Class-Oriented Spectral Partitioning for Remotely Sensed Hyperspectral Image Classification

Yi Liu, *Student Member, IEEE*, Jun Li, *Member, IEEE*, Peijun Du, *Senior Member, IEEE*, Antonio Plaza, *Fellow, IEEE*, Xiuping Jia, *Senior Member, IEEE*, and Xinchang Zhang

***Abstract*—Remotely sensed hyperspectral images exhibit very high dimensionality in the spectral domain. As opposed to band selection techniques, which extract a subset of the original spectral bands in the image, spectral partitioning (SP) techniques reassign the original bands into subgroups that are then processed separately. From a classification perspective, this strategy has the advantage that all the original information in the hyperspectral data can be retained while addressing the curse of dimensionality given by the Hughes phenomenon. Even if SP prior to classification has been widely used, the strategies adopted to perform such partitioning did not consider the diversity of spectral classes in the scene. In other words, available techniques are not driven by the information contained in the classes of interest, which can be very useful to perform the SP in a more effective manner for classification purposes. To address this issue, in this paper, we present a new class-oriented SP technique that exploits prior information about the classes by automatically ranking the spectral bands that are more useful for each specific class (instead of considering the hyperspectral image as a whole). The resulting multiple subgroups of bands with lower dimensionality are then fed to a multiple classifier system. Our experimental results, conducted with three different hyperspectral airborne images, suggest that the presented method leads to competitive results when compared to other state-of-the-art approaches in the field.**

***Index Terms*—Hyperspectral imaging, multiple classifier systems (MCS), spectral partitioning (SP).**

I. Introduction

HYPERSPECTRAL image analysis has developed significantly during the past two decades [1]. Hyperspectral data have been used in many different areas, including disaster monitoring, natural resources exploitation, environmental applications, etc. [2]–[4]. With the increasing spatial, spectral, and temporal resolutions of imaging spectrometers, the extremely high dimensionality and size of the data have become important concerns for hyperspectral data interpretation [1]. However, contrary to the rapid development of remote sensing technologies, the availability of training samples and labeled data has been quite limited. As a result, many hyperspectral image analysis methods (particularly, for supervised and semisupervised classification) have focused on addressing the existing imbalance between high spectral dimensionality and limited labeled samples [5], [6]. In order to deal with the issue, dimensionality reduction technique has been widely developed and used [7]–[10].

Classic dimensionality reduction techniques can be separated into two categories: 1) band selection (BS) and 2) feature extraction. Considering the diversity of available algorithms and techniques, feature extraction turns out to be more flexible and widely used, including well-known approaches in the literature such as principal component analysis (PCA) [11], independent component analysis [12], manifold learning [13], [14] and subspace-based approaches [15], [16]. However, feature extraction generally transforms the original information after projecting the data into a certain feature space [17], which may be a challenge for certain applications that require meaningful spectral signatures according to their physical interpretation [18], [19]. In turn, BS algorithms are more effective in preserving the original information due to their capacity for selecting the most informative spectral bands among hundreds or even thousands of bands with great correlation and redundancy with supervised or unsupervised ways [20]–[22] while, at the same time, retaining their original physical meaning (i.e., even though some bands are discarded, there is generally no transformation of the retained bands into a different feature space). As a matter of fact, the discriminative information that allows a classifier to provide good performance is usually class-dependent and the relevant information may live in weak features/bands that are usually discarded or lost after subspace transformation or BS. As a result, in practice it is challenging to use either feature extraction or BS for classification purposes.

Many techniques have been developed in order to address this problem. In [23], an efficient model selection procedure based on kernel alignment was developed for hyperspectral image classification. Resulting from this process, a weight (learned

Manuscript received January 27, 2016; revised May 5, 2016 and June 21, 2016; accepted June 28, 2016. Date of publication August 3, 2016; date of current version January 23, 2017. This work was also supported by the Chinese National Science Foundation under Project 41431178, and by Junta de Extremadura, ayudas para la realización de actividades de investigación y desarrollo tecnológico, de divulgación y de transferencia de conocimiento por los Grupos de Investigación de Extremadura, Ref. GR15005. *(Corresponding Author: Jun Li.)*

Y. Liu and A. Plaza are with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Cáceres E-10071, Spain (e-mail: yiliu@unex.es; aplaza@unex.es).

J. Li and X. Zhang are with the School of Geography and Planning and Guangdong Key Laboratory for Urbanization and Geo-Simulation, Sun Yat-sen University, Guangzhou 510275, China (e-mail: lijun48@mail.sysu.edu.cn; eeszxc@mail.sysu.edu.cn).

P. Du is with the Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing 210023, China (e-mail: dupjrs@126.com).

X. Jia is with the School of Engineering and Information Technology, University of New South Wales, Canberra, ACT 2600, Australia (e-mail: x.jia@adfa.edu.au).

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/JSTARS.2016.2588980

from the data) is assigned to each kernel so that both relevant and meaningless image features automatically emerge after training the model. In [24], an unsupervised BS method called multiobjective optimization BS was developed. The objective functions were optimized by a multiobjective evolutionary algorithm to find the best tradeoff solutions. The final selected bands evolve from multiple possible instances of BS during an iterative process, until good stable performance of classification is achieved for different datasets. In order to facilitate the subsequent automatic interpretation image objects, high-level features are jointly integrated to infer the spatial and structural information encoded in the low-level and middle-level features [25]. Classic ensemble learning techniques such as bootstrap aggregation/bagging [26], AdaBoost [27], or cross-validated committees [28] can also be interpreted as feature fusion approaches that can take advantage of the relevant information provided by each classifier [29]–[32].

An alternative strategy used in recent developments is spectral partitioning (SP), which can be seen as a kind of BS approach that aims mainly at rearranging the original spectral bands in the hyperspectral image. However, as opposed to BS, SP does not necessarily discard most of the original spectral bands to achieve lower dimensionality [33]. Instead, SP generates several groups of band subsets from the original spectral bands, so that each band subset is a so-called spectral partition containing a much lower number of spectral bands as compared with the original hyperspectral image, meanwhile the union of multiple subsets can make up to the full original image [34]. Therefore, SP effectively provides multiple views of the original hyperspectral image by obtaining several subgroups that can simultaneously exploit most of the original spectral information in the hyperspectral scene without discarding a large proportion. In other words, different subgroups of spectral bands can be used to provide different classification results [34]. The diversity of classifiers constructed with the subgroups of bands provides the possibility to obtain a very robust classifier ensemble, which can be achieved by combining the classification results obtained from each of the subgroups using different ensemble learning strategies [35], [36], such as bagging and boosting [27], [37], decision combination via majority voting [38], and multiple classifier systems (MCS) [29], [32]. The diversity obtained in the generation of the multiple views is one of the keys for successful SP prior to classification [31], [32], [39], [40].

A. Related Work and Motivations

Even if SP prior to classification has already been used, the strategies adopted to perform such partitioning generally do not consider the information about the spectral classes in the scene. In other words, available techniques are not driven by the information contained in the classes of interest, which can be very useful to perform the SP in a more effective manner for classification purposes [34], [41]. In high spatial resolution images, land-use classification can be successfully performed by interpreting the multiple spectral bands according to common classes in the scene [42]. With such class-oriented interpretation, one can also automatically perform target detection [43], [44]. However, in hyperspectral image analysis, it is generally difficult to anticipate which bands play a more relevant role for classification in a specific scene [45], [46], as this information is generally class-dependent. In [23] and [47], Tuia *et al.* developed multiple kernel frameworks for hyperspectral image classification. The weight assigned to different kernel features was carefully established and interpreted. However, the specific role of spectral bands is more difficult to substantiate since it is often application-dependent. As a result, it is important that different spectral partitions derived from the original image represent different perspectives or views on the data [34], [48]–[50]. At this point, we clarify that these multiple views refer to the lower-dimensional partitions that are obtained by the proposed SP framework.

Bearing in mind the aforementioned issues, our motivations to introduce a new SP technique in this study can be summarized as follows:

1) to achieve the benefits of dimensionality reduction without actually discarding most of the spectral bands in the original image;
2) to facilitate multiple-classifier feature learning by considering different views (spectral partitions) of the original data; and
3) to explore the relevance of class-dependent features, which are often discarded by feature extraction/selection techniques, all in the context of hyperspectral image classification.

With the aforementioned ideas in mind, our main goal in this paper is to develop a new strategy for SP that can effectively exploit the spectral bands that are more relevant to each class of interest in the scene.

The newly presented SP strategy consists of three steps. First of all, we rank the spectral bands that are more useful for each specific class (instead of considering the hyperspectral image as a whole). As a result, multiple subgroups of bands with lower dimensionality are selected in this step. Then, the obtained class-oriented spectral partitions are used as the input of a classifier ensemble strategy, MCS. For this purpose, we use two state-of-the-art classifiers: the support vector machine (SVM) [51] and the multinomial logistic regression (MLR) [52] to validate our presented method. In addition, we incorporate the concept of classification with rejection to evaluate and improve our presented classifier into the MCS [53], [54]. This approach effectively merges the classification output from the spectral partitions with the classification obtained from the original image, providing a robust classification output. The proposed approach is compared to other available SP methods, such as the one presented in [33].

The remainder of this paper is organized as follows. Section II describes our proposed approach. In Section III, we discuss experimental results obtained using three well-known hyperspectral datasets: the reflective optics system imaging spectrometer (ROSIS) Pavia University scene, the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Indian Pines scene, and the hyperspectral digital image collection experiment (HYDICE) Washington DC Mall scene. The experiments suggest that our

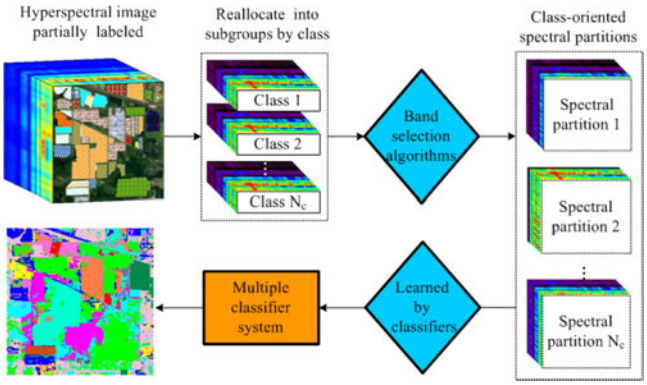

Fig. 1. Flowchart of the proposed class-oriented SP prior to classification approach.

presented method leads to competitive results when compared to other state-of-the-art approaches in the field. Conclusions and hints at plausible future research lines are given in Section IV.

II. Class-Oriented SP Method

A. Proposed SP Strategy

Let us denote by $\mathbf{x} \equiv \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in \mathbb{R}^{d \times n}$ the original spectral signatures of the hyperspectral image $\mathbf{B}^d$ with n pixels indexed by $\mathcal{S}: \{1, 2, \ldots, n\}$ and d wavebands Ω ($|\Omega| = d$). Classification is the process that assigns each pixel $\mathbf{x}_i = \{x_{i1}, x_{i2}, \ldots, x_{id}\}, i \in \{1, \ldots, n\}$ with a label $y_i \in \mathcal{L}: \{1, 2, \ldots, N_c\}$, where N_c is the number of classes of interest in the scene. SP aims at separating and reassigning the d spectral bands Ω into a group of band subsets, namely spectral partitions $\{SP_i\}, i \in \{1, 2, \ldots, M\}$, so that

$$\bigcup_{i \in \Omega} \{SP_i\} \subseteq \Omega. \tag{1}$$

This means that usually the union of the spectral partitions gives the whole set of original spectral bands of the hyperspectral image. We also note that

$$\begin{aligned} \emptyset \subseteq & \bigcap_{\forall i \neq j,\ i,j \in \Omega} \{SP_i, SP_j\}, \\ & \bigcup_{\forall i \neq j,\ i,j \in \Omega} \{SP_i, SP_j\} \supset \{SP_i\} \end{aligned} \tag{2}$$

which means that two different spectral partitions may share mutual bands or not, while any given two spectral partitions should not be equivalent (note that one spectral partition might be the subset of another). Eqs. (1) and (2) generally describe the relationship of spectral partitions, meanwhile they also induce to finite number of solutions of the SP problem. We generally need the spectral partitions to be different enough among each other in order to provide adequate diversity [35], [37]. This brings different perspectives of the original hyperspectral image [33], [34]. If we let $|SP_i|$ be the number of bands in partition SP_i, we can infer from (1) and (2) that $\sum_{i=1}^{M} |SP_i| \geq d$.

With the aforementioned ideas in mind, Fig. 1 shows a general flowchart of our proposed class-oriented SP approach prior to classification. First, training samples with labels $G = \{G_1, G_2, \ldots, G_{N_c}\}$ are separated into a certain number of subgroups G_i, where $G_i = \{\mathbf{x}_{G_i}, \mathbf{y}_{G_i}\}, i \in \{1, \ldots, N_c\}$ and the number of subgroups is given by the number of classes in the scene ($Nc = M$). Then, each subgroup G_i is used as an input to a BS algorithm that selects the most relevant bands using a specific criterion that is driven by the information contained in each class. The result is a set of Nc class-oriented spectral partitions $\{SP_i\} (i \in \{1, \ldots, N_c\})$ with much lower dimensionality as compared to the original hyperspectral image ($|SP_i| << d$). The spectral partitions are fed to an MCS that combines multiple SVM classifiers and generates a final classification result.

Herein, we remark that the presented SP strategy excludes by a large degree the influence of noise, outliers and anomalous classes, and relies on the effectiveness of the BS algorithms that are used in the SP scheme. In the following, we describe the BS techniques adopted in this study and the MCS used to generate the final classification output. In addition, we also introduce the concept of classification with rejection, as well as describe how we combine different classifiers to further improve the classification results using this concept.

B. BS Algorithms

BS techniques intend to select an appropriate band subset from the original dataset to represent the data according to some optimality criterion [55]. Generally, BS can be understood as an exhaustive searching process for all possible cases: $(L_{|\Omega_{\text{BS}}|}) = \frac{d!}{(d-|\Omega_{\text{BS}}|)!|\Omega_{\text{BS}}|!}$, with Ω_{BS} being the selected bands and L being the number of all possible subset of selected bands, given a number of bands to be selected $|\Omega_{\text{BS}}|$ [55]. A general way to perform the searching process is to solve the following optimization problem:

$$\Omega^*_{\text{BS}} = \arg \max_{\Omega_{\text{BS}} \subset \Omega,\ |\Omega_{\text{BS}}| = n_{\text{BS}}} J(\Omega_{\text{BS}}) \tag{3}$$

where $n_{\text{BS}} = |\Omega_{\text{BS}}|$ is the number of selected bands in subset Ω_{BS} and $J(\Omega_{\text{BS}})$ establishes the relative importance of a given spectral band in Ω_{BS}. We have to note that, given the large number of spectral bands in a hyperspectral image, it is almost impossible to try all possible band combinations. We also note that, for different scenes, it is difficult to decide which spectral bands play a more relevant role, or to anticipate which combination is more useful. As a result, a general strategy has been to find a group of bands that are both of high quality (e.g., with low noise or not located in the water absorption region) while exhibiting high variance [46]. Consequentially, many available BS algorithms are focused on defining target functions that calculate the priority score of a given spectral band [56]–[58].

