Remotely Sensed Image Classification Using Sparse Representations of Morphological Attribute Profiles

Benqin Song, Jun Li, Mauro Dalla Mura, *Member, IEEE*, Peijun Li, *Senior Member, IEEE*, Antonio Plaza, *Senior Member, IEEE*, José M. Bioucas-Dias, *Member, IEEE*, Jon Atli Benediktsson, *Fellow, IEEE*, and Jocelyn Chanussot, *Fellow, IEEE*

***Abstract*—In recent years, sparse representations have been widely studied in the context of remote sensing image analysis. In this paper, we propose to exploit sparse representations of morphological attribute profiles for remotely sensed image classification. Specifically, we use extended multiattribute profiles (EMAPs) to integrate the spatial and spectral information contained in the data. EMAPs provide a multilevel characterization of an image created by the sequential application of morphological attribute filters that can be used to model different kinds of structural information. Although the EMAPs' feature vectors may have high dimensionality, they lie in class-dependent low-dimensional subpaces or submanifolds. In this paper, we use the sparse representation classification framework to exploit this characteristic of the EMAPs. In short, by gathering representative samples of the low-dimensional class-dependent structures, any given sample may by sparsely represented, and thus classified, with respect to the gathered samples. Our experiments reveal that the proposed approach exploits the inherent low-dimensional structure of the EMAPs to provide state-of-the-art classification results for different multi/hyperspectral data sets.**

***Index Terms*—Extended multiattribute profiles (EMAPs), remote sensing image classification, sparse representation.**

I. Introduction

THE availability of advanced observation instruments with constantly increasing coverage of the Earth has opened important perspectives in the analysis of remotely sensed data sets [1]. Recently, sparse representations have become a widely popular approach in the context of remote sensing data processing [2]. Such representations are widely used in many areas [3], including spectral unmixing [4] and classification [5]. These approaches intend to represent most observations or image pixels with linear combinations of a small number of elementary (dictionary) samples, which are often called *atoms*, chosen from an overcomplete training dictionary. Formally, an overcomplete dictionary is a collection of atoms, such that the number of atoms exceeds the dimension of the image space, and any image pixel can be represented by more than one combination of different atoms. Under given circumstances, an image pixel can be recovered with a minimal number of atoms by solving a sparse regression problem.

An important aspect in sparse representations for remote sensing image classification is how to account for spatial information. This is because it is now commonly accepted that using the spatial and spectral information simultaneously provides significant advantages in terms of improving the performance of classification techniques (see, for instance, [6]–[11] and references herein). In the past, some efforts have been developed in this direction. For instance, in [2], two different methods are proposed to take into account the contextual information in the sparse recovery procedure. One strategy imposes a local smoothing constraint to the optimization formulation when reconstructing the pixel of interest. Another strategy adopts a joint sparsity model for neighboring pixels centered at the pixel of interest. The experimental results in [2] show that the strategies that considered the spatial-contextual information performed better in terms of classification accuracy. However, in these approaches, there is a need to define the size of a spatial window used to characterize contextual information prior to solving the joint sparse recovery problem. This involves some kind of prior knowledge about the image objects to be characterized.

In this paper, we explore the possibility of including spatial information in sparse classification of remotely sensed data using mathematical morphology (MM) [12], which is a widely used approach for modeling the spatial characteristics of the objects in remotely sensed images. Advanced MM techniques, such as derivative morphological profiles (DMPs) [13], have been successfully used for multispectral image classification by processing the panchromatic band of these instruments. This strategy has been also extended to hyperspectral image classification, by extracting the first few principal components (PCs) of the data using, for instance, PC analysis (PCA) [14]

Manuscript received January 18, 2013; revised May 14, 2013 and September 4, 2013; accepted October 1, 2013. This work was supported in part by the National Basic Research Program of China through the 973 Program under Grant 2011CB707103, by the Spanish Ministry of Science and Innovation through the Calibration of Earth Observation Satellites in Spain (CEOS-SPAIN) project under Reference AYA2011-29334-C02-02, by the Portuguese Science and Technology Foundation under Project PEst-OE/EEI/LA0008/2013, and by the Icelandic Research Fund.

B. Song and P. Li are with the Institute of Remote Sensing and Geographical Information System, School of Earth and Space Sciences, Peking University, Beijing 100871, China.

J. Li is with the School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China.

M. Dalla Mura and J. Chanussot are with the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab), Department of Image and Signal Processing (DIS), 38402 Saint Martin d'Heres Cedex, France.

A. Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, 10071 Cáceres, Spain.

J. M. Bioucas-Dias is with the Instituto de Telecomunicações, Instituto Superior Técnico, 1049-001 Lisbon, Portugal.

J. A. Benediktsson is with the Faculty of Electrical and Computer Engineering, University of Iceland, 101 Reykjavik, Iceland.

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

Digital Object Identifier 10.1109/TGRS.2013.2286953

and then building so-called extended morphological profiles (EMPs) on the first few PCA components to extract relevant features for classification [15]. Since redundant information is generally present in DMPs and EMPs with high dimensionality, feature extraction and selection techniques have been used to extract the most relevant information prior to classification [16]. Recently, morphological attribute profiles (APs) [17] have been introduced as an advanced mechanism to obtain a detailed multilevel characterization of a very high resolution image created by the sequential application of morphological attribute filters that can be used (prior to classification) to model different kinds of the structural information. According to the type of the attributes considered in the morphological attribute transformation, different parametric features can be modeled. The use of different attributes leads to the concept of extended multiattribute profiles (EMAPs) [18].

As opposed to the techniques described in [2], morphological approaches, such as EMAPs, avoid the need for prior knowledge by using multilevel operators (i.e., an image transformation with varying filter parameters) and different attributes (leading to filtering effects related to the spatial characteristics accounted by the attribute transformation) when describing the spatial context. However, the information conveyed by EMAPs can be extremely high and appeals for efficient forms of representation in order to avoid increasing the dimensionality of the problem, which can have a negative impact in the subsequent classifier due to the Hughes effect [19]. Although classifiers such as support vector machines (SVMs) [20], with reduced sensitivity to limited training samples, have been successfully used in the literature [6], [21]–[26], the information extracted by morphological approaches can be made more compact in order to further optimize the classification process. This issue was investigated in [16], [27], and [28] by applying several feature extraction and selection techniques to morphological profiles and APs prior to classification, aiming at reducing the high dimensionality of the profile by keeping only (few) relevant features. A compact representation of the profile was also proposed in [13] by defining the so-called morphological characteristic (MC), which derives from the DMP if the underlying region of each pixel is mostly darker or brighter with respect to its surroundings. An extension of the MC has been recently presented in [29], embedding the characteristic scale, saliency, and level of the DMP in a 3-D index for each pixel in the image.

In this paper, we follow a different strategy and exploit the low-dimensional structure, in which the morphological profiles lie, and its suitability to perform sparse classification using both spatial and spectral information. Specifically, we focus on EMAPs that offer the potential to model structural information in great detail according to different types of attributes, thus generating features that may appeal for efficient forms of representation. Here, we exploit the fact the EMAPs of the same class lie in a low-dimensional subspace or manifold and can, therefore, be expressed using linear sparse regression. This property, which is common to many high-dimensional signals from the real world, underlies the tremendous actual interest in sparse modeling and redundant representation of these signals [30]. Hence, combining sparse representations with morphological approaches, such as EMAPs, into a classification framework may offer the advantage of improved characterization of spatial and spectral features while, at the same time, providing a better representation in terms of the information needed to discriminate between the classes. The new classification approach proposed in this work thus combines the advantages of classification based on sparse representations and multiattribute morphological analysis. Our experiments, which are conducted with both multispectral data (a Quickbird scene collected in Beijing, China) and hyperspectral data sets, such as the well-known Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) scene over Indian Pines, IN, USA, or a scene collected by the Reflective Optics System Imaging Spectrometer (ROSIS) over the University of Pavia, Pavia, Italy, reveal that the proposed approach can exploit both the sparsity and the rich structural information provided by EMAPs to provide state-of-the-art classification results for different multi/hyperspectral data sets.

The remainder of the paper is structured as follows. Section II presents the proposed methodology, which is made up of two main ingredients: EMAPs and sparse classification. Section III describes the data sets used for evaluation purposes and performs an extensive experimental evaluation and comparison of the proposed approach with regard to state-of-the-art methods. Finally, Section IV concludes with some remarks and hints at plausible future research lines.

II. Proposed Methodology

A. EMAPs

As mentioned before, EMAPs are an extension of APs obtained using different types of attributes and stacked together. The filtering operation implemented in EMAPs is based on the evaluation of how a given attribute $\mathcal{A}$ is computed for every connected component of a grayscale image f for a given reference value λ. For a connected component C_i of the image, if the attribute meets a predefined condition (e.g., $\mathcal{A}(C_i) > \lambda$), then the region is kept unaltered; otherwise, it is set to the grayscale value of the adjacent region with closer value, thereby merging C_i to a surrounding connected component. When the region is merged to the adjacent region of a lower (or greater) gray level, the operation performed is a thinning (or thickening). Given an ordered sequence of thresholds $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, an AP is obtained by applying a sequence of attribute thinning and attribute thickening operations as follows:

$$\mathrm{AP}(f) := \{\phi_n(f), \ldots, \phi_1(f), f, \gamma_1(f), \ldots, \gamma_n(f)\} \quad (1)$$

where ϕ_i and γ_i denote the thickening and thinning transformations, respectively. Problem (1) focuses on a single feature (or spectral band) of panchromatic data. For hyperspectral images, we can perform attribute filtering on the full original data. However, hyperspectral data are very high dimensional, with a large number of spectral bands. This means that constructing the extended attribute profile (EAP) on the original spectral bands would lead to very significant computational complexity. On the other hand, it is generally observed that hyperspectral data lie in a subspace with much lower dimensionality than the original spectral space. As a result, we can perform

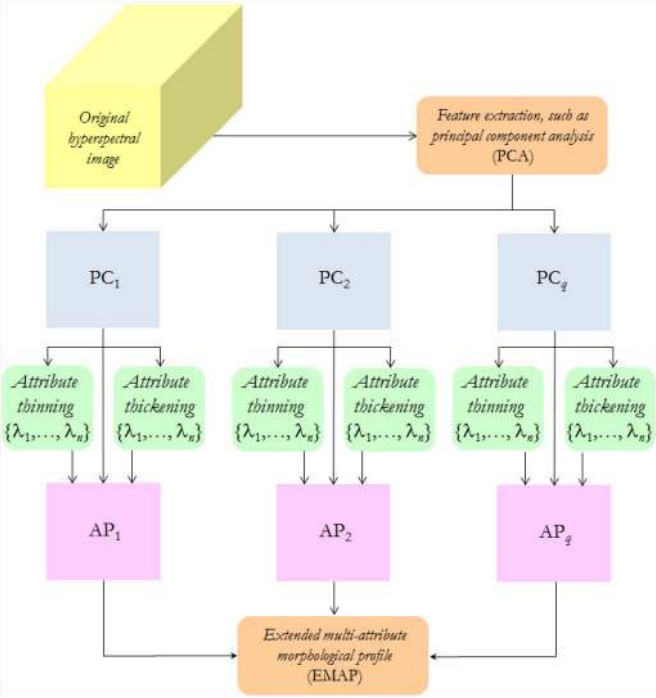

Fig. 1. Graphical illustration of the procedure adopted in order to construct an EAP from a hyperspectral image. The EMAP is a combination of EAPs obtained with different attributes.

