Subpixel Component Analysis for Hyperspectral Image Classification

Xiang Xu, *Member, IEEE*, Jun Li, *Senior Member, IEEE*, Shutao Li, *Fellow, IEEE*, and Antonio Plaza, *Fellow, IEEE*

***Abstract*— Land-cover classification with hyperspectral imagery has been an active topic in the remote sensing community. It aims at relating a unique class label to each pixel in the scene, so that it can be well defined by a given land cover type. In this paper, we explore the intrinsic characteristics of hyperspectral imagery from a subpixel-level perspective and propose a new subpixel component analysis (SCA) approach for feature extraction and land-cover classification. The core idea of SCA is that we extract a subpixel attribute component feature from the abundance maps. Compared with the abundance maps, the extracted subpixel feature image shows higher signal-to-noise level and clearer spatial distribution details. In order to deal with spectral variability, as well as obtain representative image endmember signatures and their corresponding abundance maps, we adopt a regional clustering-based spatial preprocessing (RCSPP) strategy for endmember identification, and a partial unmixing model based on mixture tuned matched filtering (MTMF) for abundance estimation. Furthermore, to highlight the spatial distribution details as well as eliminate the noise disturbance in the derived abundance maps, we perform sparse image decomposition on the obtained abundance maps, thus achieving a new subpixel feature representation for classification. Our experimental results reveal that the proposed SCA approach can obtain feature representation with explicit physical meaning, clear spatial distribution details, and better noise robustness, leading to state-of-the-art classification results.**

***Index Terms*— Hyperspectral image classification, partial unmixing, regional clustering-based spatial preprocessing (RCSPP), subpixel component analysis (SCA).**

Manuscript received September 12, 2018; revised January 28, 2019; accepted February 17, 2019. Date of publication April 3, 2019; date of current version July 22, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61771496 and Grant 61502088, in part by Guangdong Provincial Natural Science Foundation under Grant 2016A030313254, in part by the National Key Research and Development Program of China under Grant 2017YFB0502900, in part by the Open Research Fund of Key Laboratory of Spectral Imaging Technology, Chinese Academy of Sciences under Grant LSIT201708D, in part by the Zhongshan City Science and Technology Research Project under Grant 2018B1015, and in part by the MINECO Project under Grant TIN2015-63646-C5-5-R. *(Corresponding author: Jun Li.)*

X. Xu is with the Zhongshan Institute, University of Electronic Science and Technology of China, Zhongshan 528402, China.

J. Li and S. Li are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: lijun8206@hnu.edu.cn).

A. Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, 06006 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.2019.2900484

I. Introduction

LAND use/land cover (LULC) classification aims at assigning a unique class label to each pixel in the scene, so that it can be well defined by a given land cover type [1], [2]. Accurate LULC classification is of great benefit to land management and planning tasks, as well as environmental monitoring [3]. In recent decades, hyperspectral imagery has been widely used for land-cover classification purposes [4]–[6]. Hyperspectral imagery comprises hundreds of continuous and narrow spectral channels. For example, the well-known airborne visible infrared imaging spectrometer (AVIRIS) provides more than 200 spectral channels with a spectral resolution of 10 nm and a wavelength range from 0.4 to 2.5 μm [7]. Such fine spectral resolution provides strong discriminability for different land cover types. However, a very fine spectral resolution often requires a large instantaneous field of view, thus leading to limited spatial resolution. As a result, pixels in a hyperspectral image are usually a mixture of different constituent substances located in their respective pixel coverage [8]. The existence of mixed pixels poses a great challenge to accurate classification [9]. Moreover, due to the difficulty of sample collection, training samples in remote sensing are difficult to obtain, and the imbalance between the limited number of available training samples and the high dimensionality of the data often challenges the learning stage of the classification model [10], [11]. On the one hand, very few training samples generally result in an ill-posed problem for some classification methods [12]. On the other hand, high data dimensionality can easily result in an overfitting problem, thus weakening the generalization ability of the classifier. Many algorithms have been proposed for dimensionality reduction, such as the locality adaptive discriminant analysis algorithm in [13], designed to learn a representative data subspace. An optimal clustering framework is also constructed in [14] for hyperspectral band selection.

On the surface, the mixed pixels have a certain negative impact on accurate classification. However, some positive factors can be found through an in-depth analysis of the physical formation mechanism of mixed pixels, and related research has been developed. Luo and Chanussot [15] regarded abundance maps as a kind of feature for hyperspectral image classification purposes. Compared with conventional pixel-level feature extraction methods, such as principal component analysis (PCA), minimum noise fraction (MNF), independent component analysis (ICA), and so on, the abundance

features show better spatial distribution details of different ground objects and have a more explicit physical meaning. In this method, the number of endmembers is determined by an eigenvalue-based strategy [16], and the vertex component analysis (VCA) [17] is adopted for endmember identification purposes. After obtaining the set of endmember signatures, it adopts the fully constrained linear spectral unmixing (FCLSU) [18] technique for abundance estimation. Considering that a lot of noise disturbance exists in the obtained abundance maps, Luo *et al.* [19] introduced a preliminary segmentation algorithm on the obtained abundance maps and chosen the average value of each segment partition as the final features for classification. Although adopting an unsupervised classification scheme, to the best of our knowledge, this is the first framework that uses subpixel abundance information for classification purposes.

Recently, feature extraction based on subpixel abundance information has been exploited under a supervised scheme [20], [21]. In this scheme, a partial unmixing model based on the mixture-tuned matched filtering (MTMF) technique [22]–[24] is adopted for abundance estimation, which can estimate the abundance of a known target material in the presence of a composite background. In [21], different unmixing chains are proposed to identify the endmember signatures, including conventional spectral-based endmember identification methods such as VCA [17], spatial preprocessing-based unmixing methods such as spatial preprocessing (SPP) [25], clustering-based methods, and a method that calculates the average of the training samples in each class. The experimental results reveal that all the aforementioned unmixing chains can obtain abundance features with clearer spatial distribution details than conventional approaches, such as PCA and MNF. Furthermore, considering the coexistence of pure and mixed pixels in real hyperspectral scenes, Li *et al.* [26] proposed a technique for complementary integration of discriminative classification and spectral unmixing in a semisupervised classification framework. Different from the strategies presented in [21], where all pixels are handled in a similar way, Li *et al.* [26] suggested that different strategies should be applied to deal with each kind of pixel, depending on its spectral purity. If a pixel is highly mixed, it can be characterized by a partial unmixing model. If a pixel is more spectrally pure, its spectral information can be directly used for classification purposes.

From a subpixel-level perspective, the aforementioned abundance information provides a new kind of feature representation for classification purposes, and this kind of abundance features own explicit physical meaning, which can lead to competitive classification accuracy. However, due to the difficulty of sample collection and labeling as well as the prevalence of spectral variability [27], [28] in the scene, errors in endmember identification can also lead to errors in abundance estimation [29], [30]. As a result, the classification accuracy obtained by directly using such inaccurate abundance maps can still be significantly improved.

On the one hand, in order to better identify endmember signatures, the integration of spectral and spatial information in the hyperspectral image has attracted extensive attention [31]–[34]. It is based on the fact that endmember signatures are more likely to appear in spatially homogeneous regions, which exhibit the most singular signatures. Generally, the pixels in these regions have been identified by using unsupervised clustering methods [35]–[38] (e.g., ISODATA [39], K-means, or the hierarchical segmentation algorithm [40]). Then, the spectrally pure candidate pixels are selected for subsequent processing. Compared to using all image pixels, the number of candidate pixels selected this way is much smaller, and these candidates are often more spectrally pure and representative. However, unsupervised clustering methods just consider the spectral information while neglecting the spatial correlation among adjacent pixels. Recently, superpixel-based segmentation methods have been introduced to obtain homogeneous partitions in the scene, including graph-based algorithms [41], turbopixel [42], simple linear iterative clustering [43], and regional clustering-based SPP (RCSPP) [44]. These superpixel methods, while acting as unsupervised clustering methods, can obtain homogeneous regions with better spatial continuity and spectral similarity, thus promoting a more accurate identification of endmembers. In any event, and regardless of the technique used for endmember identification purposes, it is not always easy to extract a complete endmember set from the image data [45].

On the other hand, by means of the partial unmixing model, the obtained abundance maps are calculated by projecting on a particular endmember signature. Thus, the result itself reflects the spatial distribution details of a particular ground object, which improves the classification accuracy. However, due to the noise disturbance, a direct use of these abundance maps straightforwardly brings a lot of classification errors. Therefore, it is feasible and necessary to carry out further feature extraction from these noisy abundance maps.

In this paper, we propose a new subpixel component analysis (SCA) approach for feature extraction and classification of hyperspectral imagery. First, in order to deal with spectral variability, as well as obtain representative image endmember signatures and their corresponding abundance maps, we introduce an RCSPP [44] strategy for endmember identification, and a partial unmixing model based on MTMF for abundance estimation [22], [23]. Then, the SCA approach is performed on the previously obtained abundance maps for further feature extraction, thus obtaining a subpixel feature representation with clearer distribution details, stronger discriminant ability, and better noise robustness. Our newly developed SCA is derived from the theory of sparse image decomposition, which can decompose abundance feature images into a pair of attribute components (e.g., *smoothness* and *texture*). Then, the *smoothness* component is fed into the classifier to obtain better classification accuracy. Compared with the conventional hyperspectral classification methods, the main innovative contributions of our work can be summarized as follows.

1) First, we exploit subpixel-level features for hyperspectral imagery classification, while most conventional classification schemes mainly focus on pixel-level or object-level features. This kind of subpixel feature representation can provide features with explicit physical

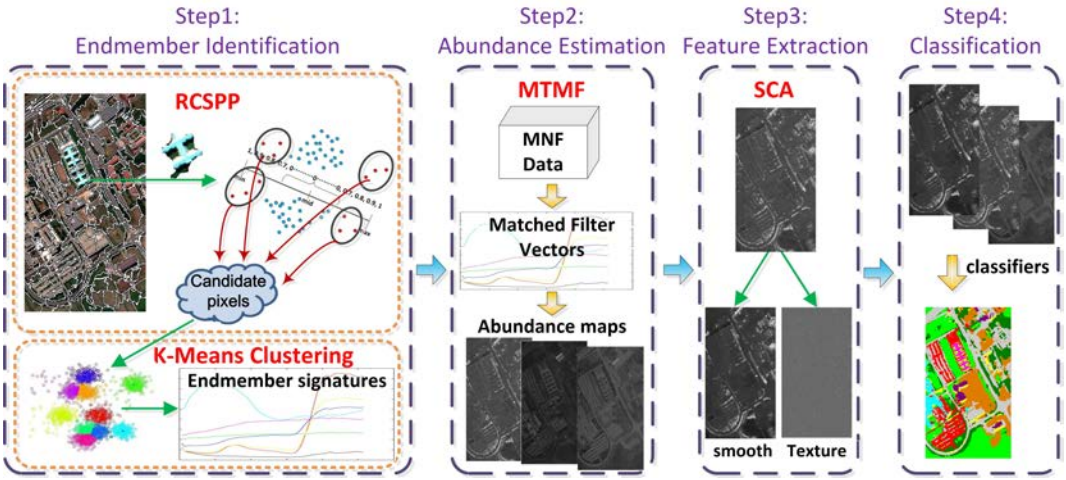


Fig. 1. Block diagram of the proposed SCA approach for hyperspectral image classification.

meaning while addressing the presence of mixed pixels in the scene.