In this study, we rely on two popular BS algorithms: signal-to-noise ratio (SNR) [59] and band dependence minimization-based linearly constrained minimum variance (BDM-LCMV) [60] for evaluation purposes. We have selected these algorithms because their strategies for BS are widely used, which makes the algorithms quite representative of existing BS techniques. In this study, both algorithms are used to perform class-oriented BS prior to classification. In the case of the SNR BS algorithm, $J(\Omega_{\text{BS}})$ is defined with a noise-adjusted principal

component-based priority score, calculated for the lth band of the image, $\mathbf{B}_l$, as follows:

$$J(\mathbf{B}_l) = \sum_{k=1}^{d} \mathbf{r}_{l,k}^2 = \rho_l^{\text{SNR}} \tag{4}$$

where $\mathbf{r}_{l,k}^2 = \sqrt{\lambda_{l,k}} \times \mathbf{v}_{l,k}$, for $l, k = 1, 2, \ldots, d$ and $\mathbf{B}_l$ is the lth spectral data. Herein $\lambda_{l,k}$ is the set of eigenvalues of the noise-adjusted covariance matrix and $\mathbf{v}_{l,k}$ denotes their associated orthonormal eigenvectors. Here, we also use d to denote the number of wavelengths of the original hyperspectral image. Given a number of bands to be selected, Ω_{SB}, the SNR algorithm selects the first Ω_{SB} spectral bands with greater $J(\mathbf{B}_l)$ values from all the available bands. At this point, we emphasize that the SNR is one of the most widely used criteria to establish band priorities. The main assumption of our SP method based on SNR (hereinafter SP-SNR) is that the spectral bands with higher SNR lead to better classification performance. This simple criterion is based on minimizing the impact of noise in the performance of the classifier. Consequentially, the SP-SNR is intended to minimize the impact of noise in classification performance by generating class-oriented spectral partitions with high SNR values.

SNR is a criterion related to single bands, while BDM-LCMV also includes a second constraint intended to minimize the similarity between the bands to be selected. From the viewpoint of classification, BDM-LCMV follows the assumption that a more informative band combination is generally given by a set of bands that are more distinct between each other. In the case of BDM-LCMV algorithm, the band priority score can be calculated as follows:

$$J(\mathbf{B}_l) = \left(\tilde{\mathbf{v}}_l^{\text{LCMV-CBS}}\right)^T \widetilde{\sum}^{-1} \tilde{\mathbf{v}}_l^{\text{LCMV-CBS}} \tag{5}$$

where

$$\tilde{\mathbf{v}}_l^{\text{LCMV-CBS}} = \widetilde{\sum}^{-1} \mathbf{B}_l \left(\mathbf{B}_l^T \widetilde{\sum}^{-1} \mathbf{B}_l\right)^{-1} \mathbf{1}_N. \tag{6}$$

Here, $\tilde{\mathbf{v}}_l^{\text{LCMV-CBS}}$ represents the weight vectors associated to each spectral band and $\widetilde{\sum} = (1/d) \sum_{l=1}^{d} \mathbf{B}_l \times \mathbf{B}_l^T$ denotes the sample band correlation matrix, in which the problem described is referred to as LCMV-based constrained BS [60]. In (6), $\mathbf{1}_N$ is an N-dimensional column vector with all 1's in its N components and N denotes the number of columns of the spectral band $\mathbf{B}_l$. As a popular BS algorithm, BDM-LCMV is also utilized in this study due to its capacity to select informative band combinations containing bands that are distinct between each other. With this issue in mind, we hold a reasonable expectation that BDM-LCMV may outperform SNR in terms of classification accuracies.

An important advantage of supervised BS techniques such as those described before is that they can reduce the interference introduced by noise, outliers, and unwanted pixels in a hyperspectral image scene. Another important advantage is that they are driven by the information contained in the classes of interest, turning the BS process into a class-oriented one. An additional advantage in the case of BDM-LCMV is that it generally decreases the size of the used samples, which makes the inverse calculation faster and more accurate.

C. Classifier Ensemble Strategy

In this section, we describe the MCS system used to provide the final classification result from the Nc partitions of the original hyperspectral image, obtained by our newly presented strategy for class-oriented SP. The base classifier used for demonstration purposes in this study is the SVM [61], which has been a state-of-the-art classifier for remotely sensed hyperspectral data. Since we are dealing with different partitions (or views) of the original hyperspectral data, we need a decision rule to fuse the individual classifications obtained by the SVM from the different spectral partitions derived by our method. Let $\mathbf{p}_m(i)$ be the probability obtained by an SVM classifier for a given pixel i and partition m. $\mathbf{p}_m(i)$ provides the degree of membership of a pixel to different classes of interest. In this study, we use a soft ensemble strategy to combine the results obtained from all the partitions. Specifically, the probabilities resulting from all the different partitions in a given pixel are modeled by

$$\widehat{\mathbf{p}}(i) = \frac{1}{N_c} \sum_{m=1}^{N_c} \mathbf{p}_m(i) \tag{7}$$

where m indexes the spectral partitions ($m \in \{1, \ldots, N_c\}$) and Nc is the number of partitions, which is equal to the number of classes, according to our interpretation in Fig. 1. The final class label for pixel i is determined by maximum probabilistic voting as follows:

$$y_i = \arg \max_{k \in \{1, \ldots, N_c\}} \widehat{p}^{(k)}(i) \tag{8}$$

where $\widehat{p}^{(k)}(i)$ is the probability corresponding to class k for a given pixel i, and $\widehat{\mathbf{p}}(i) = \{\widehat{p}^{(1)}(i), \ldots, \widehat{p}^{(N_c)}(i)\}$.

D. Classification With Rejection

In order to evaluate a set of classification outcome, four statistics are widely used: overall (OA), average (AA), and individual class accuracies, as well as the κ statistic. However, these statistics are unable to describe how confident a classifier is in the classification of a certain pixel. With limited training samples, the parametric optimization of a classifier sometimes results in high individual accuracies only for large classes (i.e., those with a significant number of pixels in the scene). However, the need for correctly classified pixels usually surpasses the need for high overall accuracies [53]. Hence, in this study, we exploit two additional statistical metrics to further refine the classification results provided by our proposed method.

Assuming that $\mathcal{R}$ is the set of rejected samples ($\bar{\mathcal{R}}$ is the set of nonrejected samples) and $\mathcal{C}$ is the set of correctly classified samples ($\bar{\mathcal{C}}$ is the set of incorrectly classified samples), we consider the *nonrejected accuracy* $\mathcal{A}$ and the *classification quality* $\mathcal{Q}$ [54], which are given as follows:

$$\mathcal{A} = \frac{|\mathcal{C} \bigcap \bar{\mathcal{R}}|}{|\bar{\mathcal{R}}|}, \tag{9}$$

and

$$\mathcal{Q} = \frac{|\mathcal{C} \bigcap \bar{\mathcal{R}} + \bar{\mathcal{C}} \bigcap \mathcal{R}|}{|\bar{\mathcal{R}} \bigcup \mathcal{R}|}. \tag{10}$$

In this study, we adopt a nonparametric way to simplify the rejection process when calculating the statistics $\mathcal{A}$ and $\mathcal{Q}$ as a function of the rejected fraction, following previous developments in [54]. By sorting the maximum probabilities of all pixels' probabilistic output, we can easily decide the rejections by accepting pixels with higher maximum probabilities and rejecting the remaining pixels. When different classifiers exhibit different confidence in the classification of the same pixels, this means that a rejected pixel by one classifier can held higher confidence by another classifier. This makes it possible to further improve the classification results by combining the pixels with confident classifications by different techniques. In order to obtain a combined classification $\widehat{\mathbf{Y}}$, we define the acceptance of a classifier m as

$$\mathcal{S}_{\text{acc}} = \text{where}(\mathbf{p}_m(i) > \lambda) \tag{11}$$

where λ is a constant decided by how many pixels we would like to reject. Then, the rejection set from classifier n can be defined as

$$\mathcal{S}_{\text{rej}} = \text{where}(\mathbf{p}_n(i) < \lambda). \tag{12}$$

Instead of using lambda, the rejection fraction (ratio of rejection) was recommended by the works [53], [54]. The main reason is that it is easier to compare different methods under the rejection fraction. As a result, in our study, we used the ratio of rejection to illustrate the obtained results. As for the selection of the rejection fraction, it is generally defined by the user according to a desired accuracy after rejection. With the above formulation in mind, two classifiers can be simply combined as follows:

$$\widehat{\mathbf{Y}} = \arg\max \ \mathbf{p}_m(\mathcal{S}_{\text{acc}}) \bigcup \mathbf{p}_n(\mathcal{S}_{\text{rej}}) \tag{13}$$

where $\widehat{\mathbf{Y}}$ is the final classification map obtained after combining the output provided by the two classifiers. In practice, we just need to select a value for the rejection fraction when combining two classification results. To do this, we search the whole solution space of rejection fractions [0.01, 0.99] with interval of 0.01 and, then, select the combination that reaches the highest overall accuracy. In this study, we use this simple concept to effectively merge the classification output of the aforementioned SP strategy with the classification obtained from the original image, thus exploiting both strategies (i.e., class-dependent and image-dependent) to provide our final classification output.

III. Experimental Results

In this section, we evaluate the presented SP method by using three well-known hyperspectral datasets: a ROSIS dataset collected over Pavia University, Italy, an AVIRIS dataset collected over the Indian Pines region, and a HYDICE dataset collected over Washington DC Mall. We first describe the three hyperspectral datasets. Then, we discuss the results obtained by using the proposed classification framework based on SP, using the aforementioned datasets.

A. Hyperspectral Images Used in Experiments

The first hyperspectral image used in our experiments [see Fig. 2(a)] was collected by ROSIS over the University of Pavia, Italy. The dataset consists of 115 spectral bands between 0.4 and 1.0 μm, with size of 610 $\times$ 340 pixels. The noisy bands had been removed, yielding 103 spectral bands. The ground-truth image in Fig. 2(b) contains nine ground-truth classes, 3921 fixed training samples distributed together with the data, and 40 002 test samples [62], [63].

The second hyperspectral image used in our experiment is the well-known AVIRIS Indian Pines[1] dataset, displayed in Fig. 2(d). It comprises 145 $\times$ 145 pixels and was collected over Northwestern Indiana in June 1992. As shown by Fig. 2(e), a total of 10 366 pixels are available in the labeled ground-truth, including 16 mutually exclusive classes. In the following experiments with this dataset, we use 640 randomly selected training samples in total to illustrate the performance of the method using limited training sets. We also conduct 20 Monte Carlo runs to ensure statistical significance. Although some of the bands are considered to be corrupted by water absorption features and noise, we will use all of them since the considered BS algorithms have the capacity to automatically screen and select the most useful spectral bands.

The third hyperspectral image used in experiments was collected by HYDICE over the Washington DC Mall. It comprises 191 bands with 3-m spatial resolution. This image originally contains 1280 lines with 307 pixels in each line. The test image used in experiments and its ground-truth are shown in Fig. 2(g) and (h), respectively. The image is available online.[2] In our experiments, we use 26 981 labeled samples collected from seven ground-truth classes.

For the three datasets, we plot the average spectral signatures of each class in Fig. 2(c), (f), and (i) to give an idea of the spectral similarity of the classes. Using the aforementioned hyperspectral datasets, we statistically evaluate the sensitivity of our methods to different percentages of training samples after 20 Monte Carlo runs. The processing times are also reported and discussed. All our experiments have been conducted by using MATLAB R2013a in a desktop PC equipped with an Intel Core i7 CPU (at 3.6 GHz) and 32 GB of RAM.

B. Experiments With Real Hyperspectral Data

Before reporting our experimental results, we emphasize that we have optimized the parameter settings in order to obtain the best performance from each individual method involved in the classification framework. The BS algorithms used in this study have a single input parameter that is the number of spectral bands to be selected. We use the same number of bands for all individual spectral partitions in the framework of the presented method for simplicity. In order to decide this number, we need

[1] https: //engineering.purdue.edu/biehl/MultiSpec/hyperspectral.html
[2] https://engineering.purdue.edu/ landgreb/Hyperspectral.Ex.html

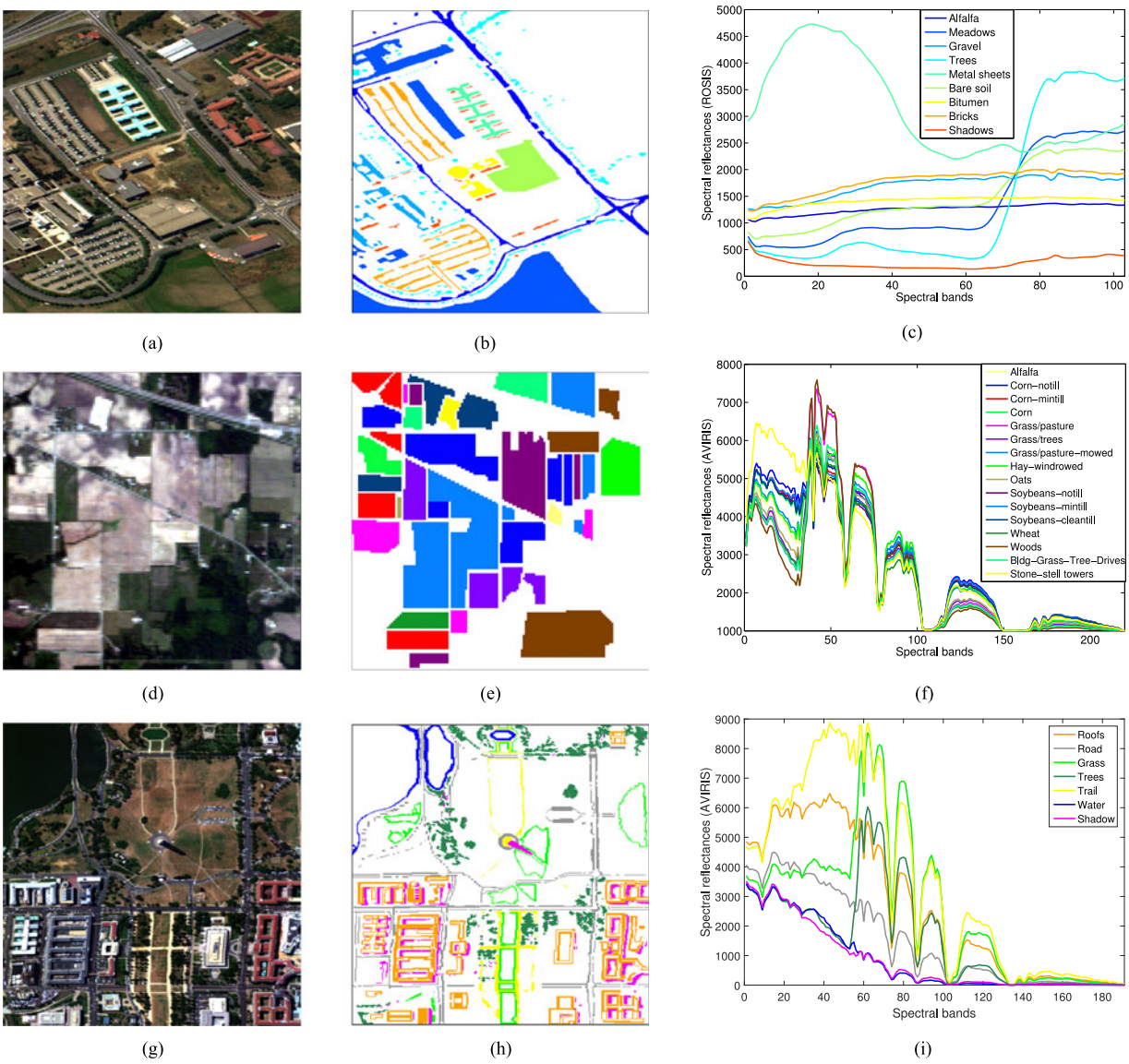

Fig. 2. Experimental hyperspectral datasets along with their ground-truth and average spectral signatures per class. (a) ROSIS Pavia University dataset. (b) Ground-truth. (c) Spectral signatures of the ROSIS data. (d) AVIRIS Indian Pines dataset. (e) Ground-truth. (f) Spectral signatures of the AVIRIS data. (g) HYDICE Washington DC Mall data set. (h) Ground-truth. (i) Spectral signatures of the Hydice data.

to bear in mind that the number of spectral bands should be enough to provide discriminative spectral details with less band overlapping among the partitions, to preserve the diversity of the views provided. After several empirical experiments, we decided to use 20 bands for the ROSIS Pavia University dataset, 60 bands for the AVIRIS Indian Pines dataset, and 40 bands for the HYDICE Washington DC Mall dataset. For the two considered classifiers: SVM and MLR, we used the Gaussian radial basis function kernel. This has been one of the most successful and widely used kernel functions in the literature [64], [65]. In our experiments, we empirically found that this kernel provides superior performance to other kernels. For the MLR classifier, the logistic regressors (assumed to be random vectors with independent Laplacian components) are learnt using the LORSAL algorithm [52].

