

Multiple Morphological Component Analysis Based Decomposition for Remote Sensing Image Classification

Xiang Xu, Jun Li, *Member, IEEE*, Xin Huang, *Senior Member, IEEE*, Mauro Dalla Mura, *Member, IEEE*, and Antonio Plaza, *Fellow, IEEE*

Abstract—Remote sensing images exhibit significant contrast and intensity regions and edges, which makes them highly suitable for using different texture features to properly represent and classify the objects that they contain. In this paper, we present a new technique based on multiple morphological component analysis (MMCA) that exploits multiple textural features for decomposition of remote sensing images. The proposed MMCA framework separates a given image into multiple pairs of morphological components (MCs) based on different textural features, with the ultimate goal of improving the signal-to-noise level and the data separability. A distinguishing feature of our proposed approach is the possibility to retrieve detailed image texture information, rather than using a single spatial characteristic of the texture. In this paper, four textural features: *content*, *coarseness*, *contrast*, and *directionality* (including *horizontal* and *vertical*), are considered for generating the MCs. In order to evaluate the obtained MCs, we conduct classification by using both remotely sensed hyperspectral and polarimetric synthetic aperture radar (SAR) scenes, showing the capacity of the proposed method to deal with different kinds of remotely sensed images. The obtained results indicate that the proposed MMCA framework can lead to very good classification performances in different analysis scenarios with limited training samples.

Index Terms—Decomposition, image separation, multinomial logistic regression (MLR), multiple morphological component analysis (MMCA), sparse representation, textural features.

Manuscript received April 24, 2015; revised October 23, 2015; accepted December 10, 2015. Date of publication January 18, 2016; date of current version March 25, 2016. (*Corresponding author: Jun Li.*)

X. Xu is with the Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, Center of Integrated Geographic Information Analysis, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China, and also with the University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan 528402, China.

J. Li is with the Guangdong Provincial Key Laboratory of Urbanization and Geo-Simulation, Center of Integrated Geographic Information Analysis, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China (e-mail: lijun48@mail.sysu.edu.cn).

X. Huang is with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China.

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

A. Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, Universidad de Extremadura, 10003 Cáceres, Spain.

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.2015.2511197

I. INTRODUCTION

REMOTE sensing image classification aims at distinguishing different categories or thematic land-cover classes using different features [1]. In the classification process, each image pixel or area is assigned into one of the several thematic categories. An important trend in remote sensing image classification is to incorporate spatial features (e.g., texture or morphology) to improve the classification results that can be obtained using the original image data alone [2]. The incorporation of spatial information is mainly performed as a spatial preprocessing or postprocessing. In addition, some methods, like discriminative random fields [3], [4], conditional random fields [5], [6], and relaxation methods [7], take the spatial information into account during the classification process. On the one hand, spatial preprocessing aims at extracting spatial features. Among the techniques based on this strategy, we can highlight the use of morphological profiles [8], [9], morphological attribute profiles [10], morphological component analysis (MCA) [11], morphological neighborhood filter-based techniques [12], empirical mode decomposition (EMD) [13], [14], wavelet filters [15], and others [16], [17]. On the other hand, postprocessing-based approaches generally perform spatial regularization after classification [18]. For instance, techniques based on partitional clustering [19], watershed transformations [20], relearning algorithms [21], graph-based classification [22], or super pixel approaches [23] have been used for this purpose.

Both preprocessing and postprocessing techniques have received great attention for remotely sensed image classification and achieved remarkable performance [24], [25]. The main difference between these two approaches is that postprocessing performs spatial regularization based on the classification result obtained from the original image, i.e., no new features are introduced. In turn, when spatial information is included at the preprocessing stage, this generally means that a new set of features is used in order to increase data separability. This has fostered significant interest in the use of preprocessing techniques. As preprocessing-based approaches, image-decomposition-based schemes have been successfully applied to different image classification tasks [26], [27]. In [11], the authors proposed an MCA-based image separation approach which constructs a sparse representation of an image and separates the image into morphological components (MCs). In [28], sparsity and morphological diversity have emerged as effective features for blind source separation. In [29], an MCA-based image separation

method was applied to remotely sensed hyperspectral image classification, exhibiting very good results in comparison with postprocessing-based approaches. The basic idea of traditional MCA-based image separation is to choose two dictionaries, i.e., *content* and *texture*, for the representation of their MCs, then compute the sparse coefficients over the images that they are serving, and finally generate the decomposed *content* and *texture* components.

As reported in previous works [11], [29], the effectiveness of MCA is mainly due to the following issues. On the one hand, via the considered decomposition approach, new features (components) are obtained, which can lead to better image separability. On the other hand, the smoothness component generally shows better signal-to-noise ratio in comparison with the original image. However, the traditional MCA decomposes an image only into *content* and *texture* components, in which new features are fixed (and limited in number), leading to limited improvements in data exploration. Furthermore, the MCA-based decomposition neglects the fact that there are many different kinds of textural features, such as *coarseness*, *directionality*, etc., which may be essential for describing the spatial information contained in the image. In order to eliminate these deficiencies and better exploit the spatial textural information contained in the image, we propose a multiple MCA (MMCA) approach for image separation that uses an approach similar to other unsupervised feature extraction approaches that use both spatial and spectral features [30]. The proposed MMCA is based on the fact that an image can be described by different textural features and then can be separated into a smoothness and a texture component for each textural feature, where five textural features: *content*, *coarseness*, *contrast*, and *directionality* (including *horizontal* and *vertical*), are considered in this work. Since remote sensing images exhibit significant contrast and intensity regions and edges, this makes them highly suitable for using different texture features to properly represent and classify the objects that they contain. A sparse representation method is then adopted for the image separation, where the dictionaries are randomly generated from the image and constructed by performing transformations based on a given textural feature. In order to evaluate the proposed MMCA, we conduct classification on the extracted features via a multinomial logistic regression (MLR) based classifier by using the variable splitting and augmented Lagrangian (LORSAL) algorithm [31].

The remainder of this paper is organized as follows. Section II presents the proposed MMCA approach and provides the details of the considered MMCA-based image separation technique, as well as the construction of the corresponding dictionaries for different textural features. Section III discusses the experimental results intended to test the performance of the proposed MMCA scheme. The experiments are conducted using two real hyperspectral data sets, respectively, collected by the Reflective Optics Spectrographic Imaging System (ROSIS) over the city of Pavia, Italy, and by the Airborne Visible/infrared Imaging Spectrometer (AVIRIS) over the Indian Pines region in Indiana. We also use polarimetric synthetic aperture radar (SAR) data sets collected by AirSAR over the Flevoland site in The Netherlands and by the Electromagnetics Institute

Synthetic Aperture Radar (EMISAR) system over Foulum, Denmark. Our experiments, focused on a comparison by using single pairs of MCs, stacked MCs (all of the MCs are stacked together), linear combinations of the MCs, and multiple kernel learning, demonstrate that the proposed approach achieves very good classification results using both hyperspectral and polarimetric SAR data. Section IV concludes this paper with some remarks and hints at plausible future research lines.

II. MMCA-BASED IMAGE SEPARATION

In this section, we describe the proposed MMCA-based image separation scheme. A first relevant issue is the dimensionality of the original data set, particularly in remotely sensed scenes with high spectral resolution such as hyperspectral data. In order to reduce the data dimensionality, we use the minimum noise fraction (MNF) [32] to retain a small number of components in comparison with the number of bands in the original data as discussed in Section II-A. The traditional MCA is outlined in Section II-B. The core of the proposed MMCA, which aims at decomposing a given image into a smoothness and a texture component by a given feature, will be presented in Section II-C. For simplicity, in this paper, we use the term “components” to refer to a new set of uncorrelated images obtained by certain transformations on the original data, while we use the term “feature” to denote the spatial textural feature extracted from each image. Section II-D introduces the details of image separation based on the considered textural features. Finally, Section II-E briefly outlines the considered classification strategy.

A. Dimensionality Reduction

Classification of remotely sensed images can be performed using all of the original image information. However, for high-dimensional data sets such as hyperspectral images, it is common that the data generally live in a subspace of much lower dimensionality in comparison with the original spectral space [33], [34]. Following our previous work [29], we propose to use the MNF [32] to reduce the dimensionality of the original data. MNF is a widely used and effective technique for dimensionality reduction which, in comparison to principal component analysis (PCA) [35], [36], considers the influence of noise. Here, we use the MNF to retain a number of components that contain 99% of the information in the original data sets. Nevertheless, it should be noted that the main goal of this step is to reduce the computational cost by reducing the data dimensionality. At this point, we emphasize that we have experimentally tested that there is no significant difference in using the MNF, PCA, or other dimensionality reduction transformation in our proposed strategy.

B. MCA-Based Image Separation

MCA is a method which allows us to separate features contained in an image when these features present different morphological aspects [11]. For an image $y \in \mathbb{R}^N$, where N

is the number of pixels in \mathbf{y} , and a given textural feature, the task is to seek the sparsest solution for the problem as follows:

$$\begin{aligned} \mathbf{x} &= \arg \min \|\mathbf{x}\|_1 \\ \text{subject to: } \mathbf{y} &= \mathbf{A}\mathbf{x} \end{aligned} \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^K$ denotes the sparse coefficients of the MC, $\mathbf{A} \in \mathbb{R}^{N \times K}$ denotes the associated dictionaries, and K denotes the number of atoms in the dictionary (typically $K > N$); $\|\cdot\|_1$ is the ℓ_1 norm, which is a tractable convex constrained optimization problem and can reduce the computational complexity by linear programming. In image textural separation, we often use MCA to decompose an image into *texture* and *content* components. In the work presented in [29], MCA-based decomposition is applied for hyperspectral image classification, where the results obtained were very promising.

C. MMCA-Based Image Separation

For a given image \mathbf{y} with N pixels, the objective of MMCA is to separate it into two components: a smoothness component \mathbf{y}_s and a texture component \mathbf{y}_t , respectively. These components represent the original image under a linear combination as follows:

$$\mathbf{y} = \mathbf{y}_s + \mathbf{y}_t + \mathbf{n} \quad (2)$$

where \mathbf{n} is the residual in the approximation of the image. The traditional MCA separates the image \mathbf{y} into *content* \mathbf{y}_s and *texture* \mathbf{y}_t components. This conventional formulation limits the exploration of the spatial information contained in remote sensing images, which are dominated by significant contrast and intensity regions and edges that can be better captured by using more than one type of textural features. In order to fully exploit the spatial texture information contained in the image, we propose an MMCA scheme to better describe the textural features. As a result, in addition to the traditional *content* feature, we consider four new textural features, namely, *coarseness*, *contrast*, *horizontal*, and *vertical* for separation [37]. Fig. 1 illustrates the proposed MMCA decomposition scheme, where a toy example with five different classes is considered. For illustrative purposes, the considered image has distinct textural features. As can be seen from Fig. 1, the proposed MMCA includes three main steps. In the first step, we randomly choose several partitions from image \mathbf{y} for the initialization of the two dictionaries. In the second step, for the two components (i.e., the smoothness \mathbf{y}_s and its texture \mathbf{y}_t), we build two corresponding dictionaries based on certain transformations on the chosen image partitions. Finally, in the last step, sparse coding is performed to learn the MC coefficients. At the same time, following [29], the associated dictionaries are iteratively updated by adopting total variation and hard threshold regularization. After separation, it is observable that different textural features lead to specific components.

For the aforementioned three steps, the most difficult part is the construction of the dictionaries involved in the second step, which are essential for the learning of the sparse coefficients. Therefore, in the following, we provide a detailed description of

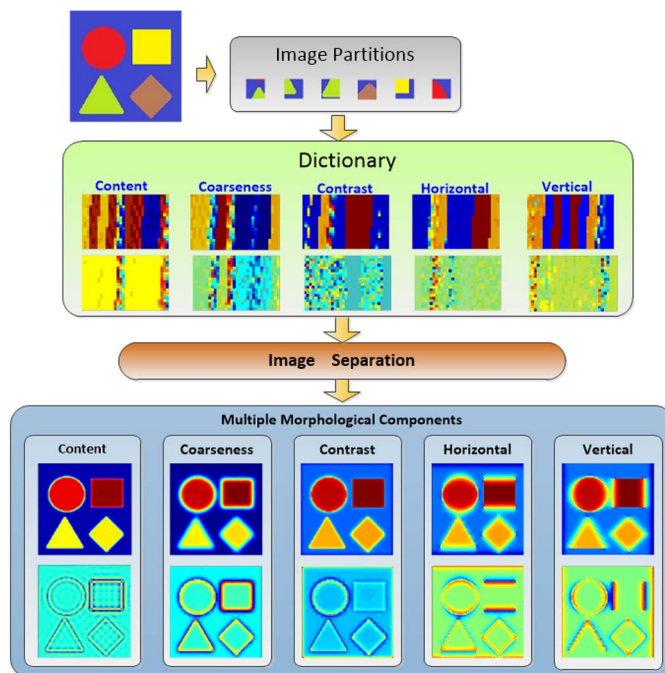


Fig. 1. Proposed MMCA-based image separation scheme for a toy example with five classes and distinct textural features.

the textural features and their corresponding dictionaries used in this work.

1) *Content Feature*: This feature represents the traditional textural feature used for MCA decomposition, resulting in the standard *cartoon* and *texture* MCs. For the content component, which allows for the extraction of anisotropic structures, smooth curves and edges of different lengths in an image can be extracted by the *curvelet* transform [38], [39], biorthogonal wavelet transform [40], undecimated wavelet transform [41], and local ridgelet transform [42], among others. Following [29], for the content component, we use a local curvelet transform to generate the dictionary from the randomly chosen image partitions. For the texture component, the local discrete cosine transform [43] or the Gabor transform [44] can be used to build a morphological dictionary. Similar to [29], a local Gabor wavelet transform is adopted to build the dictionary from the same image partitions.

2) *Coarseness Feature*: This is a relevant textural feature in an image. As the bilateral filter is a nonlinear, edge-preserving, and noise-reducing smoothing filter for images [45], we use it to build a coarseness dictionary. Bilateral filtering replaces the intensity value at each pixel in an image with a weighted average of intensity values from nearby pixels. This weight can be based on a Gaussian distribution. Crucially, the weights depend not only on the Euclidean distance between pixels but also on the radiometric differences (e.g., range differences such as color intensity or depth distance). This preserves sharp edges by systematically looping through each pixel and adjusting weights to the adjacent pixels accordingly. For the opposite component, we use a wavelet thresholding filter [46] to preserve small edges and elements while weakening strong edges and larger elements in the image.