2) Second, we introduce a new endmember identification strategy. Compared to the strategies that perform unmixing on available (limited) training sets [21], the proposed strategy, which does not require any training samples, can obtain more representative endmember signatures directly from the original image.
3) Third, we perform feature extraction on the noisy abundance maps and propose a new SCA approach. Corresponding to the abundance features, the obtained subpixel component features eliminate the noise disturbance and highlight the distribution details of ground objects, leading to higher classification accuracy.

The remainder of this paper is organized as follows. Section II presents the new SCA approach in step-by-step fashion. Section III discusses experimental results to comprehensively assess the performance of our new SCA approach. The experiments are carried out on four benchmark hyperspectral data sets, and the first two data sets are acquired by the AVIRIS sensor over the Indian Pines region in the Indiana and Kennedy Space Center (KSC), Merritt Island, FL, USA, respectively. The third data set is collected by the ROSIS sensor over the University of Pavia, Pavia, Italy. The last data set is collected by a hyperspectral data collection experiment (HYDICE) airborne flight over the Washington DC Mall, USA. Section IV concludes this paper with some remarks and hints at plausible future research lines.

II. Methodology

In this section, we describe the newly proposed SCA approach. In our proposed approach, the linear mixing model (LMM) [45] is used to model each mixed pixel. For a given hyperspectral pixel vector $\mathbf{x}$, the LMM can be written as

$$\mathbf{x} = \mathbf{M}\mathbf{a} + \mathbf{n} \tag{1}$$

where $\mathbf{M} := [\mathbf{m}_1, \ldots, \mathbf{m}_p]$ stands for the endmember matrix, $\mathbf{a} := [a_1, \ldots, a_p]^T$ stands for a fractional abundance vector, p is the number of endmembers, and $\mathbf{n}$ accounts for additive noise and model mismatches. Since the value of $\mathbf{a}$ stands for the abundance fractions, it should satisfy two constraints: abundance nonnegativity, i.e., $a_i \geq 0$, and abundance sum-to-one, i.e., $\sum_{i=1}^{p} a_i = 1$, respectively [18]. Based on the LMM, our purpose is to identify the endmember signatures from the image and then estimate their corresponding abundance maps. This kind of abundance maps can be thought of as a preliminary subpixel feature representation and can be used for further feature extraction. A block diagram of the proposed SCA approach is given in Fig. 1, including four main steps that are described as follows.

1) In step 1, the RCSPP strategy [44] is used to select a subset of candidate pixels from the set of original image pixels. Although we can perform endmember identification on the whole set of image pixels directly, the computational burden is quite high. Recent studies have indicated that the integration of spectral–spatial information usually exhibits better unmixing performance [31]. In [44], the RCSPP strategy is shown to exhibit fast computational speed and adaptivity to the spectral variability. After obtaining the candidate pixels, a K-means clustering procedure is performed on them, and only the centers of each cluster are chosen as representative endmember signatures.
2) In step 2, a partial MTMF unmixing model [23], [24] is used for abundance estimation. Compared to the FCLSU [18], which requires a full set of endmember signatures in advance, the partial unmixing model can map a known endmember signature without knowing the other background endmember signatures.
3) In step 3, subpixel component decomposition is performed on the obtained abundance features for further feature extraction, where the abundance feature images are decomposed into a pair of attribute components.
4) In step 4, the *smoothness* components are chosen as new subpixel features and fed into the classifier. In our experiments, two types of classifiers, including the support vector machine (SVM) [46], [47] and multinomial

logistic regression (MLR) [48], [49], are used for classification purposes. In the following, we provide the details of each building block.

A. Endmember Identification

In [21] and [26], two sources of information are exploited for the identification of endmember signatures: the original image pixels and the labeled training set. The former often contains a large number of mixed pixels, which poses a great challenge for the identification of pure signatures. Due to the difficulty of sample collection, it is also a great challenge to obtain complete and representative endmember sets from the latter. Therefore, in order to acquire all endmember signatures from the image accurately, we developed a rapid and reliable endmember identification strategy. This strategy is derived from the RCSPP algorithm [44]. In the RCSPP, a regional clustering procedure is introduced, which can quickly divide the original image into a series of homogeneous partitions. Then, the candidate endmembers are picked out from each partition by using a PCA-based projection technique. Compared to using all image pixels, the number of obtained candidate endmembers by means of the RCSPP is smaller, and these candidates are more pure and representative. As a result, the endmember signatures only need to be extracted from the set of candidates provided by the RCSPP. A more detailed description of the aforementioned procedure is given as follows.

Aiming at the generation of homogeneous partitions, we adopted a regional clustering procedure which transforms the traditional global search into a local search. As a result, not only the spatial correlation and the spectral similarity are intrinsically incorporated at the clustering step but also fast clustering speed is obtained. First, we need to define the initial number of clusters and each cluster center. Here, the number of clusters is determined according to the spatial resolution of target hyperspectral imagery. Concretely, we use a spatial neighborhood purity index (SNPI) [50] to measure the purity of clustered partitions. If SNPI values of clustered partitions are larger than a predefined threshold, we stop further subdivision, which means that the homogeneity of these partitions has met the requirements. Let us suppose that the number of clusters is denoted by C. We first divide the original image into C nonoverlapping subblocks with the same scale, then the initial cluster centers are set to the central position of each subblock. After finishing the initialization, the main loop of the regional clustering procedure begins to run. Here, two key points need to be emphasized as follows.

1) The first one is related to the search scope of the regional clustering. Here, we restrict the search scope within a rectangular neighborhood window around each cluster center, and the size of the neighborhood window is set to two times the size of the initial partitions. Although this setting may generate nonoptimal clustered results from a local point of view, the obtained clustering partitions exhibit good homogeneity. Moreover, due to the local search, the clustering speed is significantly improved.
2) The second key point is related to the similarity measurement between pixels. For the purpose of integrating spatial and spectral information simultaneously, we design a new similarity measure criterion which is described by a weighted sum of the spectral and spatial distance. Relevant formulas are as follows:

$$\mathrm{SC} = (1-\lambda)\times \mathrm{SC}_{\mathrm{spe}} + \lambda \times \mathrm{SC}_{\mathrm{spa}} \tag{2}$$

$$\mathrm{SC}_{\mathrm{spe}} = \arccos\left(\frac{\sum_{j=1}^{b} x_{ij} \times x_{cj}}{\sqrt{\left(\sum_{j=1}^{b} x_{ij}^2\right) \times \left(\sum_{j=1}^{b} x_{cj}^2\right)}}\right) \tag{3}$$

$$\mathrm{SC}_{\mathrm{spa}} = \sqrt{(i_x - c_x)^2 + (i_y - c_y)^2}/d \tag{4}$$

where $\mathrm{SC}_{\mathrm{spe}}$ stands for the spectral similarity measurement, which is defined by the spectral angle distance [see (3)], $\mathbf{x}_i \equiv [x_{i1}, \ldots, x_{ib}]$ is the pixel vector i, $\mathbf{x}_c \equiv [x_{c1}, \ldots, x_{cb}]$ is the vector of cluster center pixel, and b is the number of spectral bands. $\mathrm{SC}_{\mathrm{spa}}$ stands for the spatial similarity measurement, which is defined as a regularized Euclidean distance [see (4)], where (i_x, i_y) and (c_x, c_y) are the spatial coordinates of pixel i and the corresponding cluster center c, and d is the length of the diagonal for the local search scope. Finally, λ is a weighting factor used to balance the spectral and spatial information.

Once the clustering loop is finished, we can obtain a series of homogeneous partitions, and the candidate pixels will be picked out from each partition. Considering that most pure pixels generally exhibit the most singular signatures in each homogeneous partition, we adopt a PCA-based projection technique. Concretely, for each partition, we choose the first few *PCs* as projection axes. Then, all pixels in this partition are projected onto these axes and we record their projection position. Finally, pixels that fall into the extreme projection position are picked out as candidate pixels. The number of pixels in the candidate subset is much smaller than that in the original image, but it is still far greater than the number of ground categories in the image. Therefore, we use K-means clustering to divide these candidate pixels into a specified number of clusters and choose the cluster centers as the final representative endmember signatures. A detailed pseudocode of our adopted endmember identification strategy is given in Algorithm 1.

B. Abundance Estimation

After obtaining endmember signatures, the corresponding abundance information can be extracted based on the LMM. However, in real hyperspectral scenes, spectral variability intrinsically exists. If we lack a prior knowledge of land cover types, it will be a challenging issue to obtain endmember signatures of all ground types. This incomplete endmember set easily invalidates some conventional unmixing algorithms (such as FCLSU). To solve this issue, we adopted a partial unmixing model to estimate the abundance of a known endmember signature without needing to know the other background endmember signatures. This partial unmixing model

Algorithm 1 Pseudocode of Our Endmember Identification Strategy

Input:
- Hyperspectral image data cube: $\mathbf{X} \in \Re^{m\times n\times b}$;
- Initial number of clusters: C;
- Weighting factor in Eq. (2): λ;
- Number of iterations: $iter$.

Output:
- Endmember signatures: $\mathbf{M} := [\mathbf{m}_1, \cdots, \mathbf{m}_p]$.

// Initialization:
- Divide $\mathbf{X}$ into C non-overlapping sub-blocks with the same scale;
- Set initial cluster center of each sub-block.

// Main looping of Regional clustering:
for $it = 1$ to $iter$ **do**
  for each partition **do**
    - calculate SC between cluster center with neighboring pixels by using Eq. (2).
  end for
  - Update each cluster and calculate new cluster center.
end for
// Extract endmember signatures:
for each partition **do**
  - Perform PCA transform and obtain PC vectors;
  - Project all pixels onto the first few PCs;
  - Choose candidate pixels that fall to the extreme projection positions.
end for
- Perform K-Means clustering on the set of candidate pixels;
- Choose the cluster centers as the final endmember signatures $\mathbf{M}$.

is based on the MTMF [22] and includes two main phases: the first phase is a matched filtering procedure, which is used for abundance estimation, and the second phase is a mixture tuning procedure, which is used for the identification and rejection of false positives. Here, we mainly use the matched filtering part. In addition, considering that the MTMF generally requires an input data with zero mean and unit noise variance, we perform the MNF transform on the original image. Detailed descriptions are given below.

During the matched filtering procedure, the key step is the calculation of the matched filtering vectors. Then, these filtering vectors will be used as projection axes, and the matched filtering scores of each pixel are calculated by projecting the data onto them. This process can also be described as filtering the input data for good matches to the target spectrum while suppressing the remaining background spectra. First, the endmember signatures are transformed into the MNF space. Then, the transformed endmember signatures are projected onto the inverse covariance of the MNF transformed image data. Finally, the projected endmember signatures are normalized by their magnitudes. Relevant formula is as follows:

$$\mathbf{mfv_i} = \frac{[\mathrm{COV}]^{-1}\#\mathbf{m_i}}{(\mathbf{m_i})^T\#[\mathrm{COV}]^{-1}\#\mathbf{m_i}} \tag{5}$$

where $\mathbf{mfv_i}$ corresponds to the matched filtering vector of endmember signature $\mathbf{m_i}$, [COV] is the covariance matrix of the MNF-transformed input image data, $[\mathrm{COV}]^{-1}$ is the inverse matrix of [COV], and # is the projection operator. Once the matched filtering vectors are obtained, the corresponding abundances can be calculated by projecting the entire MNF-transformed image onto these projection vectors. Finally, the abundances of all endmembers are stacked together so as to form a preliminary set of subpixel features.

