

# Abundance-Indicated Subspace for Hyperspectral Classification With Limited Training Samples

Shuyuan Xu<sup>ID</sup>, Jun Li<sup>ID</sup>, Senior Member, IEEE, Mahdi Khodadadzadeh<sup>ID</sup>, Andrea Marinoni<sup>ID</sup>, Senior Member, IEEE, Paolo Gamba<sup>ID</sup>, Fellow, IEEE, and Bo Li<sup>ID</sup>

**Abstract**—The imbalance between the (often limited) number of available training samples and the high data dimensionality, together with the presence of mixed pixels, often complicates the classification of remotely sensed hyperspectral data. In this paper, we tackle these problems by developing a new method that combines spectral unmixing and classification techniques in a subspace-based approach. The proposed method is developed under the assumption that the spectral signature of a land cover class is associated with a given set of pure spectral signatures (called endmembers in spectral unmixing terminology), which define a low-dimensional subspace with clear physical meaning. We aim to exploit this relationship to learn the class-dependent subspaces and integrate them with a multinomial logistic regression procedure. Experiments on synthetic datasets and real hyperspectral images show that our method is able to obtain competitive performances in comparison with other approaches, particularly when very limited training sets are available.

**Index Terms**—Hyperspectral image classification, mixed pixels, multinomial logistic regression (MLR), spectral unmixing, subspace learning.

Manuscript received July 6, 2018; revised October 31, 2018, December 5, 2018, and January 25, 2019; accepted February 19, 2019. Date of publication March 25, 2019; date of current version April 17, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61771149, in part by the National Key Research and Development Program of China under Grant 2017YFB0502900, in part by the Guangdong Provincial Natural Science Foundation under Grant 2016A030313254, and in part by CIRFA under RCN Grant 237906. (*Corresponding author: Jun Li*)

S. Xu is with the Guangdong Provincial Key Laboratory of Urbanization and Geo-Simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China (e-mail: xushuy@mail3.sysu.edu.cn).

J. Li is with the Key Laboratory of Visual Perception and Artificial Intelligence of Hunan Province and the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: ljun48@mail.sysu.edu.cn).

M. Khodadadzadeh is with the Helmholtz Institute Freiberg for Resource Technology, Helmholtz-Zentrum Dresden-Rossendorf 09599, Freiberg, Germany (e-mail: m.khodadadzadeh@hzdr.de).

A. Marinoni is with the Earth Observation Group, Centre for Integrated Remote Sensing and Forecasting for Arctic Operations, Department of Physics and Technology, UiT-The Arctic University of Norway, 9037 Tromsø, Norway (e-mail: andrea.marinoni@uit.no).

P. Gamba is with the Telecommunications and Remote Sensing Laboratory, Department of Electrical, Computer and Biomedical Engineering, University of Pavia, I-27100 Pavia, Italy (e-mail: paolo.gamba@unipv.it).

B. Li is with the Beijing Key Laboratory of Digital Media, School of Computer Science and Engineering, and the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China (e-mail: boli@buaa.edu.cn).

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

Digital Object Identifier 10.1109/JSTARS.2019.2903940

## I. INTRODUCTION

HYPERSPECTRAL imaging is an active research area due to its advanced capabilities for Earth surface analysis [1]. It provides detailed and physically meaningful spectral information within a narrow continuous wavelength range. This high spectral resolution can be exploited to develop discriminative measures to distinguish different materials [2]. The fast development of hyperspectral imaging and processing techniques has played an increasingly important role in remote sensing applications, such as defense, disaster monitoring, precision agriculture [3]–[5], etc.

Classification is an important research topic in hyperspectral data analysis [6]–[8]. Given a set of observations from a hyperspectral image, the goal of classification is to assign a unique predefined class label to each pixel. Especially, for the precise classification of the surface of the Earth, the rich spectral information contained in hyperspectral images offers great advantages for material discrimination and thus facilitates good classification performance [9]. However, conducting classification on hyperspectral images is often challenging [10]–[12]. The first major problem is the so-called curse of dimensionality [13]. This effect is caused by the unbalance between the high dimensionality of the data and the small number of available labeled samples, which leads to ill-posed problems that are very hard to solve [1]. Another important issue is the characterization of mixed pixels, which are inherently contained in the data [14], [15]. As each pixel in the hyperspectral image is determined by the combination of different substances, it is difficult to categorize the samples in the original feature space [16]. Moreover, the noisy components generated during the transmission stage and the efficiency problems due to the large volume of hyperspectral data also receive continuous attention for practical applications [17]–[20].

These challenges in hyperspectral image classification call for the development of advanced methods [21], [22]. In the literature, there are plenty of effective feature extraction approaches to deal with the curse of dimensionality, such as topology-based methods like manifold learning [23]–[25], statistic-based approaches like principal component analysis [26], [27] and minimum noise fraction [28], and sparse representation approaches [29]. These techniques aim at solving the small training set problem by exploiting the data structures in low-dimensional subspaces. Moreover, advanced classifiers such as support vector machines (SVMs) [30]–[32] and artificial

neural networks [33] have also been adopted due to their effectiveness in addressing high dimensionality problems. Recently, the multinomial logistic regression (MLR) has shown notable success in hyperspectral classification [34], [35]. MLR-based approaches improve the class separability in nonlinear transformative spaces and model the class posterior probabilities under the Bayesian framework. A related work in [36] tackled the nonlinear (kernel) features under the modality of MLR and achieved satisfactory classification results. Another study in [37] proposed the concept of class-dependent subspace by preserving the class-dependent structure in low-dimensional space to achieve an accurate characterization for mixed pixels. A further improvement is achieved in [38] by expressing the spectral information via features defined by a linear mixing model. These methods emphasize the idea of using an inherent low-dimensional subspace to solve the problem of mixed pixel characterization, extending the well-established MLR classifier. However, these MLR-based methods generally define low-dimensional subspaces by uncorrelated components obtained by statistical criteria, without taking into account the physical meaning of the subspaces.

More recently, spectral unmixing techniques have been exploited to address the mixed pixel problem for classification purposes. As a straightforward strategy to deal with mixed pixels, spectral unmixing estimates the corresponding proportions of endmembers for each sample [39]. Another strategy that fully exploits the concept of mixed pixels is to consider unmixing as a feature extraction technique. In this case, the spectral unmixing and classification algorithms can be exploited separately [40]. As revealed in recent studies, spectral unmixing is a useful source of information for classification purposes [41], [42]. There are also pieces of evidence that unmixing prior to classification offers an effective solution for dealing with the small training set problem [43]–[45]. In [46], a quantitative assessment of unmixing-based feature extraction techniques is given and suggests that such features can enhance the classification results. In fact, the features derived from unmixing (i.e., abundance maps) show high correlation with the land cover classes and can be interpreted with a clear physical meaning, which is generally lost after applying statistical methods such as linear discriminant analysis [47] and independent component analysis [48]. Moreover, in [49] and [50], the integration of unmixing and classification has shown adequate results in pixel labeling, where the fractional abundances have been applied as a complementary source to the MLR classifier. Although these approaches are effective for classification purposes, they devote insufficient attention to the class-dependent nature of the data, i.e., these approaches rarely take comprehensive considerations of the class-dependent structure in the low-dimensional spaces. The unmixing techniques are exploited in a simple manner to retrieve abundance fractions for discriminant learning or for soft classification purposes. Regarding this point, an improvement can be expected by introducing class-dependent considerations in these approaches.

With the aforementioned issues in mind, in this work, we propose a new physically interpretable hyperspectral image classification approach, called abundance-indicated subspace

(AiSub), which is specifically designed for scenarios with limited training samples. The proposed method relies on the fact that pixels from different classes generally comprise different groups of materials, i.e., endmembers. The spectral vector of a class, therefore, is a mixture of endmembers corresponding to this class. For instance, the spectral signatures of a so-called *grass land* class may be composed of grass and soil endmembers, while a so-called *forest area* may be characterized by endmembers of the leaf surface and the soil underneath. Under this basis, we adopt a new strategy to jointly exploit spectral unmixing and classification techniques for enhanced classification. In our context, the abundance fractions are used to indicate class-related endmembers that are crucial for the construction of the subspace. In this way, the land cover classes and the corresponding class-dependent endmembers are physically linked and can be better exploited for the purpose of classification.

In summary, the main contribution of this paper is the combination of endmembers and abundances for hyperspectral classification, aiming at preserving the class-dependent nature of hyperspectral data by modeling the class-dependent subspaces with endmembers, especially designed for scenarios with limited training samples. This is a distinguishing feature with regard to other pixel-level classifiers and a novel contribution in hyperspectral image analysis. Furthermore, our method is characterized by the physical interpretation within the unmixing framework, while most of the existing methods interpret hyperspectral data mainly from statistical viewpoints. In this regard, the use of abundance maps as features for classification offers a new perspective to tackle the classification problem. This effect is especially evident when low-spatial-resolution datasets are considered, since, in this scenario, the information provided by spectral unmixing affords a great asset for description at subpixel level. Hence, the proposed approach might represent a valid alternative for hyperspectral image classification with respect to methods introduced in technical literature so far.

The remainder of this paper is organized as follows. Section II formulates the problem of hyperspectral classification and describes the proposed approach in detail. Section III reports our experimental results with both synthetic and real hyperspectral datasets, as well as a comparison to other established approaches. Section IV draws a conclusion about the merits of our proposed approach and gives plausible future research lines for further improvements.

## II. PROPOSED APPROACH

In this section, we illustrate our proposed approach, in which the information provided by spectral unmixing is consistently interpreted in class-dependent subspaces for hyperspectral image classification. Section II-A describes a subspace version of MLR and introduces the concept of the endmember-based subspace. Then, we devote Section II-B to the elaboration of the abundance-indicated strategy by means of which endmembers and classes are associated to generate the class-dependent subspaces.

### A. Multinomial Logistic Regression

Let  $\mathbf{x}$  be a pixel in hyperspectral image containing  $N$  samples with  $L$  bands and  $y \in \{1, 2, \dots, K\}$  be the label index of ground truth for  $\mathbf{x}$ , where  $K$  is the number of the classes. Given the aforementioned notations, the posterior probability distribution can be learned by MLR from the whole image under a Bayesian framework, which is generally formulated as

$$p(y = k | \mathbf{x}, \boldsymbol{\omega}) = \frac{\exp(\boldsymbol{\omega}^{(k)T} \phi(\mathbf{x}))}{\sum_{l=1}^K \exp(\boldsymbol{\omega}^{(l)T} \phi(\mathbf{x}))} \quad (1)$$

where  $\boldsymbol{\omega}^{(k)}$  is the regressor vector for class  $k$ , and  $\phi(\cdot)$  is a transformative function of the input vector. In [37] and [38], class-dependent subspaces are defined by eigendecomposition, where the orthogonal bases of the subspace are extracted from the correlation matrix of training samples.