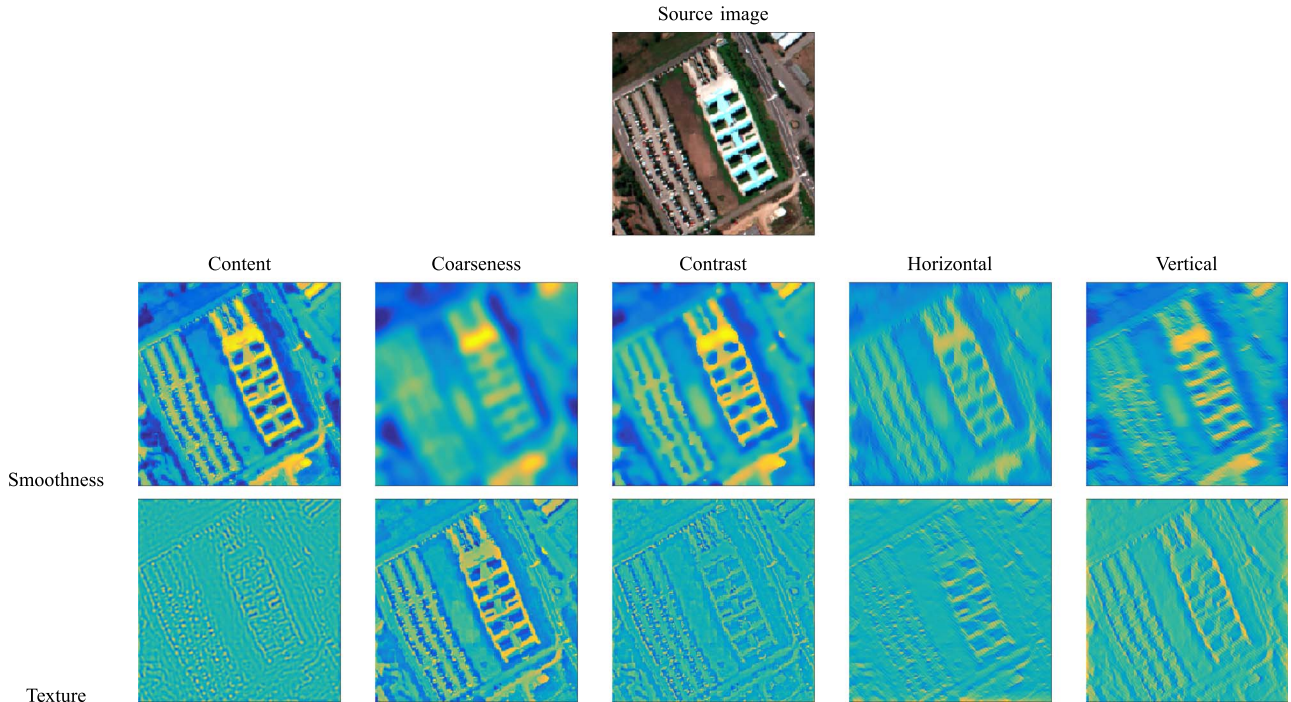


Fig. 2. Toy example: remote sensing image separation results based on the proposed MMCA framework.

3) *Contrast Feature*: This feature measures the variance of the gray-scale distribution, where high and low contrast values mean fast and slow intensity changes. In [37], the authors explored four factors for contrast, including the dynamic range of gray-levels, ratio of black and white areas, sharpness of edges, and period of repeating patterns. In this paper, we adopt the anisotropic diffusion (AD) [47] and its modification to build high-contrast and low-contrast dictionaries. On the one hand, after applying the AD, the high-contrast regional textures will be smoothed, while the low-contrast regional texture will be preserved. On the other hand, by changing the diffusion coefficient from positive to negative, we can obtain the opposite behavior for low-contrast regional textures.

4) *Directionality Feature*: This is a global property which describes the orientation of the local texture. In [37], directionality just measures the total degree of directionality, while the orientation of the texture pattern was not taken into consideration. In this paper, two directional features, i.e., horizontal and vertical, are considered. Again, for each feature, two dictionaries, one for the smoothness component and another one for the texture component, are constructed. Here, we use a wavelet thresholding filter based on the stationary wavelet transform (SWT) [48] to build the dictionaries. The SWT is a wavelet transform algorithm designed to overcome the lack of translation-invariance of the discrete wavelet transform, where we use an SWT thresholding filter to preserve different high-frequency subband coefficients so that the texture is decomposed into different direction components.

A final aspect to point out is that, although only four textural features are used for decomposition in this work, other textural features such as *line-likeness*, *regularity*, *roughness* [37], etc., can also be considered according to the image properties.

Nevertheless, in our experiments, we have empirically found out that the four considered textural features are able to cover the contextual information and produce good quality components, leading to excellent classification accuracies.

In order to illustrate the proposed MMCA framework, we present a toy experiment for image separation which is based on a real remote sensing image with 220×220 pixels, which is a portion of the well-known ROSIS Pavia University hyperspectral data set. The separation results are shown in Fig. 2. For each type of textural feature, we use the aforementioned transforms to generate the corresponding MCs, where the leftmost column of Fig. 2 gives the traditional MCA separation results (*content* feature) and the other columns correspond to *coarseness*, *contrast*, *horizontal*, and *vertical* features. The toy example demonstrates that our proposed MMCA can bring new additional feature information when compared to traditional MCA. For instance, it can be observed in Fig. 2 that the behavior of the *contrast* feature, which focuses on the density changes, is quite different from that of the *content* component. A similar observation can be made for the *directionality* features, in which *horizontal* and *vertical* features provide complementary information. Since the objects typically contained in remote sensing images (e.g., roads and buildings) have distinguishable directionality, it is essential to consider more features than simply the *content*, as it is the case with the traditional MCA method.

D. Image Separation

Let \mathbf{A}_s and \mathbf{A}_t be the dictionaries (for the smoothness and texture components, respectively) for a given textural feature. Let \mathbf{x}_s and \mathbf{x}_t be the sparse coefficients corresponding to the

smoothness \mathbf{y}_s and texture \mathbf{y}_t components, respectively. For a given image \mathbf{y} , we can obtain

$$\mathbf{y} = \mathbf{y}_s + \mathbf{y}_t + \mathbf{n} = \mathbf{A}_s \mathbf{x}_s + \mathbf{A}_t \mathbf{x}_t + \mathbf{n}. \quad (3)$$

Notice that, in the original MCA work presented in [29], the dictionaries \mathbf{A}_s and \mathbf{A}_t and the components \mathbf{y}_s and \mathbf{y}_t are all linked to the *content* textural feature. One of the main innovations of our proposed MMCA is its capacity to consider different textural features, i.e., the dictionaries \mathbf{A}_s and \mathbf{A}_t are constructed according to different image properties, leading to different image components \mathbf{y}_s and \mathbf{y}_t .

Following [49], we can obtain \mathbf{y}_s and \mathbf{y}_t by solving the following optimization problem:

$$\langle \hat{\mathbf{y}}_s, \hat{\mathbf{y}}_t \rangle = \arg \min_{\mathbf{y}_s, \mathbf{y}_t} \frac{1}{2} \|\mathbf{y} - \mathbf{y}_s - \mathbf{y}_t\|_2^2 + \lambda_1 \|\mathbf{T}_s \mathbf{y}_s\|_1 + \lambda_2 \|\mathbf{T}_t \mathbf{y}_t\|_1 \quad (4)$$

where λ_1 and λ_2 are the regularization parameters, $\mathbf{T}_s = (\mathbf{A}_s^T \mathbf{A}_s)^{-1} \mathbf{A}_s^T$ and $\mathbf{T}_t = (\mathbf{A}_t^T \mathbf{A}_t)^{-1} \mathbf{A}_t^T$ are the pseudoinverse of \mathbf{A}_s and \mathbf{A}_t , which are derived from $\mathbf{y}_s = \mathbf{A}_s \mathbf{x}_s$ and $\mathbf{y}_t = \mathbf{A}_t \mathbf{x}_t$, respectively. Concerning the sparse coding stage, in this work, problem (4) is solved via the spectral unmixing by variable splitting and augmented Lagrangian (SUnSAL) algorithm [49]. It should be noted that, at each iteration, \mathbf{A}_s and \mathbf{A}_t are updated by using total variation and hard threshold regularization.

E. Classification

After the MMCA-based image separation has been conducted, we perform classification to evaluate the quality of the obtained components. Generally, the smoothness components have been traditionally used for this purpose. In our particular case, we use the MLR classifier [31] implemented via the logistic regression via variable splitting and augmented Lagrangian (LORSAL) algorithm [31], which has been shown to be an effective and computationally efficient technique to deal with training sets with limited training samples, regardless of the number of classes.

III. EXPERIMENTAL RESULTS

In this section, we discuss the performance of the proposed method using different remote sensing images collected by hyperspectral and polarimetric SAR instruments. The hyperspectral data were collected by the ROSIS and AVIRIS instruments, while the polarimetric SAR data were collected by the AirSAR and EMISAR instruments. The main objective of our experiments is to show the ability of the MMCA-based image separation technique to exploit textural features for classification using limited training samples. We will analyze the effectiveness of the proposed framework based on multiple features, while the selection of an optimized (single) textural feature is left for future developments of this work. Nevertheless, we will also present a discussion on individual texture features for the two considered data sets. Before describing our experiments,

we first introduce the parameter settings and notations adopted in our experiments.

- 1) In our experiments, only the smoothness components are considered for classification purposes. Nine different types of results, including one from the original MNF component denoted as “raw,” five from the textural features denoted as “content,” “coarseness,” “contrast,” “horizontal,” and “vertical,” and three from the combinations of the MCs denoted as “ \sum_s ,” “ \mathbf{V}_s ,” and “ CK_s ,” are reported, where $\sum_s \equiv 1/t \sum_i \mathbf{y}_{s_i}$, t is the number of textural features, here fixed to $t = 5$, and \mathbf{y}_{s_i} is the smoothness component of the i th textural feature; $\mathbf{V}_s \equiv [\mathbf{y}_{s_1}, \dots, \mathbf{y}_{s_t}]$ is a collection of all of the smoothness components; and CK_s follows a composite kernel learning framework [50]. In our work, in order to simplify the classification complexity, only the Gaussian radial basis function kernel is considered, and every kernel is equally weighted. It should be noted that the “content” textural feature, which represents the traditional MCA approach, is implemented as in [29].
- 2) For the parameters involved in the classification, we follow the procedure described in [31]. Although not optimal, this leads to very good performance. Another reason that we use this suboptimal setting is that the objective of our experiment is mainly to evaluate the proposed MMCA scheme via classification.
- 3) For dimensionality reduction, as mentioned in Section II-A, we use the MNF to retain a number of components that contain 99% of the spectral information in the original hyperspectral data sets, resulting in ten MNF components for both the AVIRIS and ROSIS images, in which the considered spectral information is 99.7% and 99.8%, respectively.
- 4) For the construction of initial dictionaries, in our previous work [29], we illustrated that, when the dictionary size increases, the classification improvements are not relevant. However, the computational time increases significantly. Following [29], the size of the image partition for the considered hyperspectral images is set to 8×8 pixels, and we use ten partitions (randomly chosen from the original image). For the polarimetric SAR data sets, we perform an investigation on the impact of the size and number of dictionaries used for classification purposes.
- 5) For the parameters involved in the dictionary transform, we empirically selected their values after trial and error. For the curvelet transform, we split the frequency domain into $\log_2(\min(M, N)) - 3$ partitions (where M, N is the size of the image), and the coarse scale is set as 1. For the local Gabor transform, the frequency is set as $1/4$, and three scale levels with four orientations are adopted to build the Gabor filter bank. For the bilateral filtering transform, the half-size of the window and the spatial-domain standard deviation are all set as 3, and the intensity domain standard deviation is set to 10. For the AD, the number of iterations is set to 15, and the gradient modulus threshold is set to 30. Again, although

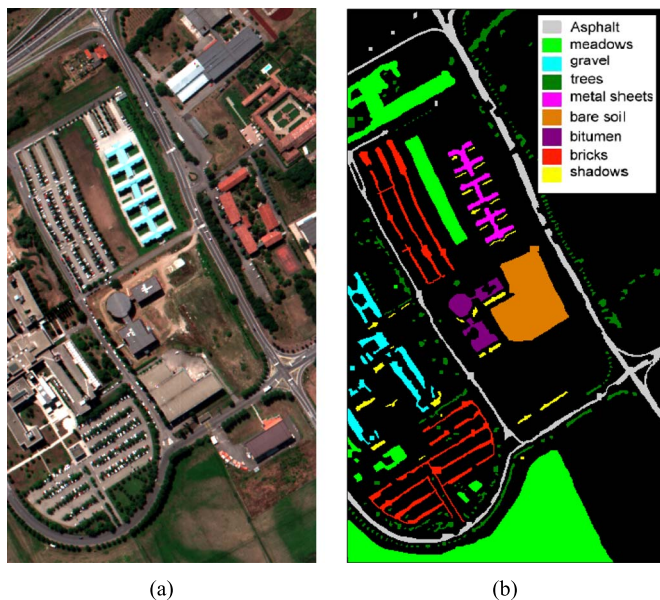


Fig. 3. (a) False color composition of the ROSIS Pavia scene. (b) Ground-truth map containing nine mutually exclusive land-cover classes.

this empirical strategy might be suboptimal, it has been observed to produce good results in practice.

- 6) The training set used for the experiments is randomly selected from the available ground-truth images. The reported overall accuracies (OAs), average accuracies (AAs), kappa statistics (κ), and class individual accuracies are obtained after conducting ten independent Monte Carlo runs with respect to the initial training set and averaging obtained results.
- 7) Finally, we would like to emphasize that all of the experiments were conducted using MATLAB R2013a in a desktop PC equipped with an Intel Core i7 CPU (at 3.6 GHz) and 16 GB of RAM.

A. Experiments With Hyperspectral Data

In this section, two hyperspectral data sets are used for evaluation. The first hyperspectral data set 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 für Luftund Raumfahrt (DLR, the German Aerospace Agency) in the framework of the HySens project, managed and sponsored by the European Union. The image size in pixels is 610×340 , with a very high spatial resolution of 1.3 m per pixel. The number of data channels in the acquired image is 103 (with a spectral range from 0.43 to 0.86 μm). Fig. 3(a) shows a false color composite of the image, while Fig. 3(b) shows the ground-truth map, which contains 42 776 samples and 9 ground-truth classes of interest, comprised of urban features, as well as soil and vegetation features.

