A New Hybrid Strategy Combining Semisupervised Classification and Unmixing of Hyperspectral Data

Inmaculada Dópido, Jun Li, *Member, IEEE,* Paolo Gamba, *Fellow, IEEE,* and Antonio Plaza, *Senior Member, IEEE*

***Abstract*—Spectral unmixing and classification have been widely used in the recent literature to analyze remotely sensed hyperspectral data. However, few strategies have combined these two approaches in the analysis. In this work, we propose a new hybrid strategy for semisupervised classification of hyperspectral data which exploits both spectral unmixing and classification in a synergetic fashion. During the process, the most informative unlabeled samples are automatically selected from the pool of candidates, thus reducing the computational cost of the process by including only the most informative unlabeled samples. Our approach integrates a well-established discriminative probabilistic classifier—the multinomial logistic regression (MLR) with different spectral unmixing chains, thus bridging the gap between spectral unmixing and classification and exploiting them together for the analysis of hyperspectral data. The effectiveness of the proposed method is evaluated using two real hyperspectral data sets, collected by the NASA Jet Propulsion Laboratory's airborne visible infrared imaging spectrometer (AVIRIS) over the Indian Pines region, Indiana, and by the reflective optics spectrographic imaging system (ROSIS) over the University of Pavia, Italy.**

***Index Terms*—Classification, hyperspectral imaging, semisupervised learning, spectral unmixing.**

I. Introduction

SPECTRAL unmixing [1] and classification [2] are two active areas of research in hyperspectral data interpretation. On the one hand, spectral unmixing is a fast growing area in which many algorithms have been recently developed to retrieve pure spectral components (*endmembers*) and to determine their abundance fractions in mixed pixels [3]. Specifically, in hyperspectral imaging, there has been a lot of interest to address the problem of mixed pixels, which arises when distinct materials are combined into a homogeneous or *intimate* mixture. This occurs independently of the spatial resolution of the sensor [4]. In order to mitigate the impact of mixed pixels, several endmember extraction [1], [5], [6] and abundance estimation algorithms [3], [7] have been developed in the literature under the assumption that a single pixel vector may comprise the response of multiple underlying materials.

On the other hand, hyperspectral image classification has also been a very active area of research in recent years [8]. Given a set of observations (i.e., possibly mixed pixel vectors), the goal of classification is to assign a unique label to each pixel, so that it is well defined by a given class [9]. Supervised classification techniques, such as the support vector machine (SVM) [10] [11] or multinomial logistic regression (MLR) [12], can deal effectively with the Hughes phenomenon [13]. However, supervised classification is generally a difficult task due to the unbalance between the high dimensionality of the data and the limited availability of labeled training samples in real analysis scenarios [14]. While the collection of labeled samples is generally difficult, expensive, and time-consuming, unlabeled samples can be generated in a much easier way [15]. This observation has fostered the idea of adopting semisupervised learning techniques in hyperspectral image classification.

A survey of semisupervised learning algorithms is available in [16]. These algorithms generally assume that a limited number of labeled samples are available *a priori*, and then enlarge the training set using unlabeled samples, thus allowing these approaches to address ill-posed problems [17]. However, in order for this strategy to be successful, several requirements need to be satisfied. First and foremost, the new (unlabeled) samples should be generated with relatively low cost/effort. Second, the number of unlabeled samples required in order for the semisupervised classifier to perform properly should not be too high in order to avoid increasing the computational complexity of the classification stage. In other words, as the number of unlabeled samples increases, it may be unbearable for the classifier to properly exploit all the available training samples due to computational issues. Further, if the unlabeled samples are not properly selected, these may confuse the classifier, thus introducing significant divergence or even reducing the classification accuracy obtained with the initial set of labeled samples. In order to address these issues, recent literature has explored several strategies [18], [19] in order to effectively design semisupervised self-learning techniques without the need for human intervention [20].

At this point, it is important to reiterate that the analysis of hyperspectral images is not an easy task. This is due to the great variability of hyperspectral signatures and the high dimensionality of the data. Another problem is the intrinsic nature of the pixels, which may be highly mixed. The most traditional approach in the literature to describe the phenomenon of the mixture at subpixel levels is the linear mixture model [1]. As opposed to nonlinear unmixing [4], which generally requires

Manuscript received January 10, 2014; revised March 27, 2014; accepted April 24, 2014. Date of publication May 21, 2014; date of current version October 03, 2014. This work was supported in part by the CARIPLO 2009-2936 "Azioni di internazionalizzazione per il post-laurea nell'ambito delle tecnologie dell ICT e biomediche," in part by the National Basic Research Program of China (973 Program, Grant 2011CB707103), and in part by the Spanish Ministry of Science and Innovation (research project CEOS-SPAIN with reference AYA2011-29334-C02-02). *(Corresponding author: Jun Li.)*

I. Dópido and A. Plaza are with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, Cáceres E-10003, Spain.

J. Li is with the School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510275, China (e-mail: junli48@mail.sysu.edu.cn).

P. Gamba is with the Telecommunications and Remote Sensing Laboratory, University of Pavia, Pavia I-27100, Italy.

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.2014.2322143

detailed information about physical properties that may not be always available, linear spectral unmixing consists of identifying the pure spectral components or endmembers. When the pure spectral signatures are identified, the proportion of each material in each pixel can be estimated and this information can be used as a form of *soft* classification. As a result, abundances provide additional information about the composition of each pixel. If this information is properly exploited, it may properly complement the results provided by traditional *hard* classification techniques.

In this work, we develop a new strategy for hyperspectral data interpretation that combines hyperspectral unmixing and classification in order to exploit both sources of information in complementary nature, thus overcoming the limitations of using these techniques in separate fashion. The result is a new hybrid method that incorporates information obtained from hyperspectral unmixing into the classification process, with the possibility to control the relative weight of unmixing with regards to classification and vice versa. Specifically, we include the outcome of spectral unmixing into a semisupervised classification process in which self-learning is used to select the most useful unlabeled samples.

In the recent literature, some efforts have been developed toward the integration of spectral unmixing and classification. For instance, in [21], spectral unmixing was used as a feature extraction strategy prior to classification. It was found that the features obtained by unmixing can be often associated to physical features in the scene as opposed to statistical approaches for feature extraction such as the principal component analysis (PCA) [9] or the minimum noise fraction (MNF) [22], in which the physical meaning of the features is generally lost. This idea was expanded in [23] by analyzing additional unmixing chains for feature extraction prior to classification, including chains which perform the unmixing based on the available training samples (used as endmembers for mixture characterization) and chains which integrate spatial and spectral information in spectral unmixing via clustering techniques. More recently, the synergetic nature of spectral unmixing and classification has been further explored in the context of a semisupervised self-learning framework [24]. This strategy provides a joint approach for hyperspectral data interpretation that considers simultaneously the output provided by both unmixing and classification, where the weight given to either technique can be controlled by the end user to adjust the classification output. As a follow-up to the work in [24], here we explore the possibility of using spectral unmixing to improve semisupervised classification. Our speculation is that spectral unmixing provides a useful source of information to classification with the ability to further interpret mixed pixels, possibly changing their classification labels accordingly, in particular, for classes that are dominated by highly mixed pixels.

The remainder of the paper is organized as follows. Section II presents the newly developed hybrid strategy for combining semisupervised classification and spectral unmixing. Section III presents our experimental setting, with emphasis on describing the hyperspectral scenes considered in experiments, and further reports the experimental results conducted with real hyperspectral scenes in order to validate the presented technique. Section IV concludes the paper with some remarks and hints at plausible future research lines.