In our approach, we assume that the class-dependent subspace can be represented by a set of endmembers related to this class. The transformation function is, therefore, defined as

$$\phi_E(\mathbf{x}) = [\|\mathbf{x}\|^2, \|\mathbf{x}^T \mathbf{E}^{(1)}\|^2, \|\mathbf{x}^T \mathbf{E}^{(2)}\|^2, \dots, \|\mathbf{x}^T \mathbf{E}^{(K)}\|^2], \quad (2)$$

where  $\mathbf{E}^{(k)} = [\mathbf{e}_1^{(k)}, \dots, \mathbf{e}_i^{(k)}, \dots, \mathbf{e}_{c^{(k)}}^{(k)}]$  is the endmember set for class  $k$ , and  $c^{(k)}$  is the number of the class-dependent endmembers. In this context, learning the class density means estimating the logistic regressors  $\boldsymbol{\omega}$ , which can be obtained by computing the MAP estimate

$$\hat{\boldsymbol{\omega}} = \arg \max_{\boldsymbol{\omega}} \ell(\boldsymbol{\omega}) + \log p(\boldsymbol{\omega}) \quad (3)$$

where

$$\ell(\boldsymbol{\omega}) \equiv \log \prod_i p(y_i | \mathbf{x}_i, \boldsymbol{\omega}) \quad (4)$$

and

$$p(\boldsymbol{\omega}) \propto \exp(-\lambda \|\boldsymbol{\omega}\|_1). \quad (5)$$

To solve the problem in (3), we resort to the variable splitting and augmented Lagrangian (LORSAL) algorithm [36]. LORSAL attempts to approximate the solution of (3) by a series of quadratic problems, which are easier to solve than the original one. In [36], [38], and [51], LORSAL is proven to be very effective even under unfavorable conditions such as limited training samples.

### B. Subspace Learning Indicated by Abundance Fractions

As shown in (2), a critical issue is how to find a proper set of endmembers for the representation of each class. However, this task is difficult when the ground-truth reference information is limited. To overcome this problem, we develop an endmember selection strategy that exploits the correspondent relationship between abundances and classes, which is an important contribution of our work.

Under the context of spectral unmixing analysis, and based on the linear mixture model, the spectral vector  $\mathbf{x}$  can be expressed as follows:

$$\mathbf{x} = \mathbf{M}\boldsymbol{\alpha} = \sum_{i=1}^c (\alpha_i \cdot \mathbf{m}_i) \quad (6)$$

where  $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_c]$  is the so-called *mixing matrix* in the unmixing approach, whose columns are spectral signals of  $c$  material components extracted from the whole image;  $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_c]^T$  is the abundance vector of  $\mathbf{M}$ , which indicates the corresponding fractional coverages in the given pixel. Two physical constraints are generally imposed onto the abundance estimation, i.e., abundance sum-to-one constraint and abundance nonnegativity constraint. Since the mixing matrix  $\mathbf{M}$  contains endmembers of other classes and, theoretically, the abundance fractions of other classes are zeros,  $\boldsymbol{\alpha}$  is expected to be sparse. Based on this assumption, the pixel  $\mathbf{x}^{(k)}$  in class  $k$  can be expressed as

$$\begin{aligned} \mathbf{x}^{(k)} &= \mathbf{E}^{(k)} \boldsymbol{\alpha}^{(k)} = \sum_{i=1}^{c^{(k)}} \alpha_i^{(k)} \cdot \mathbf{e}_i^{(k)} \\ \text{s.t. } \mathbf{E}^{(k)} &\subset \mathbf{M}, 0 < c^{(k)} \leq c \end{aligned} \quad (7)$$

where  $\boldsymbol{\alpha}^{(k)} = [\alpha_1^{(k)}, \dots, \alpha_{c^{(k)}}^{(k)}]^T$  denotes the abundance fractions corresponding to the class-dependent endmember set  $\mathbf{E}^{(k)}$ . Here, the mixing matrix  $\mathbf{M}$  (recall that it consists of  $c$  endmembers from the whole image) serves as the pool of candidates for  $\mathbf{E}^{(k)}$ .

A possible strategy to obtain the endmember set for each class is to select endmembers from the training set and associate them with their ground-truth labels. However, under the condition of limited available labeled samples, unlabeled samples are the majority. As a result, most of the endmember candidates are most likely existing in the unlabeled sample set. In this context, if we select endmembers from the training set, it will be very difficult to find the class-dependent subspaces that fully characterize the spectral signatures of the land cover classes. To overcome this problem, we propose an alternative strategy to select endmembers for classes, that is, we extract the spectrally pure pixels from the whole image, and let abundance fractions indicate the correspondence of endmembers and classes.

In this paper, we adopt the vertex component analysis (VCA) [52] to generate the endmember candidate matrix  $\mathbf{M}$  and the fully constrained least square (FCLS) [53] to retrieve the abundances. We choose VCA and FCLS due to their high efficiency and effectiveness for unmixing process [16]. In order to quantitatively evaluate the class-dependent subspaces resulted from VCA and FCLS, Fig. 1 displays the histogram of the classwise subspace dissimilarities generated from 20 runs of VCA and FCLS for the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Indian Pines dataset,<sup>1</sup> where  $d$  denotes the subspace similarity metric introduced in [54], which is given as

$$d = \sqrt{\max(m, n) - \sum_{i=1}^m \sum_{j=1}^n (\mathbf{u}_i^T \mathbf{v}_j)^2} \quad (8)$$

where  $\mathbf{u}_i$  and  $\mathbf{v}_j$  are the  $i$ th and  $j$ th orthonormal bases of subspaces  $\mathbf{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$  and  $\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ , respectively. Notice that  $1/2$  is generally assumed as the upper bound of  $d$ . From Fig. 1, it is worth noting that the obtained  $d$  in our case are much less than  $1/2$ , mostly in the range of 0.12–0.20.

<sup>1</sup>Detailed description of the AVIRIS Indian Pines dataset will be provided in the experimental section.

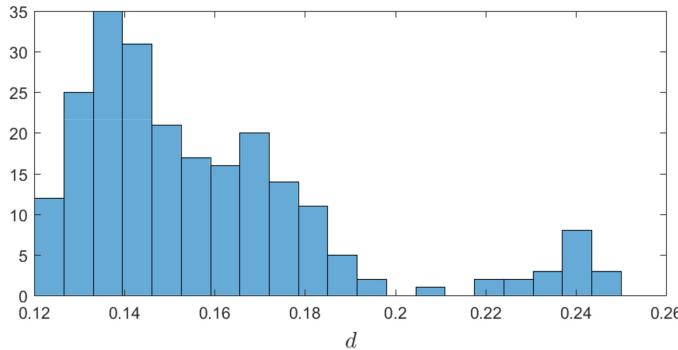


Fig. 1. Histogram of the classwise subspace similarity metric  $d$  obtained from 20 runs of VCA and FCLS for the AVIRIS Indian Pines dataset.

This result reveals that the class-dependent subspaces of the same class resulted from VCA and FCLS in different runs are similar. Nevertheless, it is proper to emphasize that the proposed approach can function along with different choices of unmixing algorithms. Other methods can also be applied for the purposes of endmember extraction and abundance estimation.

In order to ensure the selection of all potential endmembers, the candidates are obtained from the whole image. Assuming that the training samples are correctly labeled, and taking into account that samples in the same class share similar structures, we can use the abundances of the training samples in a class to indicate its representative endmembers, thus obtaining its class-dependent endmembers. Inspired by this observation, we use the average abundances, i.e., the average abundances derived from its training samples, to indicate the class-dependent endmembers. Notice that there are two main advantages of using the average abundances for indication. On the one hand, this strategy can handle the intraclass variation of the labeled samples, thus compensating the estimation bias from unmixing. On the other hand, as an averaging approach, the image noise is generally reduced. Let  $\bar{\alpha}^{(k)} = [\bar{\alpha}_1^{(k)}, \bar{\alpha}_2^{(k)}, \dots, \bar{\alpha}_c^{(k)}]$  denote the average abundance vector of class  $k$  with decreasing magnitude, i.e.,  $\bar{\alpha}_1^{(k)} > \bar{\alpha}_2^{(k)} > \dots > \bar{\alpha}_c^{(k)}$ . The first  $c^{(k)}$  abundances  $\{\bar{\alpha}_i^{(k)} \mid i = 1, \dots, c^{(k)}\}$  are considered for  $\mathbf{E}^{(k)}$ , which are defined as

$$c^{(k)} = \min \left\{ c^{(k)} : \sum_{i=1}^{c^{(k)}} \bar{\alpha}_i^{(k)} \geq \tau \right\}, \quad 0 < \tau \leq 1 \quad (9)$$

where  $\tau$  controls the amount of original information preserved in the subspaces. As shown in (9), a large value of  $\tau$  leads to precise spectral representation for class dependence, which is supposed to improve classification accuracy. This point is verified in Fig. 2, where the overall accuracy (OA) increases with  $\tau$  when the AVIRIS Indian Pines dataset is considered. It also can be observed that, with  $\tau$  close to 1, the performance decreases. This effect is expected to be caused by the presence of noise. Furthermore, the OA results are better with a larger number of endmember candidates. This aspect is consistent with our aforementioned investigation.

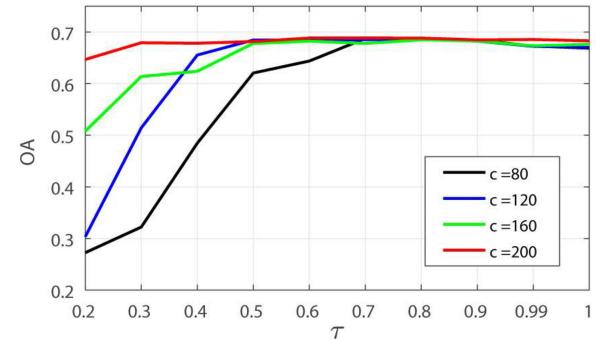


Fig. 2. OA as a function of  $\tau$ , obtained from the AVIRIS Indian Pines dataset with 15 training samples per class. The lines in red, green, blue, and black represent the results of the number of endmember candidates  $c = 200, 160, 120, 80$ , respectively.