C. Subpixel Component Decomposition

The aforementioned abundance feature images contain a stack of fractional information describing each constitute substance, which also reflect certain spatial distribution characteristics for different substances. However, due to spectral variability, one single endmember signature cannot accurately represent the corresponding constitute substance in each pixel, thus leading to errors in abundance estimation. Therefore, it is necessary to perform further feature learning on the set of abundance features. On the one hand, the noise disturbance should be eliminated as best as possible; on the other hand, some more discriminative attribute features can be extracted from the abundance features so as to further highlight the spatial distribution details of each ground type. Here, we adopt a sparse image decomposition approach called SCA. The basic idea of our newly developed SCA is derived from the MCA approach [51]. Different from conventional feature extraction methods (such as PCA or ICA), the SCA approach assumes that the target abundance image is composed of different attribute components, and each component can be expressed as the product of a particular dictionary and the corresponding sparse coefficient. Moreover, this particular dictionary is unable to provide an effective sparse representation on the other components. As a result, these dictionaries can be used to discriminate different components.

Let $\mathbf{y} \in \mathbb{R}^N$ denote the obtained abundance feature images, where N is the number of pixels in $\mathbf{y}$. Following the SCA approach, $\mathbf{y}$ can be decomposed into a *smoothness* component $\mathbf{y}_s$ and a *texture* component $\mathbf{y}_t$, respectively. These two components represent the abundance image $\mathbf{y}$ by means of a linear combination as follows:

$$\mathbf{y} = \mathbf{y}_s + \mathbf{y}_t + \mathbf{n} \tag{6}$$

where $\mathbf{n}$ is the residual in the approximation of the abundance image $\mathbf{y}$. Generally, the *smoothness* component reflects the edge contours, shape, and structural features, while the *texture* component reflects the spatial texture distribution characteristics.

In order to recover $\mathbf{y}_s$ and $\mathbf{y}_t$ from the abundance image $\mathbf{y}$, we construct the following sparse optimization problem:

$$\begin{aligned} &\arg\min_{\langle \mathbf{x}_s, \mathbf{x}_t\rangle} \|\mathbf{y} - \mathbf{y}_s - \mathbf{y}_t\|_2^2 + \lambda_1\|\mathbf{x}_s\|_1 + \lambda_2\|\mathbf{x}_t\|_1 \\ &\text{s.t.: } \mathbf{y}_s = \mathbf{A}_s x_s, \quad \mathbf{y}_t = \mathbf{A}_t x_t \end{aligned} \tag{7}$$

where $\mathbf{A}_s$ and $\mathbf{A}_t$ stand for two overcomplete dictionaries corresponding to the *smoothness* component $\mathbf{y}_s$ and *texture* component $\mathbf{y}_t$, respectively. x_s and x_t stand for the sparse

Algorithm 2 Pseudocode of the Proposed Subpixel Component Decomposition

Input:
- Abundance feature image: $\mathbf{y}$;
- Regularization parameters: λ_1, λ_2;
- Threshold of iteration residual: T_ε;
- Thresholds for *SoftThresholding()*: θ^s, θ^t;
- Number of iterations: $iter$.

Output:
- Decomposed attribute components: $\mathbf{y}_s$, $\mathbf{y}_t$.

// Initialization:
- Initial components: $\hat{\mathbf{y}}_s^0 = \hat{\mathbf{y}}_t^0 = 0$;
- Initial residuals: $\varepsilon^0 = 0, \varepsilon^1 = T_\varepsilon$;
- Randomly choose 10 image sub-blocks with size of 8×8 pixels, generate initial dictionaries A_s^0 and A_t^0.

// Component decomposition:

for $i = 1$ to $iter$ **do**

==========*smoothness* component==========

$A_s^i = DictionaryUpdate(\mathbf{y}, \mathbf{y}_s^{i-1}, A_s^{i-1})$;

$\hat{\mathbf{x}}_s^i = SUnSAL(A_s^i, \mathbf{y}_s^{i-1}, \mathbf{y}_t^{i-1}, \lambda_1)$; //sparse coding

$\hat{\mathbf{x}}_s^i = SoftThresholding(\hat{\mathbf{x}}_s^i, \theta^s)$;

$\hat{\mathbf{y}}_s^i = A_s^i \cdot \hat{\mathbf{x}}_s^i$;

$\hat{\mathbf{y}}_s^i = TotalVariation(\hat{\mathbf{y}}_s^i, \gamma_s)$;

==========*texture* component==========

$A_t^i = DictionaryUpdate(\mathbf{y}, \mathbf{y}_t^{i-1}, A_t^{i-1})$;

$\hat{\mathbf{x}}_t^i = SUnSAL(A_t^i, \mathbf{y}_s^i, \mathbf{y}_t^{i-1}, \lambda_2)$; //sparse coding

$\hat{\mathbf{x}}_t^i = SoftThresholding(\hat{\mathbf{x}}_t^i, \theta^t)$;

$\hat{\mathbf{y}}_t^i = A_t^i \cdot \hat{\mathbf{x}}_t^i$;

==========calculate residual==========

$\varepsilon^i = std(\mathbf{y} - \hat{\mathbf{y}}_s^i - \hat{\mathbf{y}}_t^i)$.

if $\varepsilon^i - \varepsilon^{i-1} < T_\varepsilon$ **then**

break;

end if

end for

coefficients corresponding to $\mathbf{y}_s$ and $\mathbf{y}_t$, respectively, λ_1 and λ_2 are the regularization parameters. Generally, instead of solving the sparse coefficient matrices $\langle \mathbf{x}_s, \mathbf{x}_t \rangle$, we can obtain $\mathbf{y}_s$ and $\mathbf{y}_t$ by solving an alternative optimization problem

$$\underset{\langle \mathbf{y}_s, \mathbf{y}_t \rangle}{\arg\min} \ \|\mathbf{y} - \mathbf{y}_s - \mathbf{y}_t\|_2^2 + \lambda_1 \|\mathbf{T}_s \mathbf{y}_s\|_1 + \lambda_2 \|\mathbf{T}_t \mathbf{y}_t\|_1 \quad (8)$$

where $\mathbf{T}_s = (\mathbf{A}_s^T \mathbf{A}_s)^{-1} \mathbf{A}_s^T$ 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, here, problem (8) is solved via the spectral unmixing by variable splitting and augmented Lagrangian (SUnSAL) algorithm [52]. A detailed pseudocode of the proposed subpixel component decomposition is given in Algorithm 2. It should be noted that, at each iteration, $\mathbf{A}_s$ and $\mathbf{A}_t$ are updated by using total variation and soft threshold regularization. Concretely, we first perform an analysis transform (such as curvelet or Gabor) on each dictionary image subblock. Then, the obtained transform coefficients are penalized via soft thresholds. Finally, the dictionary is reconstructed by the penalized coefficients and used for subsequent sparse image decomposition.

In (8), the quality of dictionaries plays an important role in the image component decomposition. For the representation of the *smoothness* component, multiscale geometric analysis and harmonic analysis techniques are widely used to construct dictionaries, including the curvelet, biorthogonal wavelet, undecimated wavelet, and local ridgelet. Among these techniques, the curvelet combines the advantages of ridgelet and wavelet and shows multiple desirable features such as anisotropy, multiscale, and multidirectional, which can optimally represent many geometrical features in the image, including lines, curves, edges, contours, and so on. With the aforementioned issues in mind, we adopted the curvelet technique to construct the *smoothness* dictionary. For the representation of *texture* component, two techniques are widely used, including the discrete cosine transform (DCT) and Gabor filtering. Among them, Gabor filtering has shown certain superiority in the task of texture feature extraction, especially exhibiting robustness to the change of intensities and contrast in the image. As a result, we resorted to the Gabor filtering to construct the *texture* dictionary.

D. Classification

Following the aforementioned procedures, we obtain subpixel component features which not only contain the fractional abundance information but also exhibit high signal-to-noise ratio (SNR) and clear spatial details. These subpixel component features are fed into a classifier for land-cover classification purposes. In our work, two widely used classifiers have been considered, both of which provide a probabilistic output. The first one is the MLR classifier [53], which takes advantage of a sparsity-inducing prior added on the regressors in order to obtain sparse estimations [54]. In its implementation, the logistic regression via SAL algorithm [52] is used to learn the MLR classifier, as this technique is able to learn directly the posterior class distributions and deal with the high dimensionality of hyperspectral data in a very effective way. The second classifier considered in this paper is a probabilistic SVM [55], which has also been widely used for hyperspectral image classification.

III. Experimental Results and Discussion

This section is organized as follows. Section III-A describes four hyperspectral data sets (hereinafter referred to as "IndianP," "KSC," "PaviaU," and "Washington DC") used to assess our newly developed SCA approach. Section III-B shows a comparative assessment between subpixel attribute component features and other commonly adopted features, including the original spectral information, PCA features, and traditional subpixel abundance features. In addition, we also perform a quantitative assessment on subpixel attribute component features, especially considering the SNR level. Section III-C shows classification experiments using the SCA approach, adopting both the MLR and SVM classifiers. Section III-D analyzes the sensitivity of the SCA approach to different parameter settings, with a particular focus on the number of

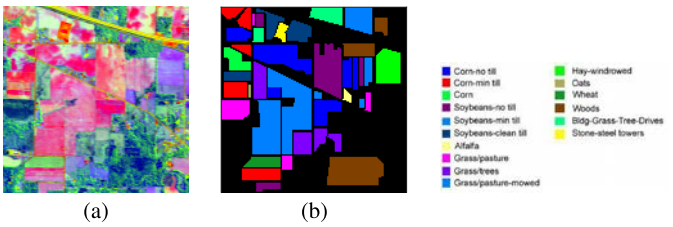
endmembers, the size of the regional search window, and the regularization parameters of the subpixel component decomposition. Finally, Section III-E analyzes the computational complexity of the SCA approach.

Before describing the experiments, we first give the parameter settings in our experiments. In the endmember identification procedure, we used the RCSPP strategy for clustering segmentation, where the number of clusters is set to 100, 300, 200, and 2000 for IndianP, KSC, PaviaU, and Washington DC, respectively. Considering the high spatial resolution of PaviaU and Washington DC, we set the weighting factor λ in (2) to 0.4, while for the IndianP and KSC, the λ is set to 0.1 due to the lower spatial resolution. Then, in each partition, we only pick 5% of all the pixels as candidates, thus the set of candidate endmembers chosen has a number of 1089, 15 803, 10 465, and 20 681 for IndianP, KSC, PaviaU, and Washington DC, respectively. After, we empirically set the number of representative endmembers to 20, 20, 15, and 10 for IndianP, KSC, PaviaU, and Washington DC, respectively. Since this parameter is sensitive to the approach, we will discuss in more detail later. In the procedure of subpixel component decomposition, most input parameters are empirically chosen after trial and error. Among them, the regularization parameters of λ_1 and λ_2 are set to the same value 1e$-$5, and the threshold of iteration residual $T_\varepsilon = 1\text{e}{-2}$. For the function of *SoftThresholding()*, the corresponding two thresholds: θ^s and θ^t are both set to 1e$-$4. For the maximum number of main iterations, we set it to 10. Finally, we would like to emphasize that all of our experiments were conducted on MATLAB *R*2015*a* in a desktop PC equipped with an Intel Core *i*7 CPU (at 3.6 GHz) and 32 GB of RAM.