1) Experiments With ROSIS Pavia University Data: In this section, our presented SP method is first tested with the ROSIS Pavia University data. Two well-established BS algorithms, SNR and BDM-LCMV, are used to construct two SP methods called SP-SNR and SP-BDM-LCMV, respectively. Our method is compared with three other strategies based on:

1) Conducting classification on the original hyperspectral image.
2) Conducting classification on the band subsets selected by the same two BS algorithms.
3) Conducting classification after applying the recently developed adaptive affinity propagation based SP method (SP-AAP) in [33].

Figs. 3(a) and 3(b), respectively, illustrate the multiple groups of selected bands obtained using the SNR and the BDM-

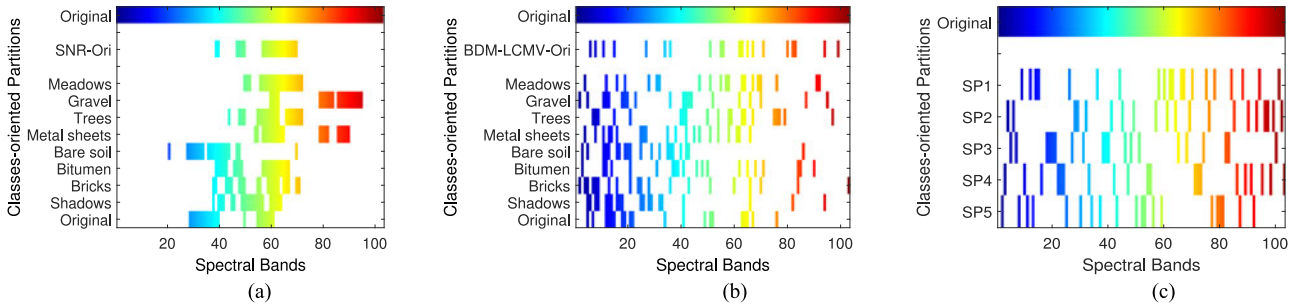

Fig. 3. Spectral partitions obtained by our presented SP method from the ROSIS Pavia University data. Our method incorporates two BS algorithms: SNR (a) and BDM-LCMV (b), respectively. The spectral partitions obtained by SP-AAP are displayed in (c). In all the plots, the x-axis denotes the set of original spectral bands, while the y-axis represents the group of selected bands (each horizontal line displays one group of selected bands).

LCMV criteria, while Fig. 3(c) shows the bands in the spectral partitions obtained after using SP-AAP. In all the plots, the x-axis denotes the set of original spectral bands while the y-axis represents the selected bands for each class. For example, from up to down along the y-axis, the first horizontal line of the plots shows all the original spectral bands of the ROSIS data and the second horizontal line (SNR-Ori) shows the selected bands by the SNR algorithm with all labeled training samples as input. Then, from the third line on, the figure lists the class-oriented spectral partitions (multiple views) consisting of the selected bands that correspond to the different classes in the scene, e.g., *Meadows*, *Gravel*, *trees*, and so forth. In the case of SP-AAP, five SPs are identified in Fig. 3(c).

Several observations can be made for the plots displayed in Fig. 3. First and foremost, in contrast with the original BS algorithms, our presented method is able to make use of more of the original spectral bands via different multiple partitions of selected bands. The union of the spectral partitioned bands gives a much larger band set in comparison to applying a BS algorithm over the original image. Another important observation is that there is great diversity in the selected bands among the spectral partitions that are obtained by using our presented SP method. We remark here that possible reason is that most of the classes in the Pavia University scene are dissimilar resulting in different BSs. And this is crucial for generating different views for the classification ensemble process [31], [40]. We also observed that the union of the obtained partitions does not give the full original hyperspectral image. This is expected, as the original hyperspectral image exhibits high redundancy and not all the bands may be useful for classification purposes. Quite opposite, the SP-AAP method generates more equally sampled spectral bands in each partition. In this case, we are still considering the full spectral information in the scene but voluntarily deciding not to use all of the bands for the construction of the multiple views. Finally, we note that the two considered BS algorithms obtained different spectral partitions, which will lead to different classification results. The key of our approach is to exploit these partitions synergistically by means of the presented MCS. In the following, we discuss the experimental results acquired after using the obtained spectral partitions for classification purposes. In order to provide a fair comparison with the SP-AAP, we set its number of spectral partitions to five with the ultimate goal of obtaining the same number of spectral bands in each partition. This is mainly because the number of bands of each spectral partition has a strong impact on the final classification performance.

Table I shows the classification results obtained by the different tested methods for the ROSIS Pavia University data. When the SNR is used for BS prior to classification, the results remain almost the same, while 1.55% increase in OA is observed when the BDM-LCMV BS algorithm is used, as compared to using the original spectral information. In comparison with the single BS strategy, the classification results are improved by the considered multiple classifier-based feature learning strategy based on multiple views provided by different spectral partitions. On the other hand, our SP method leads to improvements in classification accuracy. For example, in the case of using BDM-LCMV for SP, the OA improved by 5.33% for the SVM classifier and by 5.07% for the MLR classifier, respectively. When the SNR is used for SP, the observed increase in OA amounts to 1.64% (SVM) and 1.74% (MLR). This is consistent with the observations in Fig. 3(a) and (b), where we can see that the BDM-LCMV obtained more diversity in the spectral partitions than the SNR criterion. Also, our BDM-LCMV obtained better results than the SP-AAP with 2.09% (SVM) and 0.13% (MLR) increase in OA, respectively.

In order to further evaluate the potential of our presented method, Fig. 4 displays the classification maps after rejection of some unconfident pixels to obtain a 90% accuracy of $\mathcal{A}$ [see (9)] of the remaining pixels. It is remarkable that 96% pixels of the classification obtained using the original spectral information have to be discarded to get a more confident result of 90% accuracy, while our presented method (using the SVM classifier) only needs to discard 39% (SP-SNR) and 20% (SP-BDM-LCMV) of the pixels. It is also remarkable that the presented SP method shows great advantage over standard BS algorithms in the case of the ROSIS Pavia University data. Similar observations can be made for the MLR classifier. For illustrative purposes, Fig. 5 plots the nonrejected accuracy ($\mathcal{A}$) as well as the classification quality ($\mathcal{Q}$) as a function of the rejected fractions. From Fig. 5, we can also see that our presented method has more biased confidence toward the individual pixels to label, in comparison with the original spectral information. As a matter of fact, the partitions obtained by the proposed method appear to be more

TABLE I
OVERALL, AVERAGE AND INDIVIDUAL CLASS ACCURACIES [%] AND κ STATISTIC OBTAINED BY THE PRESENTED CLASSIFICATION FRAMEWORK IMPLEMENTED USING THE SVM AND MLR WITH THE BS ALGORITHMS: SNR AND BDM-LCMV, FOR THE ROSIS PAVIA UNIVERSITY SCENE

	SVM	BS			SP			BS + Original		SP + Original		
Class	Train/test	All	SNR	BDM-LCMV	SNR	BDM-LCMV	AAP	SNR	BDM	SNR	BDM-LCMV	AAP
Alfalfa	548/6304	84.63	83.98	83.57	77.92	84.76	83.55	84.58	84.66	84.20	84.33	84.41
Meadows	540/18146	66.04	66.04	70.69	74.80	78.61	76.32	66.85	71.45	77.23	78.89	77.36
Gravel	392/1815	73.39	68.54	67.99	56.36	69.31	62.75	71.40	70.41	72.89	72.07	67.71
Trees	524/2912	97.49	98.18	98.11	97.32	97.97	94.06	98.45	98.15	97.80	98.18	96.67
Metal sheets	265/1113	99.46	99.37	99.28	99.28	99.46	99.37	99.46	99.46	99.46	99.46	99.46
Bare soil	53/4572	94.27	92.94	92.17	90.44	91.69	90.66	94.07	93.83	93.83	92.67	92.72
Bitumen	375/981	91.13	86.54	89.91	85.32	89.50	87.36	89.09	91.34	91.34	91.03	89.91
Bricks	514/3364	92.12	91.91	93.25	86.92	93.40	91.83	92.69	93.16	92.21	92.78	92.72
Shadows	231/795	99.37	97.74	98.74	95.85	98.87	99.75	99.25	99.12	99.25	99.50	99.62
Overall accuracy		79.22	78.63	80.77	80.26	84.55	82.46	79.53	81.63	84.19	84.86	83.85
Average accuracy		88.66	87.25	88.19	84.91	89.29	87.29	88.43	89.06	89.80	89.88	88.95
κ statistic		73.82	73.12	75.57	74.71	80.05	77.40	74.21	76.63	79.65	80.45	79.18
	MLR	BS			SP			SP + Original		BS + Original		
Class	Train/test	All	SNR	BDM-LCMV	SNR	BDM-LCMV	AAP	SNR	BDM	SNR	BDM-LCMV	AAP
Alfalfa	548/6304	77.27	71.39	79.52	75.97	80.50	82.73	80.08	80.93	78.39	81.99	82.73
Meadows	540/18146	75.79	74.14	80.31	83.03	84.45	82.55	81.97	82.88	83.33	84.85	82.55
Gravel	392/1815	65.84	64.81	78.27	75.25	75.68	81.60	78.83	78.95	78.40	81.17	81.60
Trees	524/2912	93.68	86.68	92.31	92.60	91.65	94.70	91.06	92.93	91.94	91.98	94.70
Metal sheets	265/1113	99.46	99.35	99.46	98.80	99.78	99.13	99.35	99.35	98.80	99.78	99.13
Bare soil	53/4572	84.38	78.11	79.46	80.30	87.11	88.83	83.68	84.28	81.38	87.36	88.83
Bitumen	375/981	85.63	85.88	91.22	92.11	93.77	93.26	89.06	91.98	91.98	93.64	93.26
Bricks	514/3364	86.21	72.48	73.65	69.42	83.43	80.12	71.92	75.54	70.97	75.70	80.12
Shadows	231/795	97.48	100	100	100	99.83	100	100	100	100	99.83	100
Overall accuracy		80.06	76.01	81.50	81.80	85.13	85.00	82.56	83.69	82.67	85.20	85.00
Average accuracy		85.08	81.43	86.02	85.28	88.47	89.21	86.22	87.43	86.13	88.48	89.21
κ statistic		74.42	68.48	75.55	75.70	79.97	79.86	76.64	78.18	76.78	80.04	79.68

The SP-AAP SP method is also included for comparison. Here, use the accepted pixel subset of the spectral partition approaches while using the rejected pixel subset of the original classifier. The results obtained using the original spectral information and the spectral bands selected by using the SNR and BDM-LCMV are also included. In all cases, 3921 training samples have been used.

informative than those obtained using traditional BS algorithms. This opens the door to combining several classifiers when they have different preferred pixel subsets in terms of confidence, especially when both subsets share a limited common set of pixels. In this case, the SP-AAP obtained slightly lower but competitive performance when compared to our proposed methods, SP-SNR and SP-BDM-LCMV.

With the aforementioned observations in mind, we try to further improve the classification results of our presented method by using (13). Since more confidence exists in the results of our presented method, we use the accepted pixel subset of SP-SNR (or SP-BDM-LCMV) while using the rejected pixel subset of the classifier applied to the original hyperspectral image. In this case, we are effectively combining the results obtained using the original spectral information and the spectral partitions derived by our proposed approach. The results are shown in Table I. In other words, the results in this table have been optimized to find the best combination of classifiers bearing in mind the rejection fractions. The effect of this combination on the OAs% is shown in Fig. 6, as a function of the rejection fractions. Remarkably, the classification result of SP-SNR is improved by about 4% (from 80.26% to 84.19%) when using the SVM and by 6.66% (from 76.01% to 82.56%) when using the MLR.

In order to evaluate the sensitivity of our methods to different numbers of training samples, Fig. 7 plots the OA achieved by different methods as a function of the percentage of training samples used to train the classifier. From Fig. 7, it is remarkable that the proposed framework (implemented with SP-BDM-LCMV) provides better results when compared with the other methods. When the number of training samples increases, the performance also increases, implying robustness with respect to different training sample sets. Also, the SP-SNR obtained results which are close to those obtained by the SP-AAP method, even though the single SNR BS method performed comparatively worst among all methods. On the other hand, we also calculated the processing times in the considered computing environment (see Table II). As shown by Table II, our presented SP methods are time consuming as their computational time is approximately seven to ten times that of using the original spectral bands of the hyperspectral image. However, the presented SP framework performs very fast in the task of obtaining the spectral partitions, as it only considers the labeled samples as input. Also, the spectral partitions are independently learnt by the classifiers. As a result, it is very feasible to perform the classification step in parallel by resorting to multicore architectures of exploiting coprocessors such as graphics computing units. With these implementation enhancements, the complexity of

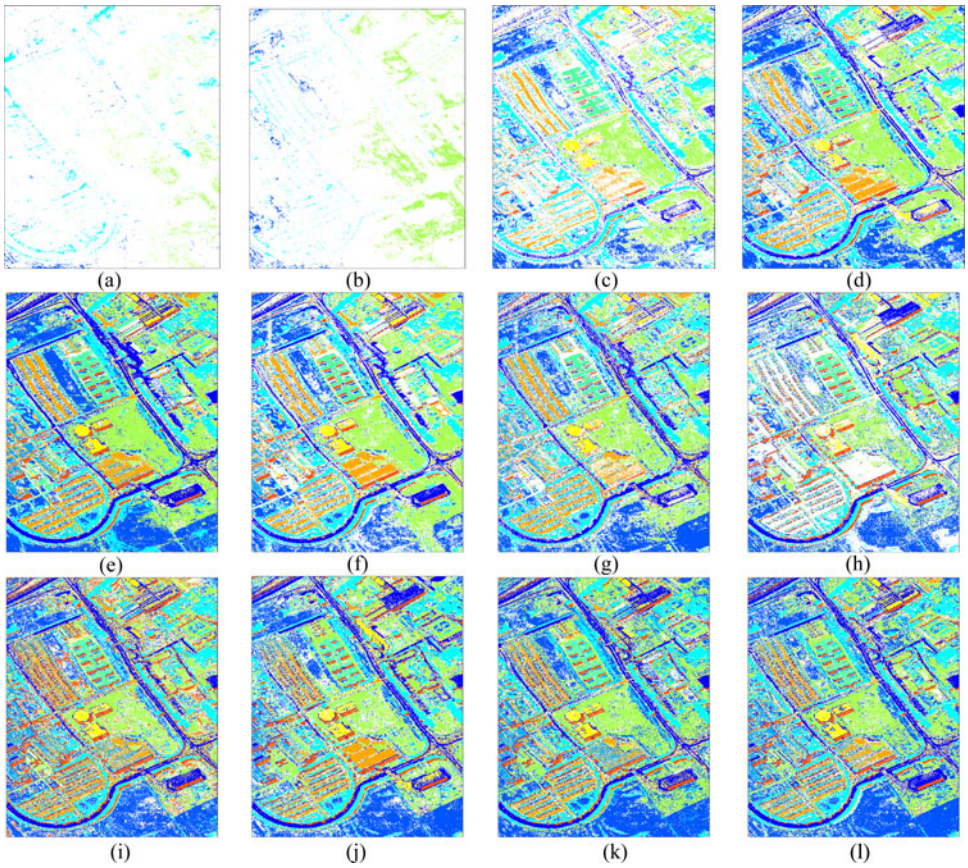

Fig. 4. Classification maps (with rejection) obtained by the proposed classification framework, using the original spectral information (a), (g), with the 20 selected bands by using SNR (b), (h) and BDM-LCMV (c), (i) algorithms, and with the spectral partitions obtained by our proposed SP approach implemented with SNR (d), (h) and BDM-LCMV (e), (k) and the SP-AAP method (f), (l) for the ROSIS Pavia University data. The percentage in the parenthesis denotes the proportion of pixels remaining after rejection. Maps (a)–(f) are obtained with the SVM classifier, while (g)–(l) are obtained with the MLR classifier. In all cases, a total of 3921 randomly selected training samples were used for training and the rest used for testing. The maps are displayed with partial pixels rejected in order to obtain a 90% classification accuracy for the remaining ones. Note that the overall accuracies for the unrejected pixels of each map are calculated by considering the labeled test samples, with the training samples excluded. (a) SVM (4%). (b) SVM-SNR (0%). (c) SVM-BDM-LCMV (41%). (d) SVM-SP-SNR (61%). (e) SVM-SP-BDM-LCMV (80%). (f) SVM-SP-AAP (65%). (g) MLR (65%). (h) MLR-SNR (52%). (i) MLR-BDM-LCMV (77%). (j) MLR-SP-SNR (78%). (k) MLR-SP-BDM-LCMV (83%). (l) MLR-SP-AAP (81%).

the method can be kept within similar levels as the classification of the original image.

