{R}_{jg}, \mathbf{R}_{jb})$, $\mathcal{A}(\mathbf{I}_i, \mathbf{I}_j) = \arccos(\mathbf{I}_{ir}\mathbf{I}_{jr} + \mathbf{I}_{ig}\mathbf{I}_{jg} + \mathbf{I}_{ib}\mathbf{I}_{jb})$ denotes the angle between the pixel vectors $\mathbf{I}_i$ and $\mathbf{I}_j$; $\sigma^2_{i\mathbf{A}}$ and $\sigma^2_{i\mathbf{Y}}$ are the variances of the angle and pixel values in the local window $\omega_i$, respectively. From the two terms of $\alpha_{ij}$, i.e., $(\mathbf{Y}_i - \mathbf{Y}_j)^2/\sigma^2_{i\mathbf{Y}}$ and $\mathcal{A}(\mathbf{I}_i, \mathbf{I}_j)^2/\sigma^2_{i\mathbf{A}}$, it can be seen that the method of Shen *et al.* [52] is based on a simple assumption: neighboring pixels in a local window having both similar intensity and chromaticity usually should have similar reflectance values. In other words, it assumes that the changes of shading values will lead to the proportional changes of its red, green, and blue color channel values. Under this assumption, the shading component could be separated from the input image. Based on (1) and (2), the resulting shading $\mathbf{S}$ and reflectance $\mathbf{R}$ components can be obtained by optimizing the following energy function:

$$E(\mathbf{R},\mathbf{S}) = \sum_{i\in\mathbf{I}} \left(\mathbf{R}_i - \sum_{j\in\omega_i}\alpha_{ij}\mathbf{R}_j\right)^2 + \sum_{i\in\mathbf{I}}(\mathbf{I}_i/\mathbf{S}_i - \mathbf{R}_i)^2. \quad (3)$$

For a complete description of the optimization method, we refer the reader to [52].

## III. Proposed Approach

Fig. 2 shows the schematic diagram of the proposed feature extraction method, which consists of the following steps: First, the dimension of the hyperspectral image is reduced with averaging-based image fusion. Then, the dimension reduced data is partitioned into several subgroups of hyperspectral bands. Finally, each subgroup is decomposed with IID to obtain the resulting spectral feature used for classification.

### A. Spectral Dimension Reduction

The first step, i.e., spectral dimension reduction has two major objectives: 1) the spectral dimension of the hyperspectral image could be reduced; and 2) the dimensionally reduced data is still directly interpretable in a physical sense. Although dimension reduction methods such as PCA and ICA can be considered to fulfill the first objective, these methods cannot ensure that the band-reduced hyperspectral image pixels are still interpretable in a physical sense. In other words, after PCA or ICA decomposition, the pixels of the dimension reduced image will not be related to the reflectance of the scene. Aside from the transformation-based methods aforementioned, feature selection and clustering-based band reduction methods [59], [60] can ensure the reduced features directly interpretable in a physical sense. However, these types of band selection and reduction-based methods usually require time-consuming clustering or optimization process. Therefore, this paper applies a simple image fusion method, i.e., the averaging method for the dimension reduction of hyperspectral data. Specifically, the $N$ dimensional hyperspectral data set is first divided into $M$ groups of equal size. The number of bands is denoted by $N_1, N_2, \ldots, N_M$ in subgroup $1, 2, \ldots, M$, respectively, with $M$ being the total number of bands in the dimension reduced

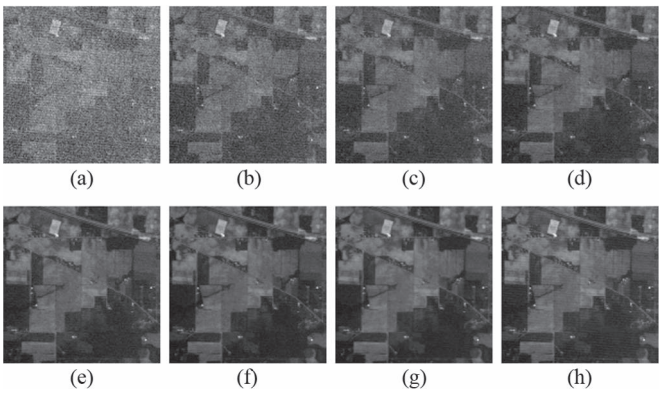



Fig. 3. Intuitive example of averaging-based image fusion for noise removal. From left-top to right-bottom: (a)–(g) first to seventh hyperspectral bands and (h) fused image of these bands obtained by averaging.

data. Then, the averaging-based image fusion is applied separately on each subgroup and the resulting fused bands are used for further processing

$$\tilde{\mathbf{I}}_m = \frac{\sum_{n=1}^{N_m} \mathbf{I}_m^n}{N_m} \tag{4}$$

where $m$ refers to the $m$th subgroup, $\mathbf{I}_m^n$ is the $n$th band in the $m$th subgroup of the original hyperspectral image, $\tilde{\mathbf{I}}_m$ is the $m$th fused band. For example, Fig. 3 shows the first to seventh bands of the Indian Pines image and their corresponding fused image obtained by averaging. It can be seen that image noise can be effectively removed by calculating the average of adjacent bands. Generally, ideal noise-free adjacent hyperspectral bands are highly correlated, and image noise is independently and identically distributed. In this situation, averaging of highly correlated pixels can yield a pixel that is similar to its noise-free value. Furthermore, since the neighboring hyperspectral bands are usually highly correlated, the fused image obtained by averaging [see Fig. 3(h)] is also able to preserve most of useful information in the original bands [see Fig. 3(a)–(g)]. In [19], the averaging method has been demonstrated a good performance in improving the classification accuracy while reducing the spectral dimension of hyperspectral images.

### B. Group IID

For IID of hyperpsectral images, a group decomposition scheme is proposed here. As shown in Fig. 2, the dimension reduced image $\tilde{\mathbf{I}}$ is first partitioned into several subgroups of adjacent bands as follows:

$$\widehat{\mathbf{I}}^k = \begin{cases} \left(\tilde{\mathbf{I}}_{(k-1)Z+1}, \ldots, \tilde{\mathbf{I}}_{((k-1)Z+Z)}\right), & k = 1, 2, \ldots, \lfloor \frac{M}{Z} \rfloor \\ (\tilde{\mathbf{I}}_{M-Z+1}, \ldots, \tilde{\mathbf{I}}_M), & k = \lceil \frac{M}{Z} \rceil \neq \lfloor \frac{M}{Z} \rfloor \end{cases} \tag{5}$$

where $\widehat{\mathbf{I}}^k$ refers to the $k$th group, $\lfloor M/Z \rfloor$ is the largest integer not greater than $M/Z$, and $\lceil M/Z \rceil$ is the smallest integer not less than $M/Z$, $Z$ represents the number of bands in each subgroup. Then, these subgroups are respectively decomposed with the optimization-based IID method [52]. Specifically, based on the descriptions in Section II, the intrinsic components $\mathbf{R}^k$ and $\mathbf{S}^k$ can be obtained by minimizing the following energy function:

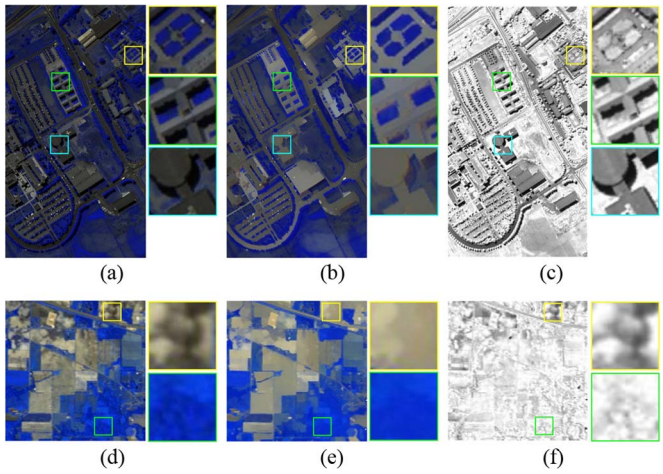



Fig. 4. Two examples for IID of hyperspectral images. (a) and (d) False color composites of the input hyperspectral images. (b) and (e) Reflectance components obtained by the proposed method. (c) and (f) Estimated shading components obtained by the proposed method.