dimensionality reduction, using techniques such as PCA [14], and then perform attribute filtering on the first few PCs, as suggested in [27], in order to reduce computational complexity. For multispectral data, since there are only a few spectral bands available, we can perform attribute filtering on the full original spectral data. In this way, the EAP is obtained by generating an AP on each of the first q PCs (or any other features retained after applying feature selection on the multi/hyperspectral image), thus building a stacked vector using the AP on each feature, as illustrated in Fig. 1. This leads to the following definition of the EAP for the pixel $\mathbf{x}_i$:

$$\text{EAP} := \{\text{AP}(f_1), \text{AP}(f_2), \ldots, \text{AP}(f_q)\} \tag{2}$$

where q is the number of retained features. From the EAP definition in (2), the consideration of multiple attributes leads to the concept of EMAP, which combines the EAPs by concatenating them in a single vector of features and improves the capability in extracting the spatial characteristics of the structures in the scene, where attribute filtering can be efficiently computed by applying a max-tree algorithm [31].

B. Sparse Representation Classification

In sparse representation models, the idea is that all the test samples can be represented by a (sparse) linear combination of atoms from an overcomplete training dictionary. Let us assume that we have a training dictionary, which is denoted by $\mathbf{A} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in \mathbb{R}^{n \times l}$, with n samples of l dimensions, comprising a total of c distinct classes, and that the dictionary is organized as $\mathbf{A} = [\mathbf{A}_1, \ldots, \mathbf{A}_c]$, where $\mathbf{A}_k = \{\mathbf{x}_{k_1}, \ldots, \mathbf{x}_{k_{n_k}}\}$ (i.e., $\mathbf{A}_k$ holds the samples of class k in its columns, n_k is the number of samples in $\mathbf{A}_k$, and $\sum_{k=1}^{c} n_k = n$). A typical scenario in supervised classification of remote sensing data is that, normally, we have a (limited) set of labeled training samples for each class, and then, we use part of this information to train a classifier, which is then tested with the remaining labeled samples. Let $\mathbf{x}_i$ be a test sample, which can be appropriately represented by a linear combination of the atoms (training samples) in the dictionary $\mathbf{A}$ as follows:

$$\begin{aligned}\mathbf{x}_i &\approx \mathbf{x}_1\alpha_1 + \mathbf{x}_2\alpha_2 + \cdots + \mathbf{x}_n\alpha_n \\ &= [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n][\alpha_1\ \alpha_2\ \ldots\ \alpha_n]^T = \mathbf{A}\boldsymbol{\alpha} + \epsilon\end{aligned} \tag{3}$$

where $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1^T, \ldots, \boldsymbol{\alpha}_c^T]^T$ is an n-dimensional sparse vector (i.e., most elements of $\boldsymbol{\alpha}$ are zero), $\boldsymbol{\alpha}_i$ is the vector of regression coefficients associated with class i, and ϵ is the representation error. As in [2] and [3], the central assumption in our approach is that $\mathbf{x}_i$ is well approximated by $\mathbf{A}_i\boldsymbol{\alpha}_i$, i.e., $\boldsymbol{\alpha}_j = 0$, for $j \neq i$. Under this condition, the classification of $\mathbf{x}_i$ amounts to detect the support of $\boldsymbol{\alpha}$.

In order to obtain a sparse representation for a test sample $\mathbf{x}_i$ (which hereinafter is assumed to be obtained using an EMAP representation), we need to obtain a sparse vector $\boldsymbol{\alpha}$ satisfying $\mathbf{x}_i = \mathbf{A}\boldsymbol{\alpha} + \epsilon$. The sparse vector $\boldsymbol{\alpha}$ can be estimated by solving the following optimization problem:

$$\hat{\boldsymbol{\alpha}} = \arg\min \|\boldsymbol{\alpha}\|_0 \quad \text{subject to} \quad \mathbf{x}_i = \mathbf{A}\boldsymbol{\alpha} \tag{4}$$

where $\|\boldsymbol{\alpha}\|_0$ denotes the ℓ_0-norm, which counts the nonzero components in the coefficient vector. Due to the presence of noise and a possible modeling error, the optimization (4) is often replaced by

$$\hat{\boldsymbol{\alpha}} = \arg\min \|\boldsymbol{\alpha}\|_0 \quad \text{subject to} \quad \|\mathbf{x}_i - \mathbf{A}\boldsymbol{\alpha}\|_2 \leq \delta \tag{5}$$

where δ is an error tolerance. The aforementioned problem is nondeterministic and NP-hard; thus, it very difficult to solve. Recently, greedy algorithms, such as basis pursuit (BP) [32] and orthogonal matching pursuit (OMP) [33], have been proposed to tackle this problem. BP replaces the ℓ_0-norm with the ℓ_1-norm. Hence, the sparse vector $\boldsymbol{\alpha}$ in (5) can be obtained using ℓ_1-norm as follows:

$$\hat{\boldsymbol{\alpha}} = \arg\min \|\boldsymbol{\alpha}\|_1 \quad \text{subject to} \quad \|\mathbf{x}_i - \mathbf{A}\boldsymbol{\alpha}\|_2 \leq \delta \tag{6}$$

where $\|\boldsymbol{\alpha}\|_1 = \sum_i |\alpha_i|$, for $i = 1, \ldots, n$. On the other hand, the OMP algorithm is a greedy strategy that performs as follows. In the first iteration of OMP, the initial residual is equal to the test sample $\mathbf{x}_i$. Then, at each iteration, the algorithm finds the index of the atom that best approximates the residual, adds this member to the matrix of atoms, updates the residual, and computes the estimate of $\boldsymbol{\alpha}$ using the set of currently obtained atoms. The OMP algorithm stops when a predefined number of atoms have been selected, or when the approximation error is below a certain prescribed error.

It should be noted that, contrary to problem (5), problem (6) is convex and can be solved using linear programming solvers. In fact, problem (6) is equivalent to the following unconstrained optimization problem:

$$\min_{\boldsymbol{\alpha}} \frac{1}{2}\|\mathbf{x}_i - \mathbf{A}\boldsymbol{\alpha}\|_2^2 + \tau\|\boldsymbol{\alpha}\|_1 \tag{7}$$

where the parameter τ is a Lagrange multiplier, which balances the tradeoff between the reconstruction error and the sparse solution, i.e., $\tau \to 0$ when $\epsilon \to 0$. In our experiments, we have observed that the nonnegativity constraint (NC) $\boldsymbol{\alpha} \geq 0$ improves the classification results. Our conjecture is that the EMAPs' feature vectors are very likely to lie in the convex cone spanned by the atoms of the respective class and, therefore, given by convex combinations thereof. Consequently, we solve the following constrained $\ell_2 - \ell_1$ optimization problem:

$$\min_{\boldsymbol{\alpha}} \frac{1}{2}\|\mathbf{x}_i - \mathbf{A}\boldsymbol{\alpha}\|_2^2 + \tau\|\boldsymbol{\alpha}\|_1, \qquad \boldsymbol{\alpha} \geq \mathbf{0}. \tag{8}$$

It should be noted that sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) is very efficient for solving NC $l_2 - l_1$ problems. From the solution of (7), which is denoted by $\widehat{\boldsymbol{\alpha}} = [\widehat{\boldsymbol{\alpha}}_1, \ldots, \widehat{\boldsymbol{\alpha}}_c]$, and according to the rationale aforementioned, the label for $\mathbf{x}_i$ corresponds to the index of the unique nonnull vector $\widehat{\boldsymbol{\alpha}}_j$ for $j = 1, \ldots, c$. Unfortunately, even when the model (3) is exact and $\epsilon = 0$, the solutions of (7) do not necessarily satisfy $\widehat{\boldsymbol{\alpha}}_j = 0$ for $j \neq i$ when $l > n$ (undertermined case). The ability to recover unique solutions in this case by solving optimization problems (6) or (7) depends on the properties of matrix $\mathbf{A}$, i.e., the so-called *restricted isometric constants* [34]. In a qualitative way, the less correlated the columns of $\mathbf{A}$, the better the quality of the solutions. Since we are just interested in detecting what $\widehat{\boldsymbol{\alpha}}_j$ is active, it is much more important to have low interclass correlation (i.e., among atoms of $\mathbf{A}_i$ and $\mathbf{A}_j$ for $i \neq j$) than low intraclass correlation (i.e., among atoms of the same $\mathbf{A}_i$). The interclass correlation is closely related with the class separability.

In the next section, we give evidence that the classification based on the EMAP features yields better results than the classification based on the spectral vectors. The explanation for this improved performance of the EMAP features is the lower interclass correlation (i.e., higher class separability) of these features compared with that of the original spectral features. A thorough treatment of this issue is, however, beyond the scope of the paper.

Finally, in order to introduce robustness with respect to the representation error ϵ and to the interclass correlation, we use the following residual-based criterion to infer the class label:

$$\widehat{\text{class}}(\mathbf{x}_i) = \arg \min_{j \in \{1,\ldots,c\}} \|\mathbf{x}_i - \mathbf{A}_j\boldsymbol{\alpha}_j\|_2.$$

III. Experimental Results

Here, we evaluate the performance of the proposed sparse classification algorithm using both simulated and real hyperspectral data sets. The main objective of the experimental validation with simulated data sets is the assessment and characterization of the algorithm in a fully controlled environment, whereas the main objective of the experimental validation with real data sets is to compare the performance of the proposed method with regard to that reported for state-of-the-art competitors in the literature.