2) Experiments With AVIRIS Indian Pines Data: In this section, we report the experimental results obtained by our SP method with the AVIRIS Indian Pines data. The same BS algorithms: SNR and BDM-LCMV are also used in this case to generate spectral partitions, which are respectively referred to as SP-SNR and SP-BDM-LCMV. For simplicity, we conduct the experiments using the same conditions for all groups. Considering the fact that most classes of the AVIRIS Indian Pines scene are given by vegetation features with spectrally similar signatures, a larger number of spectral bands is needed in order to retain enough spectral details to distinguish different classes. On the other hand, increasing the number of bands may be detrimental to the diversity for the MCS, since a larger number of bands in the partitions means higher partition overlaps [48]. Consequentially, and for simplicity purposes, we empirically set the number of selected bands to 60 for our SP methods as well as for the SP-AAP method. First of all, we perform the experiments by using 640 randomly selected labeled samples for both the SVM [61] and MLR-LORSAL classifiers [52]. Each group of results is obtained after 20 Monte Carlo runs. The average values and standard variations are reported. Then, we also reported the performance of our presented methods with regards to different percentages of labeled training samples, along with the corresponding processing times.

Fig. 8 displays the spectral bands selected by the SNR and BDM-LCMV algorithms, as well as the spectral partitions obtained by using our SP-SNR and SP-BDM-LCMV methods and the SP-AAP method. In the case of SP-SNR, it can be seen that each spectral partition consists of more diverse bands as compared to using SNR with all the training samples. It is also interesting to note that noisy and water absorption bands have low probability to be selected by SP-SNR and SP-BDM-LCMV, even though we used all of the original bands as input to the algorithms in order to allow them to select the most informative bands automatically. The SP-AAP SP method is not able to detect and screen the noisy and water absorption bands in this case. Finally, we can also see from Fig. 8 that the spectral partitions obtained exhibit diverse BSs

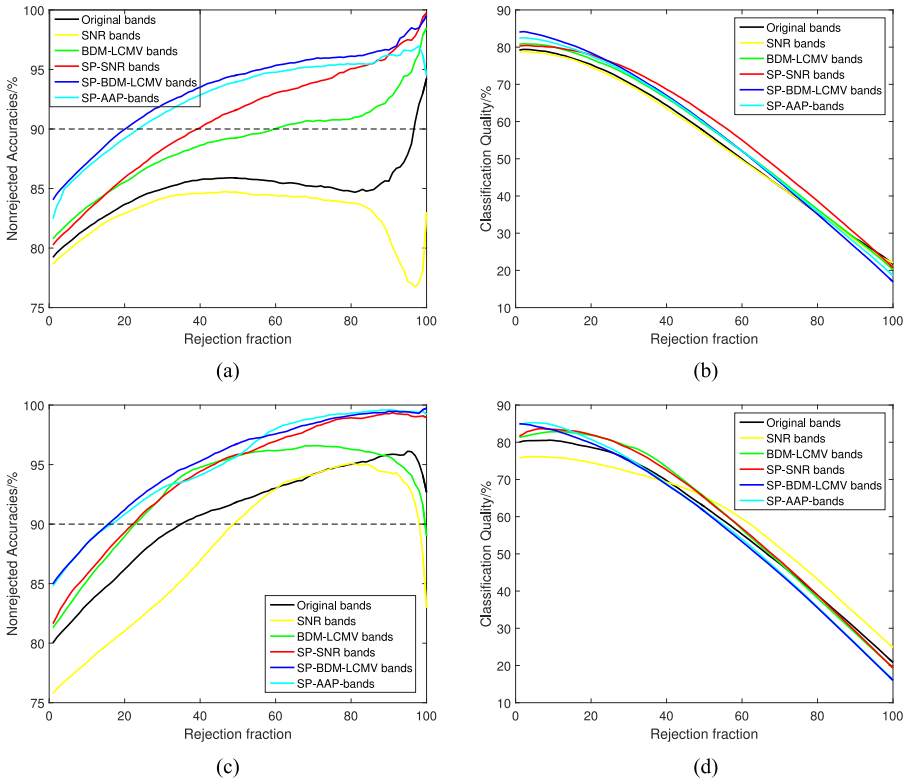

Fig. 5. Nonrejected accuracies ($\mathcal{A}$) and classification qualities ($\mathcal{Q}$) as a function of rejected fractions with the ROSIS Pavia University data set. These plots correspond to the results in Table I. (a) Nonrejected accuracies $\mathcal{A}$ (SVM). (b) Classification qualities $\mathcal{Q}$ (SVM). (c) Nonrejected accuracies $\mathcal{A}$ (MLR). (d) Classification qualities $\mathcal{Q}$ (MLR).

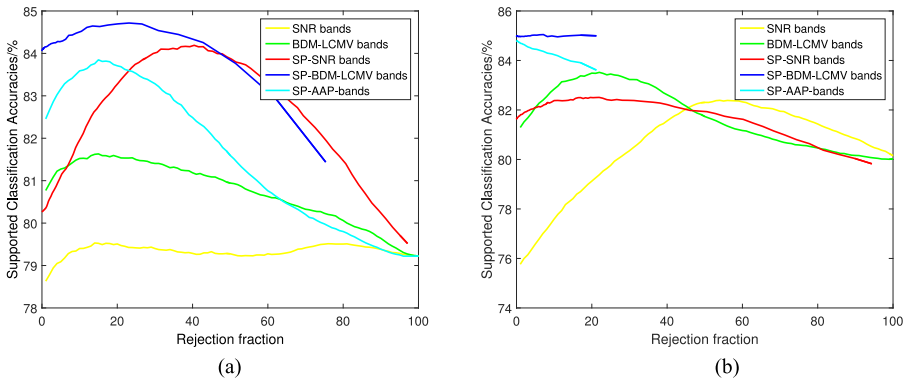

Fig. 6. Classification OAs (as a function of rejections) obtained by our proposed classification framework with the ROSIS Pavia University data, after being supported by the rejected pixels of the classifier using the original spectral information. (a) SVM. (b) MLR.

but with similar patterns across the classes. It is reasonable that the BSs share a degree of similarity considering the fact that the classes in the scene are mostly related to vegetation features [see also Fig. 2(f)].

Tables III and IV show the OA, AA, κ, and individual classification accuracies (obtained by the SVM and MLR, respectively) after using 640 randomly selected pixels (about 40 pixels per class on average) for training and the rest of the labeled samples for testing. If the number of samples available in the ground-truth image is less than 20, we take half of the total samples in that class for training and the other half for testing. Several observations can be drawn from the results reported in Tables III and IV. First of all, the presented approach leads to an increase of about 1% to 4% in classification accuracy (regardless of the BS method used) in comparison with using all the spectral bands in the original image. Although the advantage seems moderate, the individual accuracies of 13 out of 16 classes (with the SVM) and 12 out of 16 classes (with the MLR) in the scene are improved by our presented SP method. Second, when compared with standard BS (SNR and BDM-LCMV) on the original image, our presented method can further improve the classification accuracy by about 5% by providing multiple views from the original

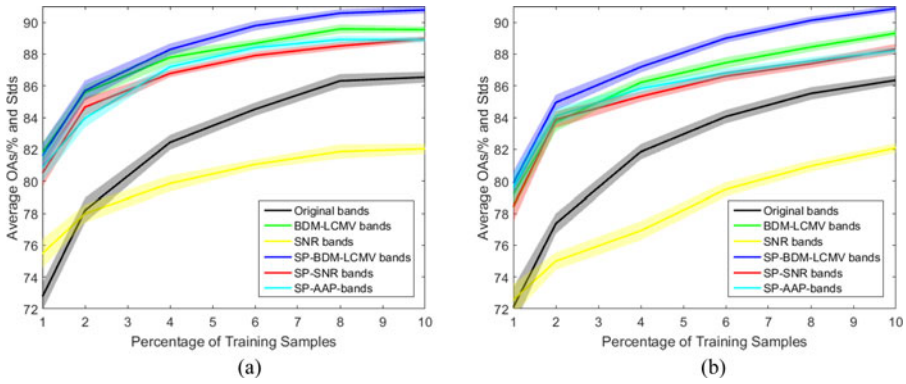

Fig. 7. Overall classification accuracies (as a function of the number of training samples) obtained by the proposed classification framework (with the original spectral information and with SP-SNR, SP-BDM-LCMV, and SP-AAP) for the ROSIS Pavia University scene. The solid lines represent the average of 20 Monte Carlo runs, whereas the colored area around the lines represent the standard deviation around the mean. Plots (a) is obtained using the SVM classifier and (b) is obtained using the MLR classifier. (a) SVM. (b) MLR.

TABLE II
PROCESSING TIMES OF DIFFERENT METHODS FOR ROSIS PAVIA UNIVERSITY SCENE

	SVM (different percentages of training samples [%])					
Time/sec	1	2	4	6	8	10
Original	1.66 ± 0.16	2.66 ± 0.32	4.65 ± 0.34	6.60 ± 0.35	8.62 ± 0.58	10.58 ± 0.6
SNR	1.48 ± 0.13	2.34 ± 0.30	4.15 ± 0.32	5.89 ± 0.34	7.61 ± 0.56	9.58 ± 0.59
BDM-LCMV	1.36 ± 0.15	2.25 ± 0.31	3.88 ± 0.32	5.28 ± 0.34	7.11 ± 0.55	9.07 ± 0.58
SP-SNR	9.99 ± 0.45	15.27 ± 0.64	27.25 ± 0.92	40.99 ± 1.48	55.80 ± 1.69	71.00 ± 2.63
SP-BDM-LCMV	8.49 ± 0.54	12.81 ± 0.72	23.75 ± 0.84	36.64 ± 1.46	51.21 ± 2.02	66.17 ± 2.46
SP-AAP	5.43 ± 0.46	8.54 ± 0.75	16.32 ± 0.80	23.01 ± 1.44	31.81 ± 1.79	39.98 ± 2.66
	MLR (different percentages of training samples [%])					
Time/sec	1	2	4	6	8	10
Original	1.20 ± 0.13	3.48 ± 0.37	10.41 ± 1.34	21.66 ± 2.07	35.60 ± 2.69	50.13 ± 3.15
SNR	1.08 ± 0.13	3.15 ± 0.38	9.38 ± 1.33	19.63 ± 2.04	32.03 ± 2.35	45.13 ± 2.95
BDM-LCMV	1.00 ± 0.12	3.07 ± 0.39	9.28 ± 1.34	19.49 ± 2.05	32.11 ± 2.46	44.57 ± 3.13
SP-SNR	10.57 ± 0.30	30.25 ± 1.08	91.63 ± 3.53	192.27 ± 5.18	311.27 ± 6.66	466.48 ± 6.90
SP-BDM-LCMV	10.23 ± 0.24	29.74 ± 0.94	91.27 ± 3.24	190.92 ± 3.67	310.65 ± 3.89	462.65 ± 4.15
SP-AAP	5.49 ± 0.28	16.43 ± 1.00	50.43 ± 3.29	104.93 ± 4.09	172.10 ± 4.73	257.27 ± 5.33

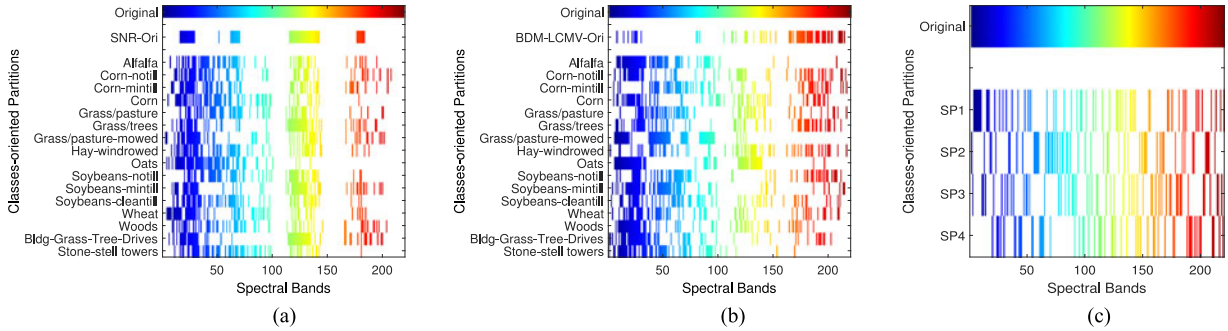

Fig. 8. Spectral partitions obtained by our presented SP method from the AVIRIS Indian Pines data. Our method incorporates two BS algorithms: SNR (a) and BDM-LCMV (b). The spectral partitions obtained by SP-AAP are displayed in (c). In all plots, the x-axis denotes the set of original spectral bands, while the y-axis represents the group of selected bands (each horizontal line displays one group of selected bands).

hyperspectral image into the MCS. This suggests that our presented SP method is able to enhance the overall statistics by improving most of the individual classes via multiple views of the original input features that are provided by different spectral partitions.