A. Hyperspectral Data Sets

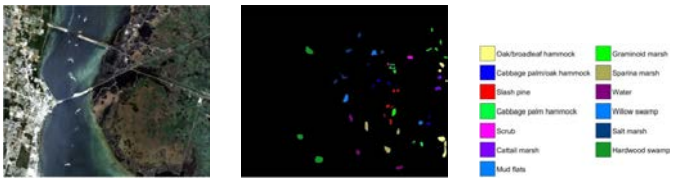
Four benchmark hyperspectral data sets have been used to assess the performance of our proposed SCA approach. The first two were collected by the NASA AVIRIS instrument. One was acquired over the Indian Pines region in Northwestern Indiana in 1992, and the other was acquired over the KSC, on March 23, 1996. The AVIRIS sensor acquires data in 224 bands of 10 nm width with center wavelengths from 0.4 to 2.5 μm. The "IndianP" data set consists of 145 $\times$ 145 pixels and was acquired over a mixed agricultural/forest area, early in the growing season. After removing the noise and water absorption bands, a total of 200 radiance channels are used in our experiments. For classification purposes, 16 mutually exclusive ground-truth classes are provided. This scene presents highly mixed pixels in all available classes, as well as an unbalanced number of labeled pixels per class. For illustrative purposes, Fig. 2(a) shows a false color composite of the image, while Fig. 2(b) shows the corresponding ground-truth map.

The KSC data consists of 512 $\times$ 614 pixels, with a spatial resolution of 18 m. After removing water absorption and low SNR bands, 176 bands were used for analysis. The training data were selected using land cover maps derived from color infrared photography provided by the KSC and Landsat Thematic Mapper imagery. The vegetation classification scheme



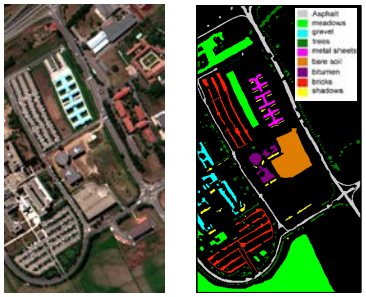
(a) (b)

Fig. 2. (a) False color composition of the AVIRIS IndianP scene. (b) Ground-truth map containing 16 mutually exclusive land-cover classes.



(a) (b)

Fig. 3. (a) False color composition of the AVIRIS KSC scene. (b) Ground-truth map containing 13 mutually exclusive land-cover classes.



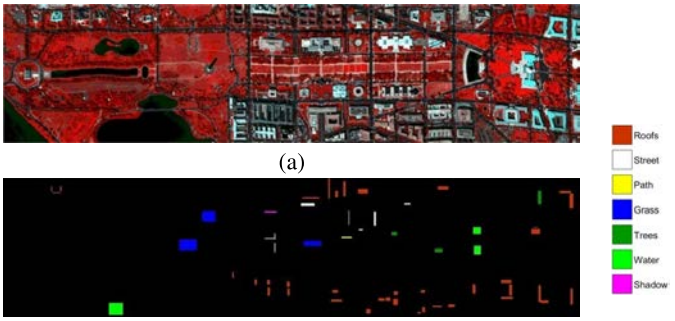
(a) (b)

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

was developed by KSC personnel in an effort to define functional types that are discernable at the spatial resolution of Landsat and these AVIRIS data. Discrimination of land cover for this scene is difficult due to the similarity of spectral signatures for certain vegetation types. Finally, 13 classes representing various land cover types that occur in this scene were defined for classification purposes. Fig. 3(a) shows a false color composite of the image, and Fig. 3(b) shows the corresponding ground-truth map.

The third hyperspectral data set was collected by the ROSIS sensor over the University of Pavia, Italy, on July 8, 2002. This image data, with a size of 610 $\times$ 340 pixels, exhibit a spatial resolution of 1.3 m and a number of 103 spectral bands (with spectral range from 0.43 to 0.86 μm). Nine ground-truth classes of interest are provided for classification, comprising urban features, soil, and vegetation. Fig. 4(a) shows a false color composite of the image, and Fig. 4(b) shows the corresponding ground-truth map.

The fourth data set was collected by an HYDICE airborne data flight over the Washington DC Mall. This image data, with a size of 1280 $\times$ 307 pixels, exhibit high resolution in both the spectral and the spatial domain (210 narrow bands



(a)

(b)

Fig. 5. (a) False color composition of the Washington DC scene. (b) (Right) Ground truth-map containing seven mutually exclusive land-cover classes.

with a spatial resolution of 2.5 m). After removing bands with water absorption and atmospheric interference, 191 bands were available for analysis. Seven ground-truth classes of interest are considered for classification purposes, comprising urban features, water, and shadow. Fig. 5(a) shows a false color composite of the image, and Fig. 5(b) shows the corresponding ground-truth map. This image is also a challenging data set for land-cover classification. First, the materials that are used to construct rooftops in the scene exhibit large diversity, so a single spectral response cannot be expected from this class; second, several groups of classes have similar spectral signatures, such as "water" and "shadow," "trees" and "grass," "roofs," "streets," and "path," which decreases class separability.

B. Comparative Experiments for SCA Feature

In this experiment, we analyze the feature extraction ability of SCA by comparing it with the other two commonly used feature extraction approaches. One is the PCA transform, a feature extraction method that can be seen as a representative data statistical method. Another one is the classic subpixel abundance features that are extracted by the MTMF partial unmixing model. Detailed descriptions of these methods are given in Section II. This kind of abundance features can be seen as representative subpixel-level features, and we note that our SCA features are also extracted on this abundance features.

First, we analyze the information contained in these different types of features. Fig. 6 shows six component images for each type of feature, extracted from the KSC data. The first row lists the first six *PCs*, ordered in terms of the energy that they comprise. The second row lists the subpixel abundance features (since each abundance feature component represents fractional information of a particular endmember signature, there is no specific order for them). The third row shows the SCA features, corresponding to each abundance feature component in the second row. Here, only the *smoothness* components are shown. At the same time, especially aiming to abundance and SCA features, we identified the ground types corresponding to each feature component, which were determined based on visual analysis and spectral matching (we compared the endmember signatures with those signatures in each training sample class). As shown in Fig. 6, some important considerations can be made as follows.

1) First, as *PCs* are derived from the transformation of the whole data, we can observe that the distribution of ground types suffers from an obvious weakening process with the decrease of energy, since only the first few *PCs* show clear spatial distribution details, while the other *PCs* are mainly occupied by the noise (such as "*PC6*" in Fig. 6). In turn, for the subpixel-level abundance features and SCA features, we can observe that each component image highlights a particular ground type and its spatial structure details, e.g., the component image of the first column shows the ground type "water," and the second and third columns show the distribution of "hardwood swamp" and "salt marsh" classes, respectively. Although these two components both represent hydrous types, the "hardwood swamp" class is mainly located in the center of the KSC scene, while the "salt marsh" class is mainly located in the river bank, upper rightmost area of the scene. The fourth column represents the distribution of "mud flats" class, and the fifth column represents some types of wetland plants, such as "oak/broadleaf hammock," "cabbage palm/oak hammock," and "cabbage palm hammock" for which we can observe clear spatial distribution details. The last column represents residential areas with man-made buildings. Although the building classes are not included in our classification scheme, we can still distinguish this ground type out of the other ground types in the component image. In summary, compared with the PCA features, the SCA features can provide clear spatial distribution details for particular ground types, based on the fact that the SCA features contain the fractional abundance of particular endmember signatures, while the PCA features are extracted by using statistical models on the data while neglecting the actual physical information.
2) Second, after comparing the component features obtained by SCA and traditional abundance features, we can observe that the SCA component features provide more clear spatial distribution details and better noise robustness. This is due to the fact that the SCA features are extracted from abundance features, and some noise disturbances are eliminated. Moreover, due to the sparse image decomposition process, the structure and distribution of ground objects have been further highlighted in the decomposed *smoothness* component. In order to make a quantitative assessment, we also estimate the noise variance in different feature images, as shown in Table I. Here, the adopted noise variance estimation strategy is derived from a fast and fully automated smoothing algorithm, which is based on a penalized least squares method by using the DCT [56]. It can be observed from Table I that both the SCA features and PCA features obtained the lowest noise variance, and SCA obtained lower noise variance. However, the PCA features are directly extracted from the whole data cube, while neglecting any prior distribution information of

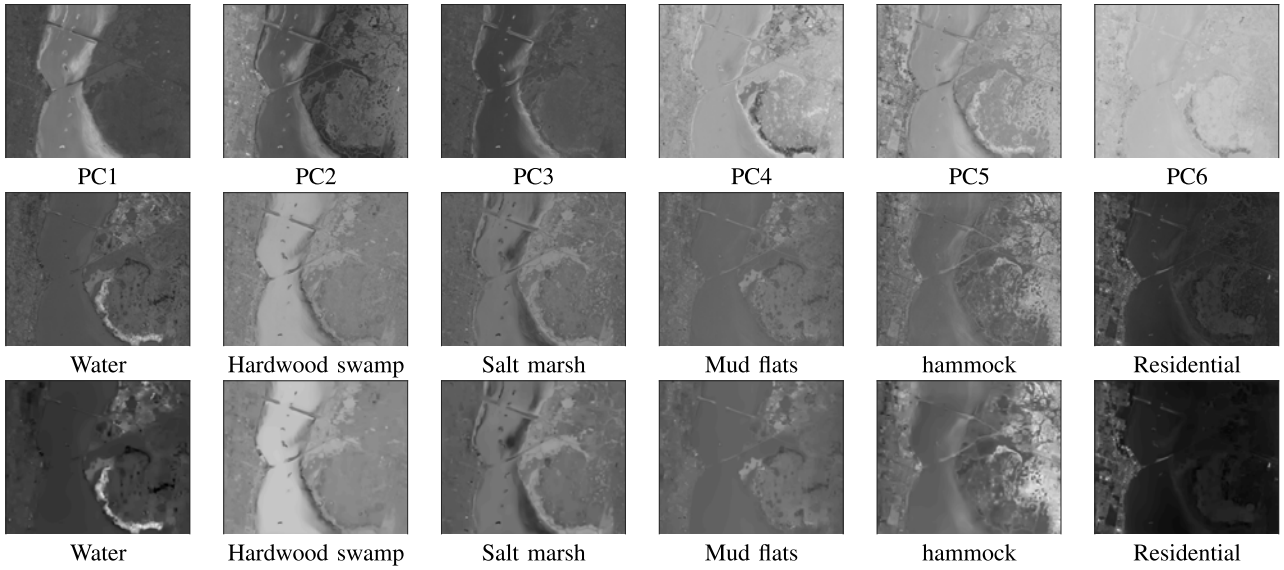


Fig. 6. Component images from three types of features from the KSC scene. The first row lists the PCA features, the second row lists six components of the abundance features, and the third row lists the SCA components corresponding to each abundance component in the second row.