It should be noted that, in all experiments, the OMP and SUnSAL techniques are used to approximately solve sparse problems (6) and (7) for the proposed classifier. For the SUnSAL algorithm, we empirically set the regularization parameter $\tau = 10^{-5}$. As it can be observed in the following experiments, although suboptimal, this parameter setting provided very good results. For completeness, we also show the case in which $\tau = 0$ in our experiments, in order to illustrate a situation in which the sparsity constraint is not considered. In addition, the SVM and the SVM based on composite kernels (SVMCK) [35] classifiers are considered for comparison, where the SVMCK takes into account the spatial and spectral information simultaneously with more than one kernel. Two additional composite kernel-based approaches are considered: $\text{SVMCK}_{\text{ori}}$ and $\text{SVMCK}_{\text{EMAP}}$. For the $\text{SVMCK}_{\text{ori}}$, the spectral kernel is built by using the original spectral bands. For the $\text{SVMCK}_{\text{EMAP}}$, the spectral kernel is built using EMAPs. Following the work in [35], in both cases, the spatial kernel is built by using the mean of the neighborhood around each pixel using a 5×5-pixel window per spectral channel. For the SVMCK, we also consider an new approach, which is denoted as $\text{SVMCK}_{\text{both}}$, in which the spectral kernel is built by using the original spectral information and the spatial kernel is built by using EMAPs. It should be noted that, in this work, the SVM parameters are obtained by cross-validation [36] and we always fix the weight between the spectral and spatial kernels to 0.5, since we have empirically found out that this value usually leads to the best results. Another important remark is that we selected the Gaussian radial basis function (RBF) after experimentation with other kernels. The RBF has been shown to be able to handle complex nonlinear class distributions. The parameters related with the SVM classifier are optimized by fivefold cross-validation.

Furthermore, in order to show that the EMAP extracts spatial information with good class separability, we also consider the 3-D wavelet transform filter (3D-DWT-based approach) for comparison [37]. Following the work in [37], we choose the same experimental setting for the comparative experiments. In order to have a statistically relevant and fair comparison with the methods in [37], the results are obtained by averaging the values obtained after 50 Monte Carlo runs. Since sparse-representation-based classification with the training samples as dictionary can be considered an extension of the k-nearest-neighbor (k-NN) method [38], we have also included this method in our comparison and carefully optimized the input parameters for this method empirically. As a result, two additional approaches are included in our comparison, i.e., k-NN_{ori}, which uses the original spectral information, and k-NN_{EMAP}, which uses the EMAP as input feature. Moreover, in order to build the EMAPs, we fix the number of PCs to be retained (q) to the number of components that contain more than 98% of the total variance of the data, where the obtained components are rescaled to the range [0, 1000] and converted to integer, in order to build the attribute filters. Specifically, the EMAPs are built using the area (related to the size of the regions) and standard deviation (which measures the homogeneity of the pixels enclosed by the regions) attributes. The threshold values are chosen in the range {50, 500}, with a stepwise increment of 50 for the area attribute. For the standard deviation, attribute values ranging from 2.5% to 20% of the mean of the feature, with a stepwise increment of 2.5%, are chosen [39]. Finally, in order to evaluate the computational complexity of the proposed method, we also report the computational time of each algorithm applied

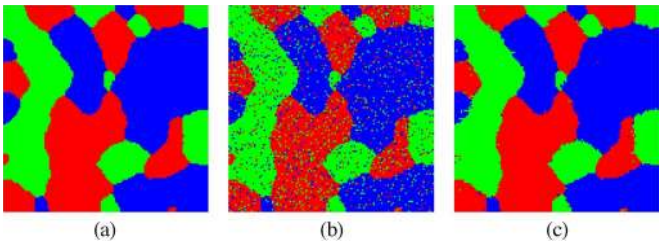

Fig. 2. (a) Image of class labels for a simulated data set made up of highly mixed pixels and noise. (b) Sparse classification (implemented using SUnSAL) based on the original spectral information (OA $= 89.34\%$). (c) Sparse classification (implemented using SUnSAL) based on EMAPs (OA $= 99.11\%$).

in each experiment. All the algorithms were implemented by using MATLAB R2011 on a desktop PC equipped with an Intel Core 2 i7-2600 (at 3.4 GHz) and 8 GB of RAM memory. Then, we run sparse-representation-based classification by using the extracted features.

In our experiments, the training dictionary is constituted by randomly selected samples from the available reference data, and the remaining samples are used for validation purposes. The overall accuracy (OA), average accuracy (AA), and the kappa statistic (κ) are used to evaluate the performance of the classification. Apart from the experiments for the comparison with the work in [37], all the other values of OA, AA, and κ are obtained by averaging the values obtained after ten Monte Carlo runs. In order to evaluate the impact of the number of training samples on the classification accuracy, we also consider different numbers of training samples in our experiments.

A. Experiments With Simulated Data

Here, a simple simulated hyperspectral scene is used to preliminarily evaluate the proposed sparse classification algorithm using some of the methods aforementioned. The simulated image is generated with a size of $N = 128 \times 128$ pixels made up of linear mixtures between three components as follows:

$$\mathbf{x}_i = \sum_{k=1}^{c} \mathbf{m}_k \mathbf{s}_i^k + \mathbf{n}_i \tag{9}$$

where $\mathbf{s} = \{\mathbf{s}_1, \ldots, \mathbf{s}_N\}$ is the fractional abundance matrix, which is generated according to the uniform distribution over the simplex. In order to ensure that only highly mixed pixels are present in the simulation, we discarded all pixels simulated with abundance fractions larger than 0.5. Furthermore, $\mathbf{m} = \{\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3\}$ is the mixing matrix, where the spectral signatures used were randomly obtained from the USGS digital library.[1] Finally, zero-mean Gaussian noise with variance $\sigma^2\mathbf{I}$, i.e., $\mathbf{n}_i \sim (0, \sigma^2\mathbf{I})$, is added to our simulated image. In our problem, we choose $\sigma = 0.3182$ (which corresponds to a signal-to-noise ratio equal to 5 dB), to make sure that we have a challenging classification problem, since we only can obtain an OA of 89.41% by using 20 training samples per class for the SVM algorithm.

The image of class labels used as reference in the experiment is shown in Fig. 2(a). This image contains the most predominant class label for each class, bearing in mind that all the pixels in each class are highly mixed. In order to construct the training dictionary, we randomly chose 20 labeled samples per class and used the remaining samples per class for validation purposes. As shown in Fig. 2, the classification results obtained by the proposed sparse classification algorithm (implemented using SUnSAL) are significantly better when EMAPs [see Fig. 2(c)] were used as input features instead of the original spectral information [see Fig. 2(b)].

The reasons why the sparse classification results based on EMAPs are significantly better than those obtained by the same sparse classifier using the original spectral information are graphically illustrated in Fig. 3. Specifically, Fig. 3(a) shows a scatterplot (constructed by randomly choosing two spectral bands in the original data to represent a projection in a 2-D space) depicting the training samples that were used in the sparse classification to represent a test sample from the *blue* class in the original spectral space. Similarly, Fig. 3(b) shows a scatterplot (which also represents a projection in a 2-D subspace) depicting the training samples used to represent the same test sample but in the EMAP space. As it can be observed by comparing Fig. 3(a) and (b), the class separability (linked with the interclass correlation) is greatly improved in the EMAP space as a result of the inclusion of spatial information. This leads to the fact that a much better sparse representation is possible in the EMAP space for the considered test sample; specifically, only four atoms (two of them belonging to the same *blue* class) are needed for the sparse representation of this sample in the EMAP space, whereas a total of 17 atoms (seven in the *blue* class, nine in the *red* class, and one in the *green* class) are needed for the sparse representation of the same sample in the original spectral space. On the other hand, Fig. 3(c) shows a scatterplot depicting the training samples that were used in the sparse classification to represent a test sample from the *red* class in the original spectral space, whereas Fig. 3(d) shows a scatterplot depicting the training samples used to represent the same test sample but in the EMAP space. Again, class separability is greatly improved in the EMAP space in which only four atoms (belonging to the same *red* class) are needed for the sparse representation of this sample [see Fig. 3(d)]; in contrast, a much higher number of samples (belonging to the three classes) are needed to explain the same sample in the original spectral space [see Fig. 3(c)].

With the aforementioned observations in mind, Table I summarizes the classification results obtained for the considered simulated scene (after ten Monte Carlo runs) by the proposed sparse classifier, using both the original spectral information $(\mathrm{SUnSAL}_{\mathrm{ori}})$ and the information provided by EMAPs $(\mathrm{SUnSAL}_{\mathrm{EMAP}})$. For comparative purposes, the classification results obtained by the SVM classifier in the same cases is also reported. As shown in Table I, the superiority of the proposed sparse classification approach in this specific example is quite remarkable. The table also reveals that it is quite important to exploit the information provided by EMAPs (instead of the original spectral information) when performing the sparse classification. Specifically, when the original spectral information was used by SUnSAL, the obtained results are inferior to those provided by the SVM. In turn, when SUnSAL was applied in EMAP space, the classification results outperformed those

[1] http://speclab.cr.usgs.gov.