In order to evaluate the classification results in Tables III and IV from the viewpoint of individual pixels, Fig. 9 displays the corresponding classification statistics: nonrejected accuracy ($\mathcal{A}$) and classification quality ($\mathcal{Q}$) as a function of the rejected fractions. It can be seen from Fig. 9(a) and (c) that all discussed

TABLE III
OVERALL, AVERAGE, AND INDIVIDUAL CLASS ACCURACIES [%] AND κ STATISTIC OBTAINED BY THE PRESENTED CLASSIFICATION FRAMEWORK IMPLEMENTED USING THE SVM WITH THE BS ALGORITHMS: SNR AND BDM-LCMV, FOR THE AVIRIS INDIAN PINES SCENE

	BS (SVM)			SP (SVM)		
Class	All	SNR	BDM-LCMV	SNR	BDM-LCMV	SP-AAP
Alfalfa	94.56 ± 5.15	96.4 ± 5.46	95.00 ± 6.99	96.76 ± 5.44	97.12 ± 5.41	93.93 ± 6.25
Corn-notill	75.66 ± 3.43	70.85 ± 6.29	71.07 ± 4.49	76.62 ± 4.27	76.82 ± 3.88	75.44 ± 3.74
Corn-mintill	73.05 ± 4.94	70.96 ± 5.70	68.46 ± 5.32	75.51 ± 5.35	75.31 ± 5.36	72.90 ± 4.54
Corn	90.78 ± 3.03	88.66 ± 3.56	88.31 ± 3.70	90.86 ± 2.95	91.27 ± 3.04	91.00 ± 3.94
Grass/pasture	92.44 ± 2.33	90.36 ± 3.01	89.45 ± 3.19	92.12 ± 2.93	91.62 ± 3.18	91.28 ± 3.26
Grass/trees	93.73 ± 2.25	91.84 ± 2.94	90.88 ± 2.10	93.67 ± 2.30	93.77 ± 2.14	94.21 ± 1.88
Grass/pasture-mowed	92.31 ± 6.11	91.54 ± 7.85	91.54 ± 6.06	91.92 ± 6.82	91.54 ± 7.01	90.32 ± 7.89
Hay-windrowed	96.52 ± 1.13	96.00 ± 1.49	95.84 ± 1.20	97.21 ± 1.30	97.27 ± 1.08	96.58 ± 1.98
Oats	92.50 ± 14.1	83.50 ± 19.81	84.00 ± 14.65	92.50 ± 13.33	93.50 ± 12.26	90.94 ± 16.83
Soybeans-notill	79.84 ± 5.21	78.04 ± 4.84	74.55 ± 5.68	83.06 ± 5.42	82.96 ± 4.63	81.85 ± 5.04
Soybeans-mintill	72.15 ± 3.13	71.24 ± 3.57	67.26 ± 3.86	75.55 ± 3.82	75.43 ± 3.58	72.19 ± 4.89
Soybeans-cleantill	87.74 ± 4.37	87.33 ± 4.10	85.69 ± 4.73	90.08 ± 3.81	89.28 ± 3.71	87.03 ± 3.61
Wheat	99.36 ± 0.50	98.66 ± 1.08	99.01 ± 0.79	99.21 ± 0.64	99.24 ± 0.54	99.36 ± 0.37
Woods	90.37 ± 2.71	91.82 ± 3.41	90.24 ± 4.32	92.62 ± 2.42	92.77 ± 1.94	91.80 ± 2.87
Bldg-Grass-Tree-Drives	75.73 ± 6.12	68.31 ± 5.48	67.00 ± 5.39	73.40 ± 6.80	73.17 ± 5.76	72.22 ± 5.75
Stone-stell towers	94.34 ± 3.08	92.97 ± 2.79	93.98 ± 2.58	93.61 ± 2.41	94.53 ± 2.64	94.97 ± 3.19
Overall accuracy	81.50 ± 1.20	79.83 ± 1.44	77.88 ± 1.29	83.34 ± 1.20	83.28 ± 1.23	83.05 ± 1.06
Average accuracy	87.57 ± 1.35	85.53 ± 1.85	84.52 ± 1.37	88.42 ± 1.53	88.48 ± 1.42	88.73 ± 1.49
κ statistic	78.98 ± 1.33	77.05 ± 1.61	74.87 ± 1.44	81.04 ± 1.34	80.98 ± 1.37	80.52 ± 1.16

The results obtained using the original spectral information and the spectral bands selected by using the SNR and BDM-LCMV are also included. SP-AAP is also included for comparison. In all cases, only 640 randomly selected training samples have been used.

TABLE IV
OVERALL, AVERAGE, AND INDIVIDUAL CLASS ACCURACIES [%] AND κ STATISTIC OBTAINED BY THE PRESENTED CLASSIFICATION FRAMEWORK IMPLEMENTED USING THE MLR WITH THE BS ALGORITHMS: SNR AND BDM-LCMV, FOR THE AVIRIS INDIAN PINES SCENE

	BS (MLR)			SP (MLR)		
Class	All	SNR	BDM-LCMV	SNR	BDM-LCMV	SP-AAP
Alfalfa	86.43 ± 6.91	86.79 ± 8.44	83.57 ± 8.39	89.21 ± 7.43	90.00 ± 8.16	82.50 ± 9.41
Corn-notill	70.86 ± 3.81	68.11 ± 3.74	67.54 ± 3.65	71.87 ± 3.44	71.50 ± 3.27	69.01 ± 3.56
Corn-mintill	58.05 ± 3.84	57.40 ± 4.68	59.01 ± 4.21	62.95 ± 3.72	58.26 ± 3.86	61.39 ± 4.44
Corn	78.60 ± 5.38	78.42 ± 3.46	73.20 ± 6.59	78.71 ± 6.39	77.85 ± 4.50	73.97 ± 6.16
Grass/pasture	83.77 ± 3.03	82.48 ± 3.49	83.16 ± 2.96	85.42 ± 3.07	87.53 ± 3.48	84.72 ± 3.31
Grass/trees	93.53 ± 1.82	92.67 ± 1.99	88.19 ± 2.50	93.16 ± 2.21	92.39 ± 1.96	89.05 ± 2.57
Grass/pasture-mowed	82.24 ± 8.67	88.46 ± 7.69	75.71 ± 10.91	83.24 ± 9.02	84.55 ± 7.90	83.40 ± 9.09
Hay-windrowed	96.56 ± 1.19	93.90 ± 1.95	91.78 ± 2.97	94.95 ± 2.55	92.96 ± 3.19	93.16 ± 3.08
Oats	82.62 ± 16.34	84.29 ± 15.70	74.85 ± 12.25	82.79 ± 13.03	90.40 ± 8.25	69.75 ± 14.00
Soybeans-notill	69.26 ± 5.81	71.07 ± 4.41	68.31 ± 4.46	76.25 ± 4.58	70.26 ± 4.55	77.06 ± 3.76
Soybeans-mintill	58.35 ± 5.19	56.52 ± 5.08	56.61 ± 4.05	66.67 ± 4.48	59.67 ± 4.51	67.64 ± 4.63
Soybeans-cleantill	78.48 ± 5.38	76.94 ± 4.04	66.24 ± 5.39	79.07 ± 3.85	74.91 ± 5.35	76.83 ± 3.94
Wheat	99.56 ± 0.26	99.53 ± 0.31	98.89 ± 0.50	99.33 ± 0.44	99.12 ± 0.75	99.06 ± 0.58
Woods	90.49 ± 2.50	87.66 ± 4.40	85.16 ± 4.03	90.83 ± 3.67	89.94 ± 2.63	90.14 ± 4.15
Bldg-Grass-Tree-Drives	71.74 ± 4.26	67.66 ± 3.44	63.99 ± 3.48	69.75 ± 4.60	68.42 ± 3.45	64.09 ± 5.85
Stone-stell towers	92.22 ± 4.18	93.23 ± 3.60	93.13 ± 3.56	94.70 ± 3.39	87.92 ± 3.70	92.41 ± 5.34
Overall accuracy	73.80 ± 1.24	72.25 ± 1.12	70.45 ± 1.12	77.37 ± 1.20	73.83 ± 1.15	76.31 ± 1.34
Average accuracy	80.80 ± 1.77	80.32 ± 1.41	76.83 ± 1.75	81.47 ± 1.43	80.98 ± 1.22	79.20 ± 1.24
κ statistic	70.33 ± 1.34	68.57 ± 1.21	66.62 ± 1.22	73.17 ± 1.33	70.30 ± 1.26	72.29 ± 1.48

The results obtained using the original spectral information and the spectral bands selected by using the SNR and BDM-LCMV are also included. SP-AAP is also included for comparison. In all cases, only 640 randomly selected training samples have been used.

methods exhibit a similar trend, with the rejection fraction increasing. A similar pattern can be observed in Fig. 9(b) and (d), which implies that similar confidence distributions toward the pixels are provided by all the methods considered in this experiment. This observation is illustrated by the classification maps with 90% of $\mathcal{A}$ after rejection in Fig. 10. Even though only about 2% to 4% increase in OA is observed after applying our proposed SP framework, about 7% and 30% increase of nonrejected accuracy $\mathcal{A}$ is observed when compared with the classifier applied to the original hyperspectral image, and also after using the considered BS algorithms. In comparison with the SP-AAP, our methods also provide higher classification accuracy and rejection confidence when using different rejection fractions. On the other hand, if we compare the five classification maps with 90% of nonrejected accuracy $\mathcal{A}$ in Fig. 10, we can observe that all the methods involved in this case hold

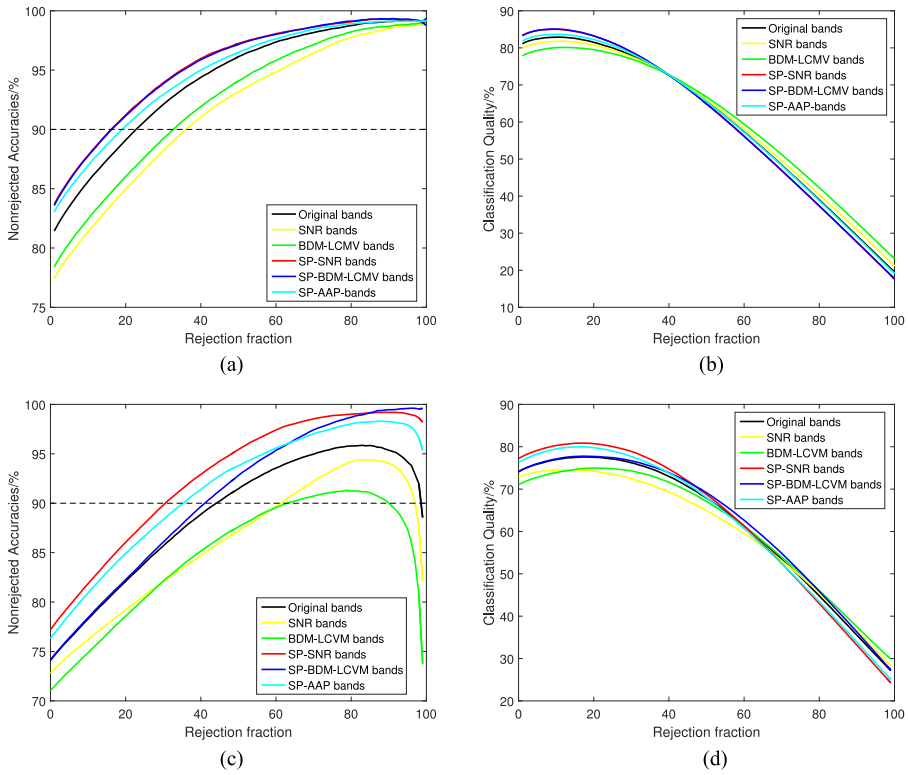

Fig. 9. Statistics of nonrejected accuracies $\mathcal{A}$ and classification qualities $\mathcal{Q}$ of the classifications as a function of rejected fractions with the AVIRIS Indian Pines data. These plots correspond to the results in Tables III–IV. (a) Nonrejected accuracies $\mathcal{A}$ (SVM). (b) Classification qualities Q (SVM). (c) Nonrejected accuracies $\mathcal{A}$ (MLR). (d) Classification qualities Q (MLR).

similar confidence toward the pixels in the scene. This is consistent with our discussion of the results in Fig. 9. Despite the presence of similar confidence distribution in the pixels, we plot the combined classification result against the classifications obtained using SVM with all the original spectral bands and our presented SP method to validate the discussions above. From Fig. 11, we can see that, although the proposed SP strategy obtains better results than standard BS algorithms and SP-AAP SP method, other extra advantages can be achieved in the views of $\mathcal{A}$ and $\mathcal{Q}$ with rejections. The improvement is also observed in the case of the ROSIS Pavia University data, after supported by the classifier using all the original spectral bands as a whole. This is consistent with the aforementioned observations and discussions, in the sense that the results of our presented SP method (embedded with BS algorithms) are further improved after being supported by the classifier using all the original spectral bands, a phenomenon that is observed for both the ROSIS Pavia University scene and the AVIRIS Indian Pines scene. Also, the similarity of the confidence distributions of the two classifications obtained for each pixel suggests that similar confidences toward individual pixels can be obtained by the different methods tested in the case of the AVIRIS data scene. This is reasonable, considering the fact that most of the classes in this scene comprise vegetation features that share very similar spectral signatures, leading to low classification diversity that SP is able to generate by providing multiple views of the original hyperspectral data.

In order to evaluate the sensitivity of our methods to different numbers of training samples, we also plot the overall accuracy statistics (mean value and standard deviation) of different methods as a function of the percentage of training samples as input (see Fig. 12). From Fig. 12, it is remarkable that our presented methods outperform those methods that only use the spectral bands of the original image or the bands selected by single-BS algorithms. This is the case for different percentages of training samples. Finally, it can also be seen that for both classifiers (SVM and MLR), our class-oriented SP methods generally obtained higher OAs when compared with the SP-AAP method. Meanwhile, we also report the processing times of the experiments reported in Fig. 12 on Table V. From Table V, we can observe that our presented SP methods required 5 to 12 times more computation than the original group and three times more than the SP-AAP. As explained before, these times can be significantly reduced by resorting to simple parallel implementation frameworks.

3) Experiments With the HYDICE Washington DC Mall Dataset: In this section, we tested our SP method with the HYDICE Washington DC Mall dataset. First of all, we used 2% randomly selected training samples from the labeled reference data for training (and the rest for testing). For simplicity, the experimental results are obtained by using the same conditions for all groups. Similar to the experiments with the previous two hyperspectral datasets, we set empirically the number of bands

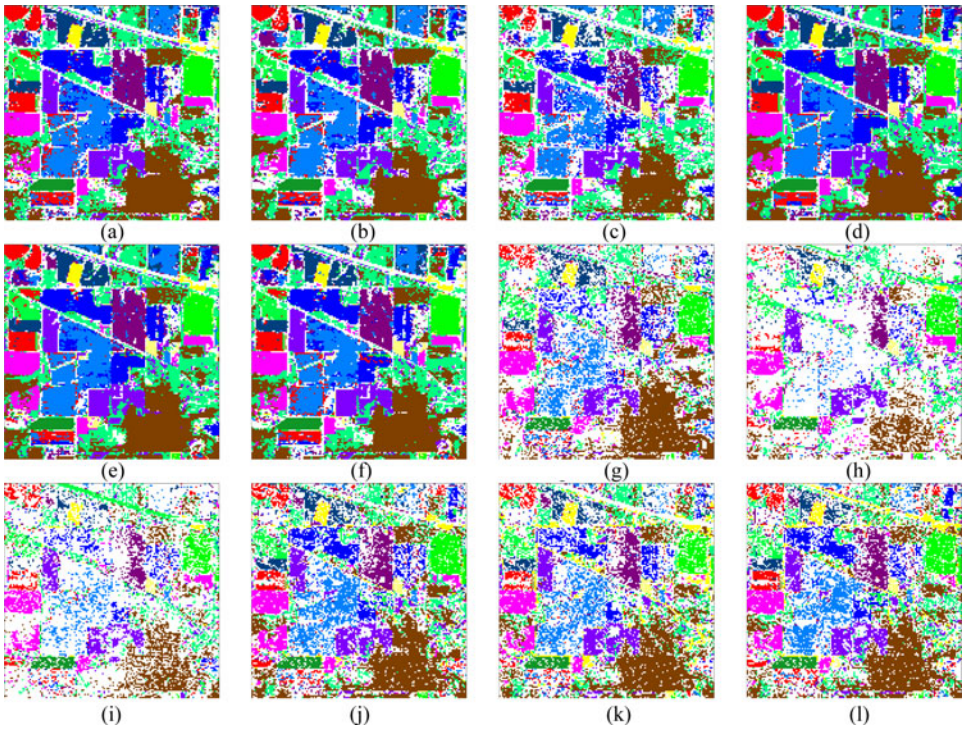

Fig. 10. Classification maps (with rejection) obtained by the proposed classification framework, using the original spectral information (a), (g), with the 60 selected bands by using SNR (b), (h) and BDM-LCMV (c), (i) algorithms, and with the spectral partitions obtained by our proposed SP approach implemented with SNR (d), (h) and BDM-LCMV (e), (k) and the SP-AAP method (f), (l) for the AVIRIS Indian Pines data. The percentage in the parenthesis denotes the proportion of pixels remaining after rejection. In all cases, a total of 640 randomly selected training samples were used for training and the rest used for testing. The maps are displayed with partial pixels rejected in order to obtain a 90% classification accuracy for the remaining ones. Note that the overall accuracies for the unrejected pixels of each map are calculated by considering the labeled test samples, with the training samples excluded. (a) SVM (79%). (b) SVM-SNR (76%). (c) SVM-BDM-LCMV (68%). (d) SVM-SP-SNR (86%). (e) SVM-SP-BDM-LCMV (86%). (f) SVM-SP-AAP (83%). (g) MLR (55%). (h) MLR-SNR (35%). (i) MLR-BDM-LCMV (37%). (j) MLR-SP-SNR (67%). (k) MLR-SP-BDM-LCMV (60%). (l) MLR-SP-AAP (62%).