II. Proposed Approach

Let $\mathcal{K} \equiv \{1, \ldots, K\}$ denote a set of K class labels, $\mathcal{S} \equiv \{1, \ldots, n\}$ a set of integers indexing the n pixels of an image, $\mathbf{x} \equiv (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^{d \times n}$ an image of d-dimensional feature vectors, $\mathbf{y} \equiv (y_1, \ldots, y_n)$ an image of labels, $\mathcal{D}_l \equiv \{(y_{l_1}, \mathbf{x}_{l_1}), \ldots, (y_{l_n}, \mathbf{x}_{l_n})\}$ a set of labeled samples, l_n the number of labeled training samples, $\mathcal{Y}_l \equiv \{y_{l_1}, \ldots, y_{l_n}\}$ the set of labels in $\mathcal{D}_l$, $\mathcal{X}_l \equiv \{\mathbf{x}_{l_1}, \ldots, \mathbf{x}_{l_n}\}$ the set of feature vectors in $\mathcal{D}_l$, $\mathcal{D}_u \equiv \{\mathcal{X}_u, \mathcal{Y}_u\}$ a set of unlabeled samples, $\mathcal{X}_u \equiv \{\mathbf{x}_{u_1}, \ldots, \mathbf{x}_{u_n}\}$ the set of unlabeled feature vectors in $\mathcal{D}_u$, $\mathcal{Y}_u \equiv \{y_{u_1}, \ldots, y_{u_n}\}$ the set of labels associated with $\mathcal{X}_u$, and u_n the number of unlabeled samples. With this notation in mind, the proposed approach consists of three main ingredients: 1) semisupervised learning; 2) self-learning; and 3) spectral unmixing, which are described before introducing our method.

A. Semisupervised Learning

For the semisupervised part of our approach, we use the MLR probabilistic classifier to model the class posterior density. This classifier is formally given by [25]

$$p(y_i = k|\mathbf{x}_i) = \frac{\exp(\boldsymbol{\omega}^{(k)^T}\mathbf{h}(\mathbf{x}_i))}{\sum_{k=1}^{K}\exp(\boldsymbol{\omega}^{(k)^T}\mathbf{h}(\mathbf{x}_i))} \tag{1}$$

where $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), \ldots, h_l(\mathbf{x})]^T$ is a vector of l fixed functions of the input, often termed features; $\boldsymbol{\omega}$ are the regressors; and $\boldsymbol{\omega} = [\boldsymbol{\omega}^{(1)^T}, \ldots, \boldsymbol{\omega}^{(K)^T}]^T$. Notice that the function $\mathbf{h}$ may be linear, i.e., $\mathbf{h}(\mathbf{x}_i) = [1, x_{i,1}, \ldots, x_{i,d}]^T$, where $x_{i,j}$ is the jth component of $\mathbf{x}_i$ or nonlinear, i.e., $\mathbf{h}(\mathbf{x}_i) = [1, K_{\mathbf{x}_i,\mathbf{x}_1}, \ldots, K_{\mathbf{x}_i,\mathbf{x}_l}]^T$, where $K_{\mathbf{x}_i,\mathbf{x}_j} = K(\mathbf{x}_i, \mathbf{x}_j)$ and $K(\cdot,\cdot)$ is some symmetric kernel function. Kernels have been largely used because they tend to improve the data separability in the transformed space. In this paper, we use a Gaussian radial basis function (RBF) $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2)$ kernel, which is widely used in hyperspectral image classification [11]. The parameter σ (spread of the Gaussian RBF kernel) was chosen by tenfold cross-validation. We selected this kernel (after extensive experimentation using other kernels, including linear and polynomial kernels) because we empirically observed that it provided very good results in our context.

Under the present setup, learning the class densities amounts to estimate the logistic regressors. Following the work in [26], we can compute $\boldsymbol{\omega}$ by obtaining the maximum *a posteriori* (MAP) estimate

$$\widehat{\boldsymbol{\omega}} = \arg\max_{\boldsymbol{\omega}} \ell(\boldsymbol{\omega}) + \log p(\boldsymbol{\omega}) \tag{2}$$

where $p(\boldsymbol{\omega}) \propto \exp(-\lambda\|\boldsymbol{\omega}\|_1)$ is a Laplacian prior to promote sparsity and λ is a regularization parameter controlling the degree of sparseness of $\widehat{\boldsymbol{\omega}}$ in [26]. In our previous work [27], it was shown that parameter λ is rather insensitive to the use of different data sets, and that there are many suboptimal values for this parameter, which lead to very accurate estimation of parameter

$\boldsymbol{\omega}$. In our experiments, we set $\lambda = 0.001$, as we have empirically found that this parameter setting provides very good performance [12]. Finally, $\ell(\boldsymbol{\omega})$ is the log-likelihood function over the training samples $\mathcal{D}_{l+u} \equiv \mathcal{D}_l + \mathcal{D}_u$, given by

$$\ell(\boldsymbol{\omega}) \equiv \sum_{i=1}^{l_n+u_n} \log p(y_i = k|\mathbf{x}_i, \boldsymbol{\omega}). \quad (3)$$

As shown in (3), labeled and unlabeled samples are integrated to learn the regressors $\boldsymbol{\omega}$. In Section II-B, we describe in more details how unlabeled samples are selected in the proposed approach.

B. Self-Learning

The considered semisupervised approach belongs to the family of self-learning approaches, where the training set $\mathcal{D}_{l+u}$ is incremented under the following criterion. Let $\mathcal{D}_{\mathcal{N}(i)} \equiv \{(\widehat{y}_{i_1}, \mathbf{x}_{i_1}), \ldots, (\widehat{y}_{i_n}, \mathbf{x}_{i_n})\}$ be the set of 4-connected neighboring samples of $(y_i, \mathbf{x}_i)$ for $i \in \{l_1, \ldots, l_n, u_1, \ldots, u_n\}$, where i_n is the number of samples in $\mathcal{D}_{\mathcal{N}(i)}$ and $\widehat{y}_{i_j}$ is the MAP estimate from the MLR classifier, with $i_j \in \{i_1, \ldots, i_n\}$. If $\widehat{y}_{i_j} = y_i$, we increment the unlabeled training set by adding $(\widehat{y}_{i_j}, \mathbf{x}_{i_j})$, i.e., $\mathcal{D}_u = \{\mathcal{D}_u, (\widehat{y}_{i_j}, \mathbf{x}_{i_j})\}$. This increment is reasonable due to the following considerations: first, from a global viewpoint, samples which have the same spectral structure likely belong to the same class and second, from a local viewpoint, it is very likely that two neighboring pixels also belong to the same class. Therefore, the newly included samples are reliable for learning the classifier. In this work, we run an iterative scheme to increment the training set, as this strategy can refine the estimates and enlarge the neighborhood set such that the set of potential unlabeled training samples is increased. Once a candidate set is inferred using spatial information, we adopt a self-learning to automatically (and intelligently) select the most informative samples from the candidate set [20]. As a result, in the proposed semisupervised self-learning scheme, our aim is to select the most informative samples without the need for human supervision.

Let $\mathcal{D}_c$ be the newly generated unlabeled training set at each iteration, which meets the criteria of the considered semisupervised algorithm. Notice that the self-learning step in the proposed approach leads to high confidence in the class labels of the newly generated set $\mathcal{D}_c$. Now we can select the most informative samples in $\mathcal{D}_c$ to obtain the set $\mathcal{D}_u$, i.e., samples with high uncertainty, such that $\mathcal{D}_u \subseteq \mathcal{D}_c$. In our study, we use the breaking ties (BT) algorithm [28] for that purpose. This algorithm relies on the smallest difference of the posterior probabilities for each sample. In a multiclass setting, the algorithm can be applied (independently of the number of classes available) by calculating the difference between the two highest probabilities. As a result, the algorithm finds the samples minimizing the distance between the first two most probable classes.