$$E(\mathbf{R}^k, \mathbf{S}^k) = \sum_{i \in \widehat{\mathbf{I}}^k} \left( \mathbf{R}_i^k - \sum_{j \in \omega_i^k} \alpha_{ij}^k \mathbf{R}_j^k \right)^2 + \sum_{i \in \widehat{\mathbf{I}}^k} \left( \frac{\widehat{\mathbf{I}}_i^k}{\mathbf{S}_i^k} - \mathbf{R}_i^k \right)^2 \tag{6}$$

where $\hat{\mathbf{I}}_i^k = (\hat{\mathbf{I}}_{i1}^k, \hat{\mathbf{I}}_{i2}^k, \ldots, \hat{\mathbf{I}}_{iZ}^k)$, $\mathbf{R}_i^k = (\mathbf{R}_{i1}^k, \mathbf{R}_{i2}^k, \ldots, \mathbf{R}_{iZ}^k)$, $\mathbf{R}_j^k = (\mathbf{R}_{j1}^k, \mathbf{R}_{j2}^k, \ldots, \mathbf{R}_{jZ}^k)$, $Z$ is the number of bands in each subgroup, $\alpha_{ij}^k = e^{-[(\mathbf{Y}_i^k - \mathbf{Y}_j^k)^2 / \sigma_{i\mathbf{Y}^k}^2 + \mathcal{A}(\widehat{\mathbf{I}}_i^k, \widehat{\mathbf{I}}_j^k)^2 / \sigma_{i\mathbf{A}^k}^2]}$, $\mathbf{Y}^k$ represents the intensity image, which is calculated by averaging all the bands in the $k$th subgroup $\widehat{\mathbf{I}}^k$, $\mathcal{A}(\widehat{\mathbf{I}}_i^k, \widehat{\mathbf{I}}_j^k)$ denotes the angle between the pixel vector $\hat{\mathbf{I}}_i^k$ and $\hat{\mathbf{I}}_j^k$, and $\sigma_{i\mathbf{A}^k}^2$ and $\sigma_{i\mathbf{Y}^k}^2$ are the variances of the angle and pixel values in the local window $\omega_i^k$, respectively. Finally, the reflectance components of different subgroups are combined together to obtain the resulting IID feature, i.e., an $M$-dimensional feature matrix $\overline{\mathbf{R}}$ used for pixel-wise classification

$$\overline{\mathbf{R}} = \begin{pmatrix} \mathbf{R}^1 \\ \mathbf{R}^2 \\ \ldots \\ \mathbf{R}^M \end{pmatrix}. \tag{7}$$

For example, Fig. 4(a) shows the color composite of three adjacent fused bands obtained in Section III-A. Fig. 4(b) and (c) shows the corresponding intrinsic components estimated by the proposed IID method. As shown in this figure, the intrinsic decomposition step has two obvious advantages. First, some invisible information occluded by shadows can be estimated reasonably [compare the middle and bottom box area of Fig. 4(a)–(c)]. Second, some texture information that are not directly related to the material spectra of objects can be effectively removed [compare the top box area of Fig. 4(a)–(c)]. Similar phenomenon can be observed by performing the experiment on another hyperspectral image [see Fig. 4(d)–(f)].

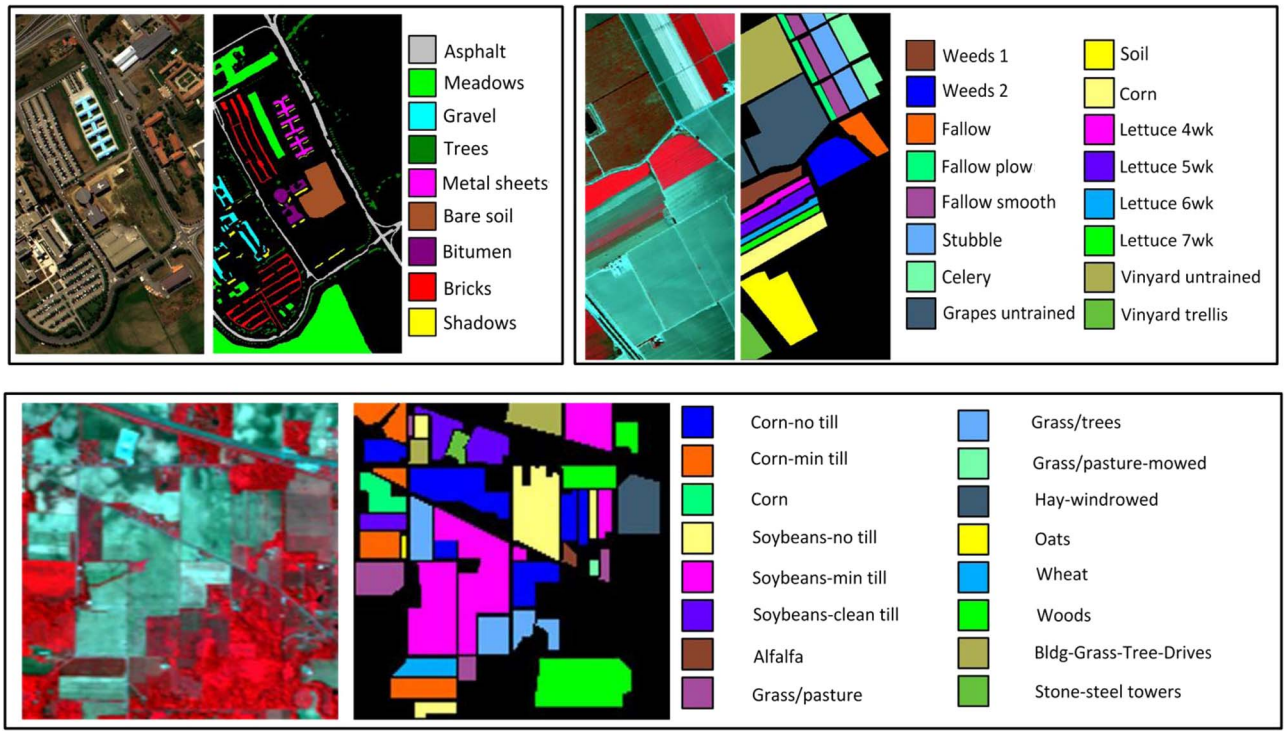



Fig. 5. Color composites of three hyperspectral images and their corresponding reference classification maps. From left-top to right-bottom are the University of Pavia, Salinas, and Indian Pines data sets, respectively.

## IV. Experiments

### A. Experimental Setup

*1) Data Sets:* Three remote sensing hyperspectral data sets, i.e., the Indian Pines image, the University of Pavia image, and the Salinas image, are utilized in our experiments.

The University of Pavia image capturing the University of Pavia, Italy was recorded by Reflective Optics System Imaging Spectrometer (ROSIS). This image contains 115 bands of size $610 \times 340$ with a spatial resolution of 1.3-m per pixel and a spectral coverage ranging from 0.43 to 0.86 $\mu$m. Before the classification, 12 noisy channels were removed, which is a standard preprocessing approach before hyperspectral image classification [15], [25].

The Salinas image was captured by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor over Salinas Valley, California, and it has a spatial resolution of 3.7-m per pixel. The Salinas image contains 224 bands of size $512 \times 217$ and 20 water absorption bands (no. 108–112, 154–167, and 224) were discarded before classification.

The Indian Pines image was acquired by the AVIRIS sensor. It captures the agricultural Indian Pine test site of Northwestern Indiana and contains 220 bands of size $145 \times 145$. Twenty water absorption bands (no. 104–108, 150–163, and 220) were removed before hyperspectral image classification. Furthermore, the spatial resolution of the Indian Pines image is 20-m per pixel, and the spectral coverage is ranging from 0.4 to 2.5 $\mu$m. Fig. 5 shows the color composites of the three images and their corresponding reference classification maps, which are all available online.[2]

[2]http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes

*2) Evaluation Metrics:* In order to evaluate the performance of image classification, three objective quality indexes, i.e., overall accuracy (OA), average accuracy (AA), and Kappa coefficient are utilized for objective evaluation. Specifically, the OA index refers to the percentage of pixels that are correctly labeled in the classification. The AA index measures the mean of the percentage of correctly labeled pixels for each class. Finally, the Kappa coefficient calculates the percentage of correctly classified pixels corrected by the number of agreements that would be expected purely by chance.

### B. Influence of Parameters

For the proposed feature extraction method, the number of features $(M)$ and the size of subgroup $(Z)$ should be determined for group intrinsic decomposition. The influence of the two parameters $M$ and $Z$ are analyzed by evaluating the accuracy and efficiency of SVM classification on the three data sets. In the experiments, the reference data of different images is first partitioned into a parameter training set (the number of samples is 50% of the reference) and a validation set. Then, the training samples used for supervised classification are selected randomly, which account for 10% (Indian Pines), 4% (University of Pavia), and 2% (Salinas) of the parameter training set, respectively. Finally, the OA, AA, and Kappa obtained by the proposed method with different $M$ and $Z$ are measured by comparing the classification result with the parameter training set. It should be mentioned that the SVM classifier is utilized as the spectral classifier because its performance is robust to the dimension of data. In related works such as [22], the SVM method was considered to provide the best pixel-wise classification results for a particular hyperspectral data