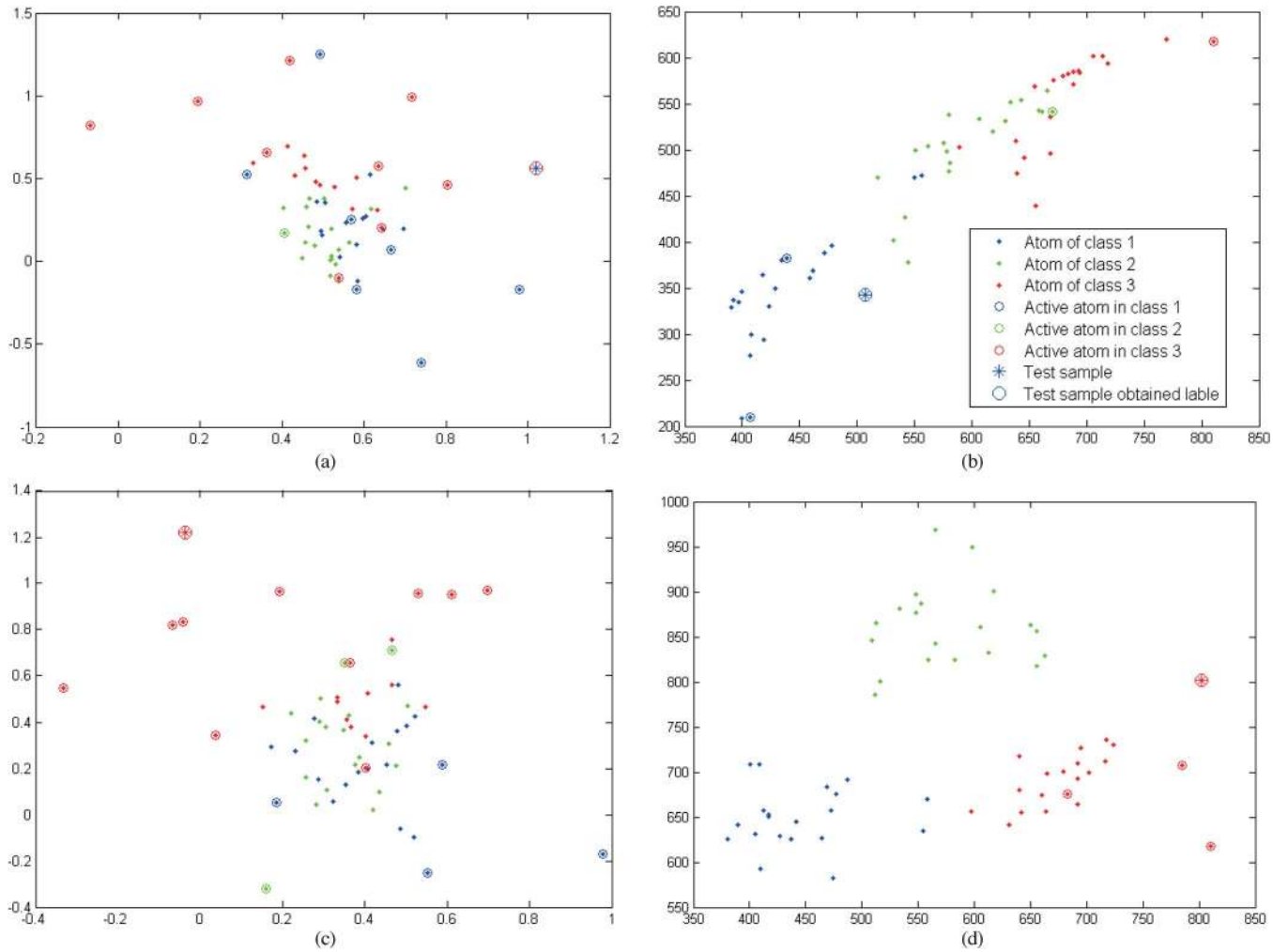

Fig. 3. Sparse representation of two different test samples from a simulated hyperspectral scene in spectral and EMAP space. (a) Sparse representation of a sample from *blue* class (spectral space). (b) Sparse representation of a sample from *blue* class (EMAP space). (c) Sparse representation of a sample from *red* class (spectral space). (d) Sparse representation of a sample from *red* class (EMAP space).

TABLE I
OA AND KAPPA STATISTIC (κ) OBTAINED AFTER CLASSIFYING THE SIMULATED IMAGE USING DIFFERENT TECHNIQUES. IN ALL CASES, TEN MONTE CARLO RUNS WERE CONDUCTED (THE STANDARD DEVIATION IS ALSO REPORTED) AND THE TRAINING SET WAS BASED ON RANDOM SELECTION OF 20 TRAINING SAMPLES PER CLASS (60 IN TOTAL)

	SVM_{ori}	SVM_{EMAP}	SUnSAL_{ori}	SUnSAL_{EMAP}
OA	89.41±1.27	92.06±4.55	87.83±0.64	99.07±0.49
κ	83.92±1.99	88.13±6.68	81.59±0.94	98.60±0.75

obtained by the SVM in the same space, indicating that a sparse representation of EMAPs can still significantly improve their exploitation for classification purposes.

B. Experiments With the AVIRIS Indian Pines Scene

The first hyperspectral image used in the experiments was collected by the AVIRIS sensor over the Indian Pines region, Northwest Indiana, USA, in 1992. This scene, with a size of 145 lines by 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-bit radiometric resolution. After an initial screening, 20 spectral bands (104–108, 150–163, and 220) were removed from the data set due to noise and water absorption phenomena. A reference map is available for the scene with 16 mutually exclusive reference classes and a total of 10 366 labeled samples. These data, including reference information, are available online,[2] a fact that has made this scene a widely used benchmark for testing the accuracy of hyperspectral data classification algorithms. This scene constitutes a challenging classification problem due to the presence of mixed pixels in all available classes and because of the unbalanced number of available labeled pixels per class.

In our first experiment with the AVIRIS Indian Pines scene, we illustrate the advantage of using a sparse representation in EMAP space for classification purposes. Here, we consider a training dictionary made up of 1028 samples (this is around 10% of the samples in the reference data) and the remaining labeled samples for testing. Using these training data, the OA

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

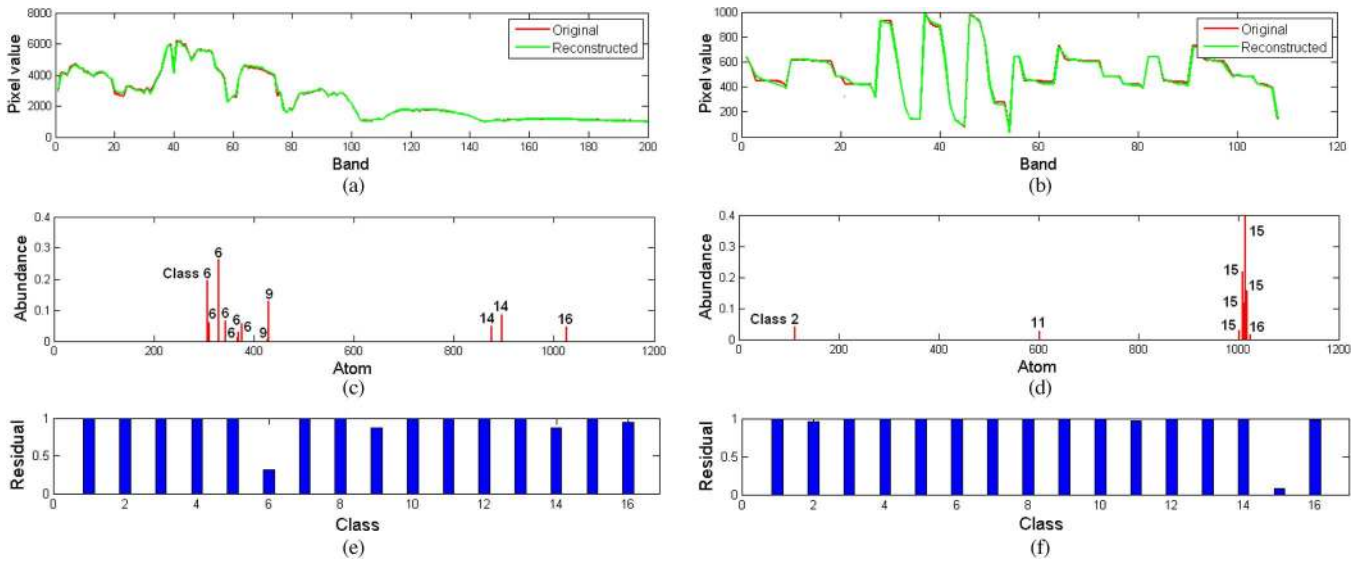

Fig. 4. Sparse representation of a sample from a highly mixed class (building–grass–tree–drives, class number 15) of the AVIRIS Indian Pines data in spectral and EMAP space. (a) Original versus reconstructed signature of a sample in class 15 (spectral space). (b) Original versus reconstructed signature of a sample in class 15 (EMAP space). (c) Sparse representation (spectral space). (d) Sparse representation (EMAP space). (e) Residuals for each class (spectral space). (f) Residuals for each class (EMAP space).

in classification accuracy obtained by $\text{SUnSAL}_{\text{EMAP}}$ (96.59%) after ten Monte Carlo runs was much higher than that obtained by $\text{SUnSAL}_{\text{ori}}$ (67.83%). However, it is interesting to note that the reconstruction of the data in the original spectral space and in the EMAP space were both very good. This is because, after the inclusion of spatial information in the EMAP space, the sparse representation is different than in the original spectral space. In the original spectral space, the observed data are mixed by materials, such that the representation is given by "mixtures." In turn, if we use classes as representees, as it is the case in the original spectral space, it may happen that a given sample can be perfectly reconstructed using atoms that do not belong to that particular class [this was already observed in Fig. 3(a) and (c)]. However, the transformed data in the EMAP space are more discriminative, in the sense of being able to separate different classes, and the representation is driven by "classes" instead of "mixtures." In this case, it is more likely that a given sample can be perfectly reconstructed by using atoms belonging to its correct class [as can be observed in Fig. 3(b) and (d)]. In order to illustrate this aspect with the AVIRIS Indian Pines data, Fig. 4 shows the sparse representation results in both the original spectral space and the EMAP space for a given sample in class number 15 (building–grass–tree–drives) of the AVIRIS Indian Pines data, which is a highly mixed class. As shown in Fig. 4, the sample can be accurately reconstructed in both the original spectral space and the EMAP space, but the classification results are different. In the original spectral space, the sample is classified as belonging to class number 6 (grass/trees), resulting from the fact that most of the samples needed to explain this sample belong to that particular class. In fact, no samples from class 15 (which is the class that the sample belongs to) are needed to explain the sample in spectral space. In turn, in the EMAP space, most of the samples needed to explain this sample belong to class 15, i.e., the class that the sample belongs to, and the number of samples needed to perfectly explain the sample is much lower. These are the reasons why the sparse classification results are significantly better in the EMAP space.

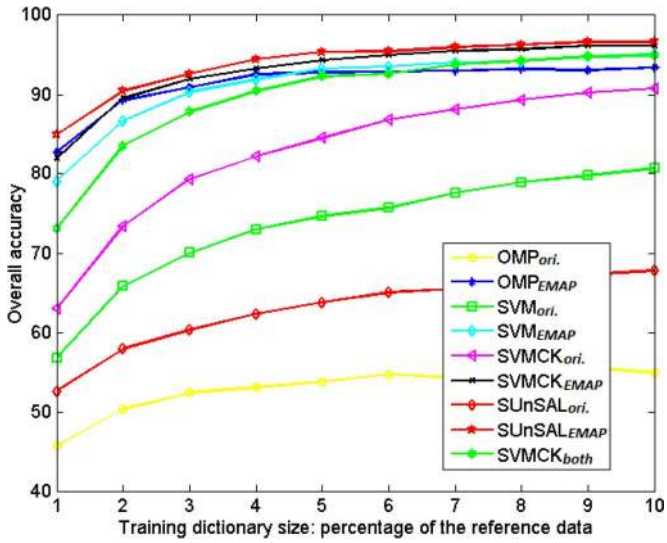

Fig. 5. Classification OAs as a function of training dictionary size (expressed in percentage of training samples for each class) for different classifiers on the AVIRIS Indian Pines data.