C. Spectral Unmixing

Several spectral unmixing chains based on the well-known linear mixture model have been used in this paper as an alternative way to estimate the probability $p(y_i = k|\mathbf{x}_i)$. The linear mixture model generally satisfies the abundance nonnegativity and abundance sum-to-one constraints at a pixel level [4]. As a result, the abundance fractions can be interpreted as probabilities for endmember classes. In the following, we outline the specific spectral unmixing chains that have been considered in this paper. These chains are described in more detail in [21], where the discussed chains are mostly based on least-squares linear spectral unmixing (FCLSU) [29]. However, the unmixing-based chains in [21] do not include spatial information, which is an important source of information since hyperspectral images exhibit spatial correlation between image features. The study in [23] suggested that partial unmixing using mixture-tuned matched filtering (MTMF) [30] could be an effective solution to deal with the likely fact that not all pure spectral constituents in the scene (needed for spectral unmixing purposes) are known *a priori*, but a more exhaustive investigation of partial unmixing (in combination with spatial information) is desirable. In this work, we consider several unmixing strategies addressing the aforementioned issues as discussed in [23].

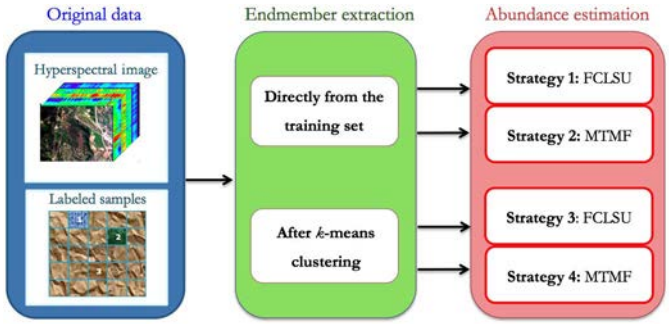

Fig. 1. Summary of different spectral unmixing chains considered in this paper.

1) *Strategy 1*. This strategy first assumes that the labeled samples are made up of spectrally pure constituents (endmembers) and then calculates their abundances by means of the FCLSU method and provides a set of fractional abundance maps (one per labeled class).
2) *Strategy 2*. This strategy also assumes that the labeled samples are made up of spectrally pure constituents (endmembers) but now calculates their abundances by means of the MTMF method, thus providing a set of fractional abundance maps (one per labeled class). In [23], it is shown that MTMF can outperform other techniques for abundance estimation such as FCLSU, since it can provide meaningful abundance maps by means of partial unmixing in case not all endmembers are available *a priori*.
3) *Strategy 3*. This strategy is intended to solve the problems highlighted by endmember extraction algorithms which are sensitive to outliers and pixels with extreme values of reflectance. Using an unsupervised clustering method such as the k-means on the available labeled samples, the endmembers extracted (from class centers) are expected to be more spatially significant. Then, FLCSU-based unmixing is conducted using the resulting endmembers.
4) *Strategy 4*. This strategy is very similar to the previous one, but in this case, MTMF-based unmixing is conducted using the resulting endmembers after k-means clustering. For illustrative purposes, Fig. 1 summarizes the different spectral unmixing chains considered in this work.

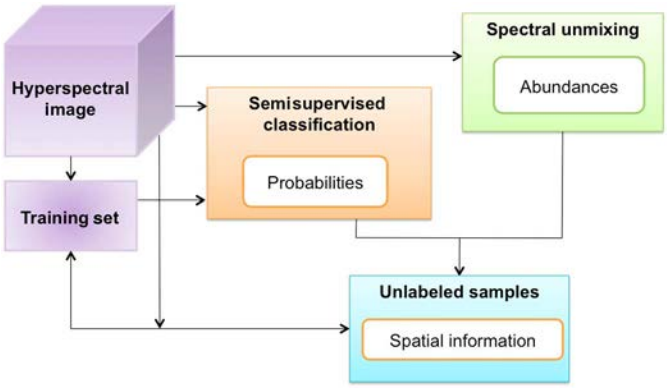

Fig. 2. General flowchart describing the newly proposed semisupervised classification method. Our methodology uses spectral unmixing to refine the results obtained by a semisupervised process based on a probabilistic classifier. The unlabeled samples are intelligently selected using spatial information.

D. Proposed Semisupervised Hybrid Classification Method

With the aforementioned modules in mind, the probabilistic estimation performed by the proposed semisupervised hybrid classification method can be simply described as follows:

$$\widehat{p}_i(y_i = k|\mathbf{x}_i) = \alpha f_c(y_i = k|\mathbf{x}_i) + (1-\alpha) f_u(y_i = k|\mathbf{x}_i) \quad (4)$$

where $\widehat{p}_i(\cdot)$ is the joint estimate for the kth class, i.e., $y_i = k$, obtained by the classification and unmixing methods given by observation $\mathbf{x}_i$, where $\widehat{p}_i(\cdot)$ will serve as the indicator, i.e., probability, for the semisupervised self-learning process. In this paper, function $f_c(\cdot)$ is the probability obtained by the classification algorithm, i.e., the MLR described in Section II-A; and function $f_u(\cdot)$ is the abundance fraction obtained by one of the spectral unmixing chains presented in Section II-C. The balance between the classification probabilities and the abundance fractions is controlled by parameter α, where $0 \leq \alpha \leq 1$. As shown in (4), if $\alpha = 1$, only classification probabilities are considered by the proposed strategy, which leads to the semisupervised self-learning strategy presented in Section II-B. On the other hand, if $\alpha = 0$, only spectral unmixing is considered in the adopted self-learning strategy. Therefore, by tuning α to a value ranging between 0 and 1, we can adjust the impact between classification and unmixing in the semisupervised self-learning process. Moreover, by means of parameter α, the proposed hybrid strategy takes advantage from both classification and unmixing, such that the new unlabeled samples selected are more informative in comparison with those samples selected only from classification or unmixing methods. For illustrative purposes, Fig. 2 provides a general flowchart summarizing the newly proposed semisupervised classification method.

III. Experimental Results

In this section, we evaluate the new methodology presented in this paper using two different hyperspectral images: AVIRIS Indian Pines and ROSIS Pavia University. In our experiments, we apply the MLR classifier (with Gaussian RBF kernel) to a normalized version of the considered hyperspectral data sets.[1] In all cases, the reported figures of overall accuracy (OA), average accuracy (AA), and κ statistic are obtained by averaging the results obtained after conducting 10 independent Monte Carlo runs with respect to the labeled training set from the ground-truth data, where the remaining samples are used for validation purposes. In order to illustrate the good performance of the proposed approach, we use very small labeled training sets on purpose.

The main difficulties that our proposed approach should circumvent can be summarized as follows. First and foremost, it is very difficult for supervised algorithms to provide good classification results, as very little information is available about the class distribution. In addition, poor generalization is also a risk when estimating class boundaries in scenarios dominated by limited training samples. Since our approach is semisupervised, we take advantage of unlabeled samples in order to improve classification accuracy. However, if the number of labeled samples is very small, increasing the number of unlabeled samples could bias the learning process. This effect is explored in the remainder of this Section III, which is organized as follows. Section III-A describes the two considered hyperspectral images. In Section III-B, we study the impact of parameter α (which balances unmixing and classification in our proposed hybrid strategy). Finally, Sections III-C and III-D describe the experiments conducted with the AVIRIS Indian Pines and ROSIS Pavia University scenes, respectively. In all cases, the results obtained by the supervised version of the considered classifier are also reported for comparative purposes.

A. Hyperspectral Data Sets

In order to evaluate the proposed approach, we use two different hyperspectral data sets. Particularly, we have selected two images that provide different characteristics in terms of spatial and spectral resolution, with the ultimate goal of validating the method in different scenarios. The scenes are collected by two different sensors: 1) airborne visible infrared imaging spectrometer (AVIRIS) and 2) reflective optics spectrographic imaging system (ROSIS). In the following, we describe the two considered scenes for experiments.