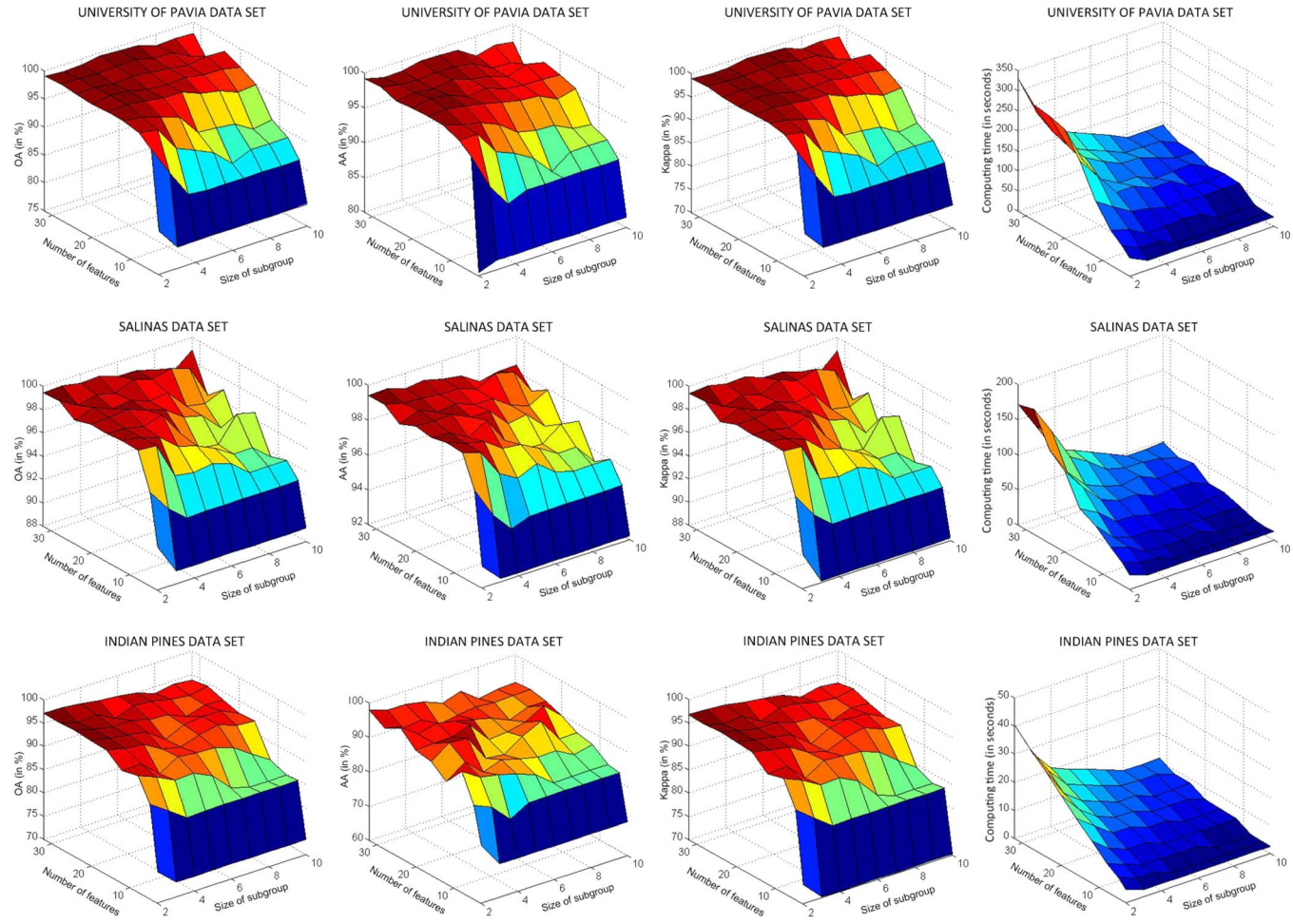



Fig. 6. Experimental results. From top to bottom are the results for the University of Pavia, Salinas, and Indian Pines data sets, respectively. Each row records the (first column) overall classification accuracy, (second column) average classification accuracy, (third column) Kappa coefficient of agreement, and (fourth column) computing time of the proposed method with respect to different number of features $M$ and different size of subgroups $Z$.

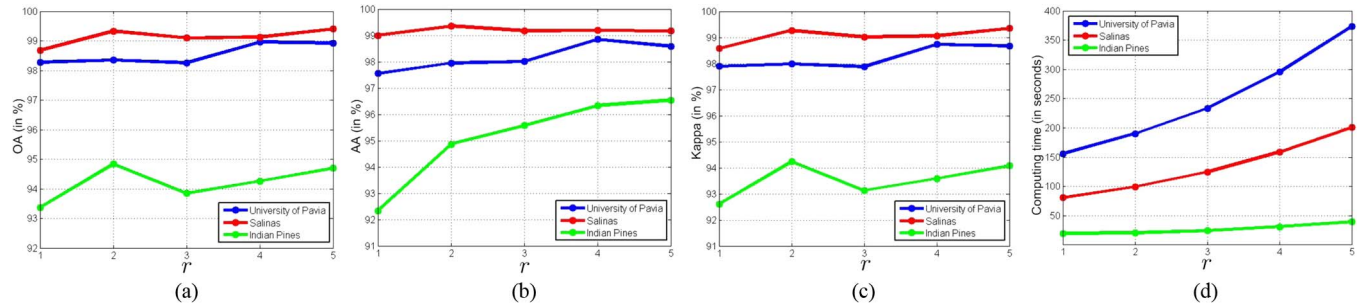



Fig. 7. (a) Overall classification accuracy, (b) AA, (c) Kappa coefficient of agreement, and (d) computing time of the proposed method with respect to different value of $r$.

set. Here, the SVM classifier is implemented in the library for SVMs [61][3] by using the Gaussian kernel with fivefold cross validation, which is a standard setup [15]. As shown in Fig. 6, when $M = 3$, the accuracies of the proposed method are relatively low. This means that when the number of features $M$ is very small, useful discriminative information will be lost in the spectral downsampling process. Furthermore, according to Fig. 6, it is shown that the proposed method can obtain stable and high accuracies when the size of subgroup is less than four. This phenomenon can also be observed for the results of the other two data sets. Moreover, from the fourth column of Fig. 6, it is shown that the computing time of the proposed feature extraction method increases when the number of features is increasing and the size of the subgroup is decreasing, which leads to a tradeoff between computing efficiency and classification accuracy. In this paper, $M = 32$ and $Z = 4$ are set to be default parameters because they can give both high classification accuracies and an acceptable computing burden. Aside from $M$ and $Z$, the radius of the local window $r$ used for IID is also an important parameter that should be considered. Fig. 7 shows the OA, AA, Kappa, and computing time for the proposed classification method with different value of $r$. It can be seen

[3]Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

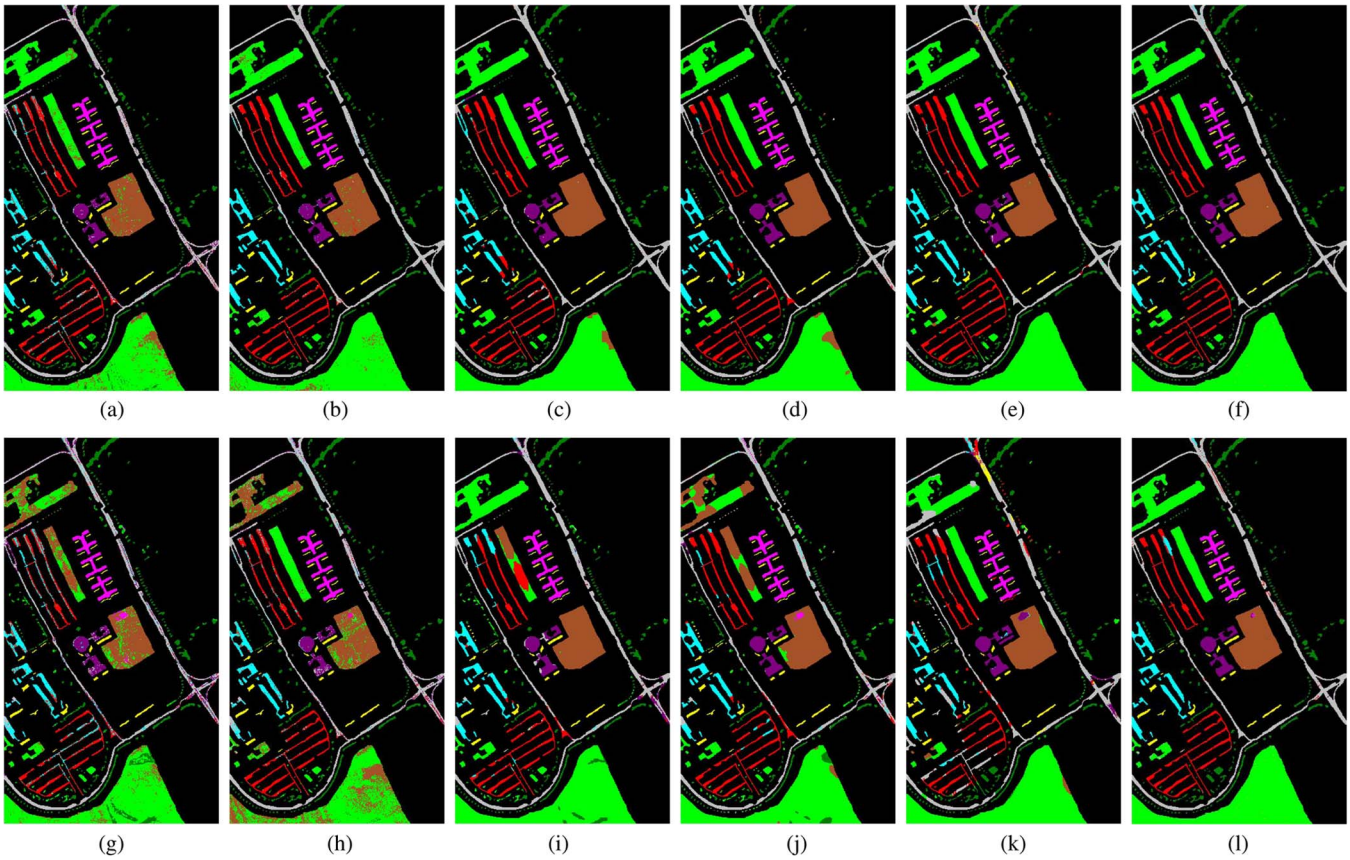



Fig. 8. Classification results (University of Paiva image) obtained by (a) and (g) the SVM method, (b) and (h) the EMP method, (c) and (i) the LMLL method, (d) and (j) the EPF method, (e) and (k) the IFRF method, and (f) and (l) the SVM-IID method. The number in the parentheses refers to the overall classification accuracy given in percent. The number of training samples is 4% and 0.4% of the reference data for (a)–(f) and (g)–(l), respectively. (a) SVM (90.17). (b) EMP (96.62). (c) LMLL (97.27). (d) EPF (97.18). (e) IFRF (98.09). (f) SVM-IID (99.54). (g) SVM (77.47). (h) EMP (78.20). (i) LMLL (90.47). (j) EPF (86.15). (k) IFRF (88.58). (l) SVM-IID (97.92).