In the second experiment, we compare the classification accuracy obtained by the proposed approach with that obtained by other classification approaches for the AVIRIS Indian Pines data. In Fig. 5, the obtained classification accuracies are plotted as a function of the size of the training dictionary. Several observations can be made in Fig. 5. First of all, the best classification accuracies (even with limited training samples) are obtained by the $\text{SUnSAL}_{\text{EMAP}}$ approach. The classification accuracies obtained by OMP_{EMAP} are very similar, and the results obtained by SVM_{EMAP} are also very competitive. This reveals that the EMAP has great ability to exploit the spatial information, which has great potential for classification purposes. The advantage is smaller with more training

TABLE II
OA, AA, KAPPA STATISTIC (κ), AND CLASS INDIVIDUAL ACCURACIES ([%]) OBTAINED BY DIFFERENT CLASSIFIERS ON THE AVIRIS INDIAN PINES DATA (HERE, WE USE A TOTAL OF 115 SAMPLES FOR TRAINING, WHICH REPRESENTS ABOUT 1% OF THE AVAILABLE LABELED DATA FOR THE SCENE)

Class	Train	Test	SVM_{ori}	SVM_{EMAP}	$SVMCK_{both}$	$SVMCK_{ori}$	$SVMCK_{EMAP}$	OMP_{ori}	OMP_{EMAP}	kNN_{ori}	kNN_{EMAP}	$SUnSAL_{ori}$	$SUnSAL_{EMAP}$	$SUnSAL_{EMAP}(\tau = 0)$
Alfalfa	3	43	54.88	94.88	88.84	79.09	94.88	34.42	96.51	74.41	88.14	51.86	**97.67**	83.95
orn-notill	14	1414	42.63	68.06	60.19	50.92	69.34	39.05	72.23	40.62	60.58	40.83	**75.57**	15.57
Corn-min	8	822	29.53	60.22	64.57	41.24	**74.04**	22.71	66.34	28.44	54.45	22.21	68.22	26.13
Corn	3	234	17.65	35.73	31.92	35.38	44.27	8.63	**51.92**	24.52	26.03	12.52	47.01	27.35
Grass/Pasture	5	478	57.51	74.10	69.27	65.69	**76.86**	55.67	74.29	59.12	67.51	58.28	71.74	41.90
Grass/Trees	7	723	81.02	91.59	91.27	83.91	93.68	72.67	88.73	72.13	92.13	83.90	**95.09**	33.98
Grass/Pasture-mowed	3	25	86.40	95.20	93.20	96.40	97.60	61.60	**98.00**	88.00	95.20	77.60	96.80	87.60
Hay-windrowed	5	473	62.37	94.52	82.85	73.81	95.58	46.70	99.98	67.51	90.95	65.60	**100**	71.21
Oats	3	17	70.59	85.88	84.71	81.18	89.41	44.12	**97.06**	58.82	89.41	59.41	91.76	58.24
Soybeans-notill	10	962	39.73	75.60	66.65	51.19	79.59	25.34	83.37	43.67	66.67	29.18	**85.25**	26.72
Soybeans-min	25	2430	73.31	87.37	80.88	72.40	88.06	54.28	88.61	60.91	79.85	64.12	**92.26**	15.29
Soybeans-clean	6	587	21.72	57.68	41.93	29.40	59.93	18.94	**70.49**	27.10	43.20	18.93	68.18	42.54
Wheat	3	202	87.28	96.44	93.66	90.54	97.97	72.87	98.61	79.80	97.67	81.19	**99.50**	76.78
Woods	13	1252	84.03	97.32	91.73	85.93	96.45	68.95	95.67	75.80	89.03	82.72	**98.43**	6.88
Bldg-Grass-Trees	4	382	17.38	65.16	57.57	34.82	74.48	20.26	75.68	18.40	61.99	17.07	**77.88**	35.68
Stone-steel towers	3	90	70.67	86.89	89.44	89.22	**98.78**	65.11	95.11	80.33	93.78	85.44	95.44	81.11
OA			56.73	79.07	73.08	62.95	81.96	45.68	82.70	52.94	72.32	52.58	**84.90**	26.71
AA			56.04	79.17	74.29	66.32	83.18	44.46	84.54	56.23	74.79	53.18	**85.05**	45.81
κ			49.88	76.08	69.29	57.52	79.47	37.82	80.26	46.33	68.43	45.40	**82.76**	19.24
Time(s)			0.49	13.39	16.52	6.07	19.28	29.01	35.94	0.35	13.28	4.93	18.03	16.39

samples available. In other words, as the number of training samples increase, the uncertainty of the classes decreases. In turn, Fig. 5 shows that both $SVMCK_{ori}$ and SVM_{ori} provide better accuracies in spectral space than the sparse classifiers $SUnSAL_{ori}$ and OMP_{ori}. This reveals the importance of using sparse representation techniques, particularly in EMAP space. Finally, it is interesting to observe that, in the original spectral space, the results obtained by the SVM are better than those obtained by the sparse-representation-based methods, i.e., OMP and SUnSAL. This is because, in the original space, sparse-representation-based methods have limited ability for classification. In turn, the SVM is a discriminative approach, which can exploit class separability in the transformed kernel space. Therefore, it is reasonable that the SVM outperforms OMP and SUnSAL for the original data. However, in the EMAP space, the class discriminability is greatly improved, such that the classification methods based on sparse representation can achieve much better results. In order to better illustrate the obtained classification accuracies with limited training samples, Table II presents the individual class accuracies obtained for the case of 1% training data in Fig. 5 (a total of 115 samples used for training). As can be observed in Table II, in most cases, the proposed sparse classification techniques $SUnSAL_{EMAP}$ and OMP_{EMAP} provide the best results in terms of individual class accuracies when compared with other methods. It is also remarkable that all the methods including spatial information obtained higher classification accuracies in comparison with the methods using only spectral information (this is also the case for the k-NN technique). Finally, it is important to emphasize that the SUnSAL method without the sparsity constraint $(\tau = 0)$ provided low classification accuracies because the generalization capability is quite poor in this particular case.

With respect to the computational cost, some important observations can be obtained in Table II. First of all, the methods based on spectral and spatial information are slower than the methods only based on the spectral information. However, as reported before, the classification accuracies achieved by spectral-spatial methods are higher. This is expected since it is more time consuming to extract the spatial information. For instance, $SUnSAL_{EMAP}$ took 18.03 s, which is worse than 4.93 s of $SUnSAL_{ori}$. However, the EMAP-based methods can provide higher classification accuracies with limited training samples. Finally, as it can be observed in Table II, SUnSAL is generally more efficient than OMP. For illustrative purposes, Fig. 6 shows the classification maps obtained for the experiments reported in Table II.

A comparison between the classification results obtained by the proposed $SUnSAL_{EMAP}$ method and by several methods that use 3D-DWT features (reported in [37]) for the AVIRIS Indian Pines scene is also given in Table III. The results reported in Table III indicate that the proposed $SUnSAL_{EMAP}$ method outperforms the methods based on using 3D-DWT features in terms of classification accuracies.

To conclude this subsection, we analyze the relationship between the number of iterations (atoms) needed to obtain the sparse representation in EMAP space and the classification accuracies obtained by the OMP and SUnSAL algorithms used in this work. For this purpose, Fig. 7 shows the OAs and the level of sparsity as a function of the number of iterations for both OMP and SUnSAL in the case in which 115

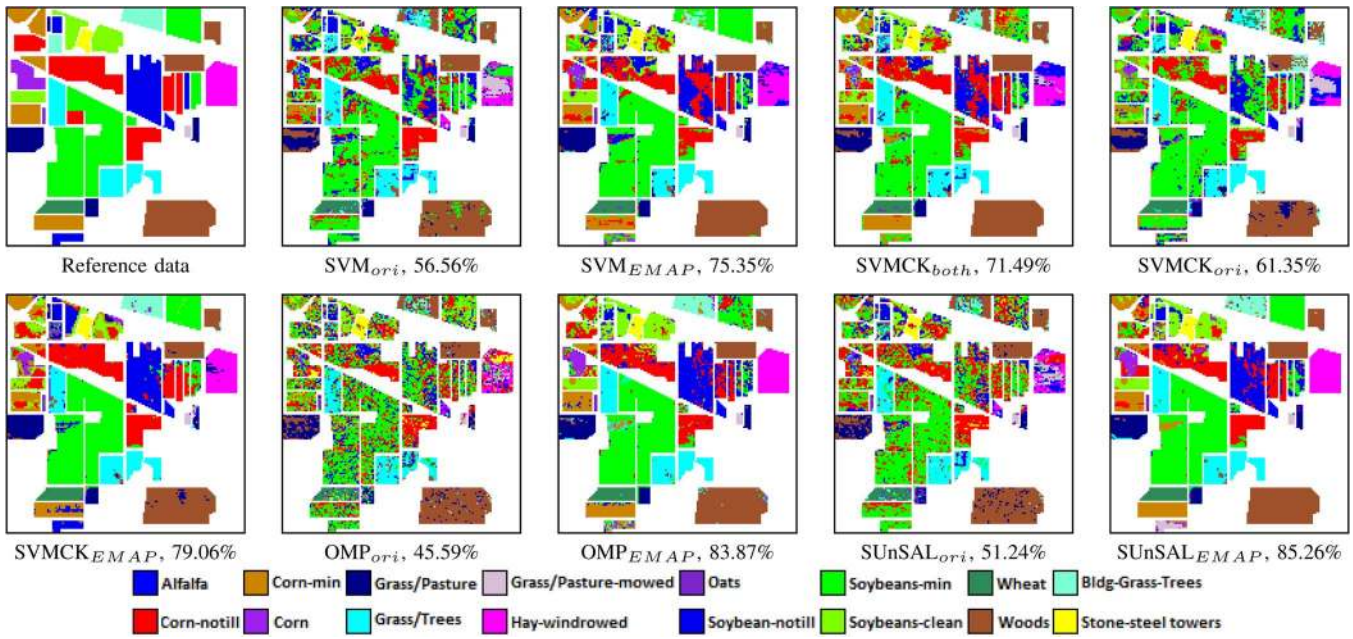

Fig. 6. Classification results obtained by different classifiers for the AVIRIS Indian Pines scene (using a total of 115 samples for training, which represents about 1% of the available labeled data for the scene).