The second hyperspectral image used in the 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 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-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. 4(a) shows a false color composition of the AVIRIS Indian Pines scene, while Fig. 4(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, in total, 10 249 samples. These data, including ground-truth information, are available online,¹ a fact which has made this scene a widely used benchmark for testing the accuracy of hyperspectral data classification algorithms. This scene constitutes a very challenging classification problem due to the significant presence of mixed pixels in all available classes and also because of the unbalanced number available labeled pixels per class.

1) Experiments With the ROSIS Pavia University Data Set:

In our first experiment with the ROSIS Pavia University data set, we estimate the quality of the MCs obtained from the proposed MMCA scheme. Let σ_n be the noise variance for a given image, which can be estimated by the fast noise variance estimation algorithm described in [51]. Fig. 5 shows the noise variance of the MCs obtained from the first MNF component. It can be observed that the noise variance is greatly improved for all of the smoothness components, which are the ones used for classification.

In the second experiment, we graphically illustrate the data separability of the MCs obtained from the proposed MMCA scheme for the ROSIS Pavia data set. Fig. 6 shows the scatterplot for classes “asphalt,” “bitumen” and “bare soil” projected on the first two MNF components. It can be seen that the three considered classes are better separated in the MC components than that in the original data set. In order to quantitatively illustrate the improvement of the class separability, we evaluate the Bhattacharyya distance [52] between different classes for the highly mixed regions in the image, as shown in Table I. Take class 1 (asphalt) and class 2 (bare soil) in the first region as an example. It can be observed that, since these two classes are dominated by very different anisotropic structures, the distance between each other is highly improved in the content space. As another example, in the second region, class 2 (metal sheets) and class 3 (shadows) are very close in the original space, while in the contrast space, the distance is greatly improved. This is expected, as these two classes present very different intensity changes. A similar observation can be obtained for most cases, which can be considered as a good indication that the class separability is consequently improved.

As another experiment, we also perform a comparison of the proposed MMCA with other spatial feature extraction methods, where extended morphological attribute profiles (EMAP) [53] and EMD [54] are included. We chose EMAP as it is a powerful tool for spatial feature extraction, where the parameters of

¹ Available online: <http://dynamo.ecn.purdue.edu/biehl/MultiSpec>.

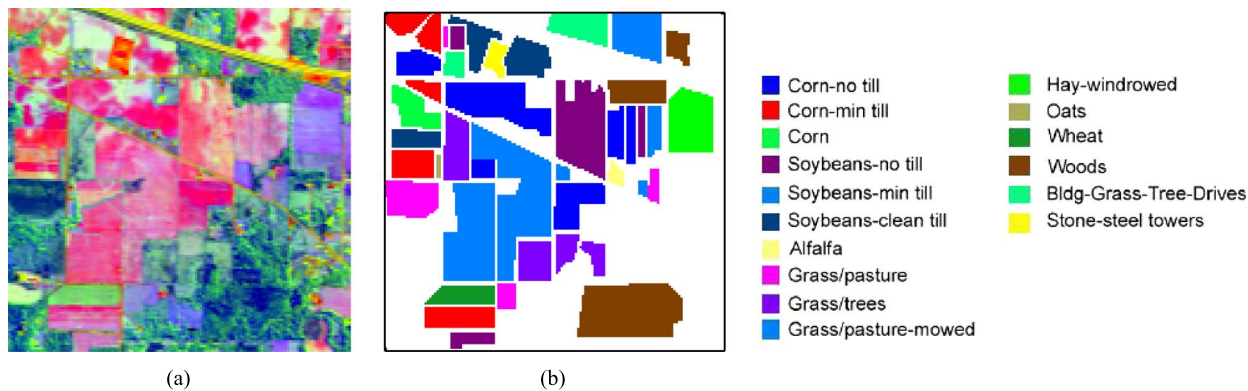


Fig. 4. (a) False color composition of the AVIRIS Indian Pines scene. (b) Ground-truth map containing 16 mutually exclusive land-cover classes (right).

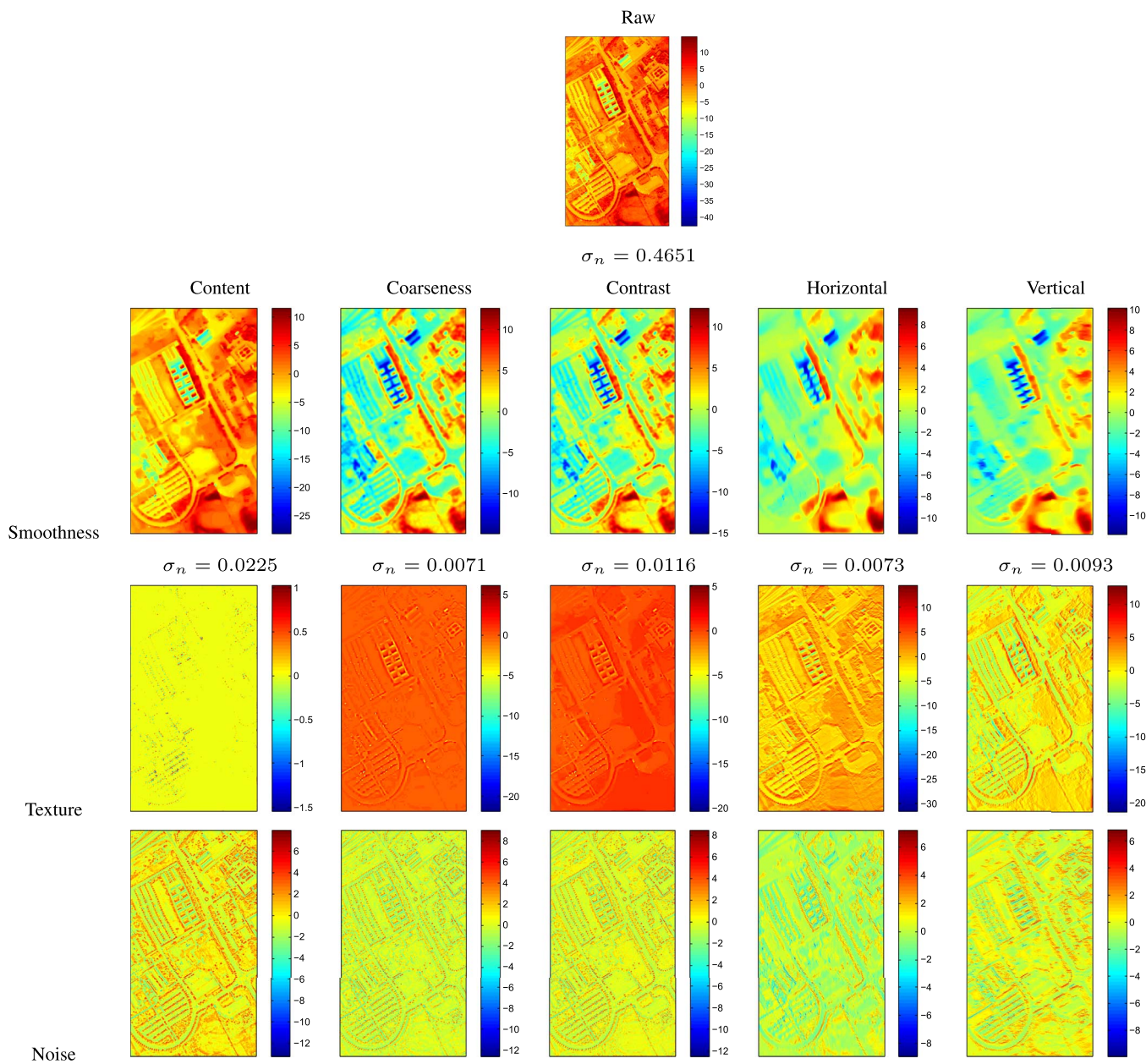


Fig. 5. Image separation results along with the noise variance σ_n for the first MNF component of the ROSIS Pavia University data set.

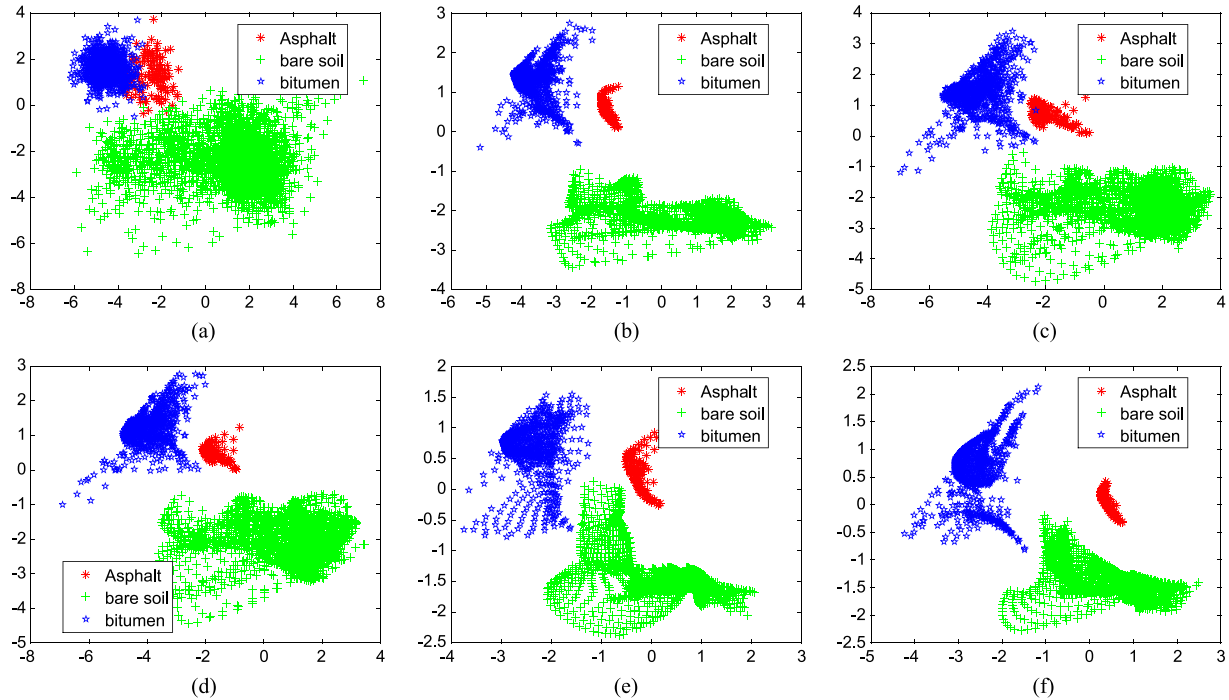


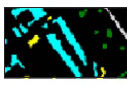
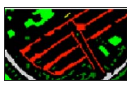


Fig. 6. Illustration of data separability by projecting the data into the first two MNF components of the ROSIS Pavia University data set. (a) Raw. (b) Content. (c) Coarseness. (d) Contrast. (e) Horizontal. (f) Vertical.

TABLE I
BHATTACHARYYA DISTANCE BETWEEN DIFFERENT CLASSES FOR THE ROSIS PAVIA UNIVERSITY DATA SET

Region	Class name	Pair of classes	Bhattacharyya distance between classes					
			Raw	Content	Coarseness	Contrast	Horizontal	Vertical
	c1: Asphalt	c1-c2	0.4289	1.1301	0.9070	1.0831	0.7837	0.5179
	c2: bare soil	c1-c3	0.3415	1.3159	1.1135	1.5505	1.2183	3.2964
	c3: bitumen	c2-c3	0.2342	1.7884	0.3677	1.4814	1.2446	2.4028
	c1: meadows	c1-c2	0.1436	1.2317	0.2057	1.6674	1.1778	1.9719
	c2: metal sheets	c1-c3	0.1730	1.2275	0.2182	1.6681	1.0875	1.8815
	c3: shadows	c2-c3	0.0726	0.0776	0.0540	0.1139	0.0430	0.1099
	c1: Asphalt	c1-c2	0.3556	0.8190	0.6451	1.9701	0.6579	1.2593
	c2: gravel	c1-c3	0.5113	1.5184	0.8429	1.5472	1.4361	2.0656
	c3: shadows	c2-c3	0.4104	0.9735	0.4129	0.6906	0.7784	0.5772
	c1: meadows	c1-c2	0.1212	0.3487	0.1357	0.3592	0.2047	0.2895
	c2: trees	c1-c3	0.0854	0.4776	0.0964	0.6397	0.2207	0.4476
	c3: bricks	c2-c3	0.0764	0.4077	0.0649	0.1879	0.1210	0.2349

the EMAP are defined according to [53] and [55]. For the EMD approach, following [14], we choose the bidimensional EMD [13]. On the one hand, we analyze the impact of the dimensionality of the MNF components, where Fig. 7(a) shows the obtained classification results as a function of the number of MNF components after using 1% of the labeled samples per class available in the ground-truth image for training and

the remaining samples for testing. Several conclusions can be obtained from Fig. 7(a). First of all, as expected, the classification accuracy increases as the number of MNF components increases. This is because more information is considered. Furthermore, when the number of MNF components is limited, the results obtained by MMCA and EMAP are comparable and superior to those obtained by EMD. Finally, as shown in Fig. 7(a),

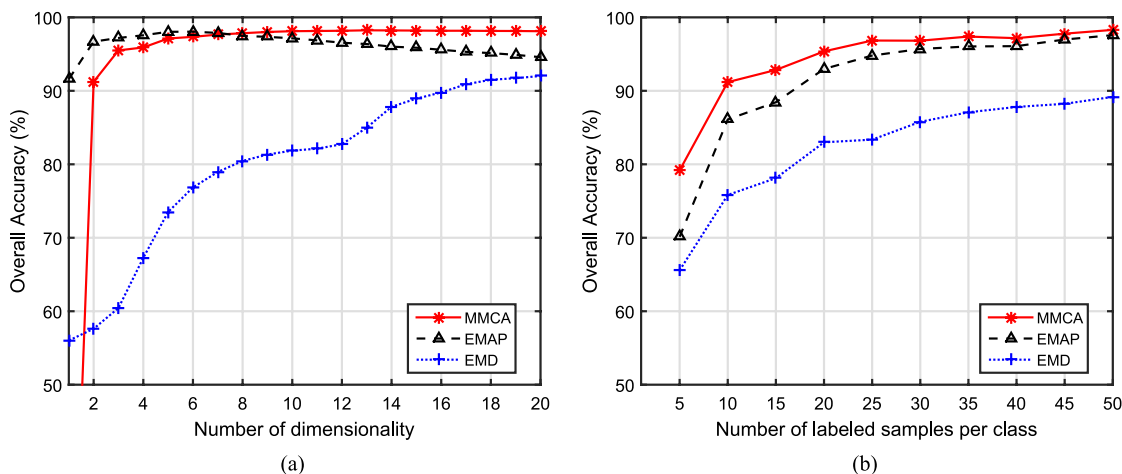


Fig. 7. OAs obtained by MMCA, EMAP, and EMD for the ROSIS Pavia University data set: (a) as a function of the number of MNF components and (b) as a function of the number of training samples per class.