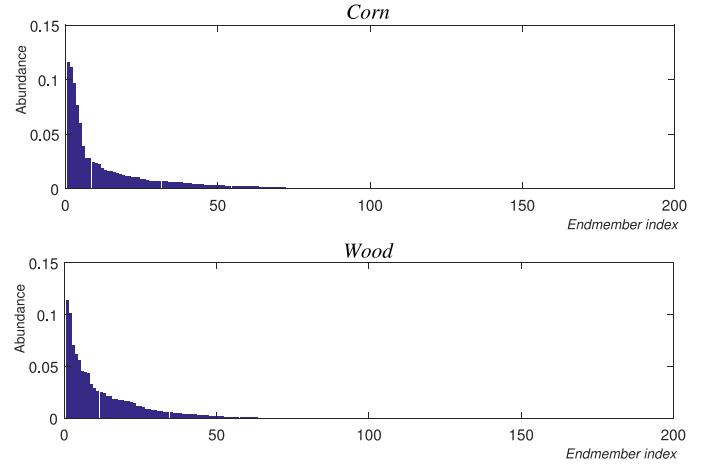


Fig. 3. Average abundances of 30 samples for *Corn* (top) and *Wood* (bottom), respectively, for the AVIRIS Indian Pines dataset. The endmember indexes are arranged in descending order.

Finally, the class-dependent subspaces  $\mathbf{E}^{(k)}$  in (2) can be obtained as follows:

$$\mathbf{E}^{(k)} = [\mathbf{m}_1, \dots, \mathbf{m}_{c^{(k)}}] \quad (10)$$

where  $\mathbf{m}_i \in \mathbf{M}$  is the specific endmember corresponding to  $\bar{\alpha}_i^{(k)}$  for class  $k$ . For illustrative purposes, Fig. 3 shows the average abundance abundances of 30 samples from the Indian Pines dataset, belonging to classes *Corn* and *Wood*. It is worth noting that, for samples in the same class, the large abundances are associated with the same endmember candidates, i.e., the representative endmembers. The class dependence of endmember can be well preserved by the indication of the average abundances. At this point, we would like to emphasize again that, as we use the average abundances to indicate the representative endmembers in the class, the uncertainty of training samples can be tackled, ensuring the certainty of the endmember-based subspaces.

For illustrative purposes, Fig. 4 gives a conceptual framework of our proposed approach, which can be summarized into two main steps. First, the class-dependent endmembers are learned based on the indication of the abundance fractions. Then, we exploit a subspace-based MLR classifier for the ultimate goal

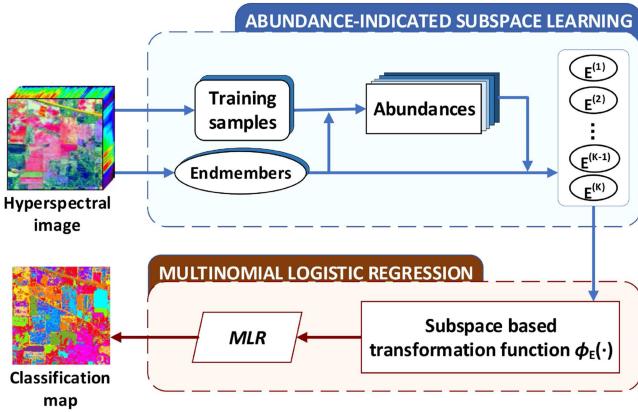


Fig. 4. Main steps of the proposed approach for hyperspectral image classification.

of hyperspectral image classification. As for the computational complexity, following [36], [52], and [53], the overall complexity of the proposed AiSub includes the unmixing process with VCA and FCLS of  $O(L^2(2N + 1))$  and the classification of  $O(K^3 + N)$ , respectively.

### III. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we evaluate the proposed approach by using both synthetic datasets and real hyperspectral images. The experimental settings are as follows.

- 1) The parameters involved in the proposed approach ( $c$  that determines the number of endmembers extracted from the whole image and  $\tau$  that controls the amount of spectral information preserved in the subspace) should be as large as possible from a theoretical point of view. Therefore, we set  $c$  as the number of bands of the considered dataset. On the other hand, due to the issue of noise, we empirically set  $\tau = 0.99$ . It should be noted that, in the proposed approach, we employ VCA [52] for endmember extraction and SUrSal [55] for abundance estimation.
- 2) For comparison, three MLR-based methods, including MLR with spectral kernel features (denoted as MLR) [56], MLRsub [37], and MLRsub<sub>mod</sub> [38], are included in the comparison. MLR exploits the spectral kernel to improve the class separability in a nonlinear feature space. MLRsub generates class-dependent subspaces by eigendecomposition and computes the posterior class density using the nonlinear functions  $\varphi_{(k)}(\mathbf{x}_i) = [\|\mathbf{x}_i\|^2, \|\mathbf{x}_i^T \mathbf{U}^{(k)}\|^2]$ . MLRsub<sub>mod</sub> defines a new nonlinear function  $\varphi(\mathbf{x}_i) = [\|\mathbf{x}_i\|^2, \|\mathbf{x}_i^T \mathbf{U}^{(1)}\|^2, \dots, \|\mathbf{x}_i^T \mathbf{U}^{(K)}\|^2]$ , which is proven to be more suitable for real cases with nonlinear mixing. Compared to MLR, MLRsub, and MLRsub<sub>mod</sub>, the main theoretical difference of the proposed method is the fact that feature subspaces contain physical meaning, while MLR, MLRsub, and MLRsub<sub>mod</sub> generate subspaces by means of statistical approaches. The parameters of MLR, MLRsub, and MLRsub<sub>mod</sub> are set according to [37], [38], and [56]. In our experiments with real hyperspectral images, we include additional comparisons with SVM [31], which is a classic spectral-based solution for hyperspectral

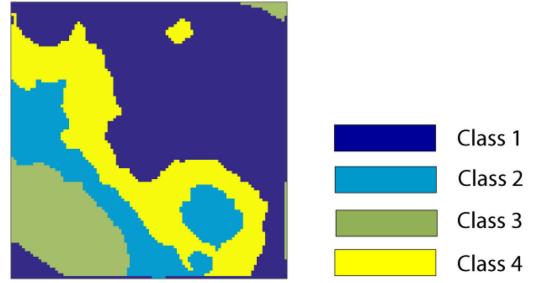


Fig. 5. Reference map of the simulated hyperspectral dataset with four classes.

image classification with limited training sets. In the implementation of SVM, we take advantage of the software package LIBSVM<sup>2</sup> [57] and use the Gaussian kernel.<sup>3</sup> Parameters are optimized by the fivefold cross validations.

- 3) Two sampling strategies are used to generate the training sets. The first one is to select an equal number of samples for each class, and the second one is to collect the same proportion of samples for each class. Quantitative metrics such as OA, average accuracy (AA), Cohen's Kappa statistic ( $\kappa$ ), and individual class accuracies are reported to evaluate the classification performance of the tested methods. All results are obtained from 20 independent Monte Carlo tests.
- 4) Finally, all our experiments have been conducted using MATLAB R2016b in a desktop PC equipped with an Intel Core i7 CPU (at 3.6 GHz) and 24 GB of RAM. In our real image experiments, the computational times are also reported. It should be noticed that, for AiSub, the time reported includes two parts: unmixing and classification.

#### A. Experiments With Synthetic Datasets

To assess the classification performance in a totally controlled environment, we generate synthetic datasets of four classes (see Fig. 5). It should be noted that the proposed approach exploits the linear mixture model, as shown in (6). However, in order to illustrate the high effectiveness of the proposed approach for challenging scenarios, we use a nonlinear mixture model for data simulation. Let  $\mathbf{x}_i^{(k)}$  be the  $i$ th sample in class  $k$ ; we have

$$\mathbf{x}_i^{(k)} = \sum_{j=0}^{c^{(k)}} \mathbf{m}^{(k+j)} \alpha_{(k+j)} + \gamma \mathbf{m}^{(i_1)} \cdot \mathbf{m}^{(i_2)} + \mathbf{n}_i \quad (11)$$

where  $\mathbf{m}^{(l)}, l = 1, \dots, 8$  are pure spectra from the U.S. Geological Survey digital spectral library,<sup>4</sup>  $\alpha_{(k+j)}$  is the corresponding abundance fraction,  $c^{(k)}$  is the number of constituents in class  $k$ , and  $\gamma$  is the parameter that controls the impact of the nonlinear phenomenon. For a certain sample  $\mathbf{x}_i$ , we assume that  $\mathbf{m}^{(k)}$  receives the maximum abundance value, which, in turn, determines the corresponding label  $y_i = k$ . In the nonlinear term  $\gamma \mathbf{m}^{(i_1)} \cdot \mathbf{m}^{(i_2)}$ , we set that

<sup>2</sup> Available online: <https://www.csie.ntu.edu.tw/~cjlin/libsvm>

<sup>3</sup>  $K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}$ , where  $\sigma$  is estimated by the median Euclidean distance computed over the training samples.