that high classification accuracy and fast implementation can be obtained when the radius of local window $r$ is fixed as 2 pixels. Therefore, in this paper, $r = 2$ is the default parameter setting used for the IID method. Here, the classification performance of the proposed feature extraction method is evaluated using the three data sets, i.e., University of Pavia, Salinas, and Indian Pines. The classification maps and accuracies of the proposed approach, as well as for five hyperspectral image classification techniques, are shown in Figs. 8–10 and Tables I–III. Among these methods, the SVM method is implemented in the LIBSVM library [61]. For the extended morphological profiles (EMP)-based method [18], the morphological profiles are constructed using the first four principal components, a circular structural element, a step size increment of two, and four openings and closings. The logistic regression and multilevel logistic (LMLL)-based method [26] is a recently proposed spectral–spatial classifier, which is based on graph-cut theory. The edge-preserving filtering (EPF)-based method [42] and image fusion and recursive filtering (IFRF) [19]-based methods are recently proposed spectral–spatial classification methods that aim at making full use of the spatial information with EPF techniques. For these methods, the default parameters given in [19], [26], [42] are adopted, respectively. For the proposed IID-based feature extraction method, the parameters given in Section IV-B, i.e., $M = 32$, $Z = 4$, and $r = 2$ are adopted in the following experiments.

### C. Classification Results of Different Methods

As shown in the first three columns of Tables I–III, two experiments are performed on the University of Pavia, Salinas, and Indian Pines images, respectively. Specifically, in the first experiment, the training samples were randomly selected to account for 4%, 2%, and 10% of the reference data, respectively. In the second experiment, only 0.4%, 0.2%, and 4% of the reference data were selected as training samples for the three images. As shown in Table III, it should be noticed that the training samples were randomly selected to account for half of the reference data for some minor classes when the total number of reference pixels in these classes is smaller than $1/N$ of the total number of training samples, which account 10% or 4% of the reference data, where $N$ is the number of classes, which is equal to 16 for the Indian Pines image. As shown in Figs. 8(a)–(f)–10(a)–(f), as compared with other spectral–spatial classification methods such as the EMP, LMLL, EPF, and IFRF methods, the SVM-IID-based method always performs better in terms of the highest overall classification accuracies, i.e., obtaining the highest accuracies in all cases, 97.75%, 99.62%, and 99.69%. More importantly, as shown in (g)–(l) of Figs. 8–10, other spectral–spatial classification methods may not perform satisfactory when there is a very limited number of training samples used. For example, for the University of Pavia image, the EMP, EPF, and IFRF methods can improve the accuracy of SVM about only 1%, 9%, and 11%, respectively, when there has only

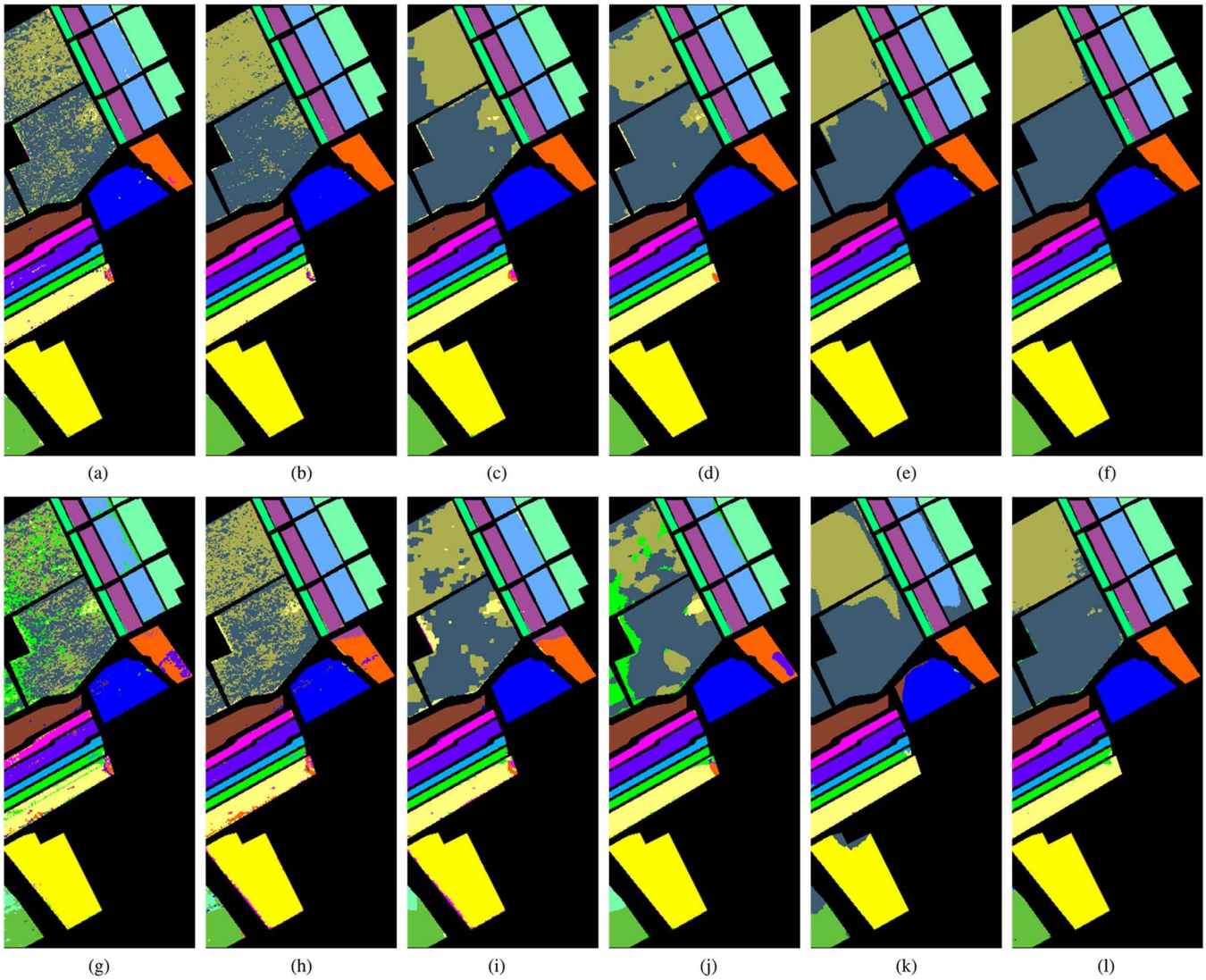



Fig. 9. Classification results (Salinas image) obtained by (a) and (g) the SVM method, (b) and (h) the EMP method, (c) and (i) the LMLL method, (d) and (j) the EPF method, (e) and (k) the IFRF method, and (f) and (l) the SVM-IID method. The number in the parentheses refers to the overall classification accuracy given in percent. The number of training samples is 2% and 0.2% of the reference data for (a)–(f) and (g)–(l), respectively. (a) SVM (88.17). (b) EMP (96.74). (c) LMLL (93.61). (d) EPF (95.22). (e) IFRF (98.73). (f) SVM-IID (99.50). (g) SVM (78.69). (h) EMP (85.35). (i) LMLL (90.06). (j) EPF (86.34). (k) IFRF (93.08). (l) SVM-IID (98.52).

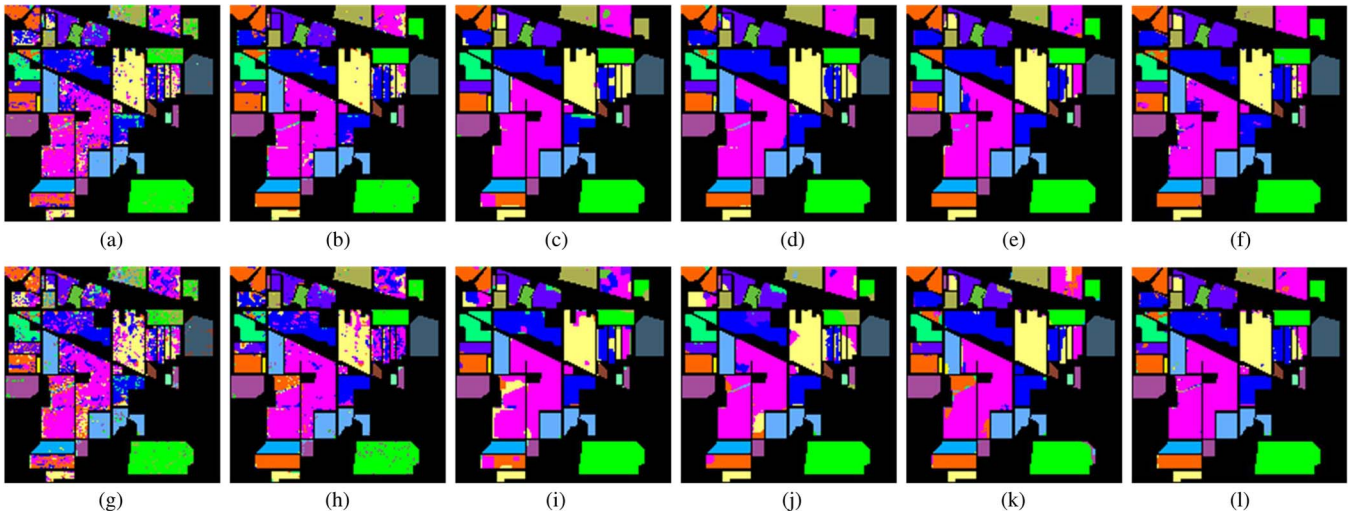



Fig. 10. Classification results (Indian Pines image) obtained by (a) and (g) the SVM method, (b) and (h) the EMP method, (c) and (i) the LMLL method, (d) and (j) the EPF method, (e) and (k) the IFRF method, and (f) and (l) the SVM-IID method. The number in the parentheses refers to the overall classification accuracy given in percent. The number of training samples is 10% and 4% of the reference data for (a)–(f) and (g)–(l), respectively. (a) SVM (79.94). (b) EMP (92.45). (c) LMLL (92.40). (d) EPF (94.26). (e) IFRF (96.86). (f) SVM-IID (97.67). (g) SVM (70.32). (h) EMP (83.75). (i) LMLL (88.63). (j) EPF (88.19). (k) IFRF (92.19). (l) SVM-IID (96.43).