1) AVIRIS Indian Pines: The first data set used in our experiments was collected by the AVIRIS sensor over the Indian Pines region in Northwestern Indiana in 1992. This scene, with a size of 145 lines $\times$ 145 samples, was acquired over a mixed agricultural/forest area, early in the growing season. The scene comprises 202 spectral channels in the wavelength range from 0.4 to 2.5 μm, nominal spectral resolution of 10 nm, moderate spatial resolution of 20 m by pixel, and 16-b radiometric resolution. After an initial screening, several spectral bands were removed from the data set due to noise and water absorption phenomena, leaving a total of 164 radiance channels to be used in the experiments. For illustrative purposes, Fig. 3(a) shows a false color composition of the AVIRIS Indian Pines scene, whereas Fig. 3(b) shows the ground-truth map available for the scene, displayed in the form of a class assignment for each labeled pixel, with 16 mutually exclusive ground-truth classes. These data, including ground-truth information, are available online,[2] a fact

[1]The normalization is simply given by $\mathbf{x}_i := \frac{\mathbf{x}_i}{(\sqrt{\sum \|\mathbf{x}_i\|^2})}$, for $i = 1, \ldots, n$, where $\mathbf{x}_i$ is a spectral vector.

[2][Online]. Available: http://dynamo.ecn.purdue.edu/biehl/MultiSpec.

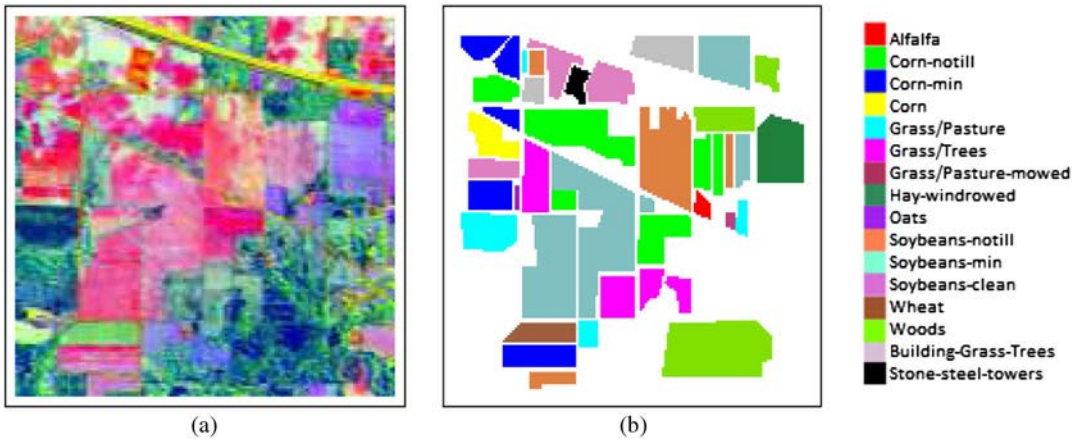

Fig. 3. (a) False color composition of the AVIRIS Indian Pines scene. (b) Ground-truth map containing 16 mutually exclusive land-cover classes (right).

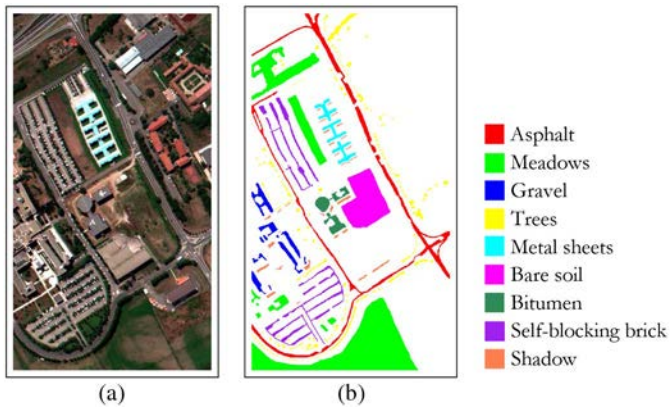

Fig. 4. (a) False color composition of the ROSIS Pavia scene. (b) Ground-truth map containing nine mutually exclusive classes.

Fig. 5. Map showing the spatial distribution of the selected $u_n = 300$ unlabeled samples (represented as black dots) in one of the classification experiments (starting with $l_n = 10$ labeled samples) conducted with the AVIRIS Indian Pines scene.

which has made this scene a widely used benchmark for testing the accuracy of hyperspectral data classification algorithms.

2) ROSIS Pavia University: The second data set used in experiments was collected by the ROSIS optical sensor over the urban area of the University of Pavia, Italy. The flight was operated by the Deutschen Zentrum for Luftund Raumfahrt (DLR, the German Aerospace Agency) in the framework of the HySens project, managed and sponsored by the European Commission. The image size in pixels is 610×340, with very high spatial resolution of 1.3 m per pixel. The number of data channels in the acquired image is 115 (with spectral range from 0.43 to 0.86 μm). Fig. 4(a) shows a false color composite of the image, whereas Fig. 4(b) shows nine ground-truth classes of interest.

B. Analysis of the Balance Between Classification and Unmixing

In this set of experiments, we evaluate the impact of parameter α controlling the relative weight of classification and unmixing in the proposed hybrid classifier. Here, the semisupervised classifier is trained using only $l_n = 5$ labeled samples per class, using the different spectral unmixing chains and testing the following values for parameter α that controls the relative weight between classification and unmixing: $\alpha = \{1, 0.75, 0.5, 0.25, 0\}$. We remind that $\alpha = 1$ means that only classification probabilities are used, whereas $\alpha = 0$ means that only unmixing information is used by the proposed method. In all cases, we execute 300 iterations to intelligently select $u_n = 300$ unlabeled samples with the BT algorithm. This is in fact the optimal configuration, as the best unlabeled sample candidate is included in each iteration. There is also a possibility to select more than one unlabeled sample per iteration in order to reduce the computational complexity of the method. However, in our experiments, we select only the best candidate per iteration. For illustrative purposes, Fig. 5 shows a map with the spatial distribution of the selected unlabeled samples for one of the experiments conducted with the AVIRIS Indian Pines image, whereas Fig. 6 shows a map with the spatial distribution of the selected unlabeled samples for one of the experiments conducted with the ROSIS Indian Pines scene. It should be noted that the selected configuration is considered to address a difficult case in which very limited training samples are used for the initial condition, whereas the semisupervised process generates a reasonable

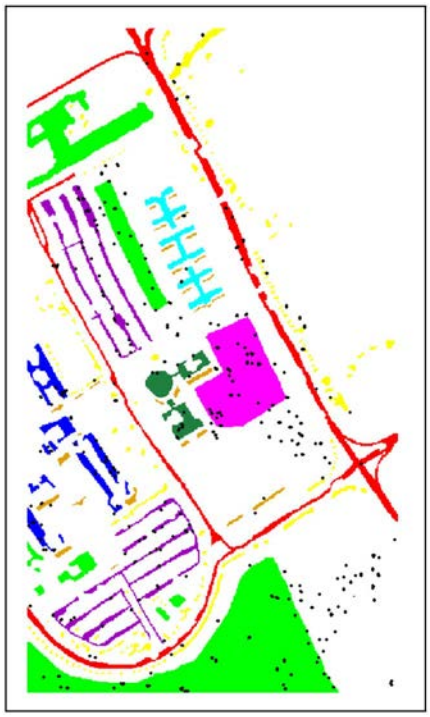

Fig. 6. Map showing the spatial distribution of the selected $u_n = 300$ unlabeled samples (represented as black dots) in one of the classification experiments (starting with $l_n = 10$ labeled samples) conducted with the ROSIS Pavia University scene.

number of unlabeled training samples using the strategy described in Fig. 2.

Table I shows the results obtained under the above-mentioned configuration for the AVIRIS Indian Pines and ROSIS Pavia University scenes. In both cases, a resonable compromise is obtained using $\alpha = 0.75$, which means that classification generally needs more weight than unmixing in order to obtain the best analysis results from our proposed hybrid classifier. This is expected, since the information provided by classification is indeed crucial in our framework. However, our experiments show that this information can be further refined by including spectral unmixing information also in the process. The experiments reported in Table I also confirm that the use of spectral unmixing information alone ($\alpha = 0$) cannot improve the results obtained by probabilistic information alone ($\alpha = 1$). The results in this experiment confirm our introspection that the joint exploitation of classification and unmixing provides advantages over the use of either techniques alone, particularly in the framework of semisupervised classification using limited training samples.