TABLE III
OA, AA, AND KAPPA STATISTIC (κ) OBTAINED AFTER EXECUTING 50 MONTE CARLO RUNS OF THE PROPOSED SUnSAL$_{\mathrm{EMAP}}$ ALGORITHM FOR THE AVIRIS INDIAN PINES DATA, IN COMPARISON WITH SEVERAL METHODS THAT USE 3D-DWT INPUT FEATURES, AS REPORTED IN [37]. SVM-LINEAR AND SVM-RBF DENOTE THE SVM CLASSIFIERS WITH LINEAR AND RBF KERNELS, RESPECTIVELY. LASSO, GLASSO, SGLASSO, AND MIXED-LASSO ARE FOUR ALGORITHMS SOLVING THE SPARSE-REPRESENTATION-BASED PROBLEM. ADDITIONAL DETAILS ABOUT THESE ALGORITHMS CAN BE FOUND IN [37]

Training Set	Input feature	3D-DWT						EMAP
	Method	SVM-linear	SVM-RBF	Lasso	Glasso	Sglasso	Mixed-lasso	SUnSAL
5% of the labeled samples	OA	88.27	90.01	87.98	87.94	87.77	88.78	94.95
	AA	70.56	78.50	80.32	79.79	79.57	82.95	94.58
	κ	86.55	88.58	86.08	86.20	86.00	87.19	94.24
10% of the labeled samples	OA	93.72	95.45	91.98	93.76	95.38	93.05	96.78
	AA	84.28	88.80	86.54	86.14	86.56	88.72	96.63
	κ	92.83	94.78	90.65	90.77	90.44	92.07	96.33

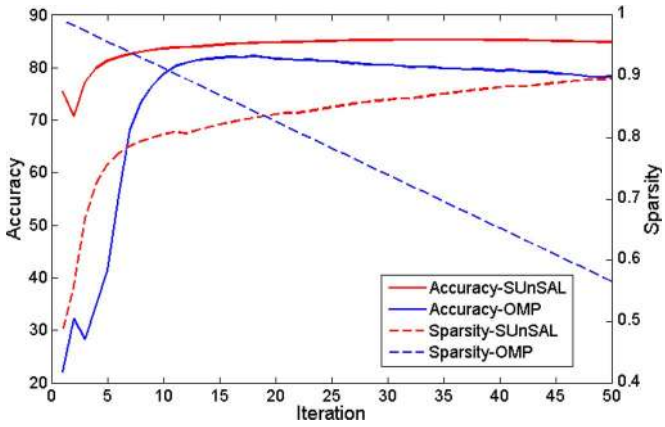

Fig. 7. Relationship between the classification accuracy (in percentage) and the level of sparsity (as a function of the number of iterations used to obtain the sparse representations in EMAP space by OMP and SUnSAL) for the AVIRIS Indian Pines data. In this experiment, 115 training samples (about 1% of the available labeled data) are used.

training samples (about 1% of the labeled data) are used for training purposes. Here, the level of sparsity is computed as $1 - (\text{number of atoms}/\text{number of training samples})$. As shown in Fig. 7, for SUnSAL, the sparsity improves as the number of iteration increases, and it reaches a degree of sparsity greater than 0.8 around ten iterations. At the same time, it is remarkable that the classification accuracy exhibits a similar trend as the level of sparsity. In other words, as the sparsity increases, the classification accuracy also increases. This is because the generalization capability improves under the sparse constraint. This result is consistent with the results presented in Table II, in which it is shown that SUnSAL with $\tau = 0$ (i.e., without the sparse constraint) achieves poor classification results. However, the sparsity of OMP decreases as the number of iterations increases. This is because OMP includes one new atom at each iteration. Furthermore, Fig. 7 shows that the accuracy of both SUnSAL and OMP is fluctuant in the first few iterations, mainly due to a poor sparse representation that can be achieved

TABLE IV
OA, AA, KAPPA STATISTIC (κ), AND CLASS INDIVIDUAL ACCURACIES ([%]) OBTAINED BY DIFFERENT CLASSIFIERS ON THE ROSIS PAVIA UNIVERSITY DATA (HERE, WE USE ONLY TEN SAMPLES PER CLASS FOR TRAINING, FOR A TOTAL OF 90 TRAINING SAMPLES)

Class	Train	Test	SVM_{ori}	SVM_{EMAP}	SVMCK_{both}	SVMCK_{ori}	SVMCK_{EMAP}	OMP_{ori}	OMP_{EMAP}	$k\text{NN}_{ori}$	$k\text{NN}_{EMAP}$	SUnSAL_{ori}	SUnSAL_{EMAP}
Asphalt	10	6631	61.19	86.82	86.96	65.66	88.73	14.73	89.31	57.70	84.87	22.99	**92.78**
Meadow	10	18649	58.78	71.08	78.99	60.38	82.92	34.77	84.74	48.26	64.53	53.82	**88.43**
Gravel	10	2099	53.23	80.50	81.24	52.35	77.72	47.53	82.99	54.31	74.54	64.29	**96.16**
Trees	10	3064	95.35	94.41	95.02	92.57	**95.63**	85.95	76.92	92.73	94.37	93.79	93.86
Metal sheets	10	1345	91.98	99.22	99.41	98.93	99.72	99.77	99.32	98.68	99.61	**99.87**	99.76
Bare soil	10	5029	55.78	85.13	84.98	47.89	87.63	57.51	88.62	58.14	72.77	71.37	**88.77**
Bitumen	10	1330	87.26	98.40	98.90	84.71	99.04	37.65	**99.70**	80.14	96.77	81.15	99.66
Bricks	10	3682	71.39	88.85	90.88	74.59	**95.08**	30.22	87.88	64.44	77.59	32.55	94.53
Shadows	10	947	96.16	**99.40**	99.21	96.19	98.14	42.07	92.08	95.34	99.44	76.38	98.43
Overall			64.99	81.20	84.92	65.62	87.44	40.53	86.61	59.38	75.28	55.45	**90.87**
Average			74.57	89.31	90.62	74.81	91.62	50.02	89.06	72.19	84.94	66.25	**93.60**
κ			56.48	76.77	80.90	57.15	84.05	29.79	82.85	50.62	69.14	46.07	**88.26**
Time(s)			1.89	124.63	146.8	26.5	147.72	225.91	345.41	1.96	124.47	33.41	153.53

with a very small number of atoms. In all cases, the SUnSAL algorithm outperforms the OMP. The classification accuracy of both approaches converges around 15 iterations, with a sparsity level of 0.826. For the OMP algorithm, the best classification result is 82.19% OA, which is about 2.8% lower than the best classification result achieved by SUnSAL for a similar sparsity level of 0.843.

C. Experiments With the ROSIS Pavia University Scene

In this experiment, we use the ROSIS Pavia University scene to evaluate the proposed approach. These data were collected by the ROSIS optical sensor over the urban area of the University of Pavia, Pavia, Italy. The flight was operated by the Deutschen Zentrum for Luftund Raumfahrt (DLR, i.e., the German Aerospace Agency) in the framework of the HySens project, which is managed and sponsored by the European Union. The image size in pixels is 610 $\times$ 340, with very high spatial resolution of 1.3 m per pixel. The number of data channels in the acquired image is 103 (with the spectral range from 0.43 to 0.86 μm). Nine thematic land-cover classes were identified in the university campus: trees, asphalt, bitumen, gravel, metal sheets, shadows, self-blocking bricks, meadows, and bare soil. For these data, a fixed training set is available comprising 3921 training samples and 42 776 test samples. Here, we randomly choose ten samples per class from the available 3921 training samples in order to build the training dictionary. Table IV shows the OA, AA, and individual class accuracies obtained for different classification algorithms. As it can be observed in Table IV, the proposed sparse classification technique $\text{SUnSAL}_{\text{EMAP}}$ provided the best results in terms of OA and individual class accuracies, followed by $\text{SVMCK}_{\text{EMAP}}$ and OMP_{EMAP}. In comparison, the use of both SVM-based and sparse classification approaches on the original spectral information provided lower classification accuracies. Concerning the computational cost, similar observations can be made as in the former experiments. In other words, the spectral-spatial methods are less computationally efficient than the spectral-based methods but provide, in turn, higher accuracies. Another observation is that SUnSAL-based methods are faster than OMP-based methods. For illustrative purposes, Fig. 8 shows some of the obtained classification maps in this experiment. Finally, Table V shows a comparison between the proposed $\text{SUnSAL}_{\text{EMAP}}$ method and several classifiers (reported in [37]) that use 3D-DWT-based input features. As shown in Table V, the proposed $\text{SUnSAL}_{\text{EMAP}}$ can provide very competitive results in comparison with other methods.

D. Experiments With a Quickbird Scene Collected Over Beijing, China

In order to illustrate the performance of the proposed approaches with a multispectral data set, in this experiment, we use a Quickbird scene collected over the city of Beijing, China, in September 2003. The scene is composed of 400 $\times$ 600 pixels, and the spatial resolution is 2.44 m. The data comprise five spectral bands: panchromatic (450–900 nm), blue (450–520 nm), green (520–600 nm), red (630–690 nm), and near IR (760–900 nm). In our experiments, we use all the available bands and built a training dictionary made up of ten samples per class, which are randomly selected from the available reference data. The remaining samples were used for validation purposes. Table VI shows the OA, AA, and individual class accuracies obtained for different classification algorithms. As in the previous experiments, the $\text{SUnSAL}_{\text{EMAP}}$ provided the best classification results, with comparable or better computational cost in comparison with its competitors. Furthermore, the SVM_{ori} could not properly model the shadow class, which is a fact that significantly decreased its OA. As shown in Fig. 9, this is related with the confusion between the water and shadow classes, which have similar spectral signatures in the multispectral image; thus, in this case, the shadow samples can be properly represented by atoms from the water class. In contrast, $\text{SUnSAL}_{\text{EMAP}}$ obtains the best classification accuracy (96.85%) for the shadow class, which further demonstrates that EMAPs can increase the separability between the shadow and water classes by means of the incorporation of spatial information in the EMAP space; hence, in this case, the shadow class can be better reconstructed using atoms from the same class in EMAP space.