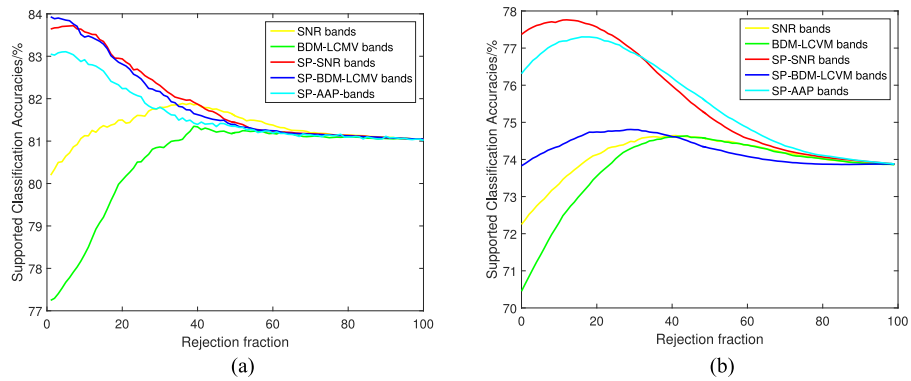

Fig. 11. Classification OAs (as a function of rejections) obtained by our proposed classification framework with the AVIRIS Indian Pines data, after being supported by the rejected pixels of the classifier using the original spectral information. Note that plot (a) is obtained from one Monte Carlo run in Table III, while plot (b) is obtained from one Monte Carlo run in Table IV. In both cases, the selected run is close to the statistical average. (a) SVM. (b) MLR.

to 40 for single BS approaches, for the partitions generated by our SP methods, and for the SP-AAP.

Fig. 13(a) and (b), respectively, plot the multiple spectral partitions obtained by using the SP-SNR and SP-BDM-LCMV methods. The SP results obtained by SP-AAP are shown in Fig. 13(c). In all the plots, the x-axis represents the set of indices of the original spectral bands while the y-axis represents the selected bands for the original image and for each class. From Fig. 13, we can observe that each class-oriented partition takes a band subset of much lower dimensionality than the original number of bands. In all cases, the selection of partitions allows us to obtain multiple views on the original data.

Table VI displays the OA, AA, κ, and individual accuracies obtained after using only 2% randomly selected samples for

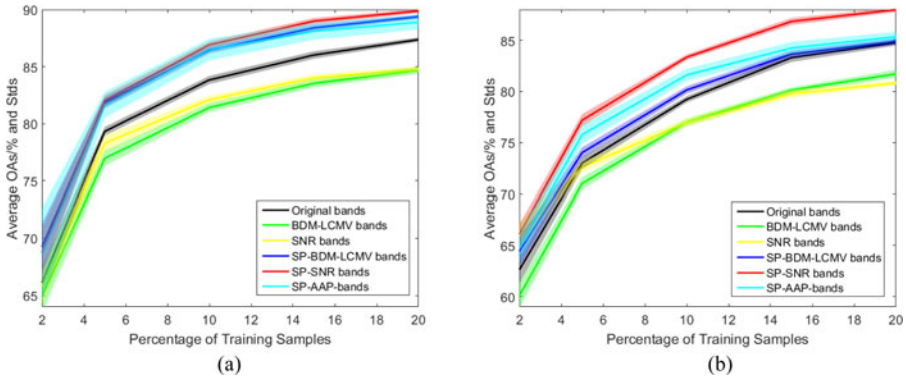

Fig. 12. Overall classification accuracies (as a function of the number of training samples) obtained by the proposed classification framework (with the original spectral information and with SP-SNR, SP-BDM-LCMV, and SP-AAP) for the AVIRIS Indian Pines scene. The solid lines represent the average of 20 Monte Carlo runs, whereas the colored area around the lines represent the standard deviation around the mean. Plots (a) is obtained using the SVM classifier and (b) is obtained using the MLR classifier.

TABLE V
PROCESSING TIMES OF DIFFERENT METHODS FOR THE AVIRIS INDIAN PINES SCENE

	SVM (different percentage of training samples [%])				
Time/sec	2	5	10	15	20
Original	1.87 ± 0.05	7.39 ± 0.24	20.96 ± 0.54	45.01 ± 1.12	77.80 ± 1.00
SNR	1.68 ± 0.05	6.37 ± 0.24	18.82 ± 0.57	40.88 ± 1.12	70.05 ± 1.00
BDM-LCMV	1.59 ± 0.05	6.29 ± 0.24	18.59 ± 0.54	40.08 ± 1.21	69.72 ± 0.99
SP-SNR	9.09 ± 0.36	37.69 ± 0.95	108.89 ± 2.04	238.03 ± 4.30	409.95 ± 3.20
SP-BDM-LCMV	8.67 ± 0.05	36.77 ± 0.80	106.92 ± 2.19	237.15 ± 2.98	417.06 ± 2.21
SP-AAP	2.91 ± 0.10	12.83 ± 0.27	34.82 ± 1.50	78.24 ± 2.11	128.57 ± 2.33
	MLR (different percentage of training samples [%])				
Time/sec	2	5	10	15	20
Original	0.81 ± 0.11	1.42 ± 0.09	2.36 ± 0.16	3.44 ± 0.27	4.48 ± 0.29
SNR	0.70 ± 0.07	1.22 ± 0.10	2.10 ± 0.16	2.97 ± 0.28	3.89 ± 0.27
BDM-LCMV	0.73 ± 0.08	1.23 ± 0.10	2.15 ± 0.12	3.06 ± 0.28	4.04 ± 0.28
SP-SNR	9.14 ± 0.24	15.19 ± 0.38	24.79 ± 0.58	36.28 ± 0.99	47.59 ± 1.30
SP-BDM-LCMV	9.21 ± 0.17	15.57 ± 0.36	25.98 ± 0.64	38.88 ± 1.12	51.44 ± 1.47
SP-AAP	3.20 ± 0.17	3.89 ± 0.24	6.99 ± 0.54	12.06 ± 0.79	17.33 ± 1.40

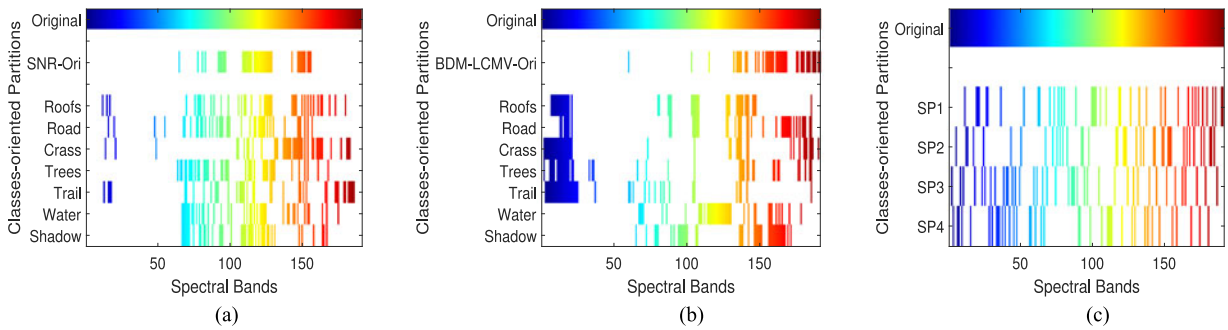

Fig. 13. Spectral partitions obtained by our presented SP method from the HYDICE DC Mall dataset. Our method incorporates two BS algorithms: SNR (a) and BDM-LCMV (b). The spectral partitions obtained by SP-AAP are displayed in (c). In all plots, the *x*-axis denotes the set of original spectral bands, while the *y*-axis represents the group of selected bands (each horizontal line displays one group of selected bands). (a) SP-SNR. (b) SP-BDM-LCMV. (c) SP-AAP.

training and the rest for testing [see Fig. 2(i)]. Several observations can be drawn from Table VI. First of all, an increase of about 1% to 6% in classification OA is obtained by our presented SP-SNR and SP-BDM-LCMV methods in comparison with those using all the spectral bands, and also with the single BS approaches. This observation is consistent with those made for the two previous hyperspectral datasets, suggesting the effectiveness of our newly proposed methods for multiple classifier-based feature learning based on multiple views that are provided by spectral partitions. Second, in comparison with the SP-AAP method, both SP-SNR and SP-BDM-LCMV obtained improved accuracies. Specifically, relevant improvements are

TABLE VI
OVERALL, AVERAGE, AND INDIVIDUAL CLASS ACCURACIES [%] AND κ STATISTIC OBTAINED BY THE PRESENTED CLASSIFICATION FRAMEWORK IMPLEMENTED USING THE SVM WITH THE BS ALGORITHMS: SNR AND BDM-LCMV, FOR THE HYDICE WASHINGTON DC MALL SCENE

	BS (SVM)			SP (SVM)		
Class	All	SNR	BDM-LCMV	SNR	BDM-LCMV	SP-AAP
Roofs	81.34 ± 2.67	83.02 ± 2.83	82.24 ± 3.26	86.78 ± 2.81	88.14 ± 2.25	85.47 ± 2.47
Road	90.48 ± 1.80	90.09 ± 1.38	92.55 ± 1.42	90.57 ± 2.09	93.56 ± 1.43	92.89 ± 1.23
Crass	94.09 ± 1.52	95.00 ± 2.37	94.86 ± 1.51	94.72 ± 2.28	95.83 ± 1.91	95.13 ± 1.77
Trees	93.16 ± 0.92	93.65 ± 1.44	93.18 ± 0.80	94.14 ± 1.51	93.94 ± 1.05	93.68 ± 0.86
Trail	90.96 ± 1.92	87.59 ± 3.22	92.17 ± 1.21	89.48 ± 2.74	93.35 ± 1.66	92.88 ± 1.74
Water	94.40 ± 1.14	90.55 ± 1.69	96.14 ± 0.89	91.78 ± 1.27	93.82 ± 1.73	94.54 ± 1.32
Shadow	97.58 ± 0.47	94.57 ± 1.58	97.36 ± 0.84	94.66 ± 1.70	97.19 ± 0.62	97.61 ± 0.45
Overall accuracy	89.77 ± 0.54	89.56 ± 0.74	90.65 ± 0.74	91.00 ± 0.64	92.71 ± 0.54	91.76 ± 0.64
Average accuracy	91.72 ± 0.35	90.64 ± 0.57	92.64 ± 0.31	91.73 ± 0.57	93.69 ± 0.44	93.17 ± 0.42
κ statistic	87.49 ± 0.65	87.22 ± 0.89	88.57 ± 0.88	88.96 ± 0.77	91.06 ± 0.65	89.91 ± 0.77
	BS (MLR)			SP (MLR)		
Class	All	SNR	BDM-LCMV	SNR	BDM-LCMV	SP-AAP
Roofs	77.51 ± 2.92	79.81 ± 2.13	81.28 ± 1.91	90.36 ± 1.95	90.01 ± 2.34	84.05 ± 3.82
Road	85.89 ± 1.49	90.49 ± 1.29	89.02 ± 1.61	91.73 ± 1.48	90.1 ± 1.58	87.26 ± 1.46
Crass	91.78 ± 1.76	94.99 ± 2.02	94.83 ± 2.07	95.4 ± 1.75	94.56 ± 1.76	92.67 ± 1.75
Trees	93.4 ± 1.43	92.33 ± 1.93	90.43 ± 1.77	94.53 ± 1.32	93.92 ± 1.01	93.95 ± 1.47
Trail	93.95 ± 1.51	91.31 ± 1.79	91.91 ± 1.42	93.72 ± 0.94	92.6 ± 1.37	93.76 ± 1.32
Water	99.05 ± 0.47	84.99 ± 7.47	70.56 ± 5.45	96.42 ± 0.51	84.28 ± 7.2	99.16 ± 0.34
Shadow	96.68 ± 1.29	95.22 ± 1.34	82.67 ± 3.34	97.08 ± 0.72	87.98 ± 4.84	95.37 ± 1.52
Overall accuracy	87.95 ± 0.75	88.56 ± 0.77	86.19 ± 0.68	93.19 ± 0.65	90.86 ± 0.81	90.10 ± 1.09
Average accuracy	91.18 ± 0.50	89.88 ± 1.12	85.81 ± 0.73	94.18 ± 0.44	90.50 ± 0.93	92.32 ± 0.64
κ statistic	85.35 ± 0.90	86.05 ± 0.93	83.18 ± 0.81	91.65 ± 0.79	88.79 ± 0.98	87.90 ± 1.31

The results obtained using the original spectral information and the spectral bands selected by using the SNR and BDM-LCMV are also included. SP-AAP is also included for comparison. In all cases, only 2% randomly selected training samples from the labeled reference data have been used.

obtained for the individual classes. This is also consistent with the observations made for the results obtained with the previous hyperspectral datasets.

In order to better illustrate this, Fig. 14 displays the classification maps obtained for different methods after rejection to pursue 97% OA of the nonrejected pixels. It is remarkable that, by using our presented SP methods (SP-SNR and SP-BDM-LCMV), an increase of 12% (SVM) and 30% (MLR) in OA are achieved in comparison with using the spectral bands of the original image, and 3% to 20% improvements in OA are obtained as compared with the SP-AAP. In order to better demonstrate this, Fig. 15 plots the classification accuracy $\mathcal{A}$ and quality $\mathcal{Q}$ as functions of the rejection fractions. In Fig. 15, it is remarkable that our presented method obtains more confidence in the classification of individual pixels than the other methods. From the previous experiments with the ROSIS and AVIRIS datasets, we could already observe that different nonrejected accuracies lead to improvements in classification when combined by the classification results obtained using the original image. Hence, in the case of the HYDICE Washington DC mall dataset, we plot the supported classification results in Fig. 16. From Fig. 16, we can see that the supported classification accuracies are improved up to 5%. It is also remarkable that, after combining two classifications, the results can be further improved, especially in the case of using the MLR classifier.

With the aforementioned experimental observations in mind, we conclude that our presented SP approaches (SP-SNR and SP-BDM-LCMV) generally obtain better classification accuracies when compared with the SP-AAP method. The obtained results are dependent on the learning rules of the classifiers as well as on the considered hyperspectral datasets. For example, the SP-BDM-LCMV generally outperforms other methods, while the SP-SNR also provides competitive performance in all cases. In turn, the SP-AAP provides comparatively worst classification accuracies when using the SVM classifier with the ROSIS Pavia University dataset. However, when the MLR classifier is used, the SP-AAP performs better than the SP-SNR and worse than the SP-BDM-LCMV (see Table I and Fig. 6).