C. Experimental Results With AVIRIS Indian Pines

In this section, the proposed approach is evaluated using the AVIRIS Indian Pines data set described in Fig. 3. In this experiment, we consider different numbers of labeled samples per class: $l_n = \{5, 10, 15\}$. Table II shows the OA, AA, and the κ statistic obtained by the supervised strategy (based on the MLR classifier) and by the proposed semisupervised approach, using two different strategies for unlabeled sample selection: 1) random sampling (RS) and 2) BT, both executed using 300 iterations to select $u_n = 300$ unlabeled samples). As a result, in Table II, BT denotes the semisupervised classifier with BT for unlabeled sample selection. RS denotes the semisupervised classifier with unlabeled samples selected using RS. Finally, Strategies 1–4 denote the semisupervised hybrid classifier integrating classification and spectral unmixing (with $\alpha = 0.75$), where unlabeled samples are selected using BT. It should be noted that the selected classification scenario represents a very challenging one. For instance, when $l_n = 5$ labeled samples are used per class, only 80 labeled samples in total are used as the initial condition for the considered classifier, which is much lower than the number of spectral bands available in the scene.

As we can observe in Table II, the inclusion of unlabeled samples significantly improved the classification results in all cases. If we compare the supervised case with the semisupervised techniques, we can observe that the unlabeled samples always significantly help to improve the accuracy results. When the proposed strategy (combining classification and spectral unmixing) was used, a significant improvement is observed over the supervised case (regardless of the unmixing strategy used).

In Fig. 7, we evaluate the impact of the number of unlabeled samples on the classification performance achieved by the presented strategy. Specifically, we plot the OAs in classification accuracy obtained by the supervised MLR (trained using only 5, 10, and 15 labeled samples per class) and by the proposed approach (based on the same classifier plus spectral unmixing) using the four considered strategies for including unmixing information, as a function of the number of unlabeled samples. For the selection of unlabeled samples, we considered two strategies: 1) RS and 2) BT. As shown in Fig. 7, the classification accuracies increase as the number of labeled training samples increases. This is important, since these labels are obtained with relatively low cost and effort, and in self-learning fashion.

For illustrative purposes, Fig. 8 shows some of the classification maps obtained for the AVIRIS Indian Pines scene. These classification maps correspond to one of the 10 Monte-Carlo runs that were averaged in order to generate the classification scores reported in Table II. The advantages obtained by adopting a semisupervised learning approach, which combines classification and unmixing concepts, can be clearly appreciated in the classification maps displayed in Fig. 8, which also reports the classification OAs obtained for each method in parentheses.

D. Experimental Results With ROSIS Pavia University

The second data set used in experiments is the ROSIS Pavia University described in Fig. 4. Table III shows the OA, AA, and κ statistic obtained by the supervised strategy (trained with $l_n = \{5, 10, 15\}$ labeled samples per class) and by the proposed semisupervised approach (based on the MLR classifier and different spectral unmixing chains), using two different strategies for selecting unlabeled samples: 1) RS and 2) BT (both executed using 300 iterations to select $u_n = 300$ unlabeled samples).

TABLE I
OA (%) OBTAINED FOR DIFFERENT VALUES OF PARAMETER α IN THE ANALYSIS OF THE AVIRIS INDIAN PINES AND ROSIS PAVIA UNIVERSITY HYPERSPECTRAL DATA SETS, USING ONLY $l_n = 5$ LABELED SAMPLES PER CLASS

	AVIRIS Indian Pines					ROSIS Pavia University				
α	1.0	0.75	0.50	0.25	0.00	1.0	0.75	0.50	0.25	0.00
Strategy 1	65.25	**69.10**	55.60	55.47	52.01	75.48	**78.05**	68.20	72.39	64.41
Strategy 2	65.25	**68.89**	66.45	68.76	58.90	75.48	**79.39**	78.80	77.76	71.42
Strategy 3	65.25	**69.83**	51.90	59.23	52.64	75.48	**79.53**	65.10	66.88	64.82
Strategy 4	65.25	**70.45**	66.34	60.50	60.71	75.48	**79.53**	73.58	71.08	69.11

The four considered spectral unmixing strategies are compared. In all cases, we execute 300 iterations to intelligently select $u_n = 300$ unlabeled samples.

TABLE II
OA, AA (%), AND κ STATISTIC OBTAINED USING DIFFERENT CLASSIFICATION TECHNIQUES WHEN APPLIED TO THE AVIRIS INDIAN PINES HYPERSPECTRAL DATA SET

	$l_n = 5$ labeled samples per class						
	Supervised	BT	RS	Strategy 1	Strategy 2	Strategy 3	Strategy 4
OA	51.78 ± 2.51	65.25 ± 5.76	60.03 ± 2.91	69.10 ± 2.66	68.89 ± 4.13	69.83 ± 1.78	70.45 ± 1.78
AA	63.82 ± 2.69	69.98 ± 2.91	66.31 ± 2.93	72.73 ± 2.24	70.81 ± 2.24	71.37 ± 2.49	72.32 ± 2.20
κ	46.26 ± 2.74	60.65 ± 6.05	54.81 ± 3.19	64.82 ± 2.94	64.54 ± 4.31	65.58 ± 2.08	66.25 ± 2.16
	$l_n = 10$ labeled samples per class						
	Supervised	BT	RS	Strategy 1	Strategy 2	Strategy 3	Strategy 4
OA	60.12 ± 3.08	68.51 ± 5.16	65.59 ± 2.94	70.91 ± 4.87	74.91 ± 2.10	75.05 ± 1.68	75.29 ± 2.40
AA	71.74 ± 1.54	75.66 ± 2.33	73.49 ± 1.83	77.13 ± 2.70	78.21 ± 2.36	78.44 ± 1.96	79.05 ± 2.00
κ	55.43 ± 3.20	64.40 ± 5.56	61.17 ± 3.16	67.08 ± 5.40	70.69 ± 2.40	71.60 ± 1.91	71.84 ± 2.75
	$l_n = 15$ labeled samples per class						
	Supervised	BT	RS	Strategy 1	Strategy 2	Strategy 3	Strategy 4
OA	66.20 ± 1.99	70.46 ± 3.16	70.29 ± 1.93	73.57 ± 3.83	74.04 ± 2.26	76.89 ± 2.56	76.68 ± 2.04
AA	77.39 ± 1.06	79.47 ± 1.81	78.76 ± 1.09	80.86 ± 1.85	79.76 ± 1.26	81.90 ± 1.46	81.94 ± 1.63
κ	62.09 ± 2.13	66.65 ± 3.38	66.41 ± 2.13	70.16 ± 4.26	70.51 ± 2.49	73.71 ± 2.85	73.53 ± 2.22

The standard deviation is also reported in each case.

Several conclusions can be obtained from the results reported in Table III. First and foremost, the use of unlabeled samples provides advantages with respect to the supervised algorithm alone. In all cases, the considered approach led to significant improvements, particularly when spectral unmixing was used. These unlabeled samples are automatically selected by the proposed approach, and represent no cost in terms of data collection or human supervision.

In Fig. 9, we can observe how the accuracy results also improves for the ROSIS Pavia University scene as the number of unlabeled samples increases. For instance, in the case with $l_n = 10$ labeled samples per class (see Table III), the supervised approach obtained an OA of 69.25%. When the proposed strategy (combining classification and spectral unmixing) was used, the classification accuracies improved to an OA of 84.14% (Strategy 3), which represents a very significant improvement over the supervised case. For illustrative purposes, Fig. 10 shows some of the classification maps obtained in our experiments with the ROSIS Pavia University scene. These maps correspond to one of the 10 Monte-Carlo runs that were averaged in order to generate the classification scores reported in Table III. The advantages obtained by adopting our semisupervised learning approach can be appreciated in these maps.