TABLE I

CLASSIFICATION ACCURACIES (UNIVERSITY OF PAVIA IMAGE) IN PERCENTAGE FOR THE DIFFERENT METHODS AS AN AVERAGE AFTER 100 REPEATED EXPERIMENTS. NUMBER IN PARENTHESIS INDICATES THE STANDARD VARIANCE OF THE REPEATED EXPERIMENTS

| Experiment 1 | The number of training samples is 4% of the reference data. | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Class | Train | Test | SVM [61] | EMP [18] | LMLL [26] | EPF [42] | IFRF [19] | SVM-IID |
| Asphalt | 190 | 6441 | 96.77(0.73) | 98.61(0.46) | 94.36(1.68) | 98.07(0.50) | 97.25(1.12) | **99.62(0.15)** |
| Meadows | 190 | 18459 | 97.48(0.40) | 98.84(0.34) | 97.79(1.29) | 99.38(0.29) | 99.68(0.11) | **99.88(0.05)** |
| Gravel | 190 | 1909 | 77.03(2.00) | 95.09(1.51) | 88.10(2.51) | 98.39(1.91) | 95.53(1.74) | **98.92(0.65)** |
| Trees | 190 | 2874 | 87.63(2.97) | 96.64(1.55) | **98.05(0.77)** | 98.45(1.52) | 96.52(3.65) | 97.54(2.76) |
| Sheets | 190 | 1155 | 97.79(1.20) | 98.00(1.56) | 99.83(0.13) | 99.51(0.59) | **99.90(0.20)** | 99.89(0.06) |
| Soil | 191 | 4838 | 78.68(3.01) | 88.33(3.01) | 99.74(0.47) | 95.35(2.42) | 98.72(1.25) | **99.91(0.10)** |
| Bitumen | 190 | 1140 | 66.69(3.04) | 93.51(2.70) | 98.14(0.73) | 97.57(4.47) | 96.69(1.67) | **99.59(0.29)** |
| Bricks | 190 | 3492 | 85.92(1.31) | 97.48(1.02) | 93.77(2.02) | 93.72(1.45) | 94.40(2.07) | **98.66(0.92)** |
| Shadows | 190 | 757 | 99.89(0.11) | 99.91(0.14) | **99.92(0.14)** | 98.23(1.51) | 92.54(3.76) | 94.84(3.93) |
| OA | | | 90.79(0.71) | 96.82(0.51) | 96.81(0.71) | 97.98(0.49) | 98.07(0.43) | **99.41(0.22)** |
| AA | | | 87.54(0.73) | 96.27(0.58) | 96.63(0.49) | 97.63(0.77) | 96.80(0.73) | **98.76(0.44)** |
| Kappa | | | 87.79(0.91) | 95.75(0.68) | 95.73(0.94) | 97.29(0.66) | 97.40(0.57) | **99.21(0.30)** |
| Experiment 2 | The number of training samples is 0.4% of the reference data. | | | | | | | |
| Asphalt | 19 | 6612 | 92.52(3.73) | 93.84(2.99) | 87.87(6.22) | **96.10(2.89)** | 81.90(7.31) | 91.40(7.50) |
| Meadows | 19 | 18630 | 89.73(2.65) | 92.19(1.85) | 84.11(7.26) | 94.60(3.79) | 97.72(1.54) | **99.19(0.67)** |
| Gravel | 19 | 2080 | 55.16(7.31) | 69.87(7.68) | 86.43(8.58) | 84.08(12.81) | 75.96(7.80) | **94.11(6.49)** |
| Trees | 19 | 3045 | 66.20(10.39) | 87.75(6.44) | **91.43(4.20)** | 71.58(15.51) | 81.61(9.78) | 88.84(7.21) |
| Sheets | 19 | 1326 | 91.52(6.76) | 95.85(2.96) | 99.30(0.45) | 95.29(5.03) | **99.49(1.74)** | 99.24(1.39) |
| Soil | 19 | 5010 | 47.80(7.54) | 51.15(5.54) | 92.06(8.61) | 62.83(13.91) | 86.36(8.75) | **98.40(3.72)** |
| Bitumen | 19 | 1311 | 50.51(6.88) | 65.15(9.48) | 96.83(2.92) | 78.88(13.71) | 78.90(8.66) | **97.28(5.30)** |
| Bricks | 19 | 3663 | 74.81(5.96) | 87.62(3.00) | **88.74(6.32)** | 85.31(7.27) | 70.89(8.38) | 82.64(6.79) |
| Shadows | 19 | 928 | 99.90(0.09) | **99.92(0.24)** | 98.39(1.98) | 96.90(1.88) | 62.03(12.80) | 92.02(10.57) |
| OA | | | 74.00(3.64) | 81.22(3.54) | 87.84(2.86) | 84.05(5.11) | 86.60(2.90) | **94.67(2.07)** |
| AA | | | 74.24(2.35) | 82.59(2.73) | 91.68(1.49) | 85.06(3.84) | 81.65(2.78) | **93.68(2.21)** |
| Kappa | | | 67.22(4.09) | 76.03(4.11) | 84.33(3.46) | 79.69(6.07) | 82.60(3.56) | **93.02(2.63)** |

TABLE II

CLASSIFICATION ACCURACIES (SALINAS IMAGE) IN PERCENTAGE FOR THE DIFFERENT METHODS AS AN AVERAGE AFTER 100 REPEATED EXPERIMENTS. NUMBER IN PARENTHESIS INDICATES THE STANDARD VARIANCE OF THE REPEATED EXPERIMENTS