To conclude this section, we evaluate the sensitivity of the compared methods to different percentages of training samples. Fig. 17 plots the overall accuracies of different methods when using different proportions of training samples as input. First of all, it is remarkable that our presented SP methods provide highly competitive results in this experiment. The processing times, reported on Table VII, indicate that our methods need about four to six times more computation than the classification of the original image and about two times more than the SP-AAP. These results can be improved by straightforward parallelization, as discussed in previous experiments.

IV. CONCLUSION AND FUTURE RESEARCH LINES

In this paper, we presented a new class-oriented SP method for hyperspectral image classification. The proposed method is shown to be effective in the task of exploiting the information contained in the specific classes by rearranging the

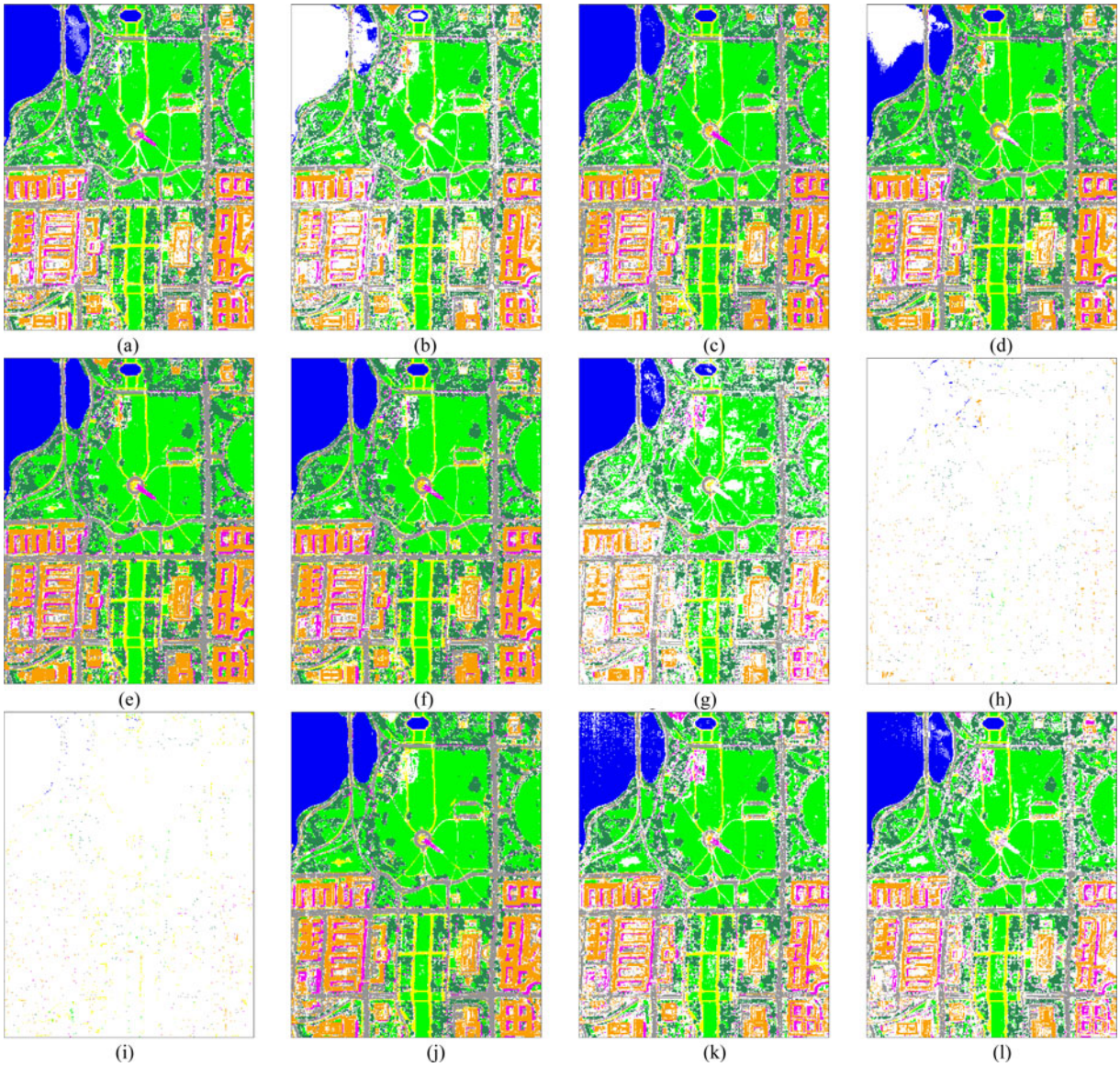

Fig. 14. Classification maps (with rejection) obtained by the proposed classification framework using the original spectral information (a), (g), the 40 selected bands by using SNR (b), (h) and BDM-LCMV (c), (i) algorithms, the spectral partitions obtained by our proposed approach implemented with SNR (d), (j) and BDM-LCMV (e), (k), and the SP-AAP method (f), (l) for the HYDICE Washington DC mall data. The number in the parenthesis denotes the proportion of pixels remaining after rejection. In all cases, a total of 2% randomly selected training samples were used for training. The maps are displayed with partial pixels rejected in order to obtain a 97% classification accuracy for the remaining ones. Note that the overall accuracies for the unrejected pixels of each map are calculated by considering the labeled test samples, with the training samples excluded. (a) SVM (75%). (b) SVM-SNR (58%). (c) SVM-BDM-LCMV (78%). (d) SVM-SP-SNR (75%). (e) SVM-SP-BDM-LCMV (87%). (f) SVM-SP-AAP (84%). (g) MLR (55%). (h) MLR-SNR (0%). (i) MLR-BDM-LCMV (0%). (j) MLR-SP-SNR (85%). (k) MLR-SP-BDM-LCMV (70%). (l) MLR-SP-AAP (65%).

original spectral information. This is mainly done to address the Hughes phenomenon by means of a MCS, while avoiding the elimination of relevant spectral bands in the original hyperspectral image that may be useful for the discrimination of the classes. Inspired by the idea of classification with rejection, we also designed a strategy that combines different classifiers for further improving the classification result. In our experiments, two SP methods are constructed based on two well-established BS algorithms: SNR and BDM-LCMV. The obtained spectral partitions keep most of the relevant spectral bands from the original image and provide different views (understood as low-dimensional partitions) for multiple feature learning, thus addressing the potential problems associated to the limited availability of training samples. Our experiments illustrated reasonable advantages in classification accuracy achieved by our presented SP framework with three well-known hyperspectral images: the ROSIS Pavia University data, the AVIRIS Indian Pines data, and the HYDICE Washington DC mall data. Furthermore, the experimental classification results indicate that the key for a successful SP lies in the capability to generate a group of spectral partitions with diverse views of the original hyperspectral image. Besides, we also observed the effectiveness of combining different classifiers with rejection (including those resulting from the multiple-classifier views generated by

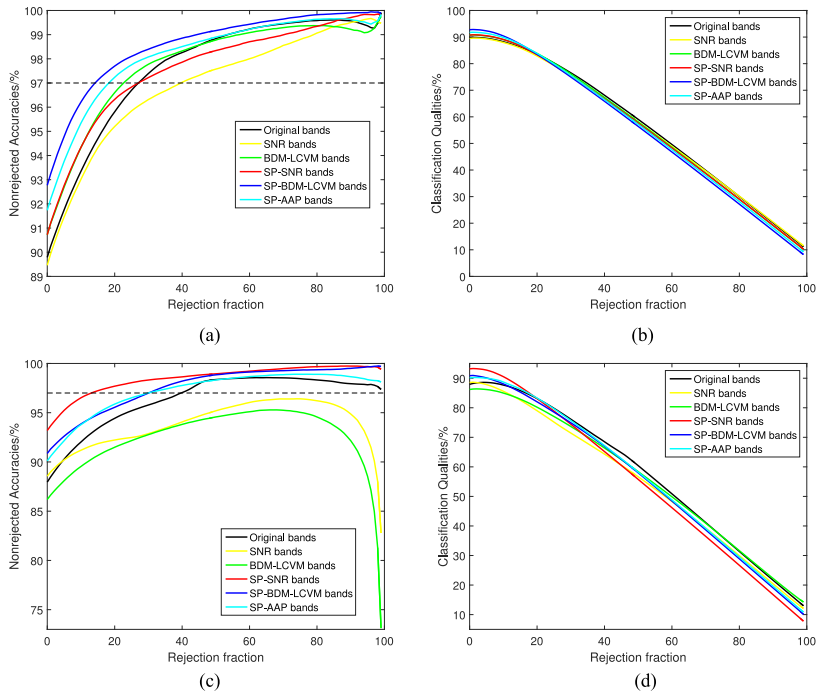

Fig. 15. Statistics of nonrejected accuracies $\mathcal{A}$ and classification qualities $\mathcal{Q}$ of the classifications as a function of rejected fractions with the HYDICE Washington DC Mall scene. These plots correspond to the results in Table VI. (a) Nonrejected accuracies $\mathcal{A}$ (SVM). (b) Classification qualities $\mathcal{Q}$ (SVM). (c) Nonrejected accuracies $\mathcal{A}$ (MLR). (d) Classification qualities Q (MLR).

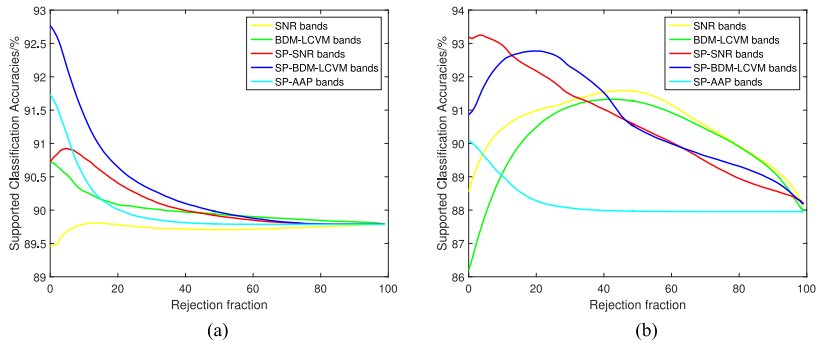

Fig. 16. Classification OAs (as a function of rejections) obtained by our proposed classification framework with the HYDICE Washington DC mall data, after being supported by the rejected pixels of the classifier using the original spectral information. Note here that the plots of both figures (a), (b) are obtained, respectively, from the same one Monte Carlo run of Table VI that is close to the statistical average. (a) SVM. (b) MLR.

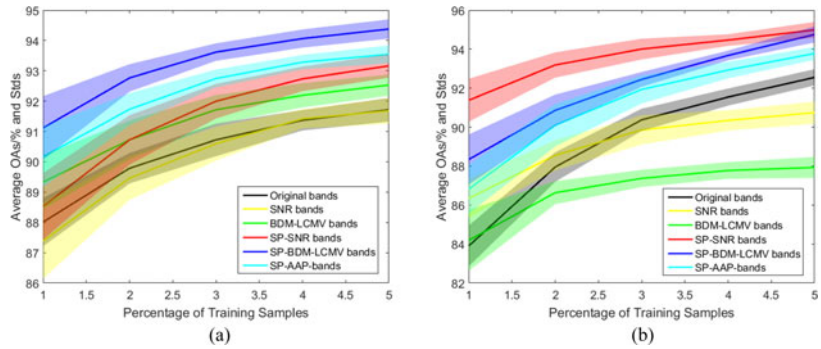

Fig. 17. Overall classification accuracies (as a function of the number of training samples) obtained by the proposed classification framework (with the original spectral information and with SNR, BDM-LCMV BS methods and the AAP SP method) for the HYDICE Washington DC mall scene. The solid lines represent the average of 20 Monte Carlo runs, whereas the colored areas around the lines represent the standard deviation around the mean. Plots (a) and (b) are respectively obtained by using SVM and MLR classifiers. (a) SVM. (b) MLR.

TABLE VII
PROCESSING TIMES OF DIFFERENT METHODS WITH HYDICE WASHINGTON DC MALL SCENE

	SVM (different percentage of training samples [%])				
Time/sec	1	2	3	4	5
Original	0.91 ± 0.19	1.74 ± 0.18	3.30 ± 0.15	5.08 ± 0.15	7.12 ± 0.33
SNR	0.80 ± 0.03	1.47 ± 0.07	2.70 ± 0.10	4.29 ± 0.21	6.12 ± 0.19
BDM-LCMV	0.79 ± 0.04	1.49 ± 0.08	2.71 ± 0.11	4.27 ± 0.16	6.14 ± 0.27
SP-SNR	5.32 ± 0.22	10.06 ± 0.52	18.80 ± 0.59	29.76 ± 1.02	42.35 ± 1.03
SP-BDM-LCMV	5.24 ± 0.23	9.96 ± 0.52	18.76 ± 0.72	29.53 ± 0.76	42.16 ± 1.39
SP-AAP	3.10 ± 0.16	5.88 ± 0.28	10.95 ± 0.42	17.30 ± 0.49	24.48 ± 0.66
	MLR (different percentage of training samples [%])				
Time/sec	1	2	3	4	5
Original	5.09 ± 0.15	6.29 ± 0.24	7.47 ± 0.24	8.49 ± 0.30	9.09 ± 0.22
SNR	3.96 ± 0.05	4.29 ± 0.08	4.67 ± 0.08	4.94 ± 0.07	5.24 ± 0.07
BDM-LCMV	3.96 ± 0.04	4.27 ± 0.07	4.55 ± 0.06	4.82 ± 0.08	5.11 ± 0.06
SP-SNR	27.27 ± 0.27	29.47 ± 0.27	31.66 ± 0.44	33.50 ± 0.61	35.52 ± 0.53
SP-BDM-LCMV	27.02 ± 0.30	28.99 ± 0.35	30.76 ± 0.43	32.31 ± 0.47	34.17 ± 0.51
SP-AAP	15.65 ± 0.16	16.82 ± 0.20	17.94 ± 0.21	18.92 ± 0.27	19.96 ± 0.32

SP and the original hyperspectral image). In our future work, we will focus on exploring the potential of generating a group of spectral partitions by multiple criteria, given by an extensive set of BS algorithms instead of using only one criterion. The potential of employing different classifiers will also be considered to enlarge the diversity of the multiple views generated through spectral partitions. We will also work toward the parallel implementation of the presented methods, which will allow us to increase their computational performance due to their inherently parallel nature.

ACKNOWLEDGMENT

The authors would like to gratefully thank the Associate Editor and the Anonymous Reviewers for their outstanding comments and suggestions, which greatly helped us to improve the technical quality and presentation of our manuscript.

REFERENCES

[1] J. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. 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] M. Govender, K. Chetty, and H. Bulcock, "A review of hyperspectral remote sensing and its application in vegetation and water resource studies," *Water SA (Pretoria)*, vol. 33, pp. 145–151, 2007.

[3] A. Plaza and C.-I. Chang, *High Performance Computing in Remote Sensing*. Boca Raton, FL, USA: Taylor & Francis, 2007.

[4] F. Dell'Acqua, P. Gamba, A. Ferrari, J. Palmason, J. Benediktsson, and K. Arnason, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 322–326, Oct. 2004.

[5] 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.

[6] S. Serpico, M. D'Inca, F. Melgani, and G. Moser, "Comparison of feature reduction techniques for classification of hyperspectral remote sensing data," in *Proc. Int. Symp. Remote Sens.*, 2003, pp. 347–358.

[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] 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.

[9] B. Waske, S. van der Linden, J. Benediktsson, A. Rabe, and P. Hostert, "Sensitivity of support vector machines to random feature selection in classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2880–2889, Jul. 2010.

[10] M. Fauvel, C. Dechesne, A. Zullo, and F. Ferraty, "Fast forward feature selection of hyperspetral images for classification with gaussian mixture models," in *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 8, no. 6, pp. 2824–2831, Jun. 2015.