<sup>4</sup> <http://speclab.cr.usgs.gov/spectral-lib.html>

TABLE I

CLASSIFICATION OVERALL ACCURACIES (IN PERCENT) OF THE SYNTHETIC IMAGE WITH SNR = 40 dB, OBTAINED BY THE CONSIDERED METHODS

OA(%)	(i) All linear mixed classes			
MLR	93.84( $\pm 0.52$ )			
MLRsub	92.21( $\pm 0.82$ )			
MLRsub <sub>mod</sub>	97.76( $\pm 0.56$ )			
AiSub	<b>98.27</b> ( $\pm 0.24$ )			
OA(%)	(ii) Linear and nonlinear mixed classes			
	$\gamma = 0.2$	$\gamma = 0.3$	$\gamma = 0.4$	$\gamma = 0.5$
MLR	96.76( $\pm 0.19$ )	95.88( $\pm 0.30$ )	95.21( $\pm 0.37$ )	94.29( $\pm 0.27$ )
MLRsub	90.68( $\pm 0.68$ )	90.22( $\pm 0.48$ )	91.31( $\pm 0.44$ )	90.51( $\pm 0.89$ )
MLRsub <sub>mod</sub>	95.07( $\pm 0.92$ )	94.52( $\pm 1.01$ )	93.19( $\pm 0.82$ )	90.76( $\pm 0.97$ )
AiSub	<b>97.64</b> ( $\pm 0.22$ )	<b>96.92</b> ( $\pm 0.33$ )	<b>96.11</b> ( $\pm 0.56$ )	<b>94.59</b> ( $\pm 1.15$ )
OA(%)	(iii) All nonlinear mixed classes			
	$\gamma = 0.2$	$\gamma = 0.3$	$\gamma = 0.4$	$\gamma = 0.5$
MLR	88.60( $\pm 0.50$ )	84.88( $\pm 0.75$ )	81.39( $\pm 0.87$ )	76.88( $\pm 1.30$ )
MLRsub	89.52( $\pm 0.62$ )	86.33( $\pm 0.61$ )	81.41( $\pm 0.77$ )	75.62( $\pm 0.81$ )
MLRsub <sub>mod</sub>	90.57( $\pm 0.90$ )	85.16( $\pm 1.38$ )	79.51( $\pm 1.03$ )	72.87( $\pm 1.25$ )
AiSub	<b>93.36</b> ( $\pm 0.36$ )	<b>89.38</b> ( $\pm 0.61$ )	<b>84.38</b> ( $\pm 0.97$ )	<b>77.84</b> ( $\pm 1.55$ )

The highest accuracies are shown in boldface.

$i_1, i_2 \in \{k, \dots, k + c^{(k)}\}$ ,  $i_1 \neq i_2$ , i.e., only two constituents are involved in the nonlinear spectral variation. Additionally, we impose the constraint  $\sum_{(j=0)}^{c^{(k)}} \alpha_{(k+j)} = 1 - \gamma$ ,  $c^{(k)} \in \{3, 4\}$  on every pixel in the synthetic image. The zero-mean Gaussian noise  $n_i \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$  is also added to the pixel  $x_i$ . In our experiments, we select 100 samples per class from the image for training and use the rest samples for testing.

In our first experiment, we use the synthetic datasets of (i) linear mixed classes, (ii) both of linear and nonlinear mixed classes, and (iii) nonlinear mixed classes, respectively, to evaluate the algorithm performance with different types of nonlinear effects. In this experiment, Gaussian additive noise with a signal-to-noise-ratio (SNR) of 40 dB is considered. Table I demonstrates the overall accuracies of the classification results as a function of parameter  $\gamma$ . It can be seen that the advantages of AiSub are remarkable in all cases. For the linearly mixed dataset, AiSub achieves an OA of 98.27%, which is much better than that of the other methods. For the other two scenarios, similar observations can be made that our proposed approach is able to achieve the highest accuracy among all compared methods. We should recall that, although our method is developed with the assumption of a linear mixture model, it obtains very promising performance in challenging scenarios with nonlinear effects, which demonstrates its high potential for real scenarios, where the mixing phenomenon is generally more complex than linearity.

In our second experiment, we aim to evaluate the impact of noise on the classification performance, in which the nonlinear factor is set to  $\gamma = 0.5$ . Fig. 6 demonstrates the OA results as a function of SNR. It can be seen that the proposed AiSub always achieves the best performance. For example, in the case of SNR = 20 dB, the OA of AiSub is 8.67%, 3.74%, and 6.71% higher than MLR, MLRsub, and MLRsub<sub>mod</sub>.

It should be noticed that the outstanding performance of AiSub is encouraged by the proposed endmember indication strategy, which uses average abundances to reduce the noise effect. As expected, with an SNR increase (i.e., the dataset contains less noise), the classification accuracies of all methods also increases.

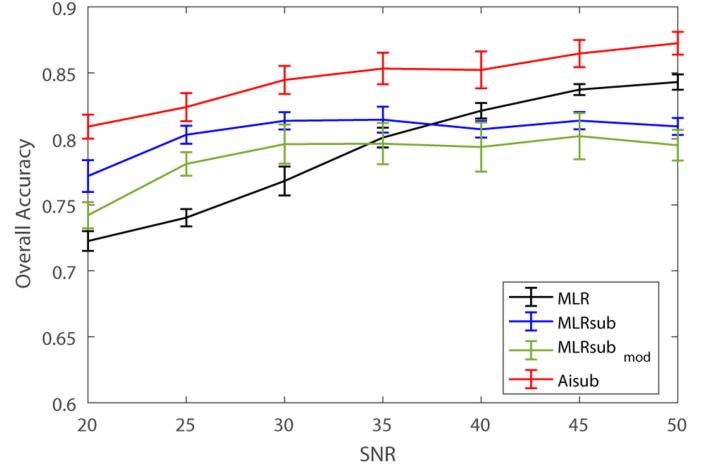


Fig. 6. OA as a function of SNR, obtained from the synthetic dataset made up of nonlinear classes with  $\gamma = 0.5$ .

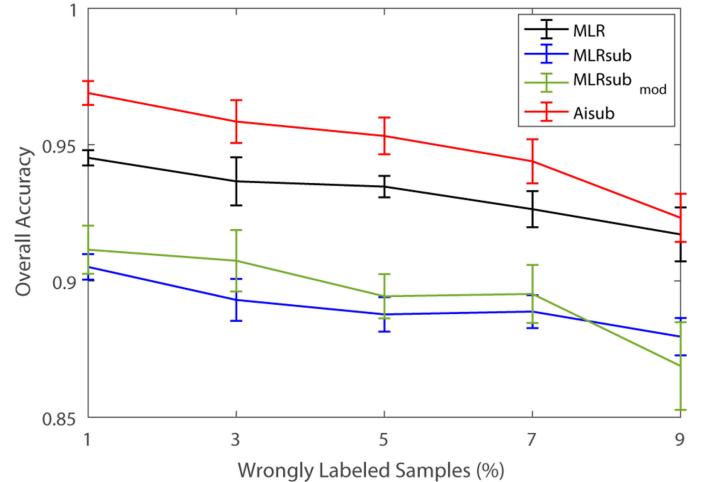


Fig. 7. OA as a function of the percentage of wrongly labeled samples in the training set, obtained from the synthetic dataset with  $\gamma = 0.2$ , SNR = 40 dB.



Fig. 8. False color image and ground truth reference of the AVIRIS scene.

In our third experiment, we use a training set with wrongly labeled samples to evaluate the performance of the methods for an image with  $\gamma = 0.2$  and SNR = 40 dB. For a wrong labeled sample, its label is associated with the second largest abundance fraction. Fig. 7 shows the OA results (along with the standard deviations) as a function of the number of wrongly labeled samples in the training set. It can be observed that AiSub achieves excellent performances in all cases. Taking advantage from the average abundances, AiSub is able to ensure its performance

TABLE II  
CLASS SPECIFIC ACCURACIES, OA, AA,  $\kappa$  STATISTIC COEFFICIENT (IN PERCENT), AND COMPUTATIONAL TIME (IN SECONDS)  
OBTAINED IN THE FIRST EXPERIMENT OF THE INDIAN PINES DATASET

	10 samples per class (160 samples in total)						
	Train	Test	SVM	MLR	MLRsub	MLRsub <sub>mod</sub>	AiSub
1.Alalfa	10	44	81.14	81.09	73.64	<b>82.61</b>	80.91
2.Corn-no till	10	1424	45.37	54.17	34.38	50.91	<b>55.11</b>
3.Corn-min till	10	824	42.79	51.71	<b>62.16</b>	46.08	41.74
4.Corn	10	224	<b>74.42</b>	73.01	69.62	67.05	65.49
5.Grass/pasture	10	487	74.67	72.72	65.64	75.32	<b>77.10</b>
6.Grass/tree	10	737	81.17	<b>89.76</b>	88.45	85.50	85.76
7.Grass/pasture-mowed	10	16	91.56	91.94	90.63	89.38	<b>94.06</b>
8.Hay-windrowed	10	479	78.63	<b>90.66</b>	90.57	86.41	89.26
9.Oats	10	10	95.00	97.50	96.50	<b>99.50</b>	94.00
10.Soybeans-no till	10	958	51.49	58.83	37.52	<b>59.67</b>	57.78
11.Soybeans-min till	10	2458	38.23	46.15	40.52	45.74	<b>47.39</b>
12.Soybean-clear till	10	604	42.73	<b>64.23</b>	42.99	60.06	59.55
13.Wheat	10	202	93.51	<b>99.53</b>	99.26	98.42	98.27
14.Woods	10	1284	79.52	80.23	<b>98.30</b>	72.09	73.10
15.Bldg.-Grass-Tree-Drives	10	370	45.54	45.19	12.49	43.92	<b>49.30</b>
16.Stone-steel towers	10	85	89.94	81.94	<b>92.65</b>	73.24	75.12
OA(%)			55.94( $\pm$ 4.79)	59.98( $\pm$ 1.98)	57.15( $\pm$ 4.94)	60.10( $\pm$ 2.84)	<b>61.06</b> ( $\pm$ 2.50)
AA(%)			69.11( $\pm$ 2.40)	70.94( $\pm$ 0.92)	68.75( $\pm$ 2.30)	70.81( $\pm$ 1.80)	<b>71.50</b> ( $\pm$ 2.12)
$\kappa$ (%)			50.92( $\pm$ 4.91)	55.74( $\pm$ 2.08)	52.02( $\pm$ 4.89)	55.33( $\pm$ 3.03)	<b>56.36</b> ( $\pm$ 2.69)
time(s)			182.58	3.42	5.78	3.07	Unmixing: 59.40 Classification: 1.06 Total: 60.46

Standard deviations are in parentheses and the highest accuracies are shown in boldface.

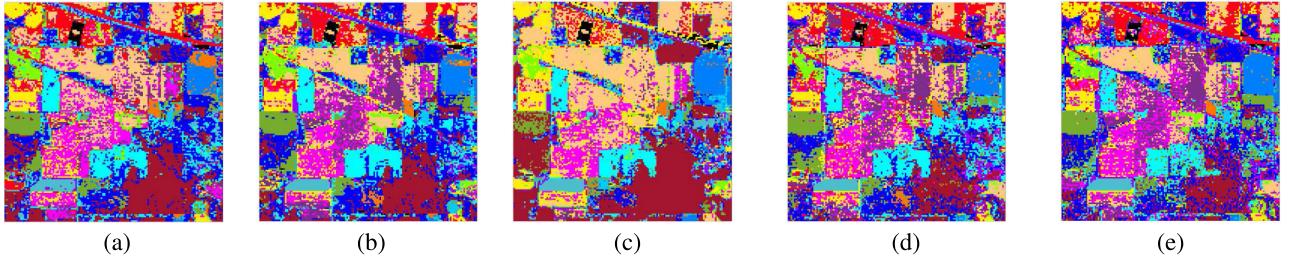


Fig. 9. Classification maps obtained in one of the 20 Monte Carlo tests by (a) SVM, (b) MLR, (c) MLRsub, (d) MLRsub<sub>mod</sub>, and (e) AiSub in the first experiment of the Indian Pines dataset.

in the classification step, even when there are wrongly labeled samples.