TABLE II
 OAs [%], AAs [%], INDIVIDUAL CLASSIFICATION ACCURACY LEVELS, AND κ [%] ALONG WITH THE STANDARD DEVIATION OF TEN MONTE CARLO (MC) RUNS FOR THE PROPOSED MMCA METHOD FOR THE ROSIS PAVIA UNIVERSITY DATA SET, USING 1% OF ALL OF THE AVAILABLE LABELED SAMPLES FOR TRAINING AND THE REST OF THE LABELED SAMPLES FOR TESTING

Class	#Samples		MCs from different textural feature						Combinations		
	Train(426)	Test(42350)	Raw	Content	Coarseness	Contrast	Horizontal	Vertical	\sum_s	V_s	CK_s
Asphalt	66	6565	87.81 ± 3.16	96.11 ± 2.45	92.57 ± 2.37	97.14 ± 0.79	92.54 ± 2.59	92.76 ± 2.50	97.00 ± 1.79	97.80 ± 1.23	97.97 ± 0.89
meadows	186	18463	95.70 ± 1.64	99.84 ± 0.31	99.65 ± 0.21	99.50 ± 0.23	99.71 ± 0.40	99.64 ± 0.50	99.89 ± 0.22	99.87 ± 0.29	99.87 ± 0.32
gravel	21	2078	66.41 ± 5.06	82.86 ± 7.17	72.62 ± 7.94	82.57 ± 8.07	85.29 ± 8.76	86.89 ± 9.04	82.47 ± 6.95	89.55 ± 7.88	90.25 ± 7.76
trees	31	3033	79.87 ± 5.30	90.06 ± 2.42	88.44 ± 2.40	86.81 ± 2.17	77.34 ± 4.65	75.12 ± 5.14	89.89 ± 2.10	93.79 ± 2.07	93.56 ± 2.14
metal sheets	13	1332	98.15 ± 0.94	99.26 ± 0.58	98.95 ± 1.89	99.27 ± 1.01	99.17 ± 0.46	98.08 ± 1.63	99.32 ± 1.54	99.49 ± 1.01	99.49 ± 0.66
bare soil	50	4979	84.70 ± 3.92	99.57 ± 0.42	98.93 ± 0.47	97.60 ± 1.10	100 ± 0.00	99.76 ± 0.51	99.57 ± 0.47	99.75 ± 0.32	99.72 ± 0.36
bitumen	13	1317	74.98 ± 7.68	96.97 ± 2.68	89.70 ± 4.75	92.82 ± 6.93	96.42 ± 3.46	97.86 ± 2.50	97.21 ± 2.10	97.82 ± 1.93	97.76 ± 1.80
bricks	37	3645	77.33 ± 3.67	90.64 ± 2.66	83.09 ± 5.80	89.87 ± 3.18	91.43 ± 2.79	93.79 ± 2.63	89.72 ± 4.78	96.72 ± 1.79	97.00 ± 1.83
shadows	9	938	96.15 ± 1.59	88.24 ± 5.26	83.50 ± 6.67	88.55 ± 6.39	72.64 ± 9.72	75.48 ± 6.06	85.45 ± 6.58	91.34 ± 3.91	90.98 ± 3.38
Average accuracy	-	-	84.57 ± 1.31	93.73 ± 0.64	89.71 ± 0.99	92.68 ± 1.19	90.50 ± 1.25	91.04 ± 1.13	93.39 ± 0.53	96.24 ± 0.92	96.29 ± 0.89
Overall accuracy	-	-	88.47 ± 1.30	96.54 ± 0.39	94.22 ± 0.50	95.89 ± 0.53	94.89 ± 0.42	95.06 ± 0.66	96.54 ± 0.32	98.06 ± 0.46	98.12 ± 0.48
κ statistic	-	-	84.67 ± 1.70	95.41 ± 0.52	92.32 ± 0.67	94.53 ± 0.70	93.21 ± 0.56	93.45 ± 0.88	95.40 ± 0.43	97.42 ± 0.61	97.50 ± 0.63

the proposed MMCA obtains results that are almost the same when the number of MNF components is greater than 8. Based on this observation, as discussed in the experimental setting, we use ten MNF components for classification purposes in the remaining experiments.

On the other hand, Fig. 7(b) reports the obtained classification accuracies as a function of the number of training samples per class with ten MNF components. It can be observed that, when the number of training samples is small, the performance of the proposed MMCA is better than that achieved by the other tested methods.

In our final set of experiments in this section, we evaluate the classification performance of the obtained MCs. Two different experiments are performed. On the one hand, we randomly choose around 1% of the labeled samples (a total of 426 samples) from the nine classes in the ground-truth for training and use the remaining 42 350 labeled samples for testing. Table II reports the obtained OAs, AAs, individual classification accuracy levels, and κ statistics, along with the standard deviation of

the ten conducted Monte Carlo runs. It can be observed that the results obtained from the MCs, which are comparable to each other, are much better than that obtained from the original MNF component. Furthermore, the results obtained from the combinations of the MCs are better than those obtained from one single type of MCs. This is again expected as more textural information is included when more textural features are considered. On the other hand, we evaluate the performance of the proposed approach under a balanced composition of the training–test samples, which provides complementary information to the one reported in the previous experiment. Around 1% of the labeled samples (about 48 samples per class) are now randomly chosen for training, and the remaining labeled samples per class are used for testing. Table III reports the OAs, AAs, individual classification accuracy levels, and κ statistics, where the standard deviations are also included. Similar observations can be obtained with regard to those reported for Table II. This experiment shows that the proposed approach can lead to very good classification accuracies for problems with limited

TABLE III

OAs [%], AAs [%], INDIVIDUAL CLASSIFICATION ACCURACY LEVELS, AND κ [%] ALONG WITH THE STANDARD DEVIATION OF TEN MONTE CARLO (MC) RUNS FOR THE PROPOSED MMCA FOR THE ROSIS PAVIA UNIVERSITY DATA SET, USING A BALANCED DISTRIBUTION WITH 1% OF THE AVAILABLE LABELED SAMPLES FOR EACH CLASS FOR TRAINING AND THE REST OF THE LABELED SAMPLES FOR TESTING

Class	#Samples		MCs from different textural feature						Combinations		
	Train(432)	Test(42344)	Raw	Content	Coarseness	Contrast	Horizontal	Vertical	\sum_s	V_s	CK_s
Asphalt	48	6583	80.29±1.41	92.42±1.35	86.91±3.06	93.57±2.41	88.19±2.86	86.10±3.45	93.04±2.08	95.18±1.32	95.21±1.59
meadows	48	18601	86.78±4.18	98.21±1.59	95.10±2.84	95.58±2.92	94.34±2.47	95.82±3.19	97.97±2.07	98.63±1.30	98.13±1.68
gravel	48	2051	76.92±5.36	93.29±2.51	83.39±2.81	91.90±1.51	93.97±2.18	95.53±2.49	90.63±1.84	96.24±1.78	96.66±1.79
trees	48	3016	89.03±3.21	95.08±0.95	92.51±2.64	93.24±2.83	83.42±2.31	86.25±3.36	94.38±1.80	96.72±1.06	96.98±1.10
metal sheets	48	1297	99.19±0.37	99.65±0.22	99.77±0.27	99.77±0.23	99.71±0.22	99.31±0.47	99.92±0.11	99.87±0.15	99.89±0.13
bare soil	48	4981	91.32±2.02	99.67±0.30	99.46±0.51	99.10±0.63	99.95±0.06	99.64±0.41	99.79±0.31	99.68±0.53	99.65±0.38
bitumen	48	1282	91.42±2.03	99.55±0.40	97.11±1.82	97.88±1.38	99.45±0.24	99.45±0.28	99.33±0.74	99.96±0.07	99.95±0.09
bricks	48	3634	76.39±4.68	91.19±2.39	82.59±3.03	90.82±2.56	92.19±3.05	95.34±2.79	89.72±3.23	98.20±0.67	98.66±0.66
shadows	48	899	99.02±0.49	97.71±1.38	95.93±1.04	98.14±0.93	92.59±1.54	92.96±2.11	96.82±1.62	99.07±0.46	99.11±0.48
Average accuracy	-	-	87.82±0.75	96.31±0.29	92.53±0.72	95.55±0.49	93.76±0.57	94.49±0.47	95.73±0.52	98.17±0.28	98.25±0.34
Overall accuracy	-	-	85.87±1.80	96.49±0.57	92.73±1.59	95.18±1.46	93.35±1.14	94.18±1.42	96.18±1.04	98.01±0.58	97.88±0.81
κ statistic	-	-	81.68±2.15	95.35±0.74	90.45±2.04	93.65±1.88	91.26±1.45	92.34±1.81	94.94±1.35	97.37±0.76	97.19±1.06

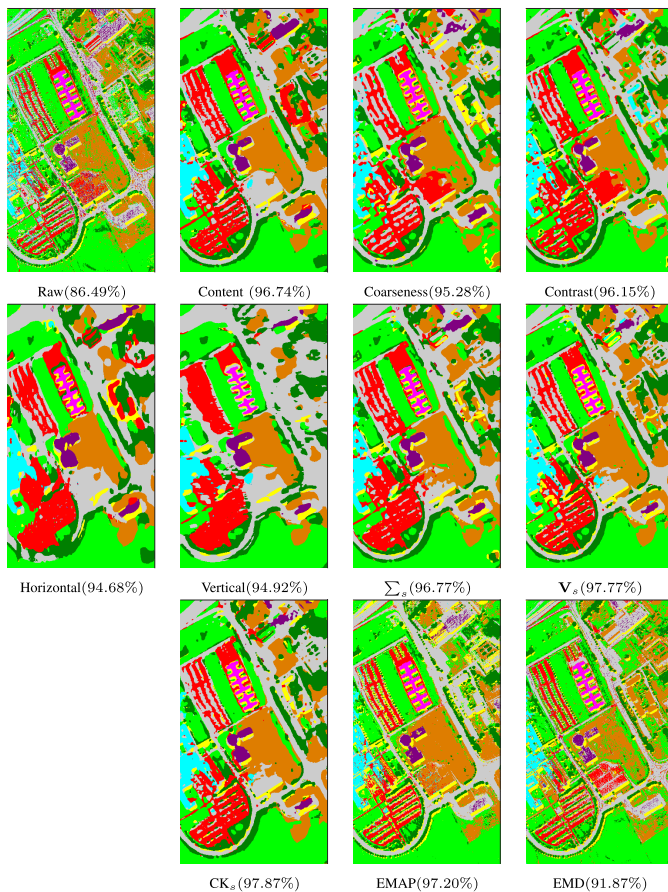


Fig. 8. Classification maps (along with the OAs) obtained by different textural feature components for the ROSIS Pavia University data set. These maps correspond to one single experiment in Table II.

training samples, regardless of whether we use a balanced or unbalanced distribution of training samples across the classes. Finally, for illustrative purposes, Fig. 8 presents some of the classification maps obtained by each different textural feature components.

2) *Experiments With the AVIRIS Indian Pines Data Set:* In our first experiment with the AVIRIS Indian Pines data, we estimate the quality of the MCs obtained from the proposed MMCA scheme. Fig. 9 shows the noise variance of the MCs obtained from the first MNF component. It can be observed that the noise variance is greatly improved for all of the smoothness components, which are the ones used in the classification.

In the second experiment, we graphically illustrate the data separability of the MCs obtained from the proposed MMCA scheme. Fig. 10 shows the scatterplot for classes “corn-no till,” “soybeans-no till,” and “soybeans-min till” projected on the first two MNF components. It can be seen that three considered classes are highly mixed with each other in the original data set, while in the obtained MCs, similar to that observed in the ROSIS data, pixels which belong to the same class tend to be more concentrated, and pixels which do not belong to the same class tend to be more separated. In order to quantitatively illustrate the improvement of the class separability, we evaluate the Bhattacharyya distance [52] between different classes for the highly mixed regions in the image. Table IV shows the Bhattacharyya distance between different classes for five different regions. It is clear that the distance between classes is greatly improved in all cases, which can be considered as an indication that the class separability is consequently improved.