| Experiment 1 | The number of training samples is 2% of the reference data. | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Class | Train | Test | SVM [61] | EMP [18] | LMLL [26] | EPF [42] | IFRF [19] | SVM-IID |
| Weeds_1 | 68 | 1941 | 99.73(0.28) | 99.93(0.11) | 99.83(0.14) | **100(0)** | 99.99(0.06) | **100(0)** |
| Weeds_2 | 67 | 3659 | 99.48(0.25) | 99.60(0.30) | 99.96(0.09) | **100(0)** | **100(0)** | 99.97(0.04) |
| Fallow | 67 | 1909 | 94.35(1.34) | 96.75(1.58) | 99.80(0.04) | 98.36(1.24) | **99.89(0.08)** | 99.85(0.09) |
| Fallow_P | 67 | 1327 | 97.71(1.04) | 98.25(0.96) | **99.22(0.72)** | 97.17(0.95) | 97.32(1.05) | 97.80(0.50) |
| Fallow_S | 67 | 2611 | 98.68(0.60) | 99.14(0.51) | 98.75(0.43) | **99.96(0.22)** | 99.94(0.13) | **99.96(0.09)** |
| Stubble | 68 | 3891 | 99.98(0.05) | 99.95(0.15) | 99.66(0.28) | 99.98(0.04) | **100(0)** | 99.98(0.04) |
| Celery | 68 | 3511 | 99.20(0.47) | 99.87(0.14) | 99.55(0.22) | **99.98(0.05)** | 99.70(0.16) | 99.90(0.08) |
| Grapes | 70 | 11201 | 78.72(1.77) | 94.79(1.64) | 85.97(2.69) | 91.78(4.25) | **99.34(0.63)** | 99.07(0.52) |
| Soil | 69 | 6134 | 99.35(0.24) | 99.41(0.29) | 99.69(0.35) | 99.65(0.21) | **99.99(0.03)** | 99.86(0.19) |
| Corn | 67 | 3211 | 86.45(3.52) | 95.10(1.99) | 94.68(1.34) | 93.88(3.13) | **99.81(0.18)** | 98.80(1.21) |
| Lettuce_4 | 67 | 1001 | 91.92(2.91) | 96.39(2.99) | 98.47(1.61) | 98.80(1.43) | 99.26(0.42) | **99.98(0.05)** |
| Lettuce_5 | 67 | 1860 | 97.06(1.00) | 98.79(0.90) | **100(0)** | 99.99(0.03) | 99.90(0.30) | 99.99(0.02) |
| Lettuce_6 | 68 | 848 | 96.85(1.25) | 98.87(1.12) | 98.48(0.87) | **100(0)** | 99.28(0.97) | **100(0)** |
| Lettuce_7 | 67 | 1003 | 92.23(3.96) | 97.06(2.33) | 98.60(1.26) | **99.19(0.83)** | 97.96(1.12) | 97.51(1.82) |
| Vinyard_U | 68 | 7200 | 63.92(3.76) | 86.67(3.09) | 78.78(5.17) | 85.50(4.22) | 97.07(2.52) | **99.52(0.95)** |
| Vinyard_T | 68 | 1739 | 97.50(1.43) | 97.31(1.47) | 99.35(0.36) | **99.99(0.07)** | 99.95(0.10) | 97.92(2.09) |
| OA | | | 88.60(0.89) | 96.11(0.62) | 93.55(0.75) | 95.59(1.08) | 99.27(0.38) | **99.45(0.19)** |
| AA | | | 93.32(0.51) | 97.37(0.37) | 96.92(0.35) | 97.76(0.47) | 99.34(0.20) | **99.38(0.20)** |
| Kappa | | | 87.32(0.98) | 95.67(0.69) | 92.80(0.84) | 95.08(1.20) | 99.18(0.42) | **99.39(0.21)** |
| Experiment 2 | The number of training samples is 0.2% of the reference data. | | | | | | | |
| Weeds_1 | 7 | 2002 | 97.30(4.45) | 98.22(4.34) | 98.88(0.89) | 98.78(3.87) | 90.91(10.51) | **100(0)** |
| Weeds_2 | 6 | 3720 | 98.73(1.06) | 98.65(1.37) | 99.30(1.07) | **99.99(0.02)** | 98.78(3.68) | 99.96(0.04) |
| Fallow | 6 | 1970 | 85.52(5.38) | 83.98(6.57) | 92.47(9.16) | 91.17(7.07) | 98.61(2.43) | **99.83(0.08)** |
| Fallow_P | 6 | 1388 | 97.02(1.22) | 97.70(0.75) | **98.88(1.26)** | 97.16(0.49) | 86.61(8.49) | 96.39(4.33) |
| Fallow_S | 8 | 2670 | 95.68(4.31) | 95.07(4.85) | 98.56(1.56) | **99.63(1.01)** | 98.35(2.38) | 99.61(1.10) |
| Stubble | 6 | 3953 | 99.89(0.75) | 99.80(0.93) | 99.11(1.41) | 99.98(0.03) | 99.95(0.42) | **100(0)** |
| Celery | 8 | 3571 | 95.97(2.34) | 98.39(1.98) | 99.44(0.42) | 96.92(3.00) | 96.73(6.04) | **99.97(0.03)** |
| Grapes | 8 | 11263 | 66.96(8.05) | 79.34(7.53) | 71.57(17.79) | 74.53(10.32) | 95.92(5.46) | **96.80(3.59)** |
| Soil | 6 | 6197 | 98.57(0.98) | 99.24(0.34) | 96.62(4.07) | 99.39(0.18) | **99.72(0.35)** | 99.18(2.01) |
| Corn | 8 | 3270 | 76.97(10.16) | 90.47(4.49) | 85.86(7.27) | 85.99(11.28) | **99.05(0.77)** | 94.43(8.36) |
| Lettuce_4 | 7 | 1061 | 82.77(10.96) | 80.30(8.70) | 93.44(2.39) | 95.46(3.45) | 96.96(3.44) | **99.48(1.47)** |
| Lettuce_5 | 6 | 1921 | 88.16(8.13) | 93.13(7.09) | 99.74(1.50) | 88.90(11.82) | 97.23(2.43) | **99.88(0.24)** |
| Lettuce_6 | 6 | 910 | 88.51(7.82) | 94.47(3.63) | **98.28(0.93)** | 94.90(6.81) | 94.32(6.37) | 92.09(11.81) |
| Lettuce_7 | 7 | 1063 | 85.72(14.54) | **96.48(5.81)** | 94.02(4.46) | 87.20(24.01) | 93.17(9.24) | 90.02(8.17) |
| Vinyard_U | 7 | 7261 | 48.47(9.31) | 63.24(11.34) | 72.00(22.26) | 51.55(21.21) | 82.90(7.55) | **90.16(6.60)** |
| Vinyard_T | 6 | 1801 | 93.02(10.14) | 93.92(6.58) | 92.81(8.23) | **99.79(0.60)** | 99.15(3.78) | 99.77(1.13) |
| OA | | | 81.02(2.25) | 87.07(2.53) | 87.97(3.73) | 85.06(2.92) | 94.51(2.21) | **96.75(1.49)** |
| AA | | | 87.46(1.80) | 91.40(1.51) | 93.19(1.57) | 91.33(3.11) | 95.52(1.63) | **97.35(1.18)** |
| Kappa | | | 78.94(2.43) | 85.63(2.80) | 86.64(4.13) | 83.38(3.27) | 93.90(2.45) | **96.38(1.65)** |

TABLE III
CLASSIFICATION ACCURACIES (INDIAN PINES IMAGE) IN PERCENTAGE FOR THE DIFFERENT METHODS AS AN AVERAGE AFTER 100 REPEATED EXPERIMENTS. NUMBER IN PARENTHESIS INDICATES THE STANDARD VARIANCE OF THE REPEATED EXPERIMENTS

| Experiment 1 | The number of training samples is 10% of the reference data. | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Class | Train | Test | SVM [61] | EMP [18] | LMLL [26] | EPF [42] | IFRF [19] | SVM-IID |
| Alfalfa | 23 | 23 | 69.52(11.75) | 89.48(6.17) | 95.40(3.51) | 98.61(6.40) | **98.65(2.06)** | 96.78(6.47) |
| Corn_N | 89 | 1339 | 75.38(2.75) | 86.63(2.84) | 90.64(4.08) | 95.03(3.33) | 95.41(2.50) | **96.90(1.80)** |
| Corn_M | 73 | 757 | 70.79(4.00) | 89.74(2.87) | 87.23(5.38) | 95.68(2.32) | 96.07(2.43) | **97.20(2.23)** |
| Corn | 66 | 171 | 48.48(3.98) | 79.14(4.44) | **99.29(1.25)** | 71.52(4.63) | 95.58(2.66) | 91.85(4.73) |
| Grass_M | 71 | 412 | 87.88(2.76) | 93.23(1.30) | 96.39(2.67) | **98.39(0.72)** | 98.07(1.87) | 97.88(0.95) |
| Grass_T | 81 | 649 | 94.02(1.54) | 97.50(0.55) | **99.90(0.22)** | 98.90(0.66) | 98.62(0.95) | 99.40(1.01) |
| Grass_P | 14 | 14 | 82.69(9.54) | 81.22(13.02) | 98.97(2.58) | **99.87(0.94)** | 98.02(10.69) | 91.17(17.81) |
| Hay_W | 71 | 407 | 98.81(0.71) | **100(0)** | **100(0)** | 99.49(0.60) | **100(0)** | **100(0)** |
| Oats | 10 | 10 | 57.87(9.26) | 68.44(12.25) | **98.96(3.99)** | 97.00(17.14) | 88.82(12.74) | 97.08(5.72) |
| Soybean_N | 76 | 896 | 70.94(2.82) | 86.18(4.24) | 93.49(5.39) | 82.09(5.43) | 95.26(2.13) | **95.49(1.81)** |
| Soybean_M | 112 | 2343 | 85.27(1.85) | 93.82(0.83) | 89.75(3.08) | 95.93(2.38) | **98.42(1.09)** | 98.03(0.59) |
| Soybean_C | 68 | 525 | 72.34(3.44) | 87.09(3.55) | **98.60(0.84)** | 92.81(4.53) | 96.72(2.20) | 97.23(1.81) |
| Wheat | 67 | 138 | 94.59(2.95) | 99.22(0.93) | 99.72(0.35) | **100(0)** | 97.40(3.02) | 99.84(0.23) |
| Woods | 89 | 1176 | 95.80(0.95) | 99.59(0.38) | 92.45(3.09) | 99.11(0.37) | 99.75(0.41) | **99.76(0.37)** |
| Buildings | 68 | 318 | 61.94(4.77) | **97.47(0.91)** | 97.19(2.95) | 80.83(9.14) | 96.22(2.61) | 96.63(1.13) |
| Stone | 47 | 46 | 88.53(6.46) | 93.95(4.01) | **98.04(2.84)** | 91.02(5.48) | 96.67(4.06) | 97.97(0.68) |
| OA | | | 80.57(0.82) | 92.21(0.83) | 93.01(1.12) | 93.55(1.17) | 97.39(0.54) | **97.69(0.51)** |
| AA | | | 78.71(1.41) | 90.17(1.23) | 96.00(0.73) | 93.52(1.49) | 96.85(1.15) | **97.08(1.28)** |
| Kappa | | | 77.81(0.92) | 91.06(0.94) | 91.98(1.28) | 92.59(1.34) | 97.00(0.62) | **97.35(0.58)** |
| Experiment 2 | The number of training samples is 4% of the reference data. | | | | | | | |
| Alfalfa | 23 | 23 | 41.22(12.21) | 80.12(17.30) | **95.69(3.06)** | 87.65(20.03) | 88.40(23.09) | 90.22(11.22) |
| Corn_N | 34 | 1394 | 64.92(5.20) | 73.06(6.66) | 79.73(6.56) | 92.17(6.48) | 89.84(4.99) | **92.57(4.58)** |
| Corn_M | 27 | 803 | 55.17(5.20) | 80.91(4.77) | 80.10(8.07) | 90.48(8.02) | 83.99(7.48) | **92.61(3.82)** |
| Corn | 25 | 212 | 39.42(5.26) | 63.21(7.67) | **97.61(3.10)** | 58.26(10.99) | 87.91(8.26) | 82.96(8.35) |
| Grass_M | 27 | 456 | 83.08(4.13) | 83.15(5.31) | 90.29(5.34) | **97.79(1.67)** | 94.28(4.76) | 95.82(3.75) |
| Grass_T | 26 | 704 | 90.18(2.20) | 97.02(1.08) | **99.23(1.48)** | 96.37(2.15) | 97.35(1.84) | 98.42(2.82) |
| Grass_P | 14 | 14 | 55.73(13.28) | 62.05(16.55) | **99.36(2.29)** | **100(0)** | 76.06(31.24) | 68.84(24.00) |
| Hay_W | 27 | 451 | 98.59(0.87) | **100(0)** | 99.92(0.18) | 99.47(0.60) | **100(0)** | **100(0)** |
| Oats | 10 | 10 | 35.83(8.30) | 46.79(8.93) | **99.90(1.00)** | 88.98(17.85) | 68.87(22.10) | 89.56(12.01) |
| Soybean_N | 30 | 942 | 62.12(4.69) | 72.39(6.10) | 82.93(8.19) | 70.71(4.77) | 87.35(5.42) | **88.77(6.14)** |
| Soybean_M | 31 | 2424 | 78.17(3.25) | 88.13(3.94) | 78.89(7.52) | 89.28(5.44) | **96.42(1.91)** | 96.37(1.69) |
| Soybean_C | 28 | 565 | 52.15(5.38) | 77.74(6.14) | **95.46(5.47)** | 81.07(7.70) | 90.58(5.96) | 89.04(6.06) |
| Wheat | 25 | 180 | 88.38(5.01) | 97.84(2.26) | **99.74(0.28)** | 99.73(0.86) | 91.06(8.52) | 99.35(1.09) |
| Woods | 30 | 1235 | 94.26(1.42) | 99.25(0.53) | 91.34(4.55) | 98.28(1.94) | 99.17(0.90) | **99.32(0.54)** |
| Buildings | 27 | 359 | 52.92(7.29) | 91.57(6.93) | 91.37(7.56) | 79.70(8.53) | 89.51(5.91) | **94.93(4.89)** |
| Stone | 26 | 67 | 85.93(8.94) | 91.75(2.71) | 96.50(3.76) | 81.14(7.31) | 95.45(4.86) | **98.73(1.17)** |
| OA | | | 71.25(1.63) | 84.88(1.88) | 86.40(1.89) | 87.37(0.95) | 92.42(1.55) | **94.31(1.15)** |
| AA | | | 67.38(1.91) | 82.11(1.98) | **92.38(1.03)** | 88.19(2.36) | 89.76(3.25) | 92.32(1.99) |
| Kappa | | | 67.52(1.80) | 82.83(2.10) | 84.57(2.11) | 85.60(1.11) | 91.37(1.75) | **93.51(1.30)** |