#### B. Experiments With Real Hyperspectral Images

We have tested our method with two publicly available real hyperspectral images.

1) *AVIRIS Indian Pines Scene*: The Indian Pines dataset was acquired by National Aeronautics and Space Administration Jet Propulsion Laboratory's AVIRIS in 1992. For illustrative purposes, Fig. 8 presents the false color composite and the ground truth reference map of this dataset. It depicts the area of Indian Pines in northwestern Indiana with 224 spectral bands in the range of 0.4–2.5  $\mu$ m with a spatial resolution of 20 m. The ground-truth map comprises 16 mutually exclusive available for this dataset. After removing the spectral bands affected by water absorption and noise, there are 200 spectral bands left for experiments.

In our first experiment, we follow the first sampling strategy to select 160 samples (i.e., ten samples per class) from the reference map for training, while the rest are used for testing. Table II gives numerical evaluations on the classification

performance, in which the advantage of our proposed method is evident. AiSub achieves results with OA, AA, and  $\kappa$  of 61.06%, 71.50%, and 56.36%, respectively, which are 5.12%, 2.39%, and 5.44% higher than those obtained by SVM. In comparison with the MLR-based methods, the AiSub provides competitive results. With respect to computational time, AiSub took 60.46 s with 59.40 s for unmixing and 1.06 s for classification, respectively. Notice that, since unmixing can be preprocessed, the computational time for classification of AiSub is just 1.06 s, while SVM, MLR, MLRsub, and MLRsub<sub>mod</sub> consume 182.58, 3.42, 5.78, and 3.07 s, respectively. In Fig. 9, it can be found that AiSub stands out among other competitors according to the presented classification maps obtained in one of 20 Monte Carlo tests.

In our second experiment with the Indian Pines dataset, we use the second sampling strategy to select 1% of the samples from the ground-truth reference (for a total of 125 samples) for training and the rest for validation. The classification results are reported in Table III. AiSub achieves significantly better OA, AA, and  $\kappa$  than other methods. For example, compared to the result of MLRsub, the OA, AA, and  $\kappa$  of AiSub are higher

TABLE III

CLASS SPECIFIC ACCURACIES, OA, AA,  $\kappa$  STATISTIC COEFFICIENT (IN PERCENT), AND COMPUTATIONAL TIME (IN SECONDS) OBTAINED IN THE SECOND EXPERIMENT OF THE INDIAN PINES DATASET

	1 percent per class (125 samples in total)						
	Train	Test	SVM	MLR	MLRsub	MLRsub <sub>mod</sub>	AiSub
1.Alalfa	5	49	55.82	79.39	80.10	<b>77.24</b>	73.98
2.Corn-no till	14	1420	35.69	58.79	25.59	52.35	<b>60.06</b>
3.Corn-min till	8	826	12.57	<b>60.91</b>	40.60	42.48	43.57
4.Corn	5	229	16.59	50.63	<b>65.63</b>	44.21	48.78
5.Grass/pasture	5	492	46.05	63.06	46.53	63.09	<b>64.22</b>
6.Grass/tree	7	740	73.62	83.30	79.14	<b>86.39</b>	82.35
7.Grass/pasture-mowed	5	21	80.24	<b>92.86</b>	91.90	87.86	89.29
8.Hay-windrowed	5	484	64.83	<b>82.46</b>	82.16	80.96	81.39
9.Oats	5	15	78.67	<b>98.33</b>	98.00	92.00	80.33
10.Soybeans-no till	9	959	16.99	49.04	49.22	46.82	<b>51.33</b>
11.Soybeans-min till	24	2444	<b>88.19</b>	72.46	32.36	65.01	64.15
12.Soybean-clear till	6	608	5.48	42.71	38.50	37.59	<b>46.63</b>
13.Wheat	5	207	89.30	95.89	98.31	<b>98.77</b>	96.01
14.Woods	12	1282	<b>97.92</b>	91.72	94.98	80.62	79.55
15.Bldg-Grass-Tree-Drives	5	375	15.12	36.72	15.60	29.08	<b>34.75</b>
16.Stone-steel towers	5	90	76.39	67.44	<b>91.56</b>	58.33	65.39
OA(%)			55.36( $\pm 2.90$ )	61.65( $\pm 1.98$ )	52.72( $\pm 5.77$ )	60.91( $\pm 2.69$ )	<b>63.39</b> ( $\pm 3.90$ )
AA(%)			53.16( $\pm 4.02$ )	65.89( $\pm 2.01$ )	65.84( $\pm 3.86$ )	64.74( $\pm 2.72$ )	<b>66.61</b> ( $\pm 4.30$ )
$\kappa$ (%)			46.89( $\pm 3.87$ )	56.73( $\pm 2.30$ )	47.69( $\pm 5.74$ )	55.30( $\pm 3.06$ )	<b>58.39</b> ( $\pm 4.28$ )
time(s)			114.25	1.84	4.14	1.54	Unmixing: 62.33 Classification: 1.03 Total: 63.36

Standard deviations are in parentheses, and the highest accuracies are shown in boldface.

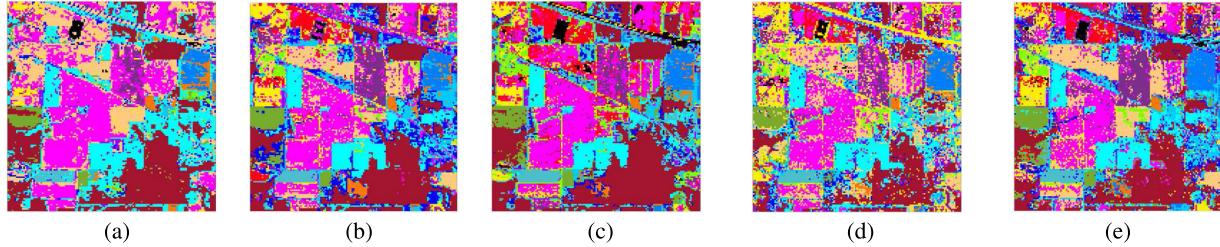


Fig. 10. Classification maps obtained in one of the 20 Monte Carlo tests by (a) SVM, (b) MLR, (c) MLRsub, (d) MLRsub<sub>mod</sub>, and (e) AiSub in the second experiment of the Indian Pines dataset.

TABLE IV  
CLASS SPECIFIC ACCURACIES, OA, AA,  $\kappa$  STATISTIC COEFFICIENT (IN PERCENT), AND COMPUTATIONAL TIME (IN SECONDS)  
OBTAINED IN THE THIRD EXPERIMENT OF THE INDIAN PINES DATASET

	50 samples per class (800 samples in total)						
	Train	Test	SVM	MLR	MLRsub	MLRsub <sub>mod</sub>	AiSub
1.Alalfa	50	4	95.00	<b>97.50</b>	91.25	93.75	96.25
2.Corn-no till	60	1374	<b>73.19</b>	71.37	52.84	67.95	71.56
3.Corn-min till	58	776	65.59	65.31	62.82	65.12	<b>68.53</b>
4.Corn	52	182	86.36	<b>88.41</b>	63.02	85.73	88.07
5.Grass/pasture	53	444	89.01	<b>93.12</b>	76.12	89.94	90.80
6.Grass/tree	58	689	<b>92.93</b>	92.79	91.40	92.05	92.46
7.Grass/pasture-mowed	13	13	<b>91.89</b>	90.92	76.83	87.24	88.85
8.Hay-windrowed	55	434	95.42	96.16	97.78	<b>98.17</b>	94.66
9.Oats	10	10	85.67	90.13	84.67	<b>91.22</b>	79.78
10.Soybeans-no till	55	913	72.96	73.07	56.47	71.13	<b>74.93</b>
11.Soybeans-min till	68	2400	64.32	68.00	51.36	63.35	<b>69.46</b>
12.Soybean-clear till	57	557	83.22	81.34	72.72	81.07	<b>83.44</b>
13.Wheat	50	162	99.53	99.25	<b>99.81</b>	99.69	98.79
14.Woods	54	1240	86.37	90.04	88.36	<b>98.56</b>	88.14
15.Bldg-Grass-Tree-Drives	54	326	<b>74.36</b>	82.62	21.88	68.97	72.15
16.Stone-steel towers	53	42	93.16	92.84	95.07	94.61	<b>95.62</b>
OA(%)			76.66( $\pm 1.50$ )	74.05( $\pm 1.01$ )	66.81( $\pm 2.31$ )	75.40( $\pm 1.31$ )	<b>78.34</b> ( $\pm 1.25$ )
AA(%)			84.31( $\pm 1.38$ )	83.86( $\pm 1.18$ )	74.95( $\pm 1.16$ )	83.11( $\pm 1.91$ )	<b>84.59</b> ( $\pm 1.54$ )
$\kappa$ (%)			73.52( $\pm 1.65$ )	73.31( $\pm 1.12$ )	62.29( $\pm 2.32$ )	72.09( $\pm 1.47$ )	<b>75.36</b> ( $\pm 1.39$ )
time(s)			1372.40	2.31	3.22	1.21	Unmixing: 92.21 Classification: 1.41 Total: 93.62

Standard deviations are in the parentheses, and the highest accuracies are shown in boldface.

