SVM-Based Unmixing-to-Classification Conversion for Hyperspectral Abundance Quantification

Fereidoun A. Mianji, *Member, IEEE*, and Ye Zhang, *Member, IEEE*

***Abstract*—Need for *a priori* knowledge of the components comprising each pixel in a scene has set the endmember determination, rather than the endmember abundance quantification, as the primary focus of many unmixing approaches. In the absence of the information about the pure signatures present in an image scene, which is often the case, the mean spectra of the pixel vectors, directly extracted from the scene, are usually used as the pure signatures' spectra. This approach which is mathematically optimized for unmixing problems with *a priori* known information ignores some statistical properties of the extracted samples and leads to a suboptimal solution for real situations. This paper proposes a novel learning-based unmixing-to-classification conversion model to treat the abundance quantification task as a classification problem. Support vector machine, as an efficient classifier, is used to realize this model. It exploits the statistical nature (endmember spectral variability) of the extracted endmember representatives from the hyperspectral scene, rather than solving the problem according to the ideal model in which only the mean spectra of each training sample set is used. Several experiments are carried out on simulated and real hyperspectral images. The obtained results validate the high performance of the proposed technique in abundance quantification which is a key subpixel information detection capability.**

***Index Terms*—Abundance quantification, hyperspectral image, key information detection, support vector machine (SVM), unmixing-to-classification conversion model.**

I. Introduction

SPECTRAL unmixing is a procedure for decomposing the measured spectrum of a mixed pixel into a collection of constituent spectra, endmembers, and the corresponding abundance fractions. Mixed pixels are more likely to happen when spectral resolution is high and/or spatial resolution is low [1]. Owing to the use of high spectral resolution by hyperspectral imaging sensors, a hyperspectral image pixel is likely to be mixed by several substances or to contain a subpixel target. Consequently, unmixing has become increasingly challenging and important in hyperspectral data exploitation.

Unmixing is a special case of the generalized inverse problem that estimates parameters describing an object using an observation(s) of a signal. The applicability of proposed models and techniques for unmixing is highly dependent on the variety of circumstances and factors that give rise to mixed pixels. The outputs of spectral unmixing are important for identifying the material composition of mixtures and also a closely related important problem in hyperspectral processing, the target detection issue. Unmixing generally refers to a process which includes two steps: 1) an endmember extraction algorithm (EEA) to define the pure (or purest) pixels in the scene followed by, 2) an endmember abundance quantification algorithm (inversion step) to define the contribution of different endmembers in each pixel. It is worth mentioning that there is a distinction between pure signature and pure pixel, i.e., an endmember is an idealized pure signature such as water or clay for a class which may come from a calibrated database or spectral library. Hence, the endmember is a pure signature which does not have to appear as a pixel in the real hyperspectral image, and when it does, it is called pure pixel [1]. This implies that many spectrally similar pure pixels which are not exactly same (due to spectral variability) may be specified by a same pure signature. Throughout this paper, as it has been widely used in the remote sensing community, we use the term "endmember" to indicate pure signature supposedly extracted by an EEA.

There are two main models to deal with an unmixing problem: 1) the macroscopic mixture [2] which models a mixed pixel as a linear combination of materials with relative concentrations, and 2) the intimate spectral mixture which proposes a nonlinear mixing model of materials [3]. The intimate model can be linearized by a method proposed in [4]; thus, the linear spectral mixture analysis (LSMA) can be reasonably applied to unmixing problems. LSMA is a versatile technique which has shown success in solving a variety of problems including mixed pixel classification [5], [6], quantification [6], [7], and subpixel detection [8], [9].

The proposed approaches for unmixing can be categorized into two main categories in terms of number of processing steps, as follows:

1) Direct or one-shot unmixing methods which carry out the endmember extraction and abundance quantification simultaneously. A group of approaches, e.g., the minimum volume transform algorithm, unmixes hyperspectral images by exploiting the geometric features of mixtures where the observations from a scene are assumed to be in a simplex and the vertices correspond to the endmembers [10]. This category is often computationally expensive and assumes the presence of at least one pure pixel of each endmember in the scene. However, recently, some pure-pixel assumption-free algorithms such as minimum volume simplex analysis [11] and simplex

Manuscript received September 6, 2010; revised July 29, 2011; accepted August 20, 2011. Date of publication October 6, 2011; date of current version October 28, 2011. This work was supported by the Natural Science Foundation of China under the grants 60972143 and 60972144, and research fund for the doctorial program of higher education of China under Grant 20092302110033.

F. A. Mianji is with the National Radiation Protection Department, Tehran, Iran (e-mail: fmianji@gmail.com).

Y. Zhang is with the Institute of Image and Information Technology, School of Electronic and Information Engineering, Harbin Institute of Technology, Harbin 150001, China (e-mail: zhye@hit.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/TGRS.2011.2166766

identification via split augmented Lagrangian [12] with low computational complexity are introduced in this category. Another group of approaches utilizes the blind source separation (BSS) premise, e.g., using independent component analysis (ICA), for unsupervised unmixing [13], [14]. ICA's capability for separation of the endmembers is limited due to violation of the abundance-independence assumption, i.e., the sum of abundance fractions associated with each pixel is constant [15], [16]. Another proposed BSS method for hyperspectral unmixing is nonnegative matrix factorization (NMF) which decomposes the data into two nonnegative matrices (the elements are constrained to be nonnegative) [17]. NMF does not need the pure-pixel assumption; nevertheless, owing to the local minima caused by the nonconvexity of the objective function [18] and the signature variability imposed by the strong atmospheric and environmental variations [19], two additional constraints are needed to be applied into NMF to efficiently unmix hyperspectral data [20]: a) the piecewise smoothness (the spectrum and abundance of each endmember vary smoothly in wavelength and spatial spaces, respectively) [21], [22], and b) sparseness constraint (any endmember does not contribute to all pixels in the scene) [23]. Despite the promising achievements presented through one-shot approaches, they rather focus on enhancing endmember determination matter as the primary objective.

2) Indirect or two-step unmixing approaches. These approaches generally consist of an EEA and an abundance quantification algorithm. The widely used EEAs are introduced in references [11], [24]–[26]. Most of these approaches use convexity principal to determine endmembers. Various strategies such as least square estimation [7], [27], maximum likelihood (ML) estimation [28] and Bayesian estimation [29] can be used for abundance quantification. Among them, a widely used approach is fully constrained least squares (FCLS) algorithm which satisfies both the abundance sum-to