0.4% of the reference data for training [compare Fig. 8(h), (j), and (k) with Fig. 8(g)]. Similar observations can be made for the Salinas and Indian Pines data sets. For example, for the Indian Pines image, the EMP, LMLL, and EPF methods cannot obtain accuracy higher than 90% when the number of training samples is 4% of the reference data. In other words, traditional spectral–spatial classification methods have the common limitation that they tend to produce smaller improvements in classification accuracies when the number of training samples is smaller. By contrast, as shown in (l) of Figs. 8–10, the SVM-IID-based method can improve the accuracies of SVM even more (77.47% to 97.92% for the University of Pavia, 78.69% to 98.52% for the Salinas, and 70.32% to 96.43% for the Indian Pines) with only a small number of labeled samples.

In order to evaluate the performance of different methods more objectively, the classification experiments have been repeated 100 times to estimate the mean and standard variance of the OA, AA, and the Kappa coefficient (Kappa) (see Tables I–III). From the top subtable of Tables I–III, it can be observed that, with the proposed SVM-IID method, the OA, AA, and Kappa for SVM can be always increased significantly (17%–19%) for the three images. More importantly, the SVM-IID method always outperforms other methods more significantly for the OA and Kappa indexes when there is a small number of training samples (see the bottom subtable of Tables I–III). Specifically, with the SVM-IID method, the OA, AA, and Kappa of SVM can be increased about 17%–22% for the three images. Although the SVM-IID method not always ranks as first for the AA index, the AA differences between the SVM-IID method and the method ranking as first (see Table III) is very small. Finally, Table IV shows the resulting $p$ values obtained in the Paired Student's $t$-test for the accuracies obtained by different methods. Specifically, the $p$ values denote the statistical significance of the difference between the two overall accuracies obtained by two different methods, i.e., (SVM&SVM-IID), (EMP&SVM-IID), (EPF&SVM-IID), and (IFRF&SVM-IID). As shown in the table, the $p$ values obtained in the Paired Student's $t$-test between the proposed method and other methods are usually much smaller than 0.01. It means that the improvement of OA obtained by the SVM-IID method can be always considered as statistically significant (at the 99% level). Therefore, it is able to convince that the mean and standard variance are meaningful in revealing the real performance of different methods. Furthermore, the Paired Student's $t$-test

TABLE IV
$p$ VALUES OBTAINED IN THE PAIRED STUDENT'S $t$-TEST BETWEEN THE PROPOSED SVM-IID METHOD AND OTHER METHODS, I.E., SVM, EMP, LMLL, EPF, AND IFRF METHODS. aE − b REFERS TO a × $10^{-b}$

| Methods | University of Pavia | | Salinas | | Indian Pines | |
|---|---|---|---|---|---|---|
| | Experiment 1 | Experiment 2 | Experiment 1 | Experiment 2 | Experiment 1 | Experiment 2 |
| SVM&SVM-IID | 1.53E-107 | 7.74E-71 | 1.89E-110 | 8.14E-77 | 4.05E-126 | 8.97E-112 |
| EMP&SVM-IID | 2.90E-71 | 2.71E-55 | 6.42E-74 | 1.14E-60 | 7.38E-86 | 9.00E-70 |
| LMLL&SVM-IID | 5.40E-58 | 1.16E-39 | 2.67E-89 | 1.15E-37 | 5.38E-63 | 6.58E-60 |
| EPF&SVM-IID | 1.35E-46 | 3.76E-35 | 7.18E-59 | 1.57E-44 | 3.55E-57 | 2.85E-55 |
| IFRF&SVM-IID | 6.69E-57 | 1.30E-53 | 4.99E-05 | 3.49E-17 | 2.93E-06 | 1.78E-20 |

TABLE V
COMPUTING TIME (IN SECONDS) OF DIFFERENT CLASSIFICATION ALGORITHMS FOR THE INDIAN PINES IMAGE OF SIZE 145 × 145 × 200

| Methods | SVM | EMP | LMLL | EPF | IFRF | SVM-IID |
|---|---|---|---|---|---|---|
| Time | 101 | 21 | 5 | 101 | 12 | 48 |

also has been performed for other quality indexes, i.e., AA, and Kappa, which demonstrate that the improvements of these accuracies are also meaningful (statistically significant at the 99% level) for all experiments.

### D. Computational Complexity

Here, the computing time of different classification methods are compared. Experiments are performed using MATLAB on a Computer with 3.5-GHz CPU and 8-GB memory. Table V records the computing time of the different classification methods (for the SVM, EMP, EPF, IFRF, and SVM-IID methods, which are based on the SVM classifier, the time of fivefold cross validation is included for a fair comparison). From this table, it can be seen that the computing time of the proposed method (48.0 s) is acceptable in comparison to the SVM method using original hyperspectral data (100.7 s) and the EPF method, which is based on the postprocessing of SVM (101.9 s). However, it is not computationally efficient compared with other methods such as LMLL, EMP, and IFRF. The reason is that the optimization-based IID method requires many iterations to find the optimal solutions. To solve this problem, designing a faster implementation of SVM-IID with GPU programing will be investigated.

## V. CONCLUSION

In this paper, IID has been proposed for feature extraction of hyperspectral images for the first time. Compared with other widely used spectral–spatial classification methods, experiments carried out on three real hyperspectral data showed the outstanding performance of the SVM-IID method in terms of classification accuracies, particularly when the number of training samples is relatively small. The reason is that the IID can effectively remove useless spatial information such as shading and texture that are not directly related to the material of different objects. However, it should be mentioned that, although the proposed method can obtain high classification accuracies with a small number of training samples, it is not computationally efficient compared with other feature extraction methods. This paper primarily aims at showing that IID, which models the perceiving function of human vision, works well for hyperspectral image feature extraction. Designing a more effective technique for IID of hyperspectral images will be a topic of future research. Furthermore, whether IID can be applied for other hyperspectral applications will also be investigated.

## ACKNOWLEDGMENT

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

## REFERENCES

[1] J. Pontius, M. Martin, L. Plourde, and R. Hallett, "Ash decline assessment in emerald ash borer-infested regions: A test of tree-level, hyperspectral technologies," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 2665–2676, May 2008.

[2] Y. Lanthier, A. Bannari, D. Haboudane, J. R. Miller, and N. Tremblay, "Hyperspectral data segmentation and classification in precision agriculture: A multi-scale analysis," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2008, vol. 2, pp. 585–588.

[3] Q. Jackson and D. A. Landgrebe, "Adaptive Bayesian contextual classification based on Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2454–2463, Nov. 2002.

[4] H. R. Bittencourt, D. A. de Oliveira Moraes, and V. Haertel, "A binary decision tree classifier implementing logistic regression as a feature selection and classification method and its comparison with maximum likelihood," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2007, pp. 1755–1758.

[5] F. Ratle, G. Camps-Valls, and J. Weston, "Semisupervised neural networks for efficient hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2271–2282, May 2010.