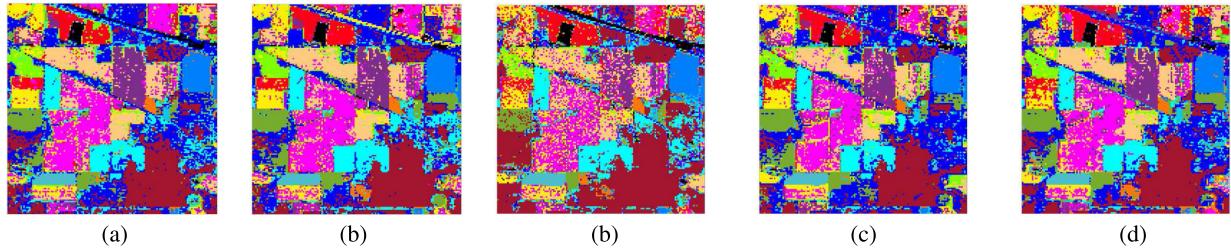


Fig. 11. Classification maps obtained in one of the 20 Monte Carlo tests by (a) SVM, (b) MLR, (c) MLRsub, (d) MLRsub<sub>mod</sub>, and (e) AiSub in the third experiment of the Indian Pines dataset.



Fig. 12. False color image and ground truth reference of the ROSIS dataset.

by 10.67%, 0.77%, and 10.70%, respectively. Regarding the computational time, AiSub takes 63.33 s with 62.33 s for unmixing and 1.03 s for classification, while SVM, MLR, MLRsub, MLRsub<sub>mod</sub> consume 114.25, 1.84, 4.14, and 1.54 s, respectively. For illustrative purposes, Fig. 10 demonstrates the classification maps obtained in one of the 20 Monte Carlo tests of this experiment.

In the third experiment with the AVIRIS Indian Pines dataset, we randomly select 50 samples per class from the reference map for training. For the classes with less than 50 samples, we just select half of the total samples in the class. The classification results are reported in Table IV, indicating the superior performance of AiSub. Regarding the computational time, AiSub consumes 93.62 s, with 92.21 s for unmixing and 1.41 s for classification, while SVM, MLR, MLRsub, and MLRsub<sub>mod</sub> consume 1372.40, 2.31, 3.22, and 1.21 s, respectively. Fig. 11 shows some of the classification maps obtained in one of the 20 Monte Carlo tests conducted in this experiment.

**2) Reflective Optics Spectrographic Imaging System (ROSIS) Pavia University Scene:** The Pavia University dataset was acquired by the ROSIS airborne instrument in 2003 over the area of Engineering School at the University of Pavia. This dataset records the spectral information ranging from 0.43 to 0.86  $\mu\text{m}$  with 115 spectral channels. Since some of the channels are heavily affected by the existence of noise, only 103 noiseless bands are used in the experiments. With a high spatial resolution of 1.3 m, this image depicts complicated classes such as urban area, soil, and vegetation with  $610 \times 340$  pixels in nine land cover classes. For this dataset, an original separated training set and testing set are provided. The false color composite and reference map of the Pavia University dataset are demonstrated in Fig. 12.

In our first experiment with the Pavia University dataset, we generated the training set by selecting 20 samples per class (180 samples in total) from the original training set. As shown in Table V, our proposed method obtained the highest OA, AA, and  $\kappa$ , outperforming the competitors even in a scenario with much higher spatial resolution than in the AVIRIS Indian Pines experiments (as expected, the advantages of our method are less significant in this case due to the increased spatial resolution of the data that reduces the impact of spectral mixing). Regarding the computational time, AiSub has a good advantage in the classification step. Fig. 13 displays the classification maps obtained in one of 20 independent Monte Carlo runs, indicating the superior performance of the proposed method in this experiment.

In our second experiment, we collect 4% samples from the original training set to build a balanced training set (152 samples in total). According to the results reported in Table VI, we can see that AiSub achieves the best classification results compared to other methods. AiSub produces the highest OA of 68.46%, which is higher than those of SVM, MLR, MLRsub, and MLRsub<sub>mod</sub> by 1.67%, 2.19%, 3.62%, and 4.93%, respectively. Similar observations can also be found regarding to AA and  $\kappa$ , indicating the outstanding performance of AiSub, even in a scenario with high spatial resolution. Regarding the computational time, AiSub takes 230.76 s with 228.29 s for unmixing and 2.47 s for classification, while SVM, MLR, MLRsub, and MLRsub<sub>mod</sub> consume 43.92, 3.76, 4.98, and 3.93 s, respectively. To further illustrate the advantage of our proposed method, we visualize some the classification maps obtained in one of 20 Monte Carlo tests in Fig. 14.

In our third experiment, conducted using the Pavia University dataset, we randomly selected 130 samples per class from the original training set (this represents about 30% of the available labeled samples) for training purposes. The numerical metrics reported in Table VII indicate that the proposed method is the best among all the compared methods. For example, AiSub produces the highest OA (77.49%), surpassing SVM, MLR, MLRsub, and MLRsub<sub>mod</sub> by 0.65%, 4.38%, 11.61%, and 7.03%, respectively. It should be noticed that, in this particular case in which abundant samples are used for training, SVM consumes much more time (1075.86 s)—due to the required parameter optimization—in order to achieve comparable performance, while MLR, MLRsub, and MLRsub<sub>mod</sub> take 28.22, 6.60, and 5.71 s, respectively. In turn, our newly developed AiSub achieves the highest classification accuracy with a reasonable processing time of 242.96 s. To illustrate the

TABLE V  
CLASS SPECIFIC ACCURACIES, OA, AA,  $\kappa$  STATISTIC COEFFICIENT (IN PERCENT), AND COMPUTATIONAL TIME (IN SECONDS)  
OBTAINED IN THE FIRST EXPERIMENT OF THE PAVIA UNIVERSITY DATASET

20 samples per class (180 samples in total)						
	Train	Test	SVM	MLR	MLRsub	MLRsub <sub>mod</sub>
1.Aspahlt	20	6631	66.76	68.77	43.84	54.56
2.Meadows	20	18649	58.37	55.69	61.20	57.50
3.Gravel	20	2099	62.77	70.01	65.74	61.68
4.Trees	20	3064	<b>94.01</b>	93.53	84.94	82.51
5.Metal sheet	20	1345	95.71	92.54	<b>99.40</b>	77.64
6.Bare soil	20	5029	67.62	<b>82.16</b>	62.11	53.85
7.Bitumen	20	1330	87.16	<b>89.27</b>	83.79	81.39
8.Bricks	20	3682	78.87	<b>81.11</b>	56.80	70.50
9.Shadows	20	947	99.08	97.89	98.56	98.10
OA(%)			68.26( $\pm 5.98$ )	68.57( $\pm 4.26$ )	62.89( $\pm 4.90$ )	62.01( $\pm 4.21$ )
AA(%)			78.93( $\pm 3.14$ )	79.22( $\pm 1.89$ )	72.93( $\pm 1.55$ )	70.86( $\pm 2.28$ )
$\kappa$ (%)			59.81( $\pm 6.87$ )	60.79( $\pm 4.50$ )	54.05( $\pm 4.89$ )	53.12( $\pm 4.35$ )
time(s)			61.67	5.25	6.08	4.94
						Unmixing: 242.86 Classification: 4.90 Total: 247.76

Standard deviations are in parentheses, and the highest accuracies are shown in boldface.

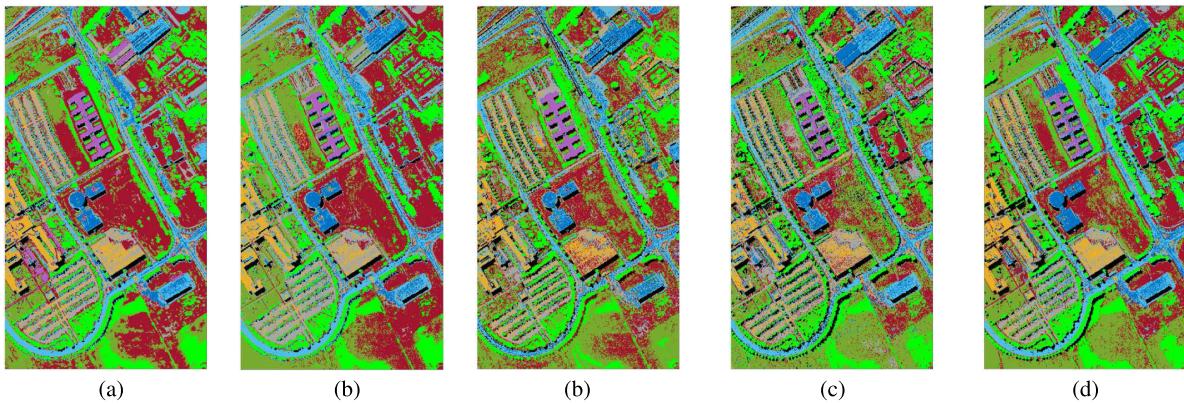


Fig. 13. Classification maps obtained in one of the 20 Monte Carlo tests by (a) SVM, (b) MLR, (c) MLRsub, (d) MLRsub<sub>mod</sub>, and (e) AiSub in the first experiment of the Pavia University dataset.

TABLE VI  
CLASS SPECIFIC ACCURACIES, OA, AA,  $\kappa$  STATISTIC COEFFICIENT (IN PERCENT), AND COMPUTATIONAL TIME (IN SECONDS)  
OBTAINED IN THE SECOND EXPERIMENT OF THE PAVIA UNIVERSITY DATASET

4 percent of 3921 samples (152 samples in total)						
	Train	Test	SVM	MLR	MLRsub	MLRsub <sub>mod</sub>
1.Aspahlt	21	6631	65.96	70.17	57.83	61.98
2.Meadows	21	18649	54.05	56.02	64.15	60.54
3.Gravel	15	2099	49.07	<b>62.29</b>	54.16	54.32
4.Trees	20	3064	<b>94.62</b>	93.92	72.16	79.22
5.Metal sheet	10	1345	86.45	75.56	98.37	78.05
6.Bare soil	21	5029	70.20	<b>79.11</b>	57.82	52.01
7.Bitumen	15	1330	75.13	84.09	65.12	73.92
8.Bricks	20	3682	<b>86.27</b>	83.79	58.74	76.57
9.Shadows	9	947	<b>93.99</b>	93.88	93.41	78.22
OA(%)			66.79( $\pm 6.25$ )	66.27( $\pm 3.02$ )	64.84( $\pm 4.99$ )	63.53( $\pm 4.80$ )
AA(%)			74.50( $\pm 4.70$ )	74.65( $\pm 1.75$ )	69.35( $\pm 2.91$ )	68.31( $\pm 3.18$ )
$\kappa$ (%)			59.50( $\pm 7.05$ )	58.58( $\pm 3.12$ )	55.44( $\pm 5.38$ )	54.44( $\pm 4.88$ )
time(s)			43.92	3.76	4.98	3.93
						Unmixing: 228.29 Classification: 2.47 Total: 230.76

Standard deviations are in parentheses, and the highest accuracies are shown in boldface.

effectiveness of our method, Fig. 15 displays the classification maps obtained in one of the 20 Monte Carlo tests conducted in this experiment.