In a third experiment, around 2% of the labeled samples (a total of 205 samples) are randomly chosen from the 16 classes for training the classifier, and the remaining 10 044 labeled samples are used for testing. Table V reports the obtained OAs, AAs, individual classification accuracies, and κ statistics, along with the standard deviation of the ten conducted Monte Carlo runs. It can be observed that the results obtained from the MCs, which are comparable to each other, are much better than those obtained from the original MNF component. Furthermore, the results obtained from the combinations of the MCs are better than those obtained from one single type of MCs. This is expected as more textural information is included when additional textural features are considered. For illustrative purposes,

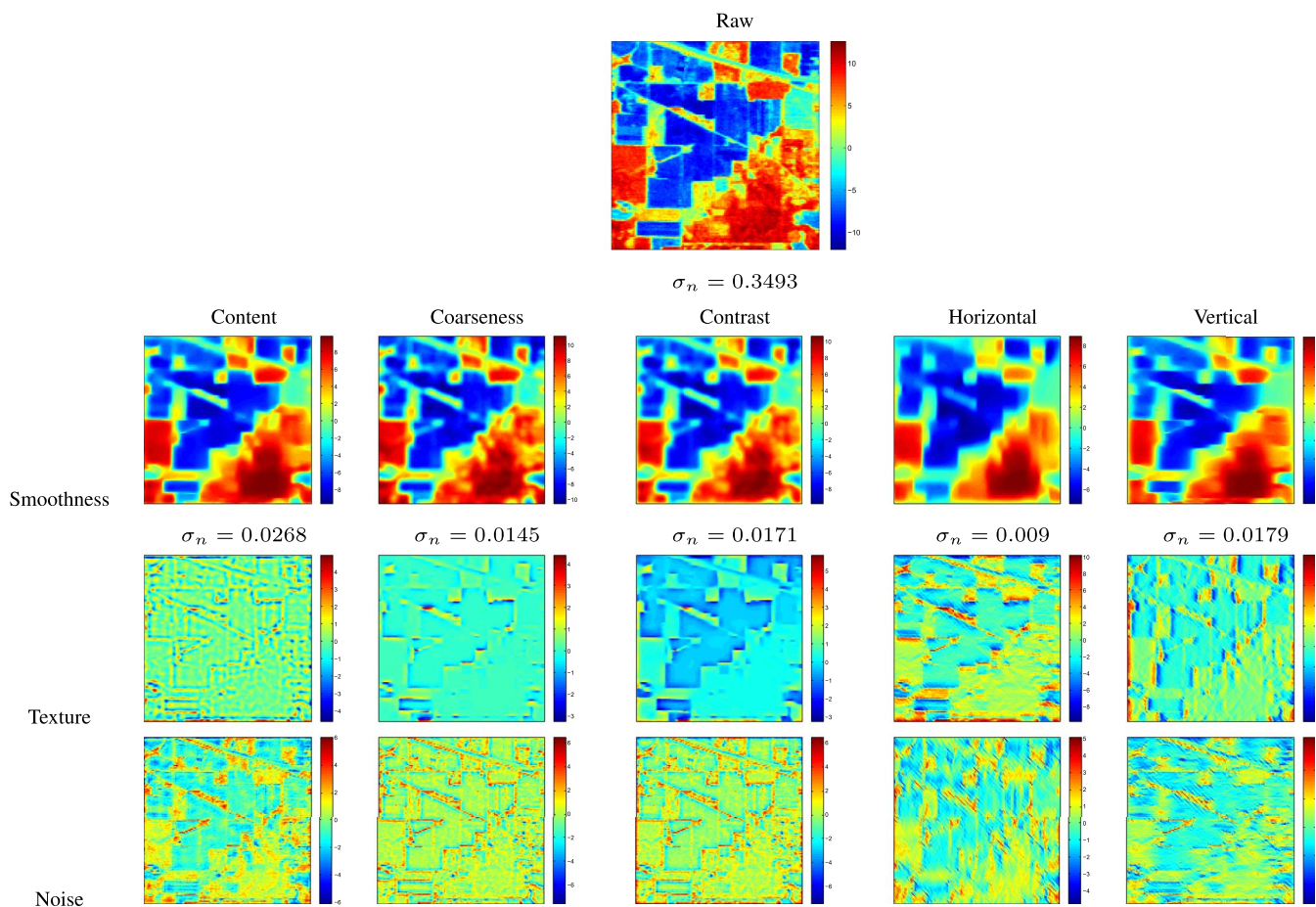


Fig. 9. Image separation results along with the noise variance σ_n for the first MNF component of the AVIRIS Indian Pines data set.

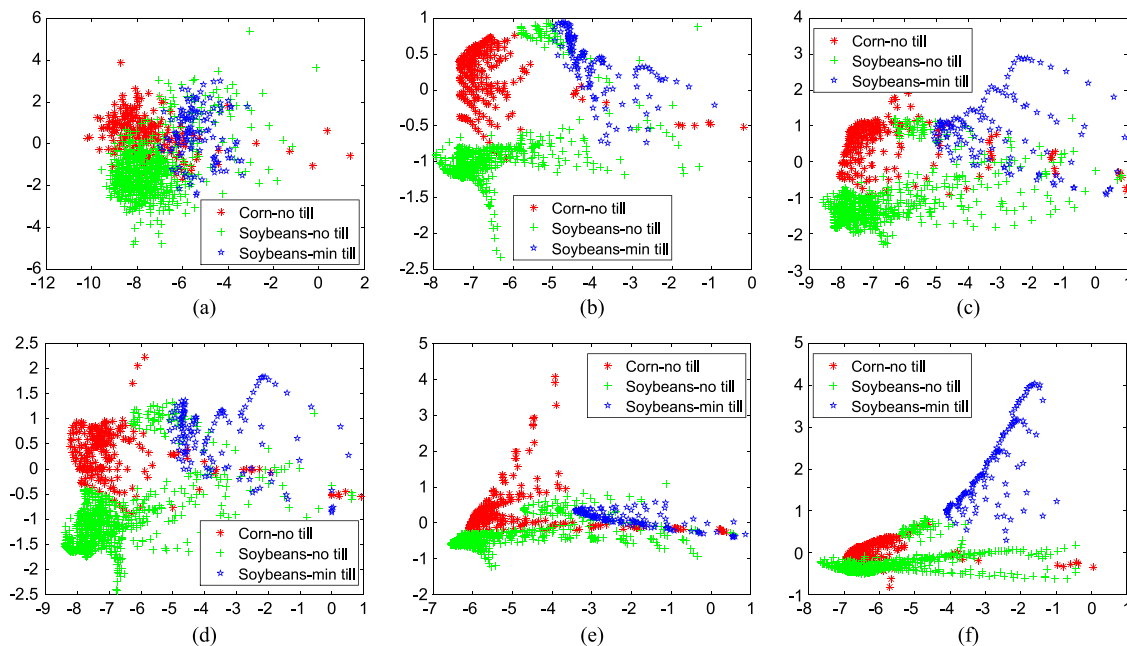


Fig. 10. Illustration of data separability by projecting the data into the first two MNF components of the AVIRIS Indian Pines data set. (a) Raw. (b) Content. (c) Coarseness. (d) Contrast. (e) Horizontal. (f) Vertical.

Fig. 11 presents the classification maps obtained by different textural feature components, which results in visually improved results. As it was already observed in our second experiment,

classes “corn-no till,” “soybeans-no till,” and “soybeans-min till” are better separated by using the MCs than the original MNF component.

TABLE IV
BHATTACHARYYA DISTANCE BETWEEN DIFFERENT CLASSES FOR THE AVIRIS INDIAN PINES DATA SET






Region	Class name	Pair of classes	Bhattacharyya distance between classes					
			Raw	Content	Coarseness	Contrast	Horizontal	Vertical
	c1: Corn-no till	c1-c2	0.3802	0.7087	0.6753	0.6955	1.1221	0.8414
	c2: Soybeans-no till	c1-c3	0.5306	0.9190	0.8226	0.6812	1.2034	0.4484
	c3: Soybeans-min till	c2-c3	0.8042	1.2661	1.8042	1.3067	2.1375	1.3655
	c1: Corn-no till	c1-c2	0.4793	0.6758	0.7460	0.7749	0.8283	0.9008
	c2: Corn-min till	c1-c3	0.5592	0.6829	0.7710	0.8922	0.8648	1.0075
	c3: Corn	c2-c3	0.1365	0.1917	0.1350	0.1831	0.2803	0.1410
	c1: Corn-min till	c1-c2	0.1617	0.2280	0.1619	0.3377	0.3687	0.4098
	c2: Grass/pasture	c1-c3	0.2134	0.4159	0.2052	0.2693	0.3966	0.1730
	c3: Soybeans-clean till	c2-c3	0.2015	0.4575	0.2334	0.3812	0.4142	0.4495
	c1: Corn-no till	c1-c2	0.3654	0.6579	1.0199	0.6756	1.2682	0.4248
	c2: Soybeans-no till	c1-c3	0.0716	0.1085	0.0575	0.1124	0.1500	0.0605
	c3: Soybeans-min till	c2-c3	0.4301	0.7961	1.2131	0.8914	1.4540	0.4155
	c1: Corn-no till	c1-c2	0.1716	0.2978	0.2561	0.2529	0.3716	0.1898
	c2: Grass/trees	c1-c3	0.4033	0.7765	0.9889	0.7734	1.3119	0.5080
	c3: Soybeans-min till	c2-c3	0.1974	0.3526	0.5553	0.3677	0.6859	0.2459

TABLE V
OAS [%], AAS [%], INDIVIDUAL CLASSIFICATION ACCURACY LEVELS, AND κ [%] ALONG WITH THE STANDARD DEVIATION OF TEN MONTE CARLO (MC) RUNS FOR THE PROPOSED MMCA METHOD FOR THE AVIRIS INDIAN PINES DATA SET, USING 2% OF ALL OF THE AVAILABLE LABELED SAMPLES FOR TRAINING AND THE REST OF THE LABELED SAMPLES FOR TESTING

Class	#Samples		MCs from different textural feature						Combinations		
	Train(205)	Test(10044)	Raw	Content	Coarseness	Contrast	Horizontal	Vertical	\sum_s	V_s	CK_s
Alfalfa	3	43	84.65±5.81	86.98±8.53	85.12±9.88	86.05±7.78	91.39±6.16	84.19±13.02	86.05±9.42	90±6.07	89.53±5.72
Corn-no till	28	1400	78.96±4.37	92.35±2.77	90.05±1.99	91.17±2.4	88.08±3.57	88.81±4.09	91±2.52	92.31±2.68	92.16±3.01
Corn-min till	17	813	63.05±8.89	87.99±7.72	88.75±6.42	85.35±6.03	86.45±6.75	83.67±7.02	88.5±5.85	90.12±6.29	90.48±6.17
Corn	5	232	49.09±13.88	82.33±11.11	84.7±6.79	75.73±12.71	81.77±6.24	93.41±4.81	88.41±7.2	91.25±6.91	91.29±5.82
Grass/pasture	10	473	89.83±3.44	95.5±2.02	92.41±3.27	92.68±3.24	93.76±2.8	89.87±4.35	94.36±2.57	95.07±2.09	95.26±2.15
Grass/trees	15	715	96.63±1.25	99.13±0.69	96.87±3.28	97.99±1.72	97.29±1.31	97.29±2.24	98.45±1.33	99.37±0.51	99.54±0.34
Grass/pasture-mowed	3	25	99.6±1.2	100±0.00	99.6±1.2	99.6±1.2	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00
Hay-windrowed	10	468	98.93±1.97	99.96±0.08	99.91±0.14	99.87±0.2	99.62±0.38	99.72±0.51	99.96±0.09	100±0.00	99.96±0.13
Oats	3	17	98.24±3.77	98.82±2.35	98.24±3.77	97.65±3.9	99.41±1.76	100±0.00	98.82±2.35	100±0.00	100±0.00
Soybeans-no till	19	953	76.89±6.41	92.35±4.55	85.11±6.76	86.64±6.17	91.87±3.45	87.27±7.39	89.07±6.45	90.38±5.09	89.84±5.22
Soybeans-min till	40	2415	77.78±4.49	91.51±3.21	91.08±3.07	90.33±2.77	91.7±2.53	89.74±3.66	92.07±2.87	91.96±2.91	92.13±2.79
Soybeans-clean till	12	581	82.01±7.89	88.67±9.4	83.96±8.75	86.59±7.56	84.03±7.15	83.79±8.64	87.73±8.73	88.06±8.41	88.14±8.41
Wheat	4	201	99.3±0.33	99.75±0.33	98.26±1.48	99.1±1.13	99.05±1.31	97.61±2.91	99±1.48	99.3±0.97	99.3±0.97
Woods	25	1240	94.89±2.49	99.12±0.5	98.82±1.52	98.58±1.51	98.13±1.36	98.01±2.94	98.73±2.3	99.5±0.52	99.48±0.7
Bldg-grass-tree-drives	8	378	62.96±9.31	86.61±9.19	82.09±11.92	82.28±11.49	87.67±9.28	90.37±9.5	86.75±11.59	86.64±10.74	87.7±10.55
Stone-steel towers	3	90	70.11±19.03	93.11±6.94	98.22±2.64	97.22±3.73	95.22±3.52	86.11±10.02	98.33±2.82	98.11±3.07	99.11±1.39
Average accuracy	-	-	82.68±1.25	93.38±1.51	92.07±1.15	91.68±1.34	92.84±1.15	91.87±1.44	93.58±1.24	94.51±1.01	94.62±0.93
Overall accuracy	-	-	81.18±1.27	93.12±1.53	91.34±0.87	91.21±0.89	91.83±1.08	90.82±1.22	92.76±1.17	93.47±1.1	93.54±1.03
κ statistic	-	-	78.52±1.44	92.15±1.75	90.12±0.99	89.97±1.03	90.69±1.23	89.54±1.38	91.75±1.33	92.56±1.25	92.64±1.17

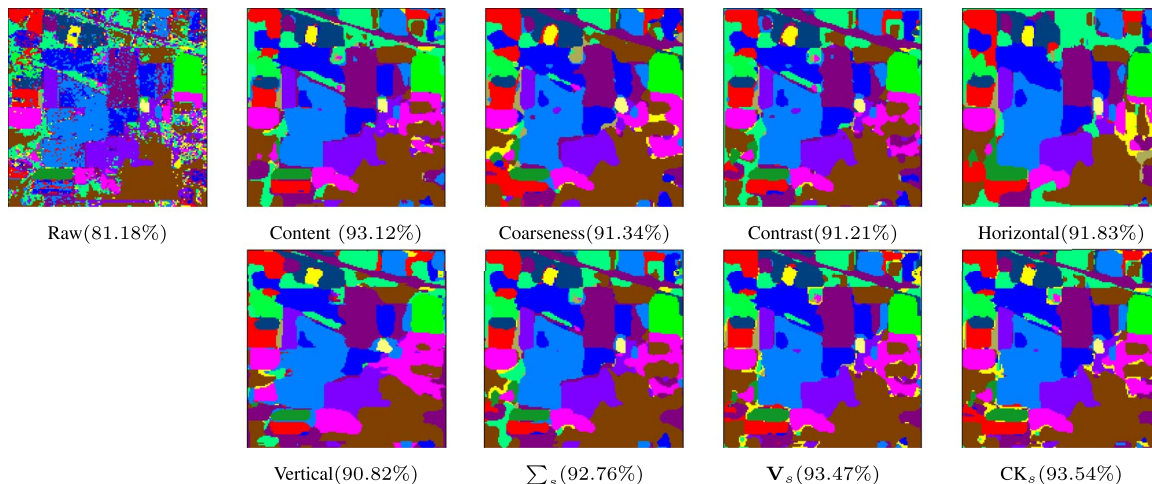


Fig. 11. Classification maps (along with the OAs) obtained by different textural feature components for the AVIRIS Indian Pines data set. These maps correspond to one single experiment in Table V.