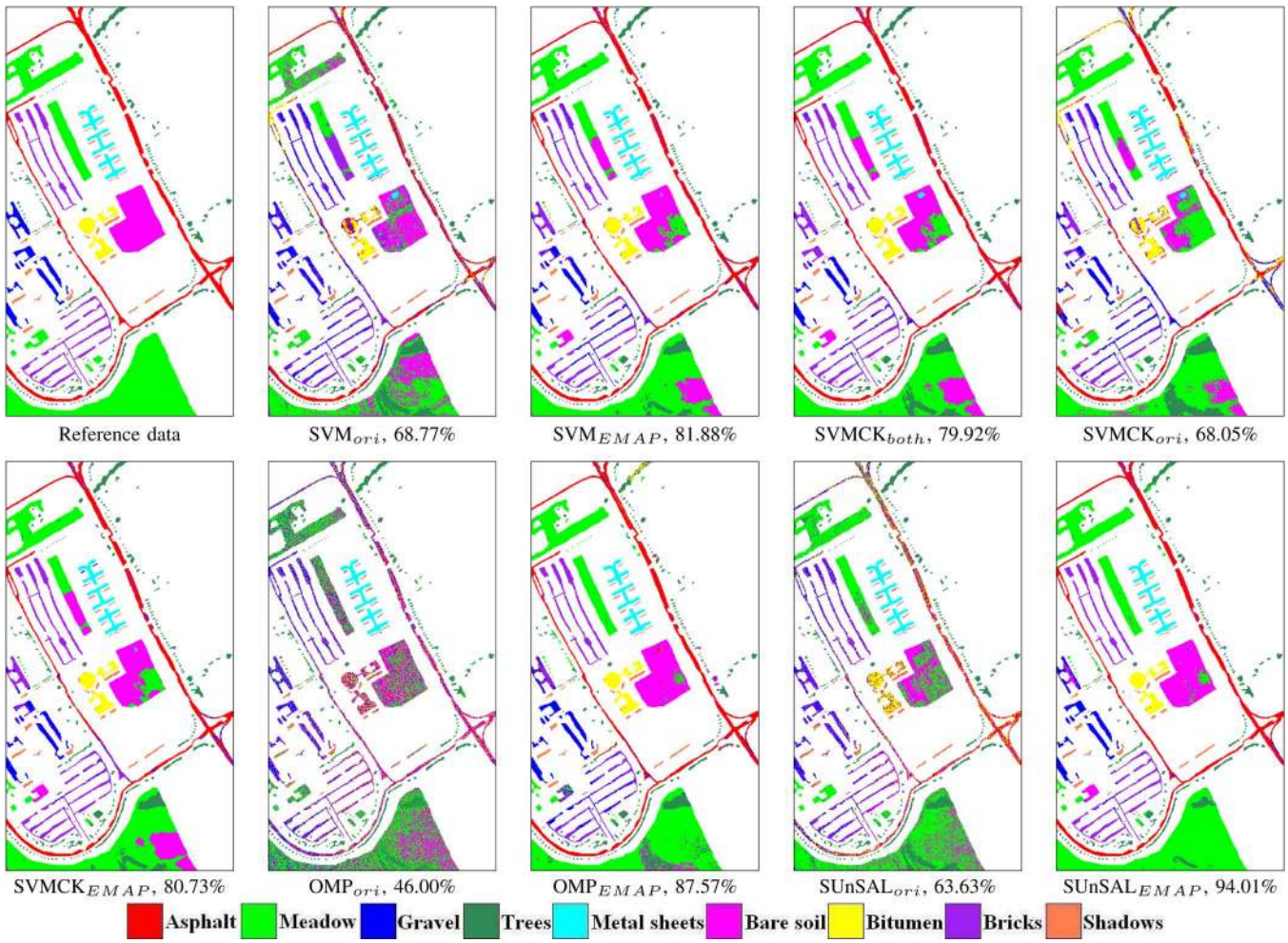

Fig. 8. Classification results obtained by different classifiers for the ROSIS Pavia University scene (only ten samples per class are used for training, for a total of 90 training samples).

TABLE V
OA, AA, AND KAPPA STATISTIC (κ) OBTAINED AFTER EXECUTING 50 MONTE CARLO RUNS OF THE PROPOSED SUnSAL$_{\text{EMAP}}$ ALGORITHM FOR THE ROSIS PAVIA UNIVERSITY DATA, IN COMPARISON WITH SEVERAL METHODS THAT USE 3D-DWT INPUT FEATURES, AS REPORTED IN [37]. SVM-LINEAR AND SVM-RBF DENOTE THE SVM CLASSIFIERS WITH LINEAR AND RBF KERNELS, RESPECTIVELY. LASSO, GLASSO, SGLASSO, AND MIXED-LASSO ARE FOUR ALGORITHMS SOLVING THE SPARSE-REPRESENTATION-BASED PROBLEM. ADDITIONAL DETAILS ABOUT THESE ALGORITHMS CAN BE FOUND IN [37]

Data Set	Input feature	3D-DWT						EMAP
	Method	SVM-linear	SVM-RBF	Lasso	Glasso	Sglasso	Mixed-lasso	SUnSAL
1% of the labeled samples	OA	71.16	71.02	74.23	75.17	75.15	77.03	85.99
	AA	72.86	75.89	78.67	79.56	79.56	79.25	88.46
	κ	63.75	63.70	92.27	68.78	68.71	70.69	81.93
10% of the labeled samples	OA	93.15	95.45	92.79	92.57	93.04	94.57	97.46
	AA	92.63	94.20	92.87	92.41	93.21	92.63	96.74
	κ	90.97	93.96	90.53	90.25	90.27	92.82	96.62

IV. CONCLUSION AND FUTURE LINES

In this paper, we have developed a new classification strategy that integrates sparse representations and EMAPs for spatial–spectral classification of remote sensing data. Our experiments reveal that the proposed approach, which combines the advantages of sparse representation and the rich structural information provided by EMAPs, can appropriately exploit the inherent sparsity present in EMAPs in order to provide

TABLE VI
OA, AA, KAPPA STATISTIC (κ), AND CLASS INDIVIDUAL ACCURACIES ([%]) OBTAINED BY DIFFERENT CLASSIFIERS ON THE QUICKBIRD BEIJING DATA (HERE, WE USE ONLY TEN SAMPLES PER CLASS FOR TRAINING, FOR A TOTAL OF 50 TRAINING SAMPLES)

Class	Train	Test	SVM_{ori}	SVM_{EMAP}	SVMCK_{both}	SVMCK_{ori}	SVMCK_{EMAP}	OMP_{ori}	OMP_{EMAP}	$k\text{NN}_{ori}$	$k\text{NN}_{EMAP}$	SUnSAL_{ori}	SUnSAL_{EMAP}
Building	10	6222	62.11	69.61	67.52	55.63	68.41	33.53	58.66	64.42	63.82	35.05	**72.15**
Vegetation	10	4734	95.93	91.31	97.32	97.13	98.17	90.86	91.71	97.43	87.43	**98.59**	98.12
Ground	10	4878	74.14	65.88	78.93	69.10	**81.81**	51.55	78.78	70.46	75.74	61.06	78.41
Shadow	10	1178	77.29	94.30	95.98	65.33	96.28	0	92.95	81.25	91.82	6.53	**96.85**
Water	10	2079	96.73	99.03	98.80	87.52	98.53	93.07	**99.98**	85.21	99.40	99.17	99.84
Overall			77.19	78.77	83.00	73.44	84.22	56.76	78.61	77.56	78.32	62.68	**85.05**
Average			79.24	84.03	87.71	74.94	88.64	53.80	84.42	79.95	83.64	60.08	**89.07**
κ			69.98	71.89	77.62	65.39	79.23	44.16	72.01	70.42	71.53	51.24	**80.24**
Time(s)			0.30	61.38	68.4	1.5	77.7	41.97	281.43	0.40	62.97	21.02	80.48

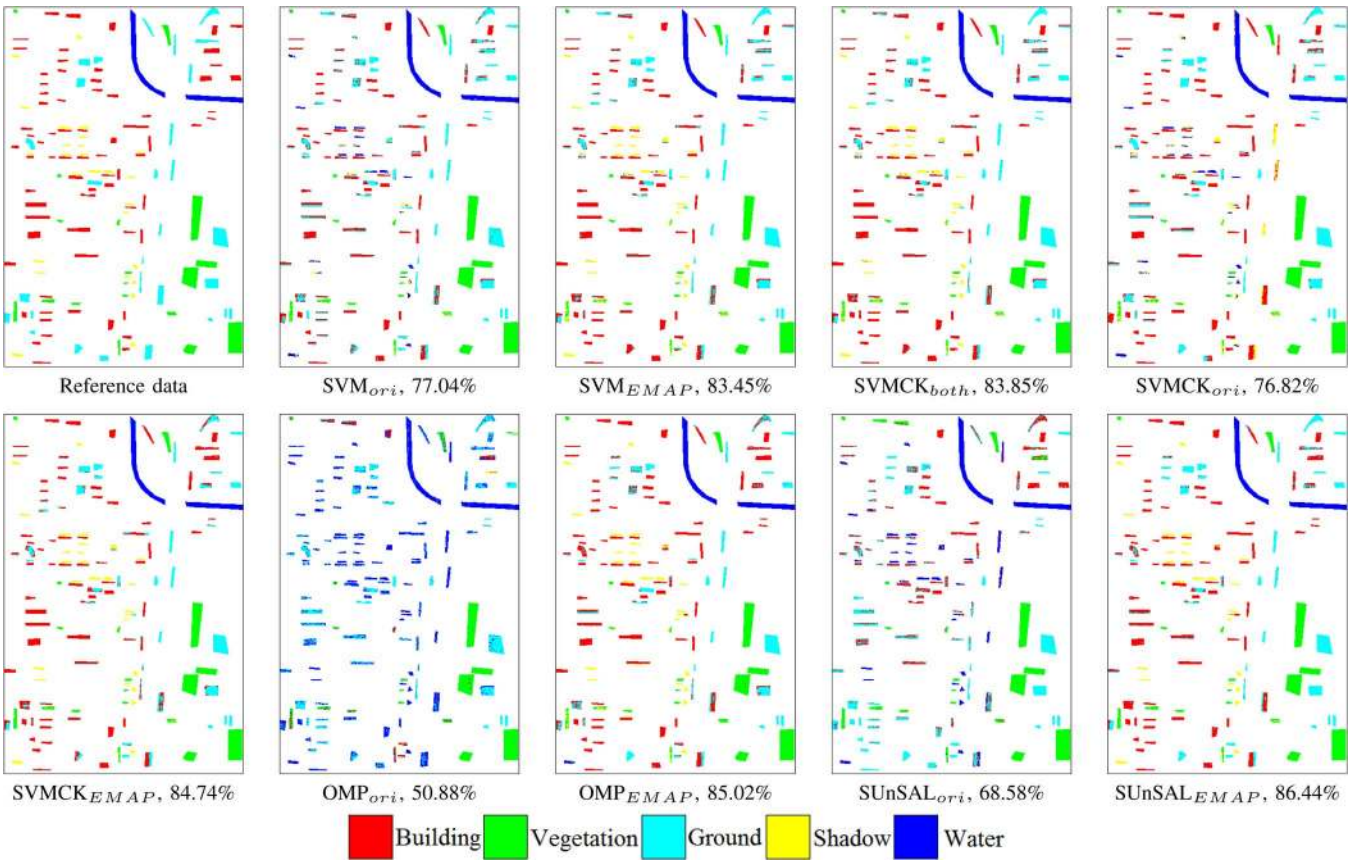

Fig. 9. Classification results obtained by different classifiers for the Quickbird Beijing scene (only ten samples per class are used for training, for a total of 50 training samples).

state-of-the-art classification results. This is mainly due to the fact that the samples in EMAP space can be approximately represented by a few number of atoms in the training dictionary after solving the optimization problem, whereas the same samples could not be represented in the original spectral space with the same level of sparsity. The proposed strategy was tested on both simulated and real multi/hyperspectral data sets. A comparison with state-of-the-art classifiers shows very promising results for the proposed approach, particularly when a very limited number of training samples are available. In this context, the SUnSAL algorithm provided excellent classification performance as compared with other techniques. Future work will be focused on the development of computationally efficient implementations for the proposed approach.