Finally, some general conclusions can be drawn from the aforementioned experiments. First of all, among MLR-based

methods, AiSub achieves the best classification results in all cases. This observation reveals that the proposed abundance-indicated strategy can provide a more accurate estimation for posterior probability distribution. Furthermore, we can observe that AiSub exhibits a competitive advantage over SVM in

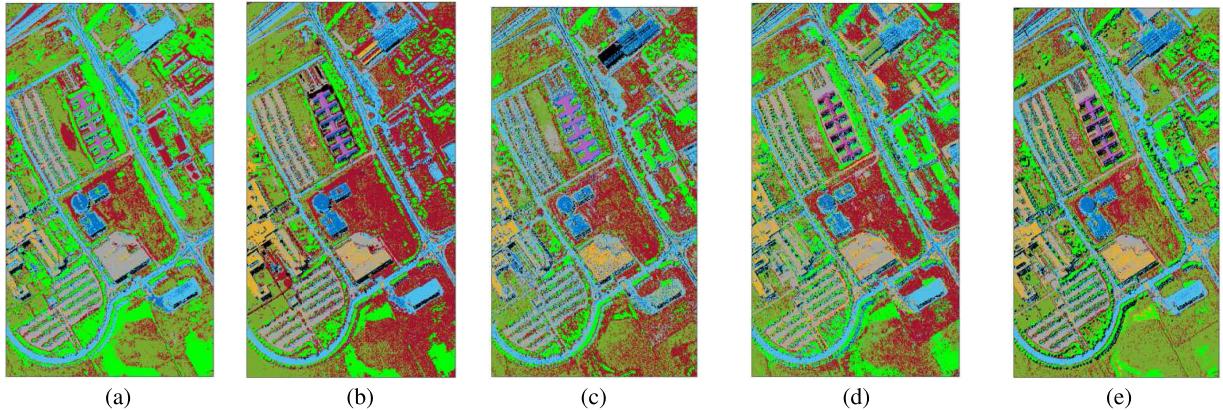


Fig. 14. Classification maps obtained in one of 20 Monte Carlo tests by (a) SVM, (b) MLR, (c) MLRsub, (d) MLRsub<sub>mod</sub>, and (e) AiSub in the second experiment of the Pavia University dataset.

TABLE VII  
CLASS SPECIFIC ACCURACIES, OA, AA,  $\kappa$  STATISTIC COEFFICIENT (IN PERCENT), AND COMPUTATIONAL TIME (IN SECONDS) OBTAINED BY AVERAGING 20 MONTE CARLO TESTS IN THE THIRD EXPERIMENT OF THE PAVIA UNIVERSITY DATASET

130 samples per class (30 percent of the original training set)						
	Train	Test	SVM	MLR	MLRsub	MLRsub <sub>mod</sub>
1.Ashphalt	130	6631	<b>78.53</b>	68.72	34.32	61.66
2.Meadows	130	18649	65.55	63.35	68.08	63.76
3.Gravel	130	2099	75.22	<b>76.44</b>	74.28	61.61
4.Trees	130	3064	95.28	88.79	70.00	89.39
5.Metal sheet	130	1345	99.46	99.95	99.65	97.90
6.Bare soil	130	5029	<b>91.41</b>	89.48	65.97	76.33
7.Bitumen	130	1330	<b>92.48</b>	87.56	88.91	86.49
8.Bricks	130	3682	<b>87.00</b>	74.42	59.49	80.75
9.Shadows	130	947	98.89	87.66	98.06	86.96
OA(%)			76.84( $\pm 1.38$ )	73.11( $\pm 0.98$ )	65.88( $\pm 2.85$ )	70.46( $\pm 3.72$ )
AA(%)			83.31( $\pm 0.75$ )	81.82( $\pm 0.68$ )	70.97( $\pm 1.69$ )	78.52( $\pm 1.52$ )
$\kappa$ (%)			70.48( $\pm 1.57$ )	66.65( $\pm 1.12$ )	63.37( $\pm 2.99$ )	66.25( $\pm 4.11$ )
time(s)			1075.68	28.22	6.60	5.71
						Unmixing: 241.22
						Classification: 1.74
						Total: 242.96

Standard deviations are in parentheses, and the highest accuracies are shown in boldface.

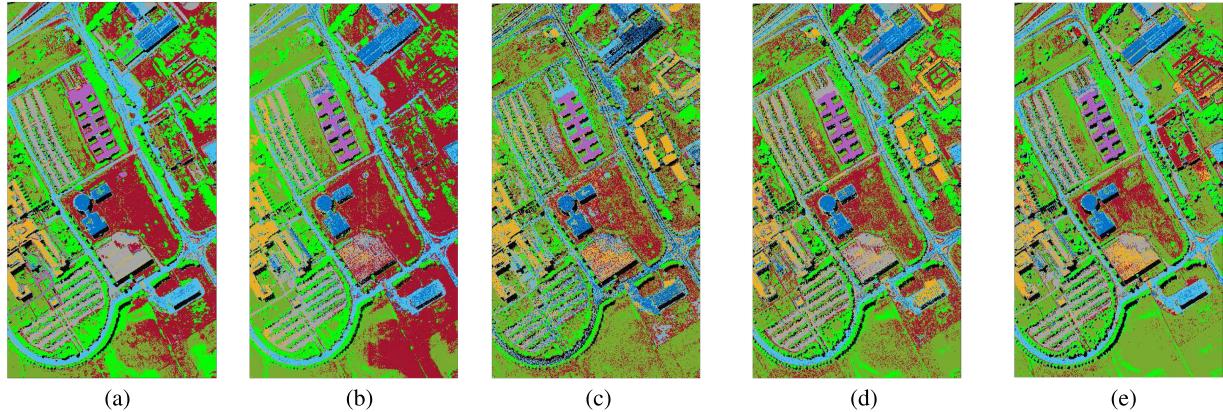


Fig. 15. Classification maps obtained in one of 20 Monte Carlo tests by (a) SVM, (b) MLR, (c) MLRsub, (d) MLRsub<sub>mod</sub>, and (e) AiSub in the third experiment of the Pavia University dataset.

terms of classification accuracy and computational time, even in scenarios dominated by high spatial resolution and abundant training samples. This suggests that the information coming from abundances can always serve as a physically meaningful complement to spectral information for classification purposes. Finally, as reported, AiSub is very fast for classification. Also, since the unmixing step can be accelerated by means of

preprocessing, AiSub has good potential for median-large-scale problems.

#### IV. CONCLUSION AND FUTURE LINES

In this paper, we present a new approach for hyperspectral image classification that integrates spectral unmixing and classi-

fication in order to characterize the phenomenon of mixed pixels with physical meaning. The proposed method is based on the assumption that a specific land cover type is related to a set of end-members representing the material components in the class. As a result, we exploit abundance fractions to indicate the endmembers for class-dependent subspaces, which are used in the MLR-based classification afterward. In the experiments conducted on synthetic and real hyperspectral images with very limited training samples, our proposed approach exhibits robustness without the need for parameter tuning, and produces results superior to other pixel-based classification methods. As expected, the proposed approach provides competitive advantages in scenarios with limited spatial resolution, since the information coming from fractional abundances can be used to characterize highly mixed pixels. In this sense, our experiments show that the proposed method can also provide better classification results in analysis scenarios characterized by high-spatial-resolution data and abundant training samples. However, further investigations should be devoted to consider unmixing approaches based on the nonlinear mixed model and with the assumption that no pure pixels are present in the scene, which are more illustrative for complicated environments. Moreover, including spatial-contextual information in our proposed method is expected to lead further improvements according to our preliminary test. Inspired by the observation of handling wrongly labeled samples, a semisupervised version of the proposed method, which jointly exploits labeled and unlabeled samples to indicate the class-dependent endmembers, will be developed in our future work.

#### ACKNOWLEDGMENT

The authors gratefully acknowledge the Editors and the Anonymous Reviewers for their detailed and constructive criticism, which has helped them improve the quality and presentation of this paper.