TABLE VI

OAS [%], AAS [%], INDIVIDUAL CLASSIFICATION ACCURACY LEVELS, AND κ [%] ALONG WITH THE STANDARD DEVIATION OF TEN MONTE CARLO (MC) RUNS FOR THE PROPOSED MMCA METHOD FOR THE AVIRIS INDIAN PINES DATA SET, USING A BALANCED DISTRIBUTION WITH 2% OF THE AVAILABLE LABELED SAMPLES FOR EACH CLASS FOR TRAINING AND THE REST OF THE LABELED SAMPLES FOR TESTING

Class	#Samples		MCs from different textural feature						Combinations		
	Train(192)	Test(10057)	Raw	Content	Coarseness	Contrast	Horizontal	Vertical	\sum_s	V_s	CK_s
Alfalfa	12	34	94.12±3.22	96.76±2.06	98.24±1.95	98.24±1.95	97.94±1.35	97.94±1.35	97.94±1.35	96.76±2.06	95.88±1.44
Corn-no till	12	1416	69.89±8.12	84.01±4.96	79.15±5.3	81.19±4.75	80.61±6.19	81.57±4.67	82.18±5.42	82.85±4.67	82.98±4.37
Corn-min till	12	818	68.23±4.69	87.6±5.16	81.94±5.8	81.3±4.2	81.38±8.74	80.98±4.25	85.21±4.72	84.07±5.04	84.22±7.09
Corn	12	225	82.18±6.52	95.6±4.71	94.04±6.83	93.82±5.75	97.78±2.43	96.4±5.13	95.96±5.33	97.02±4.76	97.33±5.08
Grass/pasture	12	471	88.34±6.17	93.65±5.47	90.93±5.1	92.4±4.76	92.95±5.12	90.93±6.38	92.12±5.05	92.89±5.09	92.59±5.00
Grass/trees	12	718	96.66±1.83	98.3±1.71	97.23±1.83	97.84±2.09	98.25±0.8	96.49±2.1	98.37±1.13	99.08±0.59	98.97±0.69
Grass/pasture-mowed	12	16	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00
Hay-windrowed	12	466	96.84±1.88	100±0.00	100±0.00	99.79±0.64	99.83±0.19	99.7±0.9	100±0.00	100±0.00	100±0.00
Oats	12	8	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00	100±0.00
Soybeans-no till	12	960	77.07±7.45	87.82±4.84	83.21±3.91	84.7±3.28	83.82±8.05	81.58±8.28	85.13±4.19	86.49±5.12	86.28±5.78
Soybeans-min till	12	2443	57.99±6.62	79.7±6.64	72.91±6.31	73.06±7.27	78.19±6.89	77.45±6.3	78.44±6.86	78.58±6.25	78.41±5.77
Soybeans-clean till	12	581	80.33±6.21	90.22±8.04	86.87±7.47	86.4±7.5	83.67±8.28	85.18±7.96	89.29±7.65	89.14±8.03	88.86±8.16
Wheat	12	193	99.74±0.26	99.95±0.16	99.74±0.42	99.9±0.21	99.59±0.69	99.95±0.16	100±0.00	100±0.00	100±0.00
Woods	12	1253	87.92±6.47	93.07±7.96	92.39±7.76	91.39±7.66	91.63±7.5	92.22±8.02	92.67±7.86	92.96±8.06	93.11±8.15
Bldg-grass-tree-drives	12	374	79.73±8.81	97.35±2.76	97.27±2.62	95.48±4.22	96.52±4.4	98.77±1.67	98.66±1.92	97.81±2.82	98.18±2.43
Stone-steel towers	12	81	89.14±3.16	99.88±0.37	99.88±0.37	99.63±0.79	99.14±1.47	98.89±1.6	99.88±0.37	100±0.00	99.88±0.37
Average accuracy	-	-	85.51±0.88	93.99±0.96	92.11±0.76	92.2±0.77	92.58±0.9	92.38±0.85	93.49±0.88	93.6±0.85	93.54±0.86
Overall accuracy	-	-	75.95±1.75	88.59±2.45	84.84±2.21	85.13±2.24	86.26±2.55	85.96±2.47	87.46±2.31	87.73±2.26	87.7±2.21
κ statistic	-	-	72.94±1.93	87.08±2.73	82.89±2.41	83.23±2.44	84.46±2.84	84.13±2.74	85.82±2.55	86.13±2.5	86.1±2.45

In a final experiment, we evaluate the performance of the proposed approach under a balanced composition of the training–test sets. As opposed to the previous experiment, in which classes with more labeled samples comprised more training samples, now around 2% of the labeled samples (about 12 samples) per class are randomly chosen for training, and the remaining labeled samples are used for testing. Table VI reports the OAs, AAs, individual classification accuracy levels, and κ statistics, where the standard deviations are also included. Similar observations can be obtained as those from Table V. Along with the third experiment, this experiment shows that the proposed approach can lead to very good classification accura-

cies for problems with limited training samples, regardless of whether the distribution of training samples across the classes is either balanced or unbalanced.

B. Experiments With Polarimetric SAR Data

In order to further validate the proposed MMCA method, two polarimetric SAR data sets are employed. PolSAR is a new form of SAR radar system, and it emits and receives multifrequency and fully polarized radar waves. For the past years, PolSAR has been widely used in land-cover classification [56] and change detection in remote sensing applications [57].

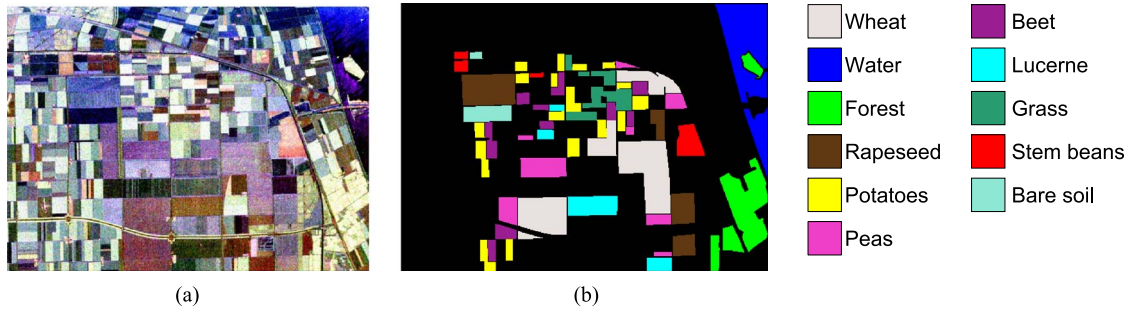


Fig. 12. (a) False color composition of the AirSAR Flevoland image. (b) Ground-truth map containing 11 mutually exclusive land-cover classes (right).

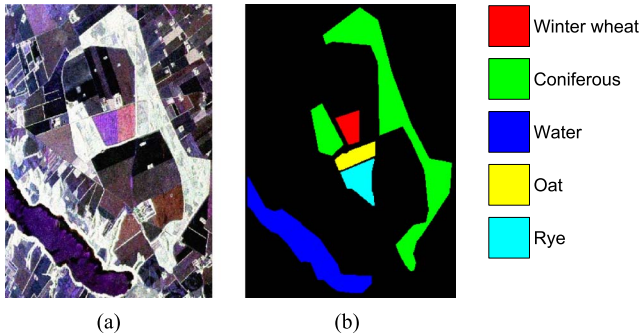


Fig. 13. (a) False color composition of the EMISAR Foulum image. (b) Ground-truth map containing five mutually exclusive land-cover classes (right).

The first data set used in our experiments is the AirSAR L-band PolSAR data set, obtained by NASA JPL over the Flevoland site in The Netherlands. These data and the ground-truth are, respectively, displayed in Fig. 12(a) and (b). The Flevoland image, with a size of 375×512 samples, contains different crop classes as well as bare soil, water, and forests. The second data set is a full polarimetric airborne SAR L-band PolSAR data set acquired by the EMISAR system over Foulum, Denmark. These data and the ground-truth are, respectively, displayed in Fig. 13(a) and (b). The Foulum image, with a size of 300×421 samples, covers a vegetated region which consists of water, coniferous, rye, oat, and winter wheat. These two data sets are very challenging due to the fact that a significant amount of speckle noise exists.

1) *Evaluation of Classification Accuracies:* In order to validate the classification performance, in a first experiment, we use 2% of the available samples per class for training in the AirSAR data and 1% of the available samples per class for training in the EMISAR data, respectively. The remaining samples are used for testing. Tables VII and VIII, respectively, report the obtained OAs, AAs, individual classification accuracy levels, and κ statistics obtained for the AirSAR and EMISAR data sets, along with the standard deviation of the ten conducted Monte Carlo runs. It can be observed from Tables VII and VIII that the results obtained from the MCs, which are comparable to each other, are much better than those obtained from the original image. This is particularly the case for the AirSAR data, in which the results obtained from the *contrast* component are better than those obtained from the other components. This is expected because the generated high-*contrast* component presents fast intensity changes which strengthen the differences

between the classes. Similar results were also found for the EMISAR data. For illustrative purposes, Figs. 14 and 15, respectively, present some of the classification maps obtained for the AirSAR and EMISAR data sets by the different considered textural feature components. Visual improvements in the obtained classification results can be clearly appreciated for the proposed MMCA approach.

2) *Comparison With State-of-the-Art Approaches:* In the second experiment, we compared the proposed MMCA approach with the widely used EMAP [53] and EMD [14] by using different numbers of training samples per class. For the EMAP approach, we considered four different attributes constructed on each MNF component: 1) area of the regions ($\lambda_a = [100, 500]$); 2) length of the diagonal of the box bounding the region ($\lambda_d = [10, 25]$); 3) moment of inertia [11] ($\lambda_i = [0.2, 0.3]$); and 4) standard deviation of the gray-level values of the pixels in the regions ($\lambda_s = [20, 30]$). For the EMD approach, we choose the first three IMFs, and we stacked them for classification. Fig. 16 presents the obtained OAs as a function of the number of training samples per class. It is noticeable that, when the number of training samples is small, the results obtained by the proposed MMCA approach are superior to those obtained by the other tested methods. With the increase in the number of training samples, the results obtained by the three considered approach are comparable.

3) *Parameter Analysis:* In a third experiment, we perform a detailed analysis on the parameters λ_1, λ_2 involved in the decomposition framework (4), along with the size of the image partitions, i.e., $U(a)$, and the number of partitions p . Fig. 17 shows the obtained OAs for the AirSAR and EMISAR data sets, respectively, in which the classification of AirSAR used 2% samples per class for training, the classification of EMISAR used 1% samples per class for training, and the remaining samples were used for validation. In this experiment, we choose the components obtained by the *contrast* feature, as it exhibits better classification performance in the former experiments. Fig. 17(a) and (c) shows the OAs (as a function of parameters λ_1 and λ_2) obtained by MMCA for the AirSAR and EMISAR data sets with fixed $U(a) = 8 \times 8$ and $p = 10$, while Fig. 17(b) and (d) shows the OAs (as a function of parameters $U(a)$ and p) obtained by MMCA for the AirSAR and EMISAR data sets with fixed $\lambda_1 = \lambda_2 = 1e - 5$. As it can be observed, the classification performance is almost insensitive to λ_1 . For λ_2 , when the value of this parameter is lower than $1e - 4$, the results are stable. Therefore, it is easy to determine a good suboptimal setting for λ_1 and λ_2 . Furthermore, we can infer that small

TABLE VII
OAS [%], AAS [%], INDIVIDUAL CLASSIFICATION ACCURACIES, AND κ [%] ALONG WITH THE STANDARD DEVIATION OF TEN MONTE CARLO (MC) RUNS FOR THE PROPOSED MMCA METHOD FOR THE FLEVLAND AIRSAR IMAGE DATA SET, USING A TOTAL OF 2% OF THE AVAILABLE LABELED SAMPLES FOR TRAINING AND THE REMAINING LABELED SAMPLES FOR TESTING

Class	#Samples		MCs from different textural feature						Combinations		
	Train(1084)	Test(53192)	Raw	Content	Coarseness	Contrast	Horizontal	Vertical	\sum_s	V_s	CK_s
Wheat	230	11252	73.72±2.09	76.49±1.80	76.30±1.84	83.61±0.98	78.31±1.87	79.53±1.66	81.49±1.48	85.63±1.63	85.67±1.66
Water	158	7767	89.45±1.83	88.08±1.32	95.40±0.85	93.12±1.03	90.42±1.23	89.50±1.28	91.70±1.26	95.33±1.31	95.26±1.29
Forest	128	6295	75.75±1.64	79.41±1.46	89.91±0.44	95.40±0.41	89.03±1.39	89.00±1.63	92.33±0.76	93.77±1.09	94.04±0.87
Rapeseed	122	6007	69.95±3.06	74.06±3.35	85.39±2.81	86.72±4.65	79.67±2.97	78.30±3.01	84.26±2.89	87.35±2.47	88.53±2.10
Potatoes	108	5271	74.98±1.93	76.66±1.86	78.79±2.74	89.28±1.46	80.58±1.92	81.21±2.72	87.31±2.02	90.43±1.99	89.96±2.30
Peas	86	4178	51.01±1.76	54.54±2.31	62.57±2.31	72.45±2.66	58.06±2.07	59.85±2.07	66.24±2.53	74.00±2.06	76.33±1.60
Beet	68	3341	57.89±2.63	63.29±2.84	54.98±3.73	78.18±3.80	70.48±3.19	65.02±4.08	74.07±3.89	77.64±3.09	76.77±3.31
Lucerne	60	2944	80.42±2.87	81.11±3.86	78.97±4.35	85.96±3.79	79.87±2.53	85.34±2.51	84.21±3.56	86.73±2.52	87.20±2.87
Grass	54	2671	50.45±3.18	54.49±2.88	60.13±3.94	73.28±3.02	56.47±3.72	60.60±2.27	68.38±2.85	83.18±3.13	82.84±2.52
Stem beans	38	1894	75.33±5.00	79.19±4.31	82.75±2.03	90.23±0.95	83.88±3.56	83.03±2.17	88.19±1.29	88.35±2.54	88.79±1.86
Bare soil	32	1572	41.23±7.26	52.19±4.62	83.11±7.62	88.86±4.63	73.03±5.06	68.00±5.83	81.04±4.52	91.25±3.05	91.25±2.65
Average accuracy	-	-	67.29±0.78	70.86±0.71	77.12±0.99	85.19±0.52	76.35±0.52	76.31±0.59	81.75±0.48	86.70±0.53	86.97±0.45
Overall accuracy	-	-	71.48±0.29	74.24±0.27	79.32±0.59	86.09±0.42	78.68±0.32	78.83±0.31	83.21±0.20	87.47±0.33	87.74±0.30
κ statistic	-	-	67.44±0.29	70.62±0.31	76.47±0.68	84.19±0.48	75.70±0.35	75.89±0.35	80.90±0.23	85.74±0.37	86.05±0.34