TABLE I

NOISE VARIANCE ESTIMATION FOR FOUR HYPERSPECTRAL DATA SETS UNDER DIFFERENT FEATURE IMAGES, WHERE "RAW" REPRESENTS THE ORIGINAL IMAGE, "PCA" REPRESENTS THE PRINCIPAL COMPONENT FEATURES, "ABUNDANCE" REPRESENTS THE ABUNDANCE FEATURES, AND "SCA" REPRESENTS THE SUBPIXEL COMPONENT FEATURES

Data set	Raw	PCA	Abundance	SCA
IndianP	1.9231	8.45e-04	0.0013	2.58e-04
KSC	4.48e+06	9.51e-04	0.0017	5.02e-04
PaviaU	4.3822	0.0133	0.0026	4.84e-04
Washington DC	2.16e+03	7.54e-04	0.0123	6.52e-04

ground types. In turn, for the SCA features, the smoothing procedure is performed on the abundance features, simultaneously leading to the elimination of noise and the preservation of the distribution details of objects.

Similar experimental results were also obtained for the other three data sets, e.g., for the Washington DC scene, we display six component images corresponding to different features in Fig. 7. The first row lists the first six PCA features. The second row lists the subpixel abundance features, which represent the classes "roof," "street," "path," "grass," "tree," and "water." The third row lists the SCA features corresponding to each abundance component in the second row (only the *smoothness* components are listed). The corresponding estimation of noise variances is shown in Table I.

C. Classification Experiments

In this experiment, we assess the classification performance of the proposed SCA approach. Two types of classifiers are adopted: MLR and SVM. For the IndianP and PaviaU data sets, we randomly choose 20 samples from each class in the training set, thus obtaining a total of 320 and 180 labeled samples, respectively. For the KSC and Washington DC data sets, we randomly choose 10 samples from each class in the training set, thus obtaining a total of 130 and 70 labeled samples. Then, the remaining labeled samples are used for testing. For comparison, we also choose five widely used feature extraction methods, including MNF, abundance features, EMAPs [57], EPF [58], and MCA [59], [60]. Concerning the MNF features, we choose the first 20 components for the IndianP and KSC data, the first 15 components for the PaviaU data, and the first 10 components for the Washington DC data (these number of components account for most of the information present in the respective data sets). Concerning the subpixel abundance features (marked as "Abundance"), we set the number of endmembers to be equal to the number of components retained by the MNF. This parameter is a preset value that needs to be determined in advance, and we will thoroughly discuss it in Section III-D. Concerning the EMAPs, we considered four different attributes constructed on each MNF component: 1) area of the regions ($\lambda_a = [100, 500, 1000, 5000]$); 2) length of the diagonal of the box bounding the region ($\lambda_d = [10, 25, 50, 100]$); 3) moment of inertia ($\lambda_i = [0.2, 0.3, 0.4, 0.5]$); and 4) standard deviation of the gray-level values of the pixels in the regions ($\lambda_s = [20, 30, 40, 50]$). Concerning the EPF, we considered the guided-filter-based EPF, where the first PCA component image is considered as the guided image, and the size of the local window is set to 3×3. Concerning the MCA, the corresponding MNF features are directly used for morphological component decomposition, and then the content components are stacked and fed into the classifiers.

Tables II–V report the obtained overall (OAs) and average (AAs) accuracies, and the κ statistic, along with the accuracy obtained for each class under the ten conducted Monte Carlo (MC) runs for all the considered hyperspectral data sets. It can be observed that the results obtained by the EMAPs, MCA, and SCA are much better than those provided by the other feature extraction approaches. This is particularly the case for some ground types that are very difficult to distinguish

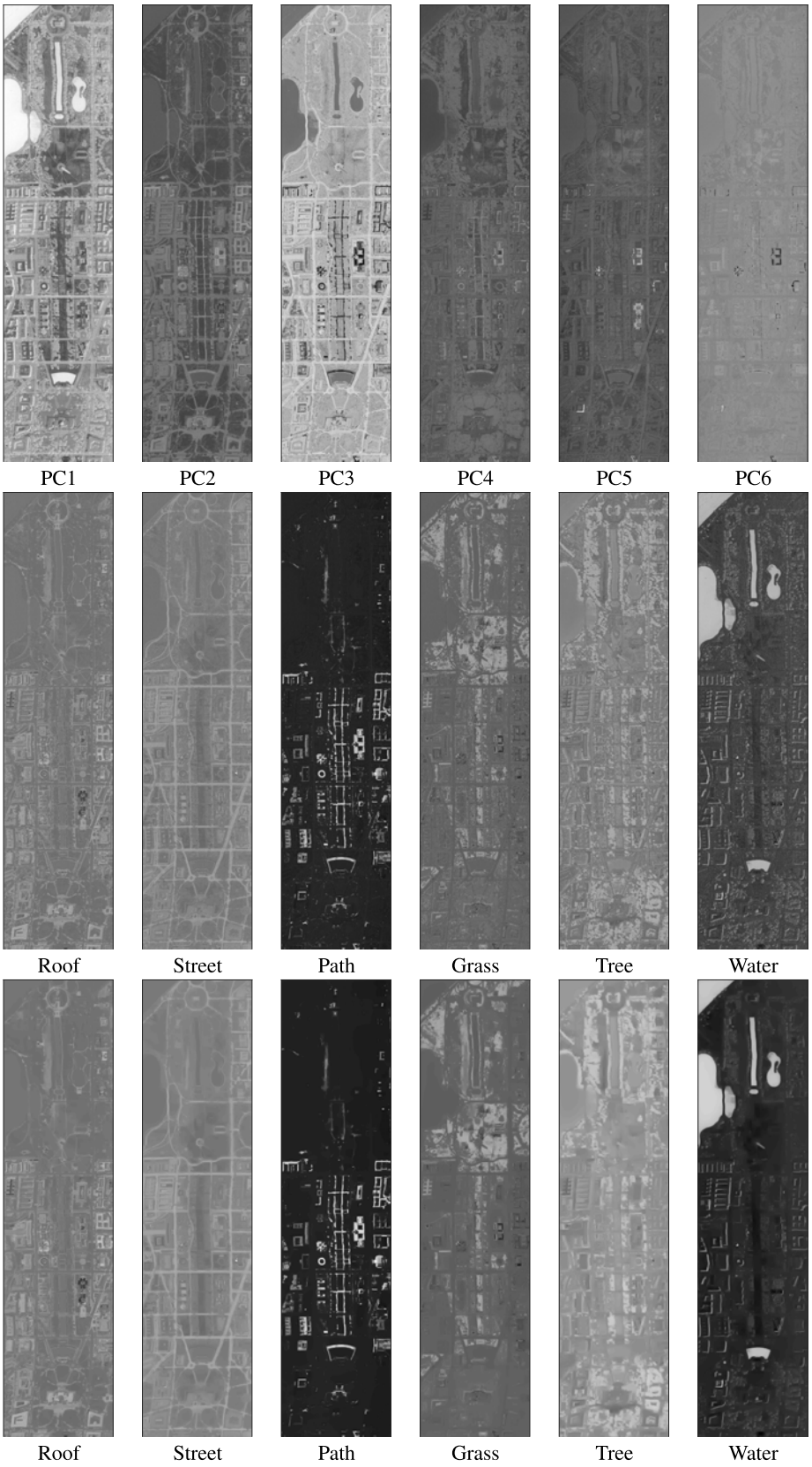


Fig. 7. Component images from three types of features from the Washington DC scene. The first row lists the PCA features, the second row lists six components of the abundance features, and the third row lists the SCA components corresponding to each abundance component in the second row.

in the MNF and abundance feature space, such as "corn" and "soybeans" in IndianP scene, "cabbage palm hammock," "scrub," and "cattail marsh" in KSC scene, and "asphalt" and "bricks" in PaviaU scene. All these examples indicate that the proposed SCA can lead to very good classification accuracies for problems with limited training samples. For illustrative

TABLE II

OAs [%], AAs [%], AND κ [%] STATISTIC AFTER TEN MC RUNS OF THE COMPARED APPROACHES FOR THE INDIANP SCENE

	MNF		Abundance		EMAPs		EPF		MCA		SCA	
	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM
alfalfa	97.83	97.39	97.83	98.26	98.48	98.48	97.6	97.22	98.26	97.17	97.61	96.96
corn-no till	75.92	74.8	77.49	77.43	90.31	87.67	76.37	77.88	91.1	88.71	89.83	91.19
corn-min till	67.13	59.9	69.07	55.1	92.06	84.9	68.91	71.62	90.96	85.95	89.98	85.87
corn	83.5	90.84	82.95	89.45	95.99	96.46	77.5	76.57	97.3	99.54	95.86	98.69
grass/pasture	93.42	95.01	91.99	92.71	94.02	95.34	95.06	94.09	96.67	97.64	95.34	96.98
grass/trees	9733	95.89	96.29	94.9	99.38	97.81	95.68	91.95	98.44	99.45	98.97	99.01
grass/pasture-mowed	100	100	100	100	100	100	94.72	89.08	100	100	100	100
hay-windrowed	97.43	97.24	97.76	98.08	100	100	97.87	90.72	100	99.92	99.96	99.96
oats	100	100	100	100	100	100	96.74	94.48	100	100	100	100
soybeans-no till	79.35	72.17	81.5	74.55	91.41	87.88	75.05	73.58	92.96	89.1	93.29	92.12
soybeans-min till	62.66	57.5	65.24	61.56	87.85	83.8	69.57	68.58	84.48	81.43	83.13	81.19
soybeans-clean till	87.39	90.25	87.03	86.1	96.1	93.61	79.6	80.55	94.49	95.51	95.43	96.04
wheat	99.76	99.66	99.66	99.61	99.76	99.61	99.47	99.32	100	99.76	100	100
woods	87.42	86.01	88.68	87.86	95.76	95.08	92.79	92.23	95.84	93.9	95.79	93.9
bldg-grass-tree-drives	85.03	85.16	85.44	84.84	98.37	97.75	73.73	71.38	98.89	98.65	98.91	98.73
stone-steel towers	95.38	94.73	93.66	93.44	100	99.89	90.11	94.79	99.89	99.89	99.89	97.74
OA	79.17	76.64	80.37	77.61	**93.04**	90.5	80.23	80	92.43	90.56	**91.87**	91.07
AA	88.09	87.28	88.41	87.12	96.22	94.89	86.3	85.25	96.21	95.41	95.87	95.52
κ	76.59	73.74	77.89	74.81	92.11	89.23	77.9	77.67	91.43	89.32	90.79	89.89