IV. CONCLUSION AND FUTURE RESEARCH LINES

In this paper, we have developed a new semisupervised hybrid technique to combine concepts of spectral unmixing and classification. This has been done by injecting spectral unmixing information in the semisupervised classification process using self-learning concepts. In this context, it is important to define what is the relative weight given to unmixing with regards to classification and vice versa. Several experiments have been performed in order to analyze this issue. Our conclusion is that there are several potential advantages of jointly considering spectral unmixing and classification, which are apparent in the classification accuracies obtained by the proposed semisupervised hybrid framework when used to analyze two popular

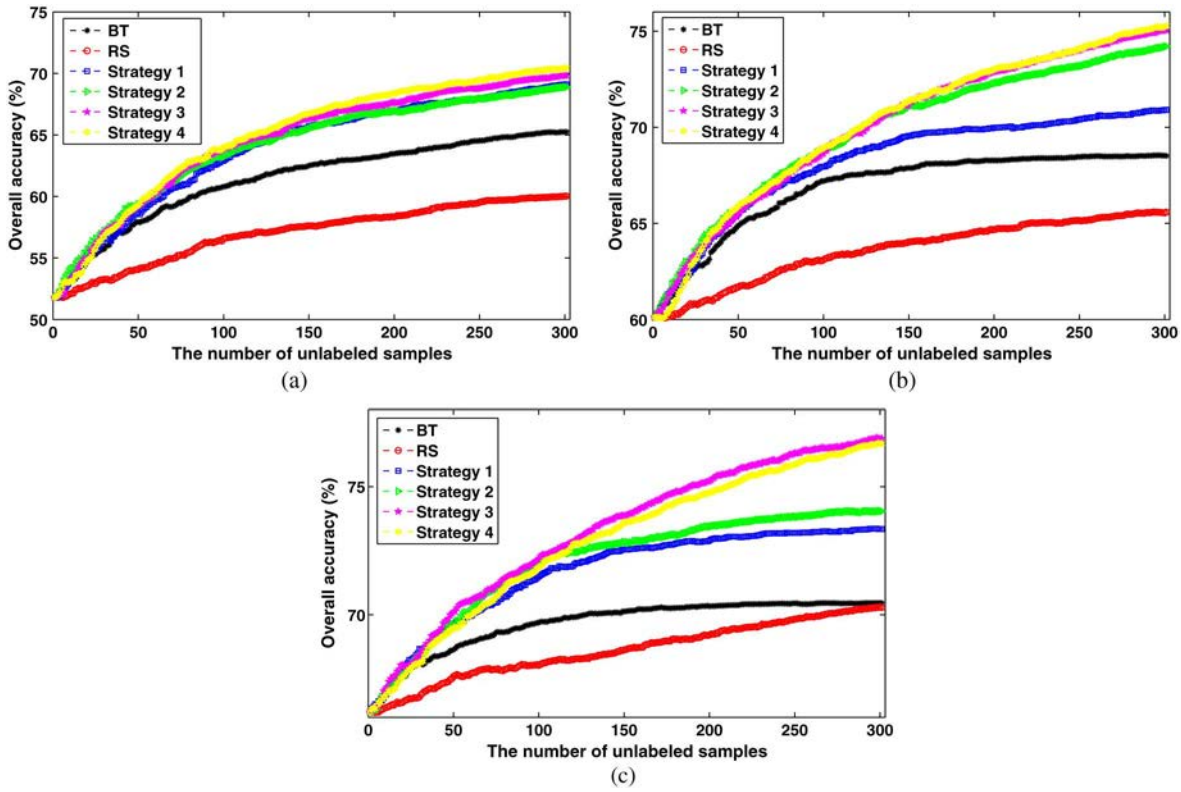

Fig. 7. OA (as a function of the number of unlabeled samples) obtained for the AVIRIS Indian Pines data set by the proposed semisupervised classifier using different strategies. For the unmixing-based strategies, $\alpha = 0.75$ in all cases. (a) $l_n = 5$ labeled samples per class. (b) $l_n = 10$ labeled samples per class. (c) $l_n = 15$ labeled samples per class.

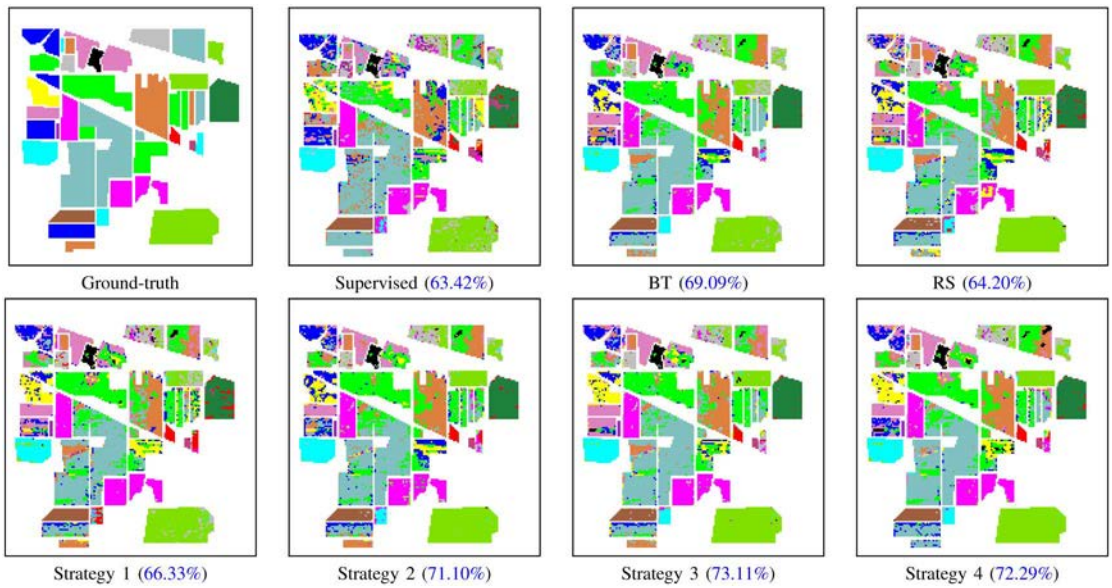

Fig. 8. Classification maps and OAs (in parentheses) obtained after applying classification techniques to the AVIRIS Indian Pines data set. In all cases, the number of labeled samples was $l_n = 10$ and the number of unlabeled samples was $u_n = 300$.

hyperspectral scenes in the community: 1) AVIRIS Indian Pines and 2) ROSIS Pavia University. To the best of our knowledge, this paper represents one of the first efforts in the literature in order to exploit unmixing and classification in a synergetic fashion. These approaches have been traditionally exploited independently. In this regard, a detailed exploration of the remaining connections and possible bridges between both techniques represent a topic that deserves future work.

In future, we will explore the possibility of developing a more intelligent stopping rule for the proposed approach. We are also planning on testing the presented developments on large-scale data repositories, in order to facilitate the processing of larger

TABLE III
OA, AA (%), AND κ STATISTIC (IN PARENTHESES) OBTAINED USING DIFFERENT CLASSIFICATION STRATEGIES WHEN APPLIED TO THE ROSIS PAVIA UNIVERSITY HYPERSPECTRAL DATA SET