[6] Y. Zhong and L. Zhang, "An adaptive artificial immune network for supervised classification of multi-/hyperspectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 894–909, Mar. 2012.

[7] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.

[8] F. Bovolo, L. Bruzzone, and L. Carlin, "A novel technique for subpixel image classification based on support vector machine," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2983–2999, Nov. 2010.

[9] Y. Chen, N. M. Nasrabadi, and T. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 217–231, Jan. 2013.

[10] A. Castrodad *et al.*, "Learning discriminative sparse representations for modeling, source separation, mapping of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4263–4281, Nov. 2011.

[11] D. G. Stavrakoudis, G. N. Galidaki, I. Z. Gitas, and J. B. Theocharis, "A genetic fuzzy-rule-based classifier for land cover classification from hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 130–148, Jan. 2012.

[12] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.

[13] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial-spectral kernel-based approach for the classification of remote-sensing images," *Pattern Recognit.*, vol. 45, no. 1, pp. 381–392, Jan. 2012.

[14] B. Guo, S. R. Gunn, R. I. Damper, and J. D. B. Nelson, "Customizing kernel functions for SVM-based hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 622–629, Apr. 2008.

[15] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. S1, pp. S110–S122, Sep. 2009.

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

[17] L. Zhang, L. Zhang, D. Tao, and X. Huang, "Tensor discriminative locality alignment for hyperspectral image spectral-spatial feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 242–256, Jan. 2013.

[18] J. A. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.

[19] 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.

[20] W. Liao, A. Pizurica, P. Scheunders, W. Philips, and Y. Pi, "Semisupervised local discriminant analysis for feature extraction in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 184–198, Jan. 2013.

[21] S. Prasad and L. M. Bruce, "Limitations of principal components analysis for hyperspectral target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 625–629, Oct. 2008.

[22] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, "Hyperspectral image classification with independent component discriminant analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 4865–4876, Dec. 2011.

[23] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving discriminant analysis in kernel-induced feature spaces for hyperspectral image classication," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 894–898, Sep. 2011.

[24] W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality-preserving dimensionality reduction and classication for hyperspectral image analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1185–1198, Apr. 2012.

[25] 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.

[26] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new Bayesian approach with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3947–3960, Oct. 2011.

[27] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial classification of hyperspectral data using loopy belief propagation and active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 844–856, Feb. 2013.

[28] W. Di and M. M. Crawford, "View generation for multiview maximum disagreement based active learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1942–1954, May 2012.

[29] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, Mar. 2012.

[30] G. Moser and S. B. Serpico, "Combining Support Vector Machines and Markov Random Fields in an integrated framework for contextual image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2734–2752, May 2013.

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

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

[33] S. Valero, P. Salembier, and J. Chanussot, "Hyperspectral image representation and processing with binary partition trees," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1430–1443, Apr. 2013.

[34] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 466–479, Mar. 2005.

[35] D. M. 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.

[36] 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.

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

[38] 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, Jul. 2010.

[39] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.

[40] K. Bernard, Y. Tarabalka, J. Angulo, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2008–2021, Apr. 2012.

[41] P. Zhong and R. Wang, "Learning conditional random fields for classification of hyperspectral images," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1890–1907, Jul. 2010.

[42] X. Kang, S. Li, and J. A. Benediktsson, "Spectral-spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.

[43] R. Ji *et al.*, "Spectral-spatial constraint hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1811–1824, Mar. 2014.

[44] A. Erturk, M. K. Gullu, and S. Erturk, "Hyperspectral image classification using empirical mode decomposition with spectral gradient enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2787–2798, May 2013.

[45] B. Demir and S. Erturk, "Empirical mode decomposition of hyperspectral images for support vector machine classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4071–4084, Nov. 2010.

[46] J. M. Duarte-Carvajalino, P. E. Castillo, and M. Velez-Reyes, "Comparative study of semi-implicit schemes for nonlinear diffusion in hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1303–1314, May 2007.

[47] R. Mendez-Rial and J. Martin-Herrero, "Efficiency of semi-implicit schemes for anisotropic diffusion in the hypercube," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2389–2398, May 2012.

[48] J. Li, H. Zhang, Y. Huang, and L. Zhang, "Hyperspectral image classification by nonlocal joint collaborative representation with a locally adaptive dictionary," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3707–3719, Jun. 2014.

[49] M. F. Tappen, W. T. Freeman, and E. H. Adelson, "Recovering intrinsic images from a single image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1459–1472, Sep. 2005.

[50] Y. Yacoob and L. S. Davis, "Segmentation using appearance of mesostructure roughness," *Int. J. Comput. Vis.*, vol. 83, no. 3, pp. 248–273, Jul. 2009.

[51] G. D. Finlayson, S. D. Hordley, C. Lu, and M. Drew, "On the removal of shadows from images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 59–68, Jan. 2006.

[52] J. Shen, X. Yang, X. Li, and Y. Jia, "Intrinsic image decomposition using optimization and user scribbles," *IEEE Trans. Cybern.*, vol. 43, no. 2, pp. 425–436, Mar. 2013.

[53] J. Barron and J. Malik, Shape, illumination, reflectance from shading, EECS Dept., Univ. California Berkeley, Berkeley, CA, USA, Tech. Rep. UCB/EECS-2013-117. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-117.html

[54] A. Hanson and E. Riseman, *Computer Vision Systems*. New York, NY, USA: Academic, 1978.

[55] A. Bousseau, S. Paris, and F. Durand, "User-assisted intrinsic images," *ACM Trans. Graph.*, vol. 28, no. 5, pp. 130:1–130:10, Dec. 2009.

[56] R. Grosse, M. K. Johnson, E. H. Adelson, and W. T. Freeman, "Ground-truth dataset and baseline evaluations for intrinsic image algorithms," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 2335–2342.

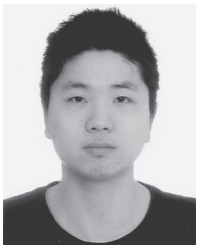
[57] L. Shen, C. Yeo, and B.-S. Hua, "Intrinsic image decomposition using a sparse representation of reflectance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2904–2915, Dec. 2013.

[58] Q. Zhao *et al.*, "A closed-form solution to retinex with nonlocal texture constraints," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1437–1444, Jul. 2012.

[59] A. Jensen and R. Solberg, "Fast hyperspectral feature reduction using piecewise constant function approximations," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 547–551, Oct. 2007.

[60] H. Su and Q. Du, "Hyperspectral band clustering and band selection for urban land cover classification," *Geocarto Int.*, vol. 27, no. 5, pp. 395–411, Aug. 2012.

[61] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, Apr. 2011.



**Xudong Kang** (S'13) received the B.Sc degree from Northeast University, Shenyang, China, in 2007. He is working toward the Ph.D. degree in electrical engineering at Hunan University, Changsha, China.

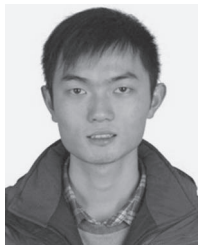
In 2012–2013, he was a Visiting Ph.D. Student in electrical engineering with the University of Iceland, Reykjavik, Iceland, working with Prof. J. A. Benediktsson. He is engaged in image fusion, pansharpening, and hyperspectral image classification.

In 2012, he was the recipient of the Scholarship Award for Excellent Doctoral Student granted by the Chinese Ministry of Education; in 2013, he was the recipient of the National Graduate Scholarship granted by Chinese Government, and the Presidential Scholarship granted by Hunan University; in 2014, he received the Best Student Paper Award (Second Prize) in International Geoscience and Remote Sensing Symposium 2014 from the IEEE Geoscience and Remote Sensing Society for the paper "Extended Random Walkers for Hyperspectral Image Classification".

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

In 2001, he joined the College of Electrical and Information Engineering, Hunan University. From May 2001 to October 2001, he was a Research Associate with the Department of Computer Science, Hong Kong University of Science and Technology, Kowloon, Hong Kong. From November 2002 to November 2003, he was a Postdoctoral Fellow with the Royal Holloway College, University of London, Surrey, U.K., working with Prof. J. Shawe-Taylor. From April 2005 to June 2005, he was a Visiting Professor with the Department of Computer Science, Hong Kong University of Science and Technology. He is currently a Full Professor with the College of Electrical and Information Engineering, Hunan University. He has authored or coauthored more than 160 refereed papers. His professional interests include compressive sensing, sparse representation, image processing, and pattern recognition.

Dr. Li is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, a member of the Editorial Board of the journal *Information Fusion and the Sensing and Imaging*. He was a recipient of two Second-Grade National Awards at the Science and Technology Progress of China in 2004 and 2006.



**Leyuan Fang** (S'10–M'14) received the B.S. degree in electrical engineering from the Hunan University of Science and Technology, Xiangtan, China, in 2008 and the Ph.D. degree in electrical and information engineering from Hunan University, China, Changsha, China, in 2008.

From September 2011 to September 2012, he was a Visiting Ph.D. Student with the Department of Ophthalmology, Duke University, Durham, NC, USA, supported by the China Scholarship Council. His research interests include sparse representation and multiresolution analysis in remote sensing and medical image processing.

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

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

Currently, he is Pro Rector of Academic Affairs and Professor of electrical and computer engineering with the University of Iceland. His research interests are in remote sensing, image analysis, pattern recognition, biomedical analysis of signals, and signal processing, and he has published extensively in those fields. He is a Cofounder of the biomedical startup company Oxymap.

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