TABLE III

OAs [%], AAs [%], AND κ [%] STATISTIC AFTER TEN MC RUNS OF THE COMPARED APPROACHES FOR THE KSC SCENE

	MNF		Abundance		EMAPs		EPF		MCA		SCA	
	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM
oak/broadleaf hammock	79.54	86.29	73.96	76.89	92.52	91.06	81.78	85.09	99.55	99.61	98.45	98.02
cabbage palm/oak hammock	87.57	91.48	85.97	86.58	98.6	97.86	80.75	87.65	96.79	95.1	97.9	97.12
slash pine	80.31	80.39	76.25	78.52	93.09	93.79	69.74	85.76	98.79	97.34	96.68	94.8
cabbage palm hammock	66.59	63.93	68.97	62.06	85.36	82.42	56.57	70.68	97.66	92.54	97.58	94.56
scrub	74.22	73.48	72.17	75.9	93.17	87.64	70.53	73.15	97.27	94.66	96.65	94.84
cattail marsh	75.41	75.2	69.04	69.74	95.33	88.78	66.4	65.25	99.52	99.34	99.3	98.12
mud flats	87.43	86.1	89.05	83.9	98.19	96.86	79.74	88.81	100	99.9	100	99.33
graminoid marsh	89.91	91.37	90.21	88.93	95.8	95.01	73.84	75.78	98.28	98.24	97.82	97.56
sparina marsh	92.94	93.94	94.27	93.35	99.15	99.12	88.71	90.49	99.9	99.96	100	99.96
water	94.95	92.35	91.61	93.86	98.99	96.36	80.38	81.27	99.75	96.78	98.39	98.32
willow swamp	89.28	89.69	89.16	89.4	93.1	92.41	90.88	89.52	95.54	95.63	94.68	93.77
salt marsh	94.16	96.5	93.6	95.09	96.6	96.5	73.58	78.29	98.41	98.67	98.35	98.15
hardwood swamp	100	100	100	100	100	100	96.15	96.22	100	100	100	100
OA	88.36	89.63	86.91	87.33	95.95	94.77	82.54	84.77	**98.82**	98.15	**98.37**	97.79
AA	85.56	86.21	84.17	84.17	95.38	93.68	78.93	81.71	98.57	97.52	98.14	97.28
κ	87.08	88.47	85.48	85.93	95.49	94.18	80.34	82.8	98.69	97.94	98.18	97.54

TABLE IV

OAs [%], AAs [%], AND κ [%] STATISTIC AFTER TEN MC RUNS OF THE COMPARED APPROACHES FOR THE PAVIAU SCENE

	MNF		Abundance		EMAPs		EPF		MCA		SCA	
	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM
asphalt	69.94	72.96	71.75	71.78	93.35	92.08	80.09	77.62	87.64	88.51	93.06	93.17
meadows	79.32	74.47	80.42	75.5	94.2	91.59	88.05	85	92.65	89.71	92.31	90.73
gravel	70.59	74.56	73.08	75.27	91.13	87.89	77.32	76.72	90.36	79.23	92.12	84.18
trees	85.15	89.63	79.51	90.02	92.63	93.15	93.41	93.31	92.53	89.76	89.89	90.34
metal sheets	98.25	98.83	96.22	98.28	99.79	99.81	97.44	97.46	99.23	98.72	99.16	98.69
bare soil	85.97	86.07	86.16	86.21	98.21	97.98	81.7	81.4	98.36	99.69	98.59	98.17
bitumen	88.24	80.21	89.73	80.4	98.94	98.01	85.8	81.01	97.92	95.7	99.4	98.75
bricks	71.3	71.57	71.58	70.04	92.08	86.7	79.25	76.92	87.81	87.09	91.52	88.08
shadows	99.16	99.87	98.99	99.96	99.78	100	98.2	96.76	96.19	97.44	99.93	99.98
OA	79.26	77.95	79.76	78.15	**94.54**	92.57	85.48	83.54	92.46	90.6	**93.52**	92.11
AA	83.1	83.13	83.05	83.05	95.57	94.13	86.8	85.13	93.63	91.76	95.11	93.57
κ	73.6	72.08	74.15	72.27	92.86	90.33	81.52	79.39	90.19	87.8	91.56	89.73

purposes, Figs. 8–11, respectively, present the classification maps obtained for the considered images by different feature extraction approaches. As we can observe, noise disturbance and misclassification phenomena still exist in the classification maps obtained using MNF and abundance features, e.g., in the KSC classification map, the "mud flats" is misclassified as "slash pine" in the MNF features and also is misclassified into "oak/broadleaf hammock" in the Abundance features. Regarding the EMAPs, EPF, and MCA, although the noise is suppressed remarkably, the distribution of ground objects shows certain oversmoothing, leading to the loss of some structural information. Finally, for the SCA classification map, the spatial distribution of ground objects exhibits smoothness and more prominent structural details, thus leading to a better classification performance. This is due to the fact that the SCA features can make full use of fractional abundance information, which can reflect the subpixel spatial distribution details of specific ground objects.

TABLE V

OAs [%], AAs [%], AND κ [%] STATISTIC AFTER TEN MC RUNS OF THE COMPARED APPROACHES FOR THE WASHINGTON DC SCENE

	MNF		Abundance		EMAPs		EPF		MCA		SCA	
	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM	MLR	SVM
roofs	86.51	89.73	86.22	89.45	91.96	91.69	80.53	84.84	94.29	92.61	93.29	92.52
street	92.62	95.75	91.92	95.5	97.69	98.73	82.6	82.18	99.95	99.78	98.41	99.06
path	96.23	96.51	98	99.2	100	100	95.78	96.2	100	100	99.09	100
grass	97.43	98.73	97.49	96.42	99.99	99.97	72.86	79.34	99.68	100	98.2	98.72
trees	95.85	95.21	96.3	96.12	99.58	99.53	71.95	82.17	99.95	100	99.51	99.56
water	96.79	96.05	96.41	96.41	99.61	98.95	97.56	97.56	99.85	99.78	99.81	99.49
shadow	95.88	96.8	94.85	95.57	99.59	99.38	68.37	84.46	100	100	100	99.9
OA	91.78	93.65	91.61	93.1	95.98	95.8	79.22	85.32	**97.19**	96.45	**96.23**	96
AA	94.47	95.54	94.45	95.53	98.35	98.32	81.38	86.68	99.1	98.88	98.33	98.46
κ	88.55	91.06	88.32	90.3	94.32	94.06	72.78	79.48	96	94.97	94.65	94.32

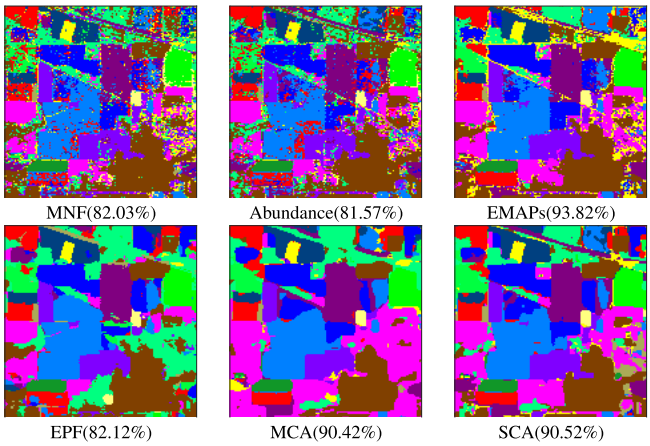


Fig. 8. Classification maps (along with the OAs) obtained by different kinds of image features for the IndianP data set. These maps correspond to a single experiment in Table II. The classifier that used here is the MLR.

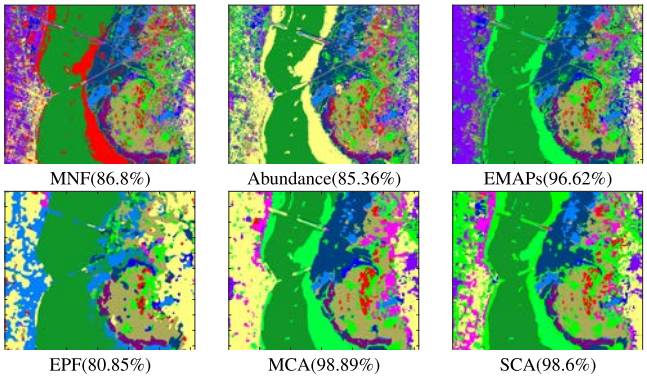


Fig. 9. Classification maps (along with the OAs) obtained by different kinds of image features for the KSC data set. These maps correspond to a single experiment in Table III. The classifier that used here is the MLR.

D. Sensitivity to Parameters

In this section, we analyze the sensitivity to parameters of the proposed SCA approach. First, concerning the size of the regional search window in the endmember identification step, we assess the clustering effectiveness and classification accuracy, respectively, as detailed in Section III-D1. Second, as described in Section III-C, the number of endmembers is a preset parameter which needs to be determined in advance. In our case, this parameter is equal to the number of feature components in the SCA approach. We perform a quantitative assessment on this parameter, as detailed in Section III-D2.

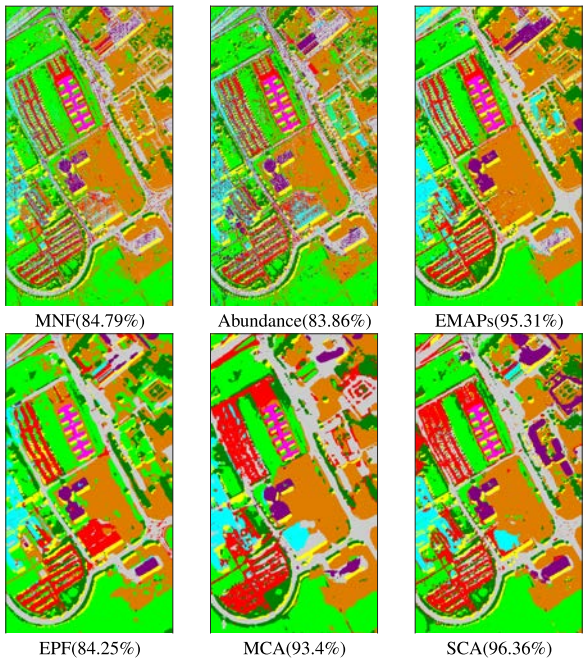


Fig. 10. Classification maps (along with the OAs) obtained by different kinds of image features for the PaviaU data set. These maps correspond to a single experiment in Table IV. The classifier that used here is the MLR.

Finally, concerning the regularization parameters in the formula of subpixel component decomposition, we focus on the influence of these parameters on classification accuracy, which is detailed in Section III-D3. It should be noted that, when one parameter is tested in our experiments below, the other parameters are fixed and set to the same values reported in Section III-C.