	$l_n = 5$ labeled samples per class						
	Supervised	BT	RS	Strategy 1	Strategy 2	Strategy 3	Strategy 4
OA	63.56 ± 4.63	75.48 ± 4.63	71.80 ± 2.16	78.05 ± 3.20	79.09 ± 1.94	79.53 ± 2.12	79.10 ± 2.34
AA	72.93 ± 2.08	79.33 ± 2.62	75.45 ± 1.86	79.62 ± 2.45	79.68 ± 3.63	79.97 ± 2.43	80.29 ± 2.67
κ	54.78 ± 4.45	68.30 ± 5.38	63.50 ± 2.12	71.54 ± 3.69	72.69 ± 2.73	73.05 ± 2.63	72.48 ± 3.03
	$l_n = 10$ labeled samples per class						
	Supervised	BT	RS	Strategy 1	Strategy 2	Strategy 3	Strategy 4
OA	69.25 ± 3.75	80.70 ± 3.07	75.72 ± 2.19	82.86 ± 2.30	83.31 ± 2.04	84.14 ± 1.97	83.93 ± 1.73
AA	78.42 ± 1.75	82.81 ± 1.55	80.52 ± 1.52	83.48 ± 1.40	83.51 ± 1.95	84.48 ± 1.04	84.83 ± 1.53
κ	61.69 ± 4.01	74.87 ± 3.75	68.95 ± 2.54	77.68 ± 2.78	78.06 ± 2.55	79.23 ± 2.37	78.97 ± 2.07
	$l_n = 15$ labeled samples per class						
	Supervised	BT	RS	Strategy 1	Strategy 2	Strategy 3	Strategy 4
OA	72.34 ± 2.22	81.00 ± 5.75	76.88 ± 2.08	83.54 ± 2.47	83.47 ± 2.18	84.83 ± 3.21	85.25 ± 2.59
AA	80.01 ± 2.29	83.35 ± 2.82	81.61 ± 2.46	83.62 ± 2.43	83.60 ± 2.34	85.54 ± 2.08	85.48 ± 1.22
κ	65.21 ± 2.23	75.44 ± 3.80	70.39 ± 2.09	78.49 ± 2.85	78.56 ± 2.68	80.15 ± 3.84	80.63 ± 3.11

The standard deviation is also reported in each case.

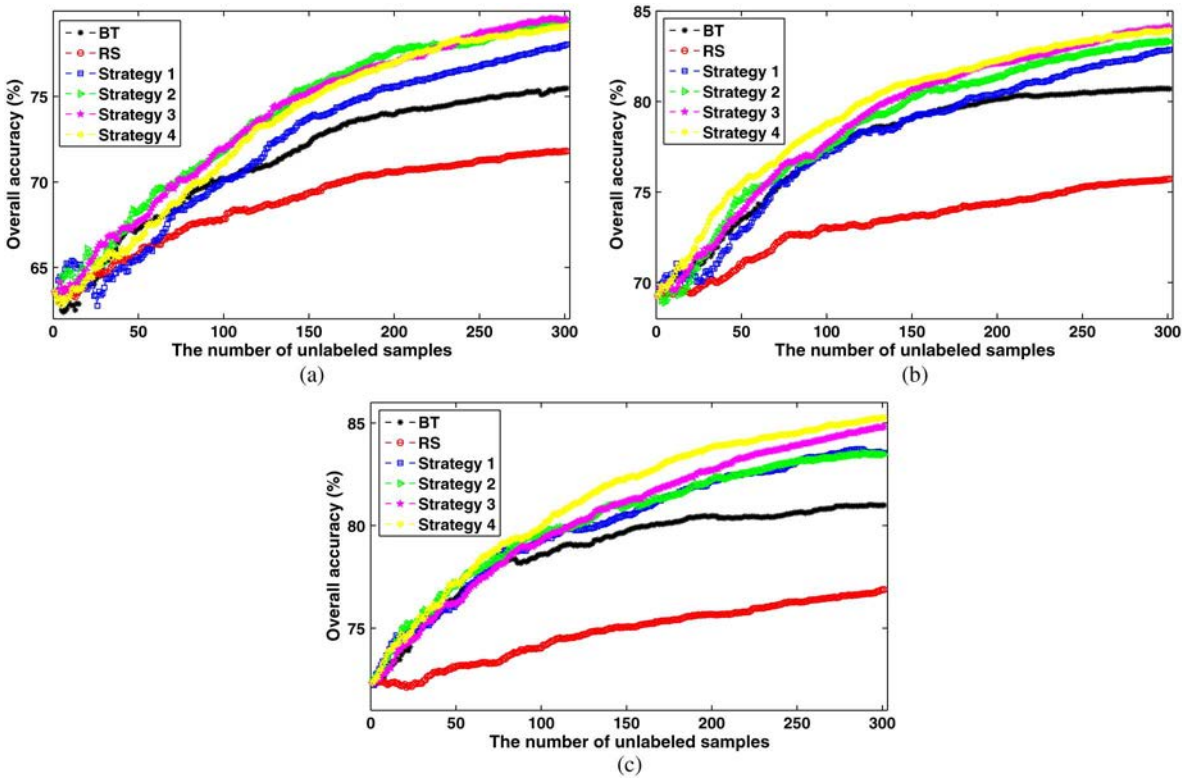

Fig. 9. OA (as a function of the number of unlabeled samples) obtained for the ROSIS Pavia University data set by the proposed semisupervised classifier using different strategies. For the unmixing-based strategies, $\alpha = 0.75$ in all cases. (a) $l_n = 5$ labeled samples per class. (b) $l_n = 10$ labeled samples per class. (c) $l_n = 15$ labeled samples per class.

volumes of data than those reported in this paper. In this case, domain adaptation techniques will be certainly needed. Although the results presented in this paper are focused on a few hyperspectral scenes only (due to the reliable ground-truth and reference information available for those scenes), the extrapolation of these techniques to larger data collections will allow a more detailed assessment of the requirements and benefits of applying the presented approaches in practical scenarios.

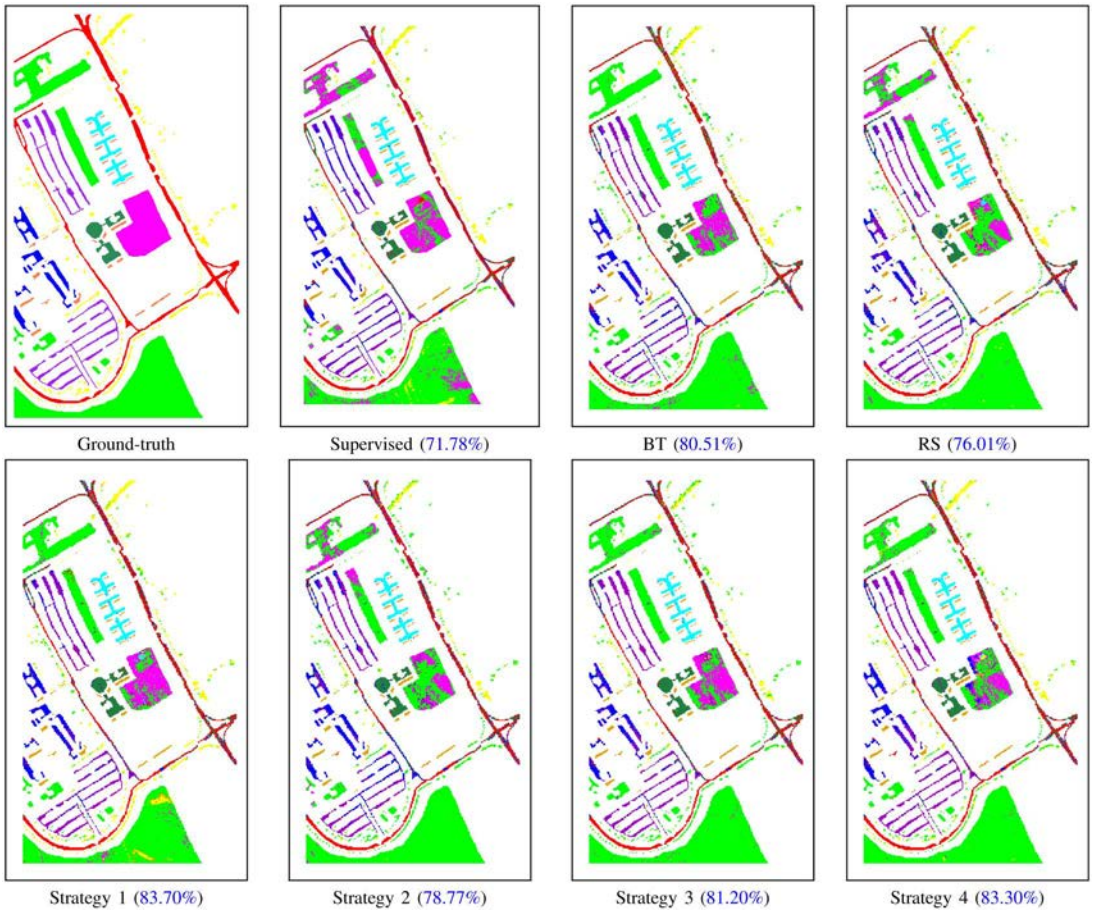

Fig. 10. Classification maps and OAs (in parentheses) obtained after applying classification techniques to the ROSIS Pavia University data set. In all cases, the number of labeled samples was $l_n = 10$ and the number of unlabeled samples was $u_n = 300$.