[11] G. Byrne, P. Crapper, and K. Mayo, "Monitoring land-cover change by principal component analysis of multitemporal landsat data," *Remote Sens. Environ.*, vol. 10, no. 3, pp. 175–184, 1980.

[12] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, vol. 46. New York, NY, USA: Wiley, 2004.

[13] L. Ma, M. Crawford, and J. Tian, "Local manifold learning-based-nearest-neighbor for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4099–4109, Nov. 2010.

[14] Y. Lu, Z. Lai, Z. Fan, J. Cui, and Q. Zhu, "Manifold discriminant regression learning for image classification," *Neurocomputing*, vol. 166, pp. 475–486, Oct. 2015.

[15] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

[16] W. Chen, J. Huang, J. Zou, and B. Fang, "Wavelet-face based subspace LDA method to solve small sample size problem in face recognition," *Int. J. Wavelets Multiresolution Inf. Process.*, vol. 7, no. 2, pp. 199–214, Mar. 2009.

[17] Z. Zhu, S. Jia, S. He, Y. Sun, Z. Ji, and L. Shen, "Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework," *Inf. Sci.*, vol. 298, pp. 274–287, Mar. 2015.

[18] P. Ceccato, N. Gobron, S. Flasse, B. Pinty, and S. Tarantola, "Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach," *Remote Sens. Environ.*, vol. 82, no. 2, pp. 188–197, 2002.

[19] D. Odermatt, A. Gitelson, V. Brando Ernesto, and M. Schaepman, "Review of constituent retrieval in optically deep and complex waters from satellite imagery," *Remote Sens. Environ.*, vol. 118, pp. 116–126, 2012.

[20] Y. Qian, F. Yao, and S. Jia, "Band selection for hyperspectral imagery using affinity propagation," *IET Comput. Vision*, vol. 3, no. 4, pp. 213–222, 2009.

[21] K. Zhao, D. Valle, S. Popescu, X. Zhang, and B. Mallick, "Hyperspectral remote sensing of plant biochemistry using bayesian model averaging with variable and band selection," *Remote Sens. Environ.*, vol. 132, pp. 102–119, 2013.

[22] V. Lombardo, M. Musacchio, and M. Buongiorno, "Error analysis of subpixel lava temperature measurements using infrared remotely sensed data," *Int. Geophys. J.*, vol. 191, no. 1, pp. 112–125, 2012.

[23] D. Tuia, G. Camps-Valls, G. Matasci, and M. Kanevski, "Learning relevant image features with multiple-kernel classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3780–3791, Oct. 2010.
[24] M. Gong, M. Zhang, and Y. Yuan, "Unsupervised band selection based on evolutionary multiobjective optimization for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 544–557, Jan. 2016.
[25] J. Han, D. Zhang, G. Cheng, L. Guo, and J. Ren, "Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 6, pp. 3325–3337, Jun. 2015.
[26] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996.
[27] R. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions," *Mach. Learn.*, vol. 37, no. 3, pp. 297–336, 1999.
[28] L. Rokach, "Ensemble-based classifiers," *Artif. Intell. Rev.*, vol. 33, nos. 1/2, pp. 1–39, 2010.
[29] J. Benediktsson, J. Chanussot, and M. Fauvel, "Multiple classifier systems in remote sensing: From basics to recent developments," in *Proc. Multiple Classifier Syst.*, 2007, pp. 501–512.
[30] W.-S. Chen, X. Dai, B. Pan, and T. Huang, "A novel discriminant criterion based on feature fusion strategy for face recognition," *Neurocomputing*, vol. 159, pp. 67–77, Jul. 2015.
[31] W. Di and M. Crawford, "Active learning via multi-view and local proximity co-regularization for hyperspectral image classification," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 618–628, Jun. 2011.
[32] M. Woźniak, M. Graña, and E. Corchado, "A survey of multiple classifier systems as hybrid systems," *Inf. Fusion*, vol. 16, pp. 3–17, 2014.
[33] Y. Liu, J. Li, A. Plaza, J. Bioucas-Dias, A. Cuartero, and P. Garcia, "Spectral partitioning for hyperspectral remote sensing image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 3434–3437.
[34] Y. Liu, J. Li, and A. Plaza, "Spectrometer-driven spectral partitioning for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 9, no. 2, pp. 668–680, Feb. 2016.
[35] T. Dietterich, "Ensemble methods in machine learning," in *Proc. 1st Int. Workshop Multiple Classifier Syst.*, 2000, pp. 1–15.
[36] Z. Zhu, S. Jia, and Z. Ji, "Towards a memetic feature selection paradigm application notes," *IEEE Comput. Intell. Mag.*, vol. 5, no. 2, pp. 41–53, May 2010.
[37] S. Kawaguchi and R. Nishii, "Hyperspectral image classification by bootstrap adaboost with random decision stumps," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3845–3851, Nov. 2007.
[38] T. Ho, J. Hull, and S. Srihari,"Decision combination in multiple classifier systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 1, pp. 66–75, Jan. 1994.
[39] X. Wang, "Learning from big data with uncertainty—Editorial," *J. Intell. Fuzzy Syst.*, vol. 28, no. 5, pp. 2329–2330, Jun. 2015.
[40] H. Wang, F. Nie, and H. Huang, "Multi-view clustering and feature learning via structured sparsity," in *Proc. 30th Int. Conf. Mach. Learn.*, 2013, pp. 352–360.
[41] P. Ghamisi, J. A. Benediktsson, and S. Phinn, "Fusion of hyperspectral and lidar data in classification of urban areas," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2014, pp. 181–184.
[42] G. Cheng, J. Han, L. Guo, Z. Liu, S. Bu, and J. Ren, "Effective and efficient midlevel visual elements-oriented land-use classification using VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4238–4249, Aug. 2015.
[43] J. Han *et al.*, "Efficient, simultaneous detection of multi-class geospatial targets based on visual saliency modeling and discriminative learning of sparse coding," *ISPRS J. Photogrammetry Remote Sens.*, vol. 89, pp. 37–48, 2014.
[44] X. Huang, Q. Lu, and L. Zhang, "A multi-index learning approach for classification of high-resolution remotely sensed images over urban areas," *ISPRS J. Photogrammetry Remote Sens.*, vol. 90, pp. 36–48, 2014.
[45] G. Zhang and X. Jia, "Feature selection using kernel based local fisher discriminant analysis for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2011, pp. 1728–1731.
[46] X. Jia, B. Kuo, and M. Crawford, "Feature mining for hyperspectral image classification," *Proc. IEEE*, vol. 101, no. 3, pp. 676–697, Mar. 2013.
[47] J. Li, M. P. Reddy, A. Plaza, J. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4816–4829, Sep. 2013.
[48] W. Di and 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.
[49] B. Waske, S. van-der Linden, J. Benediktsson, A. Rabe, and P. Hostert, "Sensitivity of support vector machines to random feature selection in classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2880–2889, Jul. 2010.
[50] J. Ham, Y. Chen, M. Crawford, and J. Ghosh, "Investigation of the random forest framework for classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 492–501, Mar. 2005.
[51] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[52] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semi-supervised hyperspectral image classification using soft sparse multinomial logistic regression," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 2, pp. 318–322, 2013.
[53] F. Condessa, J. Bioucas-Dias, C. Castro, J. Ozolek, and J. Kovacevic, "Classification with reject option using contextual information," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, 2013, pp. 1340–1343.
[54] F. Condessa, J. Bioucas-Dias, and J. Kovacevic, "Supervised hyperspectral image classification with rejection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*. 2015, pp. 2600–2603.
[55] C.-I. Chang, *Hyperspectral Data Processing: Algorithm Design and Analysis*. New York, NY, USA: Wiley, 2013.
[56] P. Bajcsy and P. Groves, "Methodology for hyperspectral band selection," *Photogrammetric Eng. Remote Sens.*, vol. 70, no. 7, pp. 793–802, 2004.
[57] N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1552–1565, Jul. 2004.
[58] E. Sarhrouni, A. Hammouch, and D. Aboutajdine, "Band selection and classification of hyperspectral images using mutual information: An algorithm based on minimizing the error probability using the inequality of fano," in *Proc. IEEE Int. Conf. Multimedia Comput. Syst.*, 2012, pp. 155–159.
[59] K. Sun, X. Geng, L. Ji, and Y. Lu, "A new band selection method for hyperspectral image based on data quality," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 7, no. 6, pp. 2697–2703, Jun. 2014.
[60] C.-I. Chang and S. Wang, "Constrained band selection for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1575–1585, Jun. 2006.
[61] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 27:1–27:27, 2011.
[62] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. S110–S122, 2009.
[63] M. Fauvel, J. Chanussot, and J. Benediktsson, "A spatial–spectral kernel-based approach for the classification of remote-sensing images," *Pattern Recognit.*, vol. 45, no. 1, pp. 381–392, 2012.
[64] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS J. Photogrammetry Remote Sens.*, vol. 66, no. 3, pp. 247–259, 2011.
[65] T. Kavzoglu and I. Colkesen, "A kernel functions analysis for support vector machines for land cover classification," *Int. J. Appl. Earth Observation Geoinformation*, vol. 11, no. 5, pp. 352–359, 2009.

Yi Liu (S'15) received the B.Sc. degree in geographic information systems from China University of Petroleum, Qingdao, China, in 2010, the M.E. degree in photogrammetry and remote sensing from China University of Mining and Technology, Xuzhou, China, in 2013. He is currently working toward the Ph.D. degree in computer engineering at the Hyperspectral Computing Laboratory (HyperComp), Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Cáceres, Spain.

His research interests include remote sensing image processing, image classification and segmentation, and Bayesian inference, with particular emphasis on spectral-spatial partitioning techniques for hyperspectral image classification.

Jun Li (M'13) received the B.S. degree in geographic information systems from Hunan Normal University, Changsha, China, in 2004, the M.E. degree in remote sensing from Peking University, Beijing, China, in 2007, and the Ph.D. degree in electrical engineering from the Instituto de Telecomunicações, Instituto Superior Técnico (IST), Universidade Técnica de Lisboa, Lisbon, Portugal, in 2011.

From 2007 to 2011, she was a Marie Curie Research Fellow with the Departamento de Engenharia Electrotécnica e de Computadores and the Instituto de Telecomunicações, IST, Universidade Técnica de Lisboa, in the framework of the European Doctorate for Signal Processing (SIGNAL). She has also been actively involved in the Hyperspectral Imaging Network, a Marie Curie Research Training Network involving 15 partners in 12 countries and intended to foster research, training, and cooperation on hyperspectral imaging at the European level. Since 2011, she has been a Postdoctoral Researcher with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Cáceres, Spain. She is currently a Professor with Sun Yat-Sen University, Guangzhou, China. Her research interests include hyperspectral image classification and segmentation, spectral unmixing, signal processing, and remote sensing.

Dr. Li is an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. She has been a Reviewer of several journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, *Pattern Recognition*, *Optical Engineering*, *Journal of Applied Remote Sensing*, and *Inverse Problems and Imaging*.

Peijun Du (M'07–SM'12) received the Ph.D. degree in geomatics engineering from China University of Mining and Technology, Xuzhou, China, in 2001.

He was a Postdoctoral Fellow with Shanghai Jiao Tong University, Shanghai, China, from February 2002 to March 2004, and was a Senior Visiting Scholar with the University of Nottingham, Nottingham, U.K., and the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab), Grenoble Institute of Technology, Grenoble, France. He is currently a Professor at the Remote Sensing Department of Geographic Information Sciences, Nanjing University, Nanjing, China, and the Deputy Director at the Key Laboratory for Satellite Mapping Technology and Applications of National Administration of Surveying, Mapping, and Geoinformation (NASG), Beijing, China. His research interests include remote sensing image processing and pattern recognition, hyperspectral remote sensing, and applications of geospatial information technologies.

Dr. Du has been the Associate Editor of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS since 2009. He was the Guest Editor of three special issues IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATION AND REMOTE SENSING. He also served as the Cochair of the Technical Committee of URBAN 2009, EORSA 2014, and IAPR-PRRS 2012, the Cochair of the Local Organizing Committee of JURSE 2009, WHISPERS 2012, and EORSA 2012, and the Member of Scientific Committee or Technical Committee of other international conferences, e.g., Spatial Accuracy 2008, ACRS 2009, WHISPERS (2010–2016), URBAN (2011, 2013, and 2015), MultiTemp (2011, 2013, and 2015), ISDIF 2011, and SPIE European Conference on Image and Signal Processing for Remote Sensing (2012–2016).

Antonio Plaza (M'05–SM'07–F'15) received the Computer Engineer degree in 1997, the M.Sc. degree in 1999, and the Ph.D. degree in 2002, all in computer engineering.

He is currently an Associate Professor (with accreditation for Full Professor) in the Department of Technology of Computers and Communications, University of Extremadura, Badajoz, Spain, where he is the Head of the Hyperspectral Computing Laboratory (HyperComp). He is the Head of the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, Caceres, Spain. His main research interests include hyperspectral data processing and parallel computing of remote sensing data. He has authored more than 500 publications, including more than 170 journal papers (more than 120 in IEEE journals), 20 book chapters, and more than 250 peer-reviewed conference proceeding papers. He has guest edited nine special issues on hyperspectral remote sensing for different journals.

Dr. Plaza is a Fellow of IEEE for contributions to hyperspectral data processing and parallel computing of Earth observation data. He received the recognition of Best Reviewers of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (in 2009) and the recognition of Best Reviewers of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (in 2010), for which he served as an Associate Editor during 2007–2012. He is also an Associate Editor for IEEE ACCESS and was a Member of the Editorial Board of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER (2011–2012) and the IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE (2013). He was also a Member of the steering committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS). He received the Best Column Award of the IEEE SIGNAL PROCESSING MAGAZINE in 2015, the 2013 Best Paper Award of the JSTARS Journal, and the most highly cited paper (2005–2010) in the Journal of Parallel and Distributed Computing. He received the best paper awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. He served as the Director of Education Activities for the IEEE GEOSCIENCE AND REMOTE SENSING SOCIETY (GRSS) in 2011–2012 and is currently serving as the President of the Spanish Chapter of IEEE GRSS. He has reviewed more than 500 manuscripts for more than 50 different journals. He is currently serving as the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING journal.

Xiuping Jia (M'93–SM'03) received the B.Eng. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 1982, and the Ph.D. degree in electrical engineering from the University of New South Wales, Kensington, Australia, in 1996.

Since 1988, she has been with the School of Information Technology and Electrical Engineering, University of New South Wales, Canberra Campus, Australia, where she is currently a Senior Lecturer. She is also a Guest Professor with China University of Petroleum, Harbin Engineering University, China, and an Adjunct Researcher at China National Engineering Research Center for Information Technology in Agriculture. She is the coauthor of the remote sensing textbook: *Remote Sensing Digital Image Analysis* (Springer-Verlag, 3rd ed., 1999 and 4th ed., 2006). Her research interests include remote sensing, image processing, and spatial data analysis.

Dr. Jia is a Subject Editor for the *Journal of Soils and Sediments*. She is also an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Xinchang Zhang received the B.Eng. degree from Wuhan Survey College, Wuhan, Hubei, China, in 1982, and the M.Eng. degree from Wuhan Survey University, Wuhan, Hubei, China, in 1990, and the Ph.D. degree from Wuhan University, Wuhan, Hubei, China, in 2004.

He has also been a Research Visitor with the London University during the year 1997–1998. Since 1998, he has been with the School of Geography and Planning, Sun Yat-Sen University, where he is currently a Professor. She is also in the University of Chinese Academy of Sciences, Beijing, China. His main research interests comprise tectonics, isotope geochemistry, field geology, philippines, neodymium, and geochronology. He has authored and coauthored more than 100 publications in both Chinese and international journals and conferences.