ACKNOWLEDGMENT

The authors would like to thank Prof. D. Landgrebe for making the AVIRIS Indian Pines hyperspectral data set available to the community and Prof. P. Gamba for providing the ROSIS data over Pavia, Italy, along with the training and test sets. Last but not least, the authors would like to thank the Associate Editor and the three anonymous reviewers for their detailed and highly constructive criticisms, which greatly helped us to improve the quality and presentation of our manuscript.

REFERENCES

[1] J. M. 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] Y. Chen, N. Nasrabadi, and T. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3973–3985, Oct. 2011.

[3] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[4] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[5] Q. Sami ul Haq, L. Tao, F. Sun, and S. Yang, "A fast and robust sparse approach for hyperspectral data classification using a few labeled samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2287–2302, Jun. 2012.

[6] L. Bruzzone and C. Persello, "A novel context-sensitive semisupervised SVM classifier robust to mislabeled training samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2142–2154, Jul. 2009.

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

[8] J. Li, J. Bioucas-Dias, and A. Plaza, "Semi-supervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4085–4098, Nov. 2010.

[9] Y. Tarabalka, J. Chanussot, and J. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recog.*, vol. 43, no. 7, pp. 2367–2379, Jul. 2010.

[10] 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.

[11] J. S. Borges, J. M. Bioucas-Dias, and A. R. S. Marcal, "Bayesian hyperspectral image segmentation with discriminative class learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2151–2164, Jun. 2011.

[12] P. Soille, *Morphological Image Analysis, Principles and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2003.

[13] M. Pesaresi and J. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.

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

[15] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.

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

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

[18] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Extended profiles with morphological attribute filters for the analysis of hyperspectral data," *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5975–5991, Jul. 2010.

[19] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. New York, NY, USA: Wiley, 2003.

[20] 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.

[21] M. Chi and L. Bruzzone, "A semi-labeled-sample driven bagging technique for ill-posed classification problems," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 1, pp. 69–73, Jan. 2005.

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

[23] M. Chi and L. Bruzzone, "An ensemble-driven k-NN approach to ill-posed classification problems," *Pattern Recog. Lett.*, vol. 27, no. 4, pp. 301–307, Mar. 2006.

[24] M. Chi and L. Bruzzone, "Semi-supervised classification of hyperspectral images by SVMs optimized in the primal," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1870–1880, Jun. 2007.

[25] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J.C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. Suppl. 1, pp. 110–122, Sep. 2009.

[26] 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.

[27] P. R. Marpu, M. Pedergnana, M. Dalla Mura, S. Peeters, J. A. Benediktsson, and L. Bruzzone, "Classification of hyperspectral data using extended attribute profiles based on supervised and unsupervised feature extraction techniques," *Int. J. Image Data Fusion*, vol. 3, no. 3, pp. 269–298, Sep. 2012.

[28] M. Pedergnana, P. R. Marpu, M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "A novel technique for optimal feature selection in attribute profiles based on genetic algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3514–3528, Jun. 2013.

[29] M. Pesaresi, G. K. Ouzounis, and L. Gueguen, "A new compact representation of morphological profiles: Report on first massive VHR image processing at the JRC," in *Proc. SPIE*, 2012, vol. 8390, pp. 839 025-1–839 025-6.

[30] M. Elad, "Sparse and redundant representation modeling—What next?" *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 922–928, Dec. 2012.

[31] P. Salembier, A. Oliveras, and L. Garrido, "Anti-extensive connected operators for image and sequence processing," *IEEE Trans. Image Process.*, vol. 7, no. 4, pp. 555–570, Apr. 1998.

[32] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 43, no. 1, pp. 129–159, 2001.

[33] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[34] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[35] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.

[36] 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.

[37] Y. Qian, M. Ye, and J. Zhou, "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features," *IEEE Trans. Geosci. Remote Sens.*,, vol. 51, no. 4, pp. 2276–2291, Apr. 2013.

[38] T. M. Mitchell, *Machine Learning*, 1st ed. New York, NY, USA: McGraw-Hill, 1997.

[39] P. R. Marpu, M. Pedergnana, M. D. Mura, J. A. Benediktsson, and L. Bruzzone, "Automatic generation of standard deviation attribute profiles for spectral-spatial classification of remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 2, pp. 293–297, Mar. 2013.

Benqin Song received the B.E. degree in electronic and information engineering from China University of Geosciences, Beijing, China, in 2007. He is currently working toward the Ph.D. degree in the School of Earth and Space Sciences, Peking University, Beijing, China.

He was a Visiting Student with the Hyperspectral Computing Laboratory (HyperComp) Research Group, Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain. His current research interests include remote sensing image classification and segmentation, urban remote sensing analysis, and decision fusion.

Jun Li 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 been also actively involved in the Hyperspectral Imaging Network, which is 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 is 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 was a recipient of the 2012 Best Reviewer Award of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

Mauro Dalla Mura (S'08–M'11) received the B.E. (*laurea*) and M.E. (*laurea specialistica*) degrees in telecommunication engineering from the University of Trento, Trento, Italy, in 2005 and 2007, respectively, and the joint Ph.D. degree in information and communication technologies (telecommunications area) and in electrical and computer engineering from the University of Trento and the University of Iceland, Reykjavik, Iceland, respectively, in 2011.

In 2011, he was a Research Fellow with the Fondazione Bruno Kessler, Trento, Italy, conducting research on computer vision. He is currently an Assistant Professor with the Grenoble Institute of Technology (Grenoble INP), Grenoble, France. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab), Saint Martin d'Heres, France. His main research activities are in the fields of remote sensing, image processing, and pattern recognition. In particular, his interests include mathematical morphology, classification, and multivariate data analysis.

Dr. Dalla Mura was the recipient of the IEEE GRSS Second Prize in the Student Paper Competition of the 2011 IEEE International Geoscience and Remote Sensing Symposium 2011 (Vancouver, CA, July 2011). He is a Reviewer for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND REMOTE SENSING, IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, *Pattern Recognition Letters*, International Society for Photogrammetry and Remote Sensing *Journal of Photogrammetry and Remote Sensing*, and *Photogrammetric Engineering and Remote Sensing*.

Peijun Li (M'02–SM'04) received the Ph.D. degree in geology from the Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, in 1995.

He is currently a Professor with Peking University, Beijing. Since 2008, he has been a Co-Chair of International Society for Photogrammetry and Remote Sensing Working Group VII/4 "Methods for land-cover classification." He was a Visiting Researcher/Professor with several institutions, including Seoul National University, Seoul, Korea; University of Bonn, Bonn, Germany; and Columbia University, New York, NY, USA. His current research interests include land-cover classification and change detection, very high resolution image analysis, urban remote sensing analysis, and geologic remote sensing.

Dr. Li is a Reviewer for many international journals.

Antonio Plaza (M'05–SM'07) was born in Cáceres, Spain, in 1975. He received the M.S. and Ph.D. degrees in computer engineering from the University of Extremadura, Cáceres.

He is an Associate Professor (with accreditation for Full Professor) with the Department of Technology of Computers and Communications, University of Extremadura, where he is the Head of the Hyperspectral Computing Laboratory (HyperComp). During 2007–2011, he was the Coordinator of the Hyperspectral Imaging Network, which is a European project with total funding of 2.8 M. He authored more than 370 publications, including more than 100 Journal Citation Reports (JCR) journal papers (60 in IEEE journals), 20 book chapters, and over 230 peer-reviewed conference proceeding papers (90 in IEEE conferences). He has guest edited seven special issues on JCR journals (three in IEEE journals).

Dr. Plaza served as the Chair for the IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, in 2011. He was a recipient of the recognition of Best Reviewers of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2009 and of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2010, a journal for which he served as an Associate Editor in 2007–2012. He is also an Associate Editor for the IEEE Geoscience and Remote Sensing Magazine. He was a member of the Editorial Board of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER in 2011–2012 and a member of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING in 2012. In 2011–2012, he served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS), where has been the President of the Spanish Chapter since November 2012. He has been the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING since January 2013.

José M. Bioucas-Dias (S'87–M'95) received the EE, M.Sc., Ph.D., and "Agregado" degrees from the Instituto Superior Técnico (IST), Technical University of Lisbon (TULisbon), Lisbon, Portugal, in 1985, 1991, 1995, and 2007, respectively, all in electrical and computer engineering.

Since 1995, he has been with the Department of Electrical and Computer Engineering, IST, where he served as an Assistant Professor from 1995 to 2007 and an Associate Professor since 2007. Since 1993, he has been a Senior Researcher with the Pattern and Image Analysis Group, Instituto de Telecomunicaçoes, IST, which is a private nonprofit research institution. His research interests include inverse problems, signal and image processing, pattern recognition, optimization, and remote sensing.

Dr. Bioucas-Dias was an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS in 1997–2000. He is currently an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He was a Guest Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING for the Special Issue on Spectral Unmixing of Remotely Sensed Data and of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING for the Special Issue on Hyperspectral Image and Signal Processing. He is currently a Guest Editor of the IEEE Signal Processing Magazine for the Special Issue on Signal and Image Processing in Hyperspectral Remote Sensing. He was the General Co-Chair of the 3rd IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS'2011) and has been a member of program/technical committees of several international conferences.

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

He is currently a Prorector for Academic Affairs and a Professor of electrical and computer engineering with the University of Iceland. He is also with the biomedical startup company Oxymap, Reykjavik, Iceland, where he was a co-founder. His research interests are in remote sensing, biomedical analysis of signals, pattern recognition, image processing, and signal processing, publishing extensively in those fields.

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

Jocelyn Chanussot (M'04–SM'04–F'12) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995 and the Ph.D. degree from Savoie University, Annecy, France, in 1998.

In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA, i.e., French National Defense Department). Since 1999, he has been with Grenoble INP, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is currently conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab), Saint Martin d'Heres, France. His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing.

Dr. Chanussot is the Founding President of the French chapter of the IEEE Geoscience and Remote Sensing Society (GRSS), where he served from 2007 to 2010 and received the 2010 IEEE GRSS Chapter Excellence Award "for excellence as a Geoscience and Remote Sensing Society chapter demonstrated by exemplary activities during 2009." He was a recipient of the 2011 IEEE GRSS Symposium Best Paper Award. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). In 2009–2011, he was the Chair of the GRS Data Fusion Technical Committee, where he was the Cochair in 2005–2008. He was the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing in 2009. He was an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING LETTERS in 2005–2007 and for *Pattern Recognition* in 2006–2008. Since 2007, he has been an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. Since 2011, he has been the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.