Acknowledgment

The authors would like to thank the Associate Editor who handled our manuscript and the three Anonymous Reviewers for their interesting comments and suggestions, which improved the technical content and presentation of our manuscript.

References

[1] J. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[2] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[3] A. Plaza, Q. Du, J. Bioucas-Dias, X. Jia, and F. Kruse, "Foreword to the special issue on spectral unmixing of remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4103–4110, Nov. 2011.

[4] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[5] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 650–663, Mar. 2004.

[6] Q. Du, N. Raksuntorn, N. Younan, and R. King, "End-member extraction for hyperspectral image analysis," *Appl. Opt.*, vol. 47, no. 28, pp. 77–84, 2008.

[7] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Norwell, MA, USA: Kluwer/Plenum, 2003.

[8] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. Hoboken, NJ, USA: Wiley, 2003.

[9] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. New York, NY, USA: Springer-Verlag, 2006.

[10] 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.

[11] 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.

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

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

[14] B. Shahshahani and D. A. Landgrebe, "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1087–1095, Sep. 1994.

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

[16] X. Zhu, "Semi-supervised learning literature survey," Computer Sciences, Univ. Wisconsin-Madison, Madison, WI, USA, Tech. Rep. 1530, 2005.

[17] J. Muñoz-Marí, D. Tuia, and G. Camps-Valls, "Semisupervised classification of remote sensing images with active queries," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3751–3763, Oct. 2012.

[18] D. Tuia, M. Volpi, L. Copa, M. Kanevski, and J. Muñoz-Marí, "A survey of active learning algorithms for supervised remote sensing image classification," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 606–617, Jun. 2011.

[19] W. Di and M. 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.

[20] I. Dopido *et al.*, "Semi-supervised self-learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4032–4044, Jul. 2013.
[21] I. Dopido, M. Zortea, A. Villa, A. Plaza, and P. Gamba, "Unmixing prior to supervised classification of remotely sensed hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 760–764, Jul. 2011.
[22] A. Green, M. Berman, P. Switzer, and M. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1988.
[23] I. Dopido, A. Villa, A. Plaza, and P. Gamba, "A quantitative and comparative assessment of unmixing-based feature extraction techniques for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 421–435, Apr. 2012.
[24] I. Dopido, J. Li, A. Plaza, and P. Gamba, "Semi-supervised classification of hyperspectral data using spectral unmixing concepts," in *Proc. Tyrrhenian Workshop Adv. Radar Remote Sens. (TyWRRS)*, 2012, vol. 1, pp. 353–358.
[25] D. Böhning, "Multinomial logistic regression algorithm," *Ann. Inst. Stat. Math.*, vol. 44, pp. 197–200, 1992.
[26] B. Krishnapuram, L. Carin, M. Figueiredo, and A. Hartemink, "Sparse multinomial logistic regression: Fast algorithms and generalization bounds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 6, pp. 957–968, Jun. 2005.
[27] J. Li, J. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new Bayesian approach with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3947–3960, Oct. 2011.
[28] T. Luo *et al.*, "Active learning to recognize multiple types of plankton," *J. Mach. Learn. Res.*, vol. 6, pp. 589–613, 2005.
[29] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
[30] J. Boardman, "Leveraging the high dimensionality of AVIRIS data for improved subpixel target unmixing and rejection of false positives: Mixture tuned matched filtering," in *Proc. 5th JPL Geosci. Workshop*, 1998, pp. 55–56.

Inmaculada Dópido received the B.S. and M.S. degrees in telecommunications and the M.S. and Ph.D. degrees in computer engineering from the University of Extremadura, Caceres, Spain.

She is a Member of the Hyperspectral Computing Laboratory (HyperComp) coordinated by Prof. Antonio Plaza. Her research interests include remotely sensed hyperspectral imaging, pattern recognition and signal and image processing, with particular emphasis on the development of new techniques for unsupervised and supervised classification, and spectral mixture analysis of hyperspectral data.

Dr. Dópido is a Reviewer for the IEEE Journal of Selected topics in Applied Earth Observations and Remote Sensing, IEEE Transaction on Geoscience and Remote Sensing, and IEEE Geoscience and Remote Sensing Letters.

Jun Li received the B.S. degree in geographic information systems from the Hunan Normal University, Changsha, China, in 2004, the M.E. degree in remote sensing from the 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. Her research interests include hyperspectral image classification and segmentation, spectral unmixing, signal processing, and remote sensing.

Dr. Li 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*. She received the 2012 Best Reviewer Award of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

Paolo Gamba (SM'00–F'13) received the Laurea (*cum laude*) and Ph.D. degrees in electronic engineering from the University of Pavia, Pavia, Italy, in 1989 and 1993, respectively.

He is an Associate Professor of Telecommunications at the University of Pavia, where he also leads the Telecommunications and Remote Sensing Laboratory. He has been invited to give keynote lectures and tutorials in several occasions about urban remote sensing, data fusion, EO data, and risk management. He published more than 100 papers in international peer-review journals and presented more than 250 research works in workshops and conferences.

Dr. Gamba served as Editor-in-Chief of the IEEE Geoscience and Remote Sensing Letters from 2009 to 2013, and as Chair of the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society from October 2005 to May 2009. Currently, he is the Chair of the Chapters' Committee of the same Society. He has been the Organizer and Technical Chair of the biennial GRSS/ISPRS Joint Workshops on "Remote Sensing and Data Fusion over Urban Areas" since 2001. The latest edition, JURSE 2015, will be held in Lausanne (Switzerland) in April 2015. He also served as a Technical Co-Chair of the 2010 IEEE Geoscience and Remote Sensing Symposium, Honolulu, Hawaii, July 2010, and will serve as a Technical Co-Chair of the 2015 IEEE Geoscience and Remote Sensing Symposium, Milan, Italy. He has been the Guest Editor of special issues of IEEE Transactions on Geoscience and Remote Sensing, IEEE Journal of Selected Topics in Remote Sensing Applications, *ISPRS Journal of Photogrammetry and Remote Sensing, International Journal of Information Fusion and Pattern Recognition Letters on the Topics of Urban Remote Sensing, Remote Sensing for Disaster Management, Pattern Recognition in Remote Sensing Applications*.

Antonio Plaza (M'05–SM'07) is an Associate Professor (with accreditation for Full Professor) with 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 was the Coordinator of the Hyperspectral Imaging Network, a European project with total funding of 2.8 MEuro. He authored more than 400 publications, including 119 JCR journal papers (71 in IEEE journals), 20 book chapters, and over 240 peer-reviewed conference proceeding papers (94 in IEEE conferences). He has guest edited seven special issues on JCR journals (three in IEEE journals).

Dr. Plaza has been a Chair for the IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (2011). He is a recipient of the recognition of Best Reviewers of the IEEE Geoscience and Remote Sensing Letters (in 2009) and a recipient of the recognition of Best Reviewers of the IEEE Transactions on Geoscience and Remote Sensing (in 2010), a journal for which he has served as an Associate Editor from 2007 to 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 (2012). He served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS), from 2011 to 2012, and is currently serving as a President of the Spanish Chapter of IEEE GRSS (since November 2012). He is currently serving as the Editor-in-Chief of the IEEE Transactions on Geoscience and Remote Sensing journal (since January 2013).