TABLE VIII
OAS [%], AAS [%], INDIVIDUAL CLASSIFICATION ACCURACIES, AND κ [%] ALONG WITH THE STANDARD DEVIATION OF TEN MONTE CARLO (MC) RUNS FOR THE PROPOSED MMCA METHOD FOR THE FOULUM EMISAR IMAGE DATA SET, USING A TOTAL OF 1% OF THE AVAILABLE LABELED SAMPLES FOR TRAINING AND THE REMAINING LABELED SAMPLES FOR TESTING

Class	#Samples		MCs from different textural feature						Combinations		
	Train(237)	Test(23336)	Raw	Content	Coarseness	Contrast	Horizontal	Vertical	\sum_s	V_s	CK_s
Winter wheat	11	1077	94.97±2.78	97.29±1.55	99.55±1.34	100±0.00	99.03±0.73	99.82±0.16	99.98±0.06	99.98±0.06	100±0.00
Coniferous	129	12722	99.81±0.24	99.94±0.05	100±0.00	100±0.00	99.99±0.01	99.99±0.01	100±0.00	100±0.00	100±0.00
Water	67	6628	91.82±1.85	93.29±1.51	98.88±0.44	98.34±0.51	97.32±0.48	97.18±0.91	98.44±0.89	99.18±0.72	99.34±0.69
Oat	12	1160	79.08±9.68	85.78±6.34	87.6±8.35	98.41±1.18	90.98±4.89	88.47±4.26	96.06±3.21	94.31±5.13	97±2.13
Rye	18	1749	94.31±3.91	96.95±3.41	92.82±2.6	99.95±0.1	97.07±4.2	99.05±0.84	99.19±1.21	100±0.00	100±0.00
Average accuracy	-	-	91.99±2.38	94.65±1.41	95.77±1.33	99.34±0.22	96.88±1.46	96.9±0.79	98.74±0.66	98.69±0.99	99.27±0.45
Overall accuracy	-	-	95.87±0.34	96.99±0.23	98.51±0.34	99.45±0.13	98.53±0.44	98.54±0.19	99.3±0.25	99.48±0.27	99.66±0.23
κ statistic	-	-	93.27±0.57	95.11±0.37	97.56±0.56	99.09±0.22	97.59±0.72	97.62±0.31	98.86±0.4	99.16±0.45	99.45±0.37

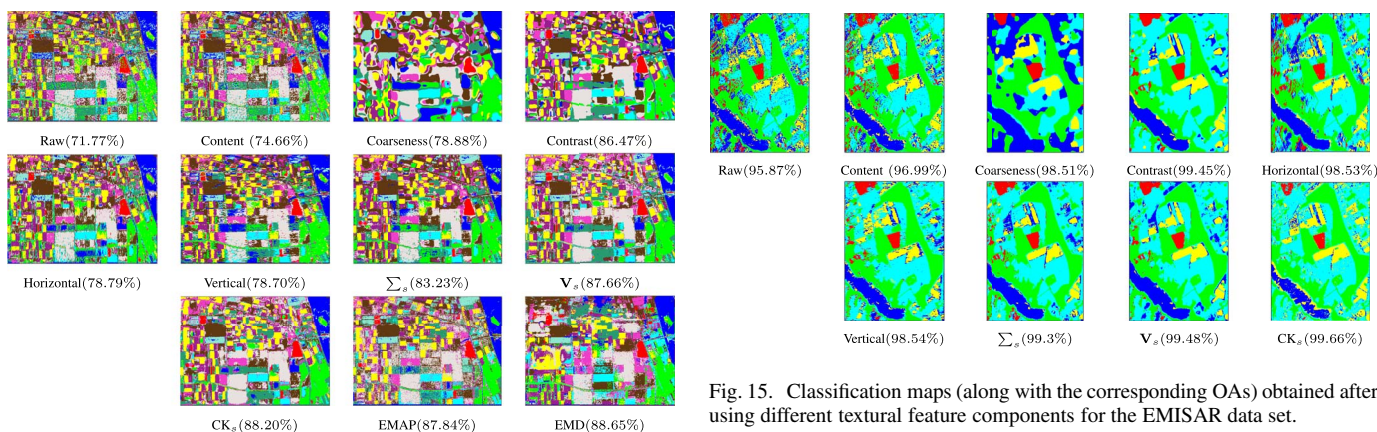


Fig. 14. Classification maps (along with the corresponding OAs) obtained after using different textural feature components for the AirSAR data set.

sizes of $U(a)$ and p also bring very good results. Therefore, we can choose relatively small dictionaries for the MMCA, thus alleviating computational cost in the experiments.

Fig. 15. Classification maps (along with the corresponding OAs) obtained after using different textural feature components for the EMISAR data set.

4) *Statistical Significance Results When Using Multiple Textural Features:* In our last experiment in this section, we perform an analysis of the statistical significance of differences among all of the considered textural features by using McNemar’s test [58]. In this experiment, the value of $|z| > 1.96$

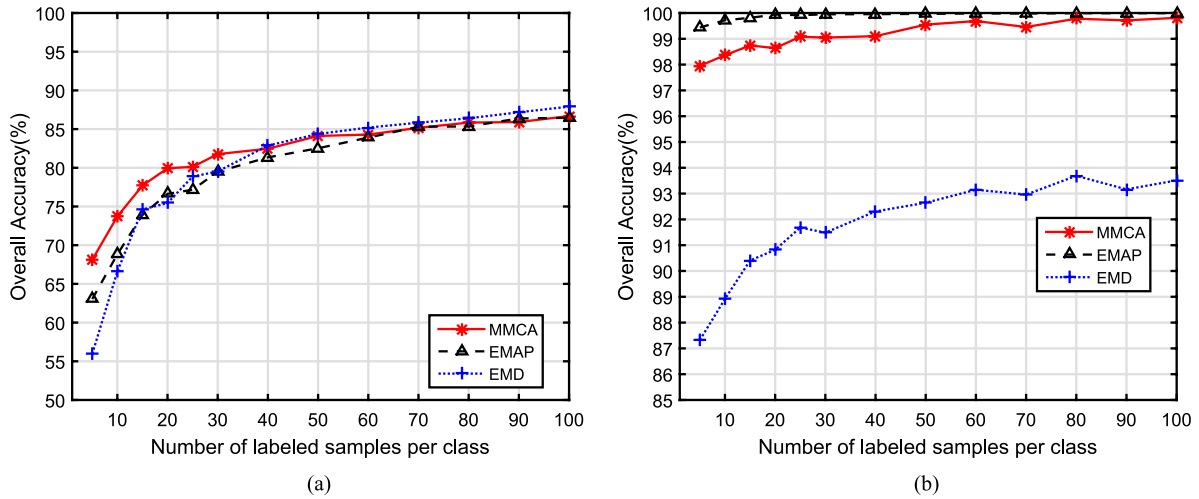


Fig. 16. OAs (as a function of the number of training samples per class) obtained by MMCA, EMAP, and EMD for two polarimetric SAR data sets. (a) For the AirSAR data set. (b) For the EMISAR data set.

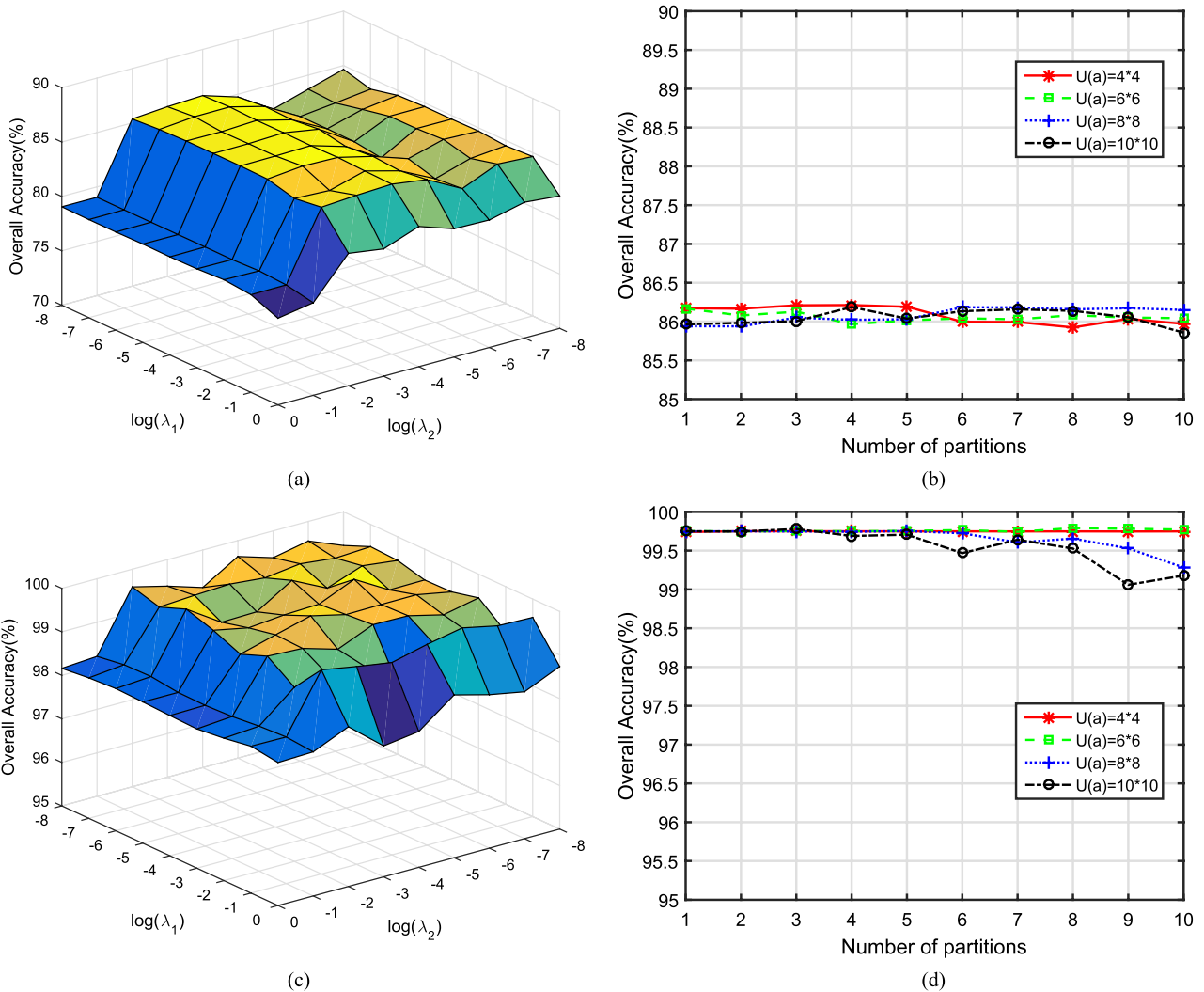


Fig. 17. Investigation of parameter settings: (a) and (b) refer to the AirSAR data set, while (c) and (d) refer to the EMISAR data set. (a) OAs as a function of λ_1 and λ_2 for AirSAR. (b) OAs as a function of the size of image partition $U(a)$ and the number of atoms for AirSAR. (c) OAs as a function of λ_1 and λ_2 for EMISAR. (d) OAs as a function of the size of image partition $U(a)$ and the number of atoms for EMISAR.

TABLE IX
STATISTICAL SIGNIFICANCE OF THE DIFFERENCES IN CLASSIFICATION ACCURACIES (MEASURED USING MCNEMAR'S TEST IN [58]) FOR THE PROPOSED MMCA FRAMEWORK, USING DIFFERENT TEXTURAL FEATURES EXTRACTED FROM THE AIRSAR DATA SET

	Value of z calculated by the McNemar's test				
	Content	Coarseness	Contrast	Horizontal	Vertical
Content	-	-16.5831	-24	-18.6815	-16.4012
Coarseness	16.5831	-	-17.3494	-8.6023	2.4495
Contrast	24	17.3494	-	15.0665	17.5214
Horizontal	18.6815	8.6023	-15.0665	-	8.9442
Vertical	16.4012	-2.4495	-17.5214	-8.9442	-

TABLE X
STATISTICAL SIGNIFICANCE OF THE DIFFERENCES IN CLASSIFICATION ACCURACIES (MEASURED USING MCNEMAR'S TEST IN [58]) FOR THE PROPOSED MMCA FRAMEWORK, USING DIFFERENT TEXTURAL FEATURES EXTRACTED FROM THE EMISAR DATA SET

	Value of z calculated by the McNemar's test				
	Content	Coarseness	Contrast	Horizontal	Vertical
Content	-	9.3808	-5.1962	2.6458	-2
Coarseness	-9.3808	-	-10.7238	-9	-9.5917
Contrast	5.1962	10.7238	-	5.8310	4.7958
Horizontal	-2.6458	9	-5.8310	-	-3.3166
Vertical	2	9.5917	-4.7958	3.3166	-

indicates the significant difference in accuracy between two classification methods. Tables IX and X, respectively, provide the results obtained for five textural features with the AirSAR and EMISAR data sets. As we can observe, these textural features have significant differences in classification accuracies (all of the values of $|z| > 1.96$) since the sign of z is a criterion to indicate the priority between two methods ($|z| > 0$ indicates that the first classifier is better than the second classifier or vice versa). The obtained results are also in accordance with the classification results from Tables VII and VIII. It is clear that the differences of the different textural features are statistically significant. Therefore, we conclude that it is essential to exploit different textural features for classification.

IV. CONCLUSION AND FUTURE LINES

In this paper, we have proposed a new method for advanced classification of remotely sensed images based on MMCA and sparse representation. The proposed MMCA method decomposes the original image into several pairs of MCs, where different MCs represent different image textural features, and corresponding dictionaries are generated to calculate morphological coefficients. Then, we have performed sparse-representation-based image decomposition. The proposed approach for advanced morphological feature extraction is then combined with a widely used classifier in order to perform clas-

sification of remotely sensed images. Our experimental results, conducted using a variety of hyperspectral and polarimetric SAR images, indicate that the proposed approach is suitable for analyzing multiple kinds of remote sensing data, leading to better classification performance than those exhibited by competitors in different scenarios, with particular emphasis in case studies dominated by limited training samples (with either balanced or unbalanced distribution of the samples used for training). The capacity of the proposed method to deal with different kinds of remotely sensed images results from its ability to incorporate multiple texture features in order to fully retrieve the image texture information, rather than using a single spatial characteristic of the texture. Since different data sets contain different structures and different images may require different textural features for adequately representing their content, in the future, we will explore the important topic of how to select an optimized textural feature for a given image. Another important topic deserving future research is a more detailed evaluation of the methodology with different types of remote sensing data such as light detection and ranging (LiDAR). We will also exploit the proposed classification approach with multiple data sources in the context of data fusion.

ACKNOWLEDGMENT

The authors would like to thank Prof. D. Landgrebe for making the AVIRIS Indian Pines hyperspectral data set available to the community, Prof. P. Gamba for providing the ROSIS data over Pavia, Italy, along with the training and test sets, and the editors and the anonymous reviewers for their detailed comments and suggestions, which greatly helped us to improve the clarity and presentation of this paper.

REFERENCES

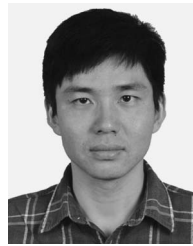
- [1] L. Jimenez and D. Landgrebe, "Supervised classification in high dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 28, no. 1, pp. 39–54, Feb. 1998.
- [2] L. Jimenez, J. Rivera-Medina, E. Rodríguez-Díaz, E. Arzuaga-Cruz, and M. Ramírez-Vélez, "Integration of spatial and spectral information by means of unsupervised extraction and classification for homogenous objects applied to multispectral and hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 844–851, Apr. 2005.
- [3] S. Kumar and M. Hebert, "Discriminative random fields," *Int. J. Comput. Vis.*, vol. 68, no. 2, pp. 179–201, Jun. 2006.
- [4] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semi-supervised discriminative random field for hyperspectral image classification," in *Proc. 4th WHISPERS*, 2012, pp. 1–4.
- [5] P. Zhong and R. Wang, "Learning conditional random fields for classification of hyperspectral images," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1890–1907, Jul. 2010.
- [6] G. Zhang and X. Jia, "Simplified conditional random fields with class boundary constraint for spectral-spatial based remote sensing image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 5, pp. 856–860, Sep. 2012.
- [7] J. A. Richards and X. Jia, "Context classification using evidential relaxation," in *Proc IEEE IGARSS*, 2004, vol. 1, pp. 532–535.
- [8] 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.
- [9] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.

- [10] M. D. Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3747–3762, Oct. 2010.
- [11] J. L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.
- [12] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial-spectral kernel-based approach for the classification of remote-sensing images," *Pattern Recognit.*, vol. 45, no. 1, pp. 381–392, Jan. 2012.
- [13] J. C. Nunes, Y. Bououane, E. Delechelle, O. Niang, and P. Bunel, "Image analysis by bidimensional empirical mode decomposition," *Image Vis. Comput.*, vol. 21, no. 12, pp. 1019–1026, Nov. 2003.
- [14] B. Demir and S. Erturk, "Empirical mode decomposition of hyperspectral images for support vector machine classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4071–4084, Nov. 2010.
- [15] Y. Tang, Y. Lu, and H. Yuan, "Hyperspectral image classification based on three-dimensional scattering wavelet transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2467–2480, May 2015.
- [16] J. M. Bioucas-Dias *et al.*, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [17] 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.
- [18] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. S1, pp. 110–122, Sep. 2009.
- [19] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitioned clustering techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.
- [20] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognit.*, vol. 43, no. 7, pp. 2367–2379, Jul. 2010.
- [21] X. Huang, Q. Lu, L. Zhang, and A. Plaza, "New postprocessing methods for remote sensing image classification: A systematic study," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7140–7159, Nov. 2014.
- [22] J. Bai, S. Xiang, and C. Pan, "A graph-based classification method for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 803–817, Feb. 2013.
- [23] G. Zhang, X. Jia, and J. Hu, "Superpixel-based graphical model for remote sensing image mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 5861–5871, Nov. 2015.
- [24] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new Bayesian approach with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3947–3960, Oct. 2011.
- [25] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, Mar. 2012.
- [26] S. Velasco-Forero and J. Angulo, "Classification of hyperspectral images by tensor modeling and additive morphological decomposition," *Pattern Recognit.*, vol. 46, no. 2, pp. 566–577, Feb. 2013.
- [27] C. Yu, Q. Qiu, Y. Zhao, and X. Chen, "Satellite image classification using morphological component analysis of texture and cartoon layers," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1109–1113, Sep. 2013.
- [28] J. Bobin, J. L. Starck, J. Fadili, and Y. Moudden, "Sparsity and morphological diversity in blind source separation," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2662–2674, Nov. 2007.
- [29] Z. Xue, J. Li, L. Cheng, and P. Du, "Spectral-spatial classification of hyperspectral data via morphological component analysis-based image separation," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 70–84, Jan. 2015.
- [30] X. Jia, B. C. Kuo, and M. Crawford, "Feature mining for hyperspectral image classification," *Proc. IEEE*, vol. 101, no. 3, pp. 676–697, Mar. 2013.
- [31] J. M. Bioucas-Dias and M. Figueiredo, "Logistic Regression via Variable Splitting and Augmented Lagrangian Tools," Inst. Superior Técnico, Tech. Univ. Lisbon, Lisbon, Portugal, Tech. Rep., 2009.
- [32] 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.
- [33] P. Ready and P. Wintz, "Information extraction, SNR improvement, and data compression in multispectral imagery," *IEEE Trans. Commun.*, vol. COM-21, no. 10, pp. 1123–1131, Oct. 1973.
- [34] J. B. Lee, A. S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 3, pp. 295–304, May 1990.
- [35] X. Jia and J. A. Richards, "Efficient hyperspectral remote sensing image display and classification using segmented principal components transformation," in *Proc. 7th Australasian Remote Sens. Conf.*, 1994, pp. 351–358.
- [36] X. Jia and J. A. Richards, "Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 538–542, Jan. 1999.
- [37] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, no. 6, pp. 460–473, Jun. 1978.
- [38] E. J. Candes and D. L. Donoho, *Curvelets: A Surprisingly Effective Non-adaptive Representation for Objects With Edges*, DTIC Document, 1999.
- [39] D. L. Donoho and M. R. Duncan, "Digital curvelet transform: Strategy, implementation, and experiments," in *Proc. SPIE Aerosense, Wavelet Appl. VII*, 2000, vol. 4056, pp. 12–29.
- [40] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Process.*, vol. 1, no. 2, pp. 205–220, Apr. 1992.
- [41] J. L. Starck, J. Fadili, and F. Murtagh, "The undecimated wavelet decomposition and its reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 297–309, Feb. 2007.
- [42] E. J. Candes and J. Emmanuel, "Ridgelets: Theory and Applications," Ph.D. dissertation, Dept. Stat., Stanford Univ., Stanford, CA, USA, 1998.
- [43] E. Feig and S. Winograd, "Fast algorithms for the discrete cosine transform," *IEEE Trans. Signal Process.*, vol. 40, no. 9, pp. 2174–2193, Sep. 1992.
- [44] T. S. Lee, "Image representation using 2D Gabor wavelets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 10, pp. 1–13, Oct. 1996.
- [45] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vis.*, Bombay, India, 1998, pp. 839–846.
- [46] P. S. Addison, *The Illustrated Wavelet Transform Handbook*. Bristol, U.K.: Institute of Physics Publishing, 2002.
- [47] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
- [48] H. Demirel and G. Anbarjafari, "Image resolution enhancement by using discrete and stationary wavelet decomposition," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1458–1460, May 2011.
- [49] J. M. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd WHISPERS*, 2010, pp. 1–4.
- [50] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Marí, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.
- [51] J. Immerkaer, "Fast noise variance estimation," *Comput. Vis. Image Understand.*, vol. 64, no. 2, pp. 300–302, Sep. 1996.
- [52] M. Basseville, "Distance measures for signal processing and pattern recognition," *Signal Process.*, vol. 18, no. 4, pp. 349–369, Dec. 1989.
- [53] M. D. Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 542–546, May 2011.
- [54] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. A, Math. Phys. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
- [55] X. Huang *et al.*, "Multiple morphological profiles from multicomponent-base images for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 12, pp. 4653–4669, Dec. 2014.
- [56] J. Lee, M. R. Grunes, and E. Pottier, "Quantitative comparison of classification capability: Fully-polarimetric versus dual and single-polarization SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2343–2351, Nov. 2001.
- [57] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 4–19, Jan. 2003.
- [58] G. M. Foody, "Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 5, pp. 627–633, May 2004.



Xiang Xu received the B.S. and M.E. degrees from Sun Yat-sen University, Guangzhou, China, in 1999 and 2002, respectively, where he is currently working toward the Ph.D. degree in the Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, Center of Integrated Geographic Information Analysis, School of Geography and Planning.

He is currently a Lecturer with the University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan, China. His research interests include hyperspectral image classification, hyperspectral unmixing, pattern recognition, and machine learning.



Xin Huang (M'13–SM'14) received the Ph.D. degree in photogrammetry and remote sensing from the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIEMARS), Wuhan University, Wuhan, China, in 2009.

He is currently a Full Professor with Wuhan University, where he teaches remote sensing, photogrammetry, image interpretation, etc. He is the Founder and Director of the Institute of Remote Sensing Information Processing (IRSIP), School of Remote Sensing and Information Engineering, Wuhan University. He has published more than 60 peer-reviewed articles in international journals. His research interests include hyperspectral data analysis, high-resolution image processing, pattern recognition, and remote sensing applications.

Prof. Huang was the recipient of the Top-Ten Academic Star of Wuhan University in 2009, the Boeing Award for the Best Paper in Image Analysis and Interpretation from the American Society for Photogrammetry and Remote Sensing (ASPRS) in 2010, the New Century Excellent Talents in University from the Ministry of Education of China in 2011, the National Excellent Doctoral Dissertation Award of China in 2012, and the China National Science Fund for Excellent Young Scholars in 2015. In 2011, he was recognized by the IEEE Geoscience and Remote Sensing Society (GRSS) as a Best Reviewer of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (GRSL). He was the winner of the IEEE GRSS 2014 Data Fusion Contest. He was the Lead Guest Editor of the issue on information extraction from high spatial resolution optical remotely sensed imagery for the IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND REMOTE SENSING (May 2015). Since 2014, he has served as an Associate Editor of the IEEE GRSL.



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

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

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



Mauro Dalla Mura, (S'08–M'11) received the Laurea (B.E.) and Laurea Specialistica (M.E.) 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) from the University of Trento and in electrical and computer engineering from the University of Iceland, Reykjavik, Iceland, in 2011.

In 2011, he was a Research Fellow with Fondazione Bruno Kessler, Trento, conducting research on computer vision. He is currently an Assistant Professor with the Grenoble Institute of Technology (Grenoble INP), Grenoble, France. He conducts his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). 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 is a member of the Geoscience and Remote Sensing Society (GRSS) and IEEE GRSS Data Fusion Technical Committee and the Secretary of the IEEE GRSS French Chapter (2013–2016). He was a Lecturer at the RSSS12—Remote Sensing Summer School 2012 (organized by the IEEE GRSS), Munich, Germany. He was the recipient of the IEEE GRSS Second Prize in the Student Paper Competition of the 2011 IEEE IGARSS and coreipient of the Best Paper Award of the *International Journal of Image and Data Fusion* for the year 2012–2013 and the Symposium Paper Award for IEEE IGARSS 2014. He is a Reviewer of the 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*, *ISPRS Journal of Photogrammetry and Remote Sensing*, and *Photogrammetric Engineering and Remote Sensing*.



Antonio Plaza (M'05–SM'07–F'15) was born in Cáceres, Spain, in 1975. He received the Computer Engineer degree in 1997, the M.Sc. degree in 1999, and the Ph.D. degree in 2002, all in Computer Engineering. He is an Associate Professor (with accreditation for Full Professor) with the Department of Technology of Computers and Communications, University of Extremadura, Cáceres, where he is the Head of the Hyperspectral Computing Laboratory (HyperComp), one of the most productive research groups working on remotely sensed hyperspectral

data processing worldwide. He has been the advisor of 12 Ph.D. dissertations and more than 30 M.Sc. dissertations. He was the Coordinator of the Hyperspectral Imaging Network, a European project with total funding of 2.8 million Euro. He is the author of more than 500 publications, including 152 journal papers (more than 100 in IEEE journals), 22 book chapters, and over 240 peer-reviewed conference proceeding papers (94 in IEEE conferences). He has edited a book on High-Performance Computing in Remote Sensing for CRC Press/Taylor and Francis and guest edited nine special issues on hyperspectral remote sensing for different journals. His main research interests comprise hyperspectral data processing and parallel computing of remote sensing data.

Dr. Plaza is a Fellow of the IEEE “for contributions to hyperspectral data processing and parallel computing of Earth observation data.” He served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) in 2011–2012, and he is currently serving as the President of the Spanish Chapter of IEEE GRSS (since November 2012). He has served as a Proposal Evaluator for the European Commission, the National Science Foundation, the European Space Agency, the Belgium Science Policy, the Israel Science Foundation, and the Spanish Ministry of Science and Innovation. He has reviewed more than 500 manuscripts for over 50 different journals. He is currently serving as the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING journal. He was 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 served as an Associate Editor in 2007–2012. He is also an Associate Editor of IEEE ACCESS, and he was a member of the Editorial Board of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER (2011–2012) and the IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE (2013). He was also a member of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS). He was a recipient of the 2013 Best Paper Award of the JSTARS journal and a recipient of the most highly cited paper (2005–2010) in the *Journal of Parallel and Distributed Computing*. He received best paper awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. He was a recipient of the Best Ph.D. Dissertation award at the University of Extremadura, a recognition also received by six of his Ph.D. students.