#### REFERENCES

- [1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [2] Chein-I Chang, *Hyperspectral Data Exploitation: Theory and Applications*. New York, NY, USA: Wiley-Interscience, 2007.
- [3] C. Chion, J. A. Landry, and L. Da Costa, "A genetic-programming-based method for hyperspectral data information extraction: Agricultural applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2446–2457, Aug. 2008.
- [4] Y. L. Chang, L. S. Liang, C. C. Han, J. P. Fang, W. Y. Liang, and K. S. Chen, "Multisource data fusion for landslide classification using generalized positive boolean functions," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1697–1708, Jun. 2007.
- [5] L. Zhang, L. Zhang, D. Tao, X. Huang, and B. Du, "Hyperspectral remote sensing image subpixel target detection based on supervised metric learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4955–4965, Aug. 2014.
- [6] 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.
- [7] 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. 2018.
- [8] Y. Gu, J. Chanussot, X. Jia, and J. A. Benediktsson, "Multiple kernel learning for hyperspectral image classification: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6547–6565, Nov. 2017.
- [9] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, no. 4704, pp. 1147–1153, 1985.
- [10] Q. Du, L. Zhang, B. Zhang, X. Tong, P. Du, and J. Chanussot, "Foreword to the special issue on hyperspectral remote sensing: Theory, methods, and applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 459–465, Apr. 2013.
- [11] G. Gao and Y. Gu, "Tensorized principal component alignment: A unified framework for multimodal high-resolution images classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 46–61, Jan. 2019.
- [12] X. Jin and Y. Gu, "Superpixel-based intrinsic image decomposition of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4285–4295, Aug. 2017.
- [13] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 1, pp. 55–63, Jan. 1968.
- [14] T. Imbiriba, J. C. M. Bermudez, C. Richard, and J. Y. Tourneret, "Nonparametric detection of nonlinearly mixed pixels and endmember estimation in hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1136–1151, Mar. 2016.
- [15] M. Khodadadzadeh, J. Li, A. Plaza, H. Ghassemian, J. M. Bioucas-Dias, and X. Li, "Spectral-spatial classification of hyperspectral data using local and global probabilities for mixed pixel characterization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6298–6314, Oct. 2014.
- [16] 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.
- [17] Z. Wu, Y. Li, A. Plaza, J. Li, F. Xiao, and Z. Wei, "Parallel and distributed dimensionality reduction of hyperspectral data on cloud computing architectures," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2270–2278, Jun. 2016.
- [18] Z. Wu *et al.*, "GPU parallel implementation of spatially adaptive hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1131–1143, Apr. 2018.
- [19] L. Zhang, W. Wei, Y. Zhang, C. Shen, A. van den Hengel, and Q. Shi, "Cluster sparsity field: An internal hyperspectral imagery prior for reconstruction," *Int. J. Comput. Vis.*, vol. 126, no. 8, pp. 797–821, Aug. 2018.
- [20] W. Wei, L. Zhang, Y. Jiao, C. Tian, C. Wang, and Y. Zhang, "Intracluster structured low-rank matrix analysis method for hyperspectral denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 866–880, Feb. 2019.
- [21] 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.
- [22] L. Zhang, W. Wei, C. Bai, Y. Gao, and Y. Zhang, "Exploiting clustering manifold structure for hyperspectral imagery super-resolution," *IEEE Trans. Image Process.*, vol. 27, no. 12, pp. 5969–5982, Dec. 2018.
- [23] Y. Y. Tang, H. Yuan, and L. Li, "Manifold-based sparse representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 12, pp. 7606–7618, Dec. 2014.
- [24] D. Hong, N. Yokoya, and X. X. Zhu, "Learning a robust local manifold representation for hyperspectral dimensionality reduction," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 6, pp. 2960–2975, Jun. 2017.
- [25] D. Lunga, S. Prasad, M. M. Crawford, and O. Ersoy, "Manifold-learning-based feature extraction for classification of hyperspectral data: A review of advances in manifold learning," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 55–66, Jan. 2014.
- [26] J. Xia, J. Chanussot, P. Du, and X. He, "(Semi-) supervised probabilistic principal component analysis for hyperspectral remote sensing image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2224–2236, Jun. 2014.
- [27] Y. Ren, L. Liao, S. J. Maybank, Y. Zhang, and X. Liu, "Hyperspectral image spectral-spatial feature extraction via tensor principal component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1431–1435, Sep. 2017.
- [28] M. D. Yang, K. S. Huang, Y. F. Yang, L. Y. Lu, Z. Y. Feng, and H. P. Tsai, "Hyperspectral image classification using fast and adaptive bidimensional empirical mode decomposition with minimum noise fraction," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1950–1954, Dec. 2016.

- [29] Y. Xu, Z. Wu, J. Li, A. Plaza, and Z. Wei, "Anomaly detection in hyperspectral images based on low-rank and sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 1990–2000, Apr. 2016.
- [30] 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.
- [31] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.
- [32] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM- and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.
- [33] F. Ratle, G. Camps-Valls, and J. Weston, "Semisupervised neural networks for efficient hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2271–2282, May 2010.
- [34] 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.
- [35] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4085–4098, Nov. 2010.
- [36] 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.
- [37] 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.
- [38] M. Khodadadzadeh, J. Li, A. Plaza, and J. M. Bioucas-Dias, "A subspace-based multinomial logistic regression for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2105–2109, Dec. 2014.
- [39] F. van der Meer, *Image Classification Through Spectral Unmixing*. Dordrecht, The Netherlands: Springer, 2002, pp. 185–193.
- [40] A. Plaza *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, no. Suppl. 1, pp. S110–S122, 2009.
- [41] Q. Meng, Y. Zhang, W. Wei, and L. Zhang, *Spectral Unmixing for Hyperspectral Image Classification with an Adaptive Endmember Selection*. Berlin, Germany: Springer, 2013, pp. 359–367.
- [42] M. Xu, B. Du, and L. Zhang, "Spatial-spectral information based abundance-constrained endmember extraction methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2004–2015, Jun. 2014.
- [43] A. Villa, J. Li, A. Plaza, and J. M. Bioucas-Dias, "A new semi-supervised algorithm for hyperspectral image classification based on spectral unmixing concepts," in *Proc. 3rd Workshop Hyperspectral Image Signal Process.: Evolution Remote Sens.*, Jun. 2011, pp. 1–4.
- [44] 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.
- [45] P. Chen, J. D. B. Nelson, and J. Y. Tourneret, "Toward a sparse Bayesian Markov random field approach to hyperspectral unmixing and classification," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 426–438, Jan. 2017.
- [46] 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.
- [47] A. J. Izenman, *Linear Discriminant Analysis*. New York, NY, USA: Springer, 2008, pp. 237–280.
- [48] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.
- [49] J. Li, I. Dopido, 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.
- [50] I. Dopido, 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.
- [51] M. Khodadadzadeh, J. Li, A. Plaza, and J. M. Bioucas-Dias, "Hyperspectral image classification based on union of subspaces," in *Proc. Joint Urban Remote Sens. Event*, Mar. 2015, pp. 1–4.
- [52] J. M. P. Nascimento and J. M. B. 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.
- [53] R. Heylen, D. Burazerovic, and P. Scheunders, "Fully constrained least squares spectral unmixing by simplex projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4112–4122, Nov. 2011.
- [54] L. Wang, X. Wang, and J. Feng, "Subspace distance analysis with application to adaptive Bayesian algorithm for face recognition," *Pattern Recognit.*, vol. 39, no. 3, pp. 456–464, 2006.
- [55] 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.: Evolution Remote Sens.*, Jun. 2010, pp. 1–4.
- [56] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4816–4829, Sep. 2013.
- [57] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, May 2011, Art. no. 27.



**Shuyuan Xu** received the B.S. degree in 2013 from Sun Yat-sen University, Guangzhou, China, where she is currently working toward the Ph.D. degree with the School of Geography and Planning.

Her research interests include hyperspectral feature extraction, image analysis, and machine learning with applications to remote sensing data analysis.



**Jun Li** (M'12–SM'16) received the Geographical Information Systems degree 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 Tecnico, Technical University of Lisbon, Lisbon, Portugal, in 2011.

From 2011 to 2012, she was a Postdoctoral Researcher with the Department of Technology of Computers and Communications, University of Extremadura, Badajoz, 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, China. She has obtained several prestigious funding grants at the national and international level. She has published a total of 69 journal citation report papers, 48 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 include remotely sensed hyperspectral image analysis, signal processing, supervised/semisupervised learning, and active learning.

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



**Mahdi Khodadadzadeh** received the B.S. degree in electrical engineering, electronics, from the Sadjad University of Technology, Mashhad, Iran, in 2008, the M.Sc. degree in electrical engineering, communications, from Tarbiat Modares University, Tehran, Iran, in 2011, and the Ph.D. degree in technology of computers and communications from Escuela Politécnica, University of Extremadura, Cáceres, Spain, under the joint supervision of Prof. A. Plaza and Prof. J. Li, in 2015.

From 2015 to 2017, he was a Postdoctoral Researcher with the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento, Trento, Italy. Since 2017, he has worked as a Machine Learning Specialist with the Exploration Technology Division, Helmholtz Institute Freiberg for Resource Technology, Freiberg, Germany. His research interests include machine learning and image processing with applications to remote sensing data analysis, with particular emphasis on hyperspectral data classification.

Dr. Khodadadzadeh is a Manuscript Reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.



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

He is currently a Professor of Telecommunications with the University of Pavia, where he leads the Telecommunications and Remote Sensing Laboratory and serves as a Deputy Coordinator of the Ph.D. School in Electronics and Computer Science.

Dr. Gamba has been invited to give keynote lectures and tutorials in several occasions about urban remote sensing, data fusion, Earth observation data, and risk management. He has been the Organizer and Technical Chair of the biennial GRSS/ISPRS Joint Workshops on Remote Sensing and Data Fusion over Urban Areas since 2001. He was the Chair of the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society from 2005 to 2009. He serves as a GRSS Vice President for Professional Activities, the Chair of the Chapters Committee, and the Chair of the Symposium Awards Committee. He was the Editor-in-Chief of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS from 2009 to 2013. He has served as the Guest Editor of special issues of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN REMOTE SENSING APPLICATIONS, the ISPRS Journal of Photogrammetry and Remote Sensing, the International Journal of Information Fusion, and the Pattern Recognition Letters on the topics of Urban Remote Sensing, Remote Sensing for Disaster Management, and Pattern Recognition in Remote Sensing Applications.



**Andrea Marinoni** (S'07–M'11–SM'16) received the B.S., M.Sc. (*cum laude*), and Ph.D. degrees in electronic engineering from the University of Pavia, Pavia, Italy, in 2005, 2007, and 2011, respectively.

He is an Associate Professor with the Earth Observation Group, Centre for Integrated Remote Sensing and Forecasting for Arctic Operations, Department of Physics and Technology, The Arctic University of Norway, Tromsø, Norway. From 2015 to 2017, he was a Visiting Researcher with the Earth and Planetary Image Facility, Ben-Gurion University of the Negev, Be'er Sheva, Israel; with the School of Geography and Planning, Sun Yat-Sen University, Guangzhou, China; with the School of Computer Science, Fudan University, Shanghai, China; with the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China; and with the Instituto de Telecomunicações, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal. His main research interests include efficient information extraction from multimodal remote sensing, nonlinear signal processing applied to large-scale heterogeneous records, Earth observation interpretation and Big Data mining, and analysis and management for human–environment interaction assessment.

Dr. Marinoni serves as a Reviewer for several major journals and conferences sponsored by the IEEE Geoscience and Remote Sensing Society, the IEEE Communications Society, the IEEE Signal Processing Society, the IEEE Information Theory Society, the International Association of Pattern Recognition, the International Society on Urban Health, and the American Medical Informatics Association.



**Bo Li** received the B.S. degree in computer science from Chongqing University, Chongqing, China, in 1986, the M.S. degree in computer science from Xi'an Jiaotong University, Xi'an, China, in 1989, and the Ph.D. degree in computer science from Beihang University, Beijing, China, in 1993.

He then joined the School of Computer Science and Engineering, Beihang University, where he is currently the Director of the Beijing Key Laboratory of Digital Media and the Director of the Professional Committee of Multimedia Technology of the China Computer Federation. He has authored or coauthored more than 100 academic papers in diverse research fields, including intelligent perception, data mining, remote sensing image fusion, and intelligent hardware.