1) Size of Regional Search Window: In order to analyze the size of the regional search window in the endmember identification step, we set this parameter to a value of 2, 3, and 4 times the initial partition size, respectively. After finishing the clustering segmentation, we assessed the spatial homogeneity of the obtained clustering partitions by using two SNPI indices, as shown in Table VI. Here, the "AvgPI" is calculated by the correlation coefficient of each pixel in a partition to their average value, as defined in [50]. This "AvgPI" index can quantitatively measure the average variance of all pixels in a partition to their average value, which is in the range $[-1, 1]$. A larger "AvgPI" denotes better spatial homogeneity. The "SVDPI" is calculated by the ratio of the

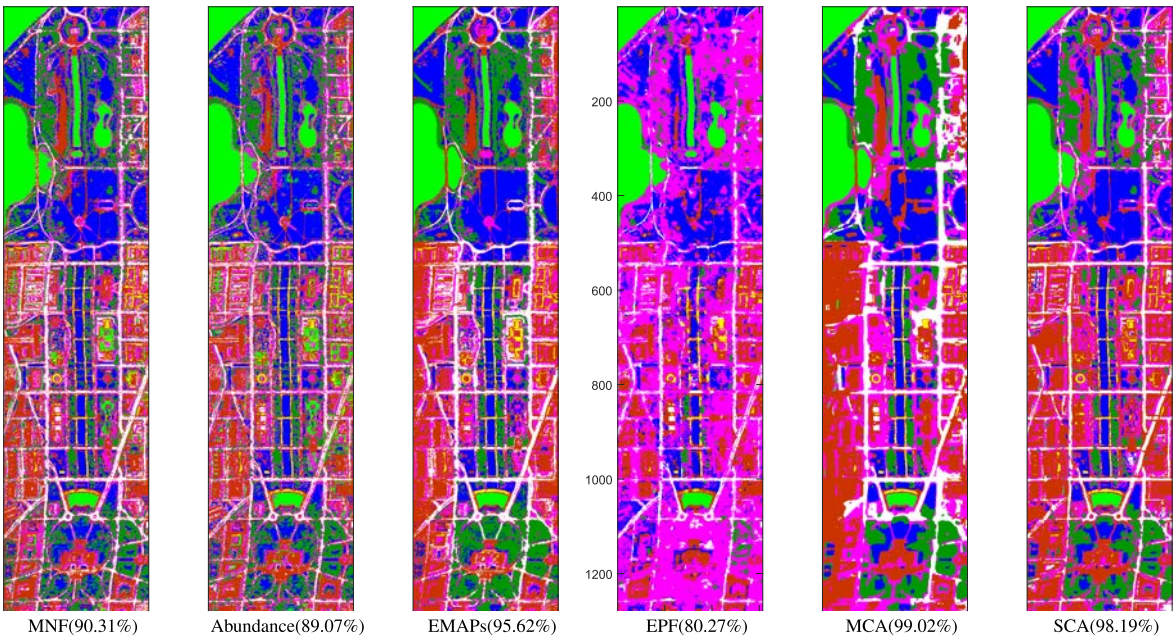


Fig. 11. Classification maps (along with the OAs) obtained by different kinds of image features for the Washington DC data set. These maps correspond to one single experiment in Table V. The classifier that used here is the MLR.

TABLE VI

ASSESSMENT OF THE SNPI INDICES AND CLASSIFICATION ACCURACIES WITH DIFFERENT SIZES OF THE REGIONAL SEARCH WINDOW. FOR THE TWO SNPI INDICES, WE GIVE THE MINIMUM AND MAXIMUM VALUES OBTAINED FOR EACH WINDOW SIZE

		IndianP	KSC	PaviaU	Washington DC
$AvgPI$	2	[0.995,0.999]	[0.38,0.99]	[0.942,0.999]	[0.149,0.998]
	3	[0.993,0.999]	[0.42,0.982]	[0.941,0.999]	[0.149,0.998]
	4	[0.994,0.999]	[0.437,0.981]	[0.941,0.999]	[0.149,0.998]
$SVDPI$	2	[0.826,0.923]	[0.237,0.753]	[0.58,0.847]	[0.193,0.867]
	3	[0.813,0.925]	[0.229,0.659]	[0.578,0.847]	[0.193,0.867]
	4	[0.824,0.931]	[0.236,0.692]	[0.578,0.847]	[0.193,0.867]
OA	2	91.87	98.37	93.52	96.23
	3	91.99	98.11	94.17	94.9
	4	91.70	97.9	93.60	96.28

maximum singular value to the sum of all singular values by performing the SVD on a partition, which represents how dominant is the first feature component. A larger "SVDPI" denotes better spatial purity in this partition.

It can be observed from Table VI that the homogeneity of the obtained partitions is not sensitive to the size of the search window. In Table VI, we also provide the classification accuracies with regards to different sizes of the search window, which are insensitive to the window size. However, we emphasize that the computational burden becomes heavier when the size of search window used to implement the searching strategy becomes larger. As a result, we recommend to use two times the size of the initial partition in practice.

2) Number of Endmembers: In this experiment, we report the plots of classification accuracy based on different number of endmembers by using KSC and Washington DC data sets. For the KSC image data, we set the number of endmembers to a value between 10 and 30, which cover the number of classes in this scene. For the Washington DC image data, we set the number of endmembers to a value between 5 and 20

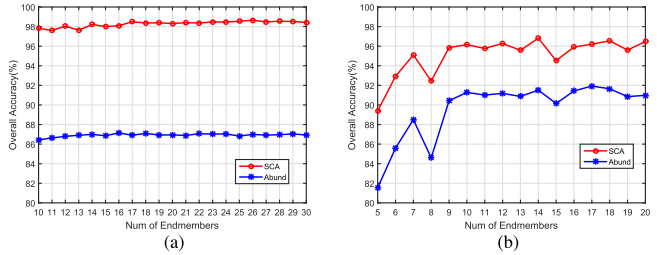


Fig. 12. Plots of classification accuracy (along with the OAs) obtained by different numbers of endmembers for two hyperspectral data sets. (a) KSC. (b) Washington DC.

for the same reason. Once the number of endmembers is given, we perform K-means clustering on the set of candidate endmembers so as to obtain the final endmember signatures. Then, these endmember signatures are used for abundance estimation and further subpixel component decomposition. Here, the adopted classifier is the MLR classifier, and the plots of classification accuracy by using abundance and SCA approaches are given in Fig. 12.

It can be observed from Fig. 12 that, as long as the number of endmembers reaches a minimum threshold, the classification accuracy is not particularly sensitive to this parameter. Here, the threshold can be set to 17 for the KSC data and to 9 for the Washington DC data. It is expected that the obtained endmember signatures cover the majority of ground objects present in the scene.

3) Regularization Parameters: In this experiment, we perform a detailed analysis of the regularization parameters λ_1 and λ_2 involved in the subpixel component decomposition formula (7). Here, we consider the IndianP and PaviaU data sets, and the values of λ_1 and λ_2 are all set to [1,1e−1,1e−2,1e−3,1e−4,1e−5,1e−6,1e−7]. Then, subpixel component decomposition is conducted by using different

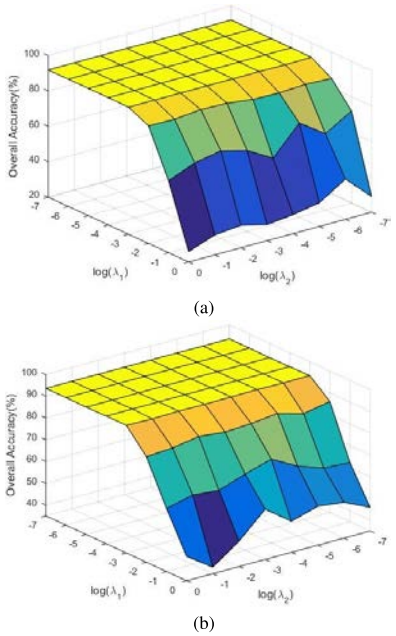


Fig. 13. OA as a function of λ_1 and λ_2 for IndianP and PaviaU data sets. (a) IndianP. (b) PaviaU.

combinations of λ_1 and λ_2, and the obtained smoothness components are stacked and fed for classification purposes. Fig. 13 shows the OA as a function of parameters λ_1 and λ_2 obtained under the ten conducted MC runs. As it can be observed, the classification performance is almost insensitive to λ_2. Regarding λ_1, when the value 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.

E. Computational Complexity Analysis

In this section, we perform an analysis of the computational complexity of our proposed SCA approach. The SCA approach includes three main steps: 1) endmember identification; 2) abundance estimation; and 3) subpixel component decomposition, and these steps are performed in step-by-step fashion. After obtaining the subpixel component features, we feed them to a classifier for land-cover classification purposes. For the endmember identification step, we adopted the RCSPP strategy to pick out the candidate pixels. As described in [44], the RCSPP strategy has a linear computational complexity which is linear with regards to the number of image pixels. Then, for the extraction of candidate endmembers, a PCA projection technique is performed on all partitions. In each partition, the computational complexity is $O(n_i^3)$ with n_i being the number of pixels in the ith partition. If we assume that the number of partitions is c, and $n_i \approx n/c$, where n is the number of all image pixels, the computational complexity of the endmember candidates extraction step is $O(c \times (n_i^3)) \approx O(n^3/c^2)$. Finally, the K-means clustering is performed, which has a computational complexity of $O(m \times k \times t)$, where m is the number of candidates, k is the number of clusters, and t is the number of iterations. For the abundance estimation step, we adopted an MTMF technique to generate the abundance information of each endmember signature. In this step, the main computational complexity lies in two projection processes, where the first one refers to the original endmember vector that is projected onto the inverse covariance of the MNF transformed image data, and the second one refers to the entire MNF-transformed image cube, which is projected onto the matched filtering vectors. These two projection operations involve a multiplication of matrices and vectors, so the computational complexity is $O(n \times b)$, where n is the number of pixels in the image, and b is the number of spectral bands. For the subpixel component decomposition step, which is the most time-consuming step in our SCA approach, we solved the optimization equation (8) via the SUnSAL algorithm [52], and the computational complexity of SUnSAL is $O(n^2)$ per iteration.

IV. Conclusion

In this paper, we developed a new subpixel-level feature extraction approach, called SCA, for advanced land-cover classification of remotely sensed hyperspectral images. Our newly proposed SCA approach contains a two-layer feature learning framework, where a preliminary feature representation is derived from the subpixel abundance maps. Then, we further perform feature extraction on the aforementioned abundance features and obtain a subpixel attribute component feature representation. Specifically, we exploit different attribute features contained in these abundance feature images and decompose the abundance feature images into a pair of attribute components (*smoothness* and *texture*). Our experimental results, conducted using four real hyperspectral images, indicate that the proposed SCA approach can obtain features with explicit physical meaning, clear spatial distribution details, and better noise robustness, leading to state-of-the-art classification accuracy, regardless of the presence of pure signatures in the original image data. In the future, we will develop techniques for efficient implementation of the proposed SCA approach in high-performance computing architectures.

Acknowledgment

The authors would like to thank the editors and the anonymous reviewers for their outstanding comments and suggestions for improvement, which greatly helped them to improve the technical quality and presentation of this paper.

References

[1] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, "Advances in hyperspectral image classification: Earth monitoring with statistical learning methods," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 45–54, Jan. 2014.

[2] K. Karalas, G. Tsagkatakis, M. Zervakis, and P. Tsakalides, "Land classification using remotely sensed data: Going multilabel," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3548–3563, Jun. 2016.

[3] J. Cihlar, "Land cover mapping of large areas from satellites: Status and research priorities," *Int. J. Remote Sens.*, vol. 21, nos. 6–7, pp. 1093–1114, 2000.

[4] S. Prasad, W. Liao, M. He, and J. Chanussot, "Foreword to the special issue on hyperspectral remote sensing and imaging spectroscopy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1019–1021, Apr. 2018.

[5] S. Jia, K. Wu, J. Zhu, and X. Jia, "Spectral–spatial Gabor surface feature fusion approach for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1142–1154, Feb. 2019.

[6] M. E. Paoletti, J. M. Haut, R. Fernandez-Beltran, J. Plaza, A. J. Plaza, and F. Pla, "Deep pyramidal residual networks for spectral–spatial hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 740–754, Feb. 2019.

[7] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. GRS-26, no. 1, pp. 65–74, Jan. 1988.

[8] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[9] I. Dópido, J. Li, P. Gamba, and A. Plaza, "A new hybrid strategy combining semisupervised classification and unmixing of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 8, pp. 3619–3629, Aug. 2014.

[10] L. He, J. Li, C. Liu, and S. Li, "Recent advances on spectral-spatial hyperspectral image classification: An overview and new guidelines," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1579–1597, Mar. 2017.

[11] B. Pan, Z. Shi, and X. Xu, "Multiobjective-based sparse representation classifier for hyperspectral imagery using limited samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 239–249, Jan. 2019.

[12] L. Ma, M. Li, X. Ma, L. Cheng, P. Du, and Y. Liu, "A review of supervised object-based land-cover image classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 277–293, Aug. 2017.

[13] Q. Wang, Z. Meng, and X. Li, "Locality adaptive discriminant analysis for spectral–spatial classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 2077–2081, Nov. 2017.

[14] Q. Wang, F. Zhang, and X. Li, "Optimal clustering framework for hyperspectral band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5910–5922, Oct. 2018.

[15] B. Luo and J. Chanussot, "Unsupervised classification of hyperspectral images by using linear unmixing algorithm," in *Proc. 16th IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 2877–2880.

[16] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.

[17] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[18] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[19] B. Luo, J.-F. Aujol, and Y. Gousseau, "Local scale measure from the topographic map and application to remote sensing images," *Multiscale Model. Simul.*, vol. 8, no. 1, pp. 1–29, 2009.

[20] I. Dopido, M. Zortea, A. Villa, A. Plaza, and P. Gamba, "Unmixing prior to supervised classification of remotely sensed hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 760–764, Jul. 2011.

[21] I. Dopido, A. Villa, A. Plaza, and P. Gamba, "A quantitative and comparative assessment of unmixing-based feature extraction techniques for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 421–435, Apr. 2012.

[22] J. T. Mundt, A. County, W. Pest, D. R. Streutker, and N. F. Glenn, "Partial unmixing of hyperspectral imagery: Theory and methods," in *Proc. ASPRS Annu. Conf.*, Jan. 2007.

[23] J. W. Boardman and F. A. Kruse, "Analysis of imaging spectrometer data using N-dimensional geometry and a mixture-tuned matched filtering approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4138–4152, Nov. 2011.

[24] J. W. Boardman, "Leveraging the high dimensionality of AVIRIS data for improved sub-pixel target unmixing and rejection of false positives: Mixture tuned matched filtering," in *Proc. Summaries 7th JPL Airborne Geosci. Workshop JPL*, 1998, vol. 97, no. 1, pp. 55–56.

[25] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2679–2693, Aug. 2009.

[26] J. Li, I. Dópido, P. Gamba, and A. Plaza, "Complementarity of discriminative classifiers and spectral unmixing techniques for the interpretation of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2899–2912, May 2015.

[27] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sens. Environ.*, vol. 115, no. 7, pp. 1603–1616, 2011.

[28] A. Zare and K. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.

[29] D. T. Anderson and A. Zare, "Spectral unmixing cluster validity index for multiple sets of endmembers," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 5, no. 4, pp. 1282–1295, Aug. 2012.

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

[31] C. Shi and L. Wang, "Incorporating spatial information in spectral unmixing: A review," *Remote Sens. Environ.*, vol. 149, pp. 70–87, Jun. 2014.

[32] N. Jamshidpour, A. Safari, and S. Homayouni, "Spectral–spatial semisupervised hyperspectral classification using adaptive neighborhood," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 9, pp. 4183–4197, Sep. 2017.

[33] S. Patra, K. Bhardwaj, and L. Bruzzone, "A spectral-spatial multicriteria active learning technique for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 12, pp. 5213–5227, Dec. 2017.

[34] Q. Wang, S. Liu, J. Chanussot, and X. Li, "Scene classification with recurrent attention of VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1155–1167, Feb. 2019.

[35] G. Martín and A. Plaza, "Region-based spatial preprocessing for endmember extraction and spectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 745–749, Jul. 2011.

[36] G. Martin and A. Plaza, "Spatial-spectral preprocessing prior to endmember identification and unmixing of remotely sensed hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 380–395, Apr. 2012.

[37] F. Kowkabi, H. Ghassemian, and A. Keshavarz, "A fast spatial–spectral preprocessing module for hyperspectral endmember extraction," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 6, pp. 782–786, Jun. 2016.

[38] F. Kowkabi, H. Ghassemian, and A. Keshavarz, "Enhancing hyperspectral endmember extraction using clustering and oversegmentation-based preprocessing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2400–2413, Jun. 2016.

[39] G. H. Ball and D. J. Hall, "ISODATA, a novel method of data analysis and pattern classification," DTIC, Fort Belvoir, VA, USA, Tech. Rep., 1965.

[40] J. C. Tilton, "Analysis of hierarchically related image segmentations," in *Proc. IEEE Workshop Adv. Techn. Anal. Remotely Sensed Data*, Oct. 2003, pp. 60–69.

[41] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.

[42] A. Levinshtein, A. Stere, K. N. Kutulakos, D. J. Fleet, S. J. Dickinson, and K. Siddiqi, "TurboPixels: Fast superpixels using geometric flows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2290–2297, Dec. 2009.

[43] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[44] X. Xu, J. Li, C. Wu, and A. Plaza, "Regional clustering-based spatial preprocessing for hyperspectral unmixing," *Remote Sens. Environ.*, vol. 204, pp. 333–346, Jan. 2018.

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

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

[47] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2004.

[48] 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.

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

[50] S. Mei, M. He, Z. Wang, and D. Feng, "Spatial purity based endmember extraction for spectral mixture analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3434–3445, Sep. 2010.
[51] 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.
[52] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Jun. 2010, pp. 1–4.
[53] D. Böhning, "Multinomial logistic regression algorithm," *Ann. Inst. Stat. Math.*, vol. 44, no. 1, pp. 197–200, 1992.
[54] B. Krishnapuram, L. Carin, M. A. T. Figueiredo, and A. J. Hartemink, "Sparse multinomial logistic regression: Fast algorithms and generalization bounds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 6, pp. 957–968, Jun. 2005.
[55] J. Platt, "Probabilities for support vector machines," in *Advances in Large Margin Classifiers*, Cambridge, MA, USA: MIT Press, 2000.
[56] D. Garcia, "Robust smoothing of gridded data in one and higher dimensions with missing values," *Comput. Statist. Data Anal.*, vol. 54, no. 4, pp. 1167–1178, Apr. 2010.
[57] 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.
[58] X. Kang, S. Li, and J. A. Benediktsson, "Spectral–spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.
[59] 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.
[60] X. Xu, J. Li, X. Huang, M. D. Mura, and A. Plaza, "Multiple morphological component analysis based decomposition for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 3083–3102, May 2016.

Xiang Xu (S'14–M'18) received the B.S., M.S., and Ph.D. degrees from Sun Yat-sen University, Guangzhou, China, in 1999, 2002, and 2018, respectively.

In 2004, he joined the University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan, China, where he is currently an Associate Professor. His current research interests include hyperspectral image classification, hyperspectral unmixing, pattern recognition, and machine learning.

Jun Li (SM'16) received the degree in geographical information systems from Hunan Normal University, Changsha, China, in 2004, the M.Sc. degree in remote sensing and photogrammetry from Peking University, Beijing, China, in 2007, and the Ph.D. degree in electrical and computer engineering from the Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal, in 2011.

From 2011 to 2012, she was a Post-Doctoral Researcher with the Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain. From 2014 to 2018, she was a Professor with the School of Geography and Planning, Sun Yat-sen University, Guangzhou, China. She is currently with the College of Electrical and Information Engineering, Hunan University, Changsha. Since then, she has obtained several prestigious funding grants at the national and international level. She has published a total of 69 journal citation report (JCR) papers, 48 conference international conference papers, and one book chapter. She has received a significant number of citations to her published works, with several papers distinguished as Highly Cited Papers in Thomson Reuters' Web of Science-Essential Science Indicators. Her main research interests comprise remotely sensed hyperspectral image analysis, signal processing, supervised/semisupervised learning, and active learning.

Dr. Li's students have also obtained important distinctions and awards at international conferences and symposia. She has served as the Guest Editor for a Special Issue in the prestigious PROCEEDINGS OF THE IEEE journal. She has also served as the Guest Editor of a Special Issue in the prestigious ISPRS *Journal of Photogrammetry and Remote Sensing* journal. She has been serving as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING since 2014.

Shutao Li (M'07–SM'15–F'19) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Hunan University, Changsha, China, in 1995, 1997, and 2001, respectively.

In 2001, he joined the College of Electrical and Information Engineering, Hunan University. In 2001, he was also a Research Associate with the Department of Computer Science, The Hong Kong University of Science and Technology, Hong Kong. From 2002 to 2003, he was a Post-Doctoral Fellow with the Royal Holloway College, University of London, Surrey, U.K. In 2005, he was a Visiting Professor with the Department of Computer Science, The Hong Kong University of Science and Technology, Hong Kong. He is currently a Full Professor with the College of Electrical and Information Engineering, Hunan University. He is also a Chang-Jiang Scholar Professor appointed by the Ministry of Education of China. In 2013, he was granted the National Science Fund for Distinguished Young Scholars in China. He has authored or co-authored over 180 refereed papers. His research interests include compressive sensing, sparse representation, image processing, and pattern recognition.

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

Antonio Plaza (M'05–SM'07–F'15) received the M.Sc. and Ph.D. degrees in computer engineering from the University of Extremadura, Cáceres, Spain, in 1999 and 2002, respectively.

He is currently the Head of the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura. He has authored more than 600 publications, including 200 JCR journal papers (145 in IEEE journals), 23 book chapters, and 285 peer-reviewed conference proceeding papers. 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 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 recognition of Best Reviewers of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2009 and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2010, for which he served as an Associate Editor from 2007 to 2012. He was a recipient of the Best Column Award of the IEEE *Signal Processing Magazine* in 2015, the 2013 Best Paper Award of the JSTARS journal, and the most highly cited paper from 2005 to 2010, in the *Journal of Parallel and Distributed Computing*. He received the Best Paper Awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. He served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) from 2011 to 2012, and as the President of the Spanish Chapter of the IEEE GRSS from 2012 to 2016. He has reviewed more than 500 manuscripts for over 50 different journals. He has guest edited 10 special issues on hyperspectral remote sensing for different journals. He is also an Associate Editor for the IEEE ACCESS, and was a member of the Editorial Board of the IEEE *Geoscience and Remote Sensing Newsletter* from 2011 to 2012 and the IEEE *Geoscience and Remote Sensing Magazine* in 2013. He previously served as the Editor-in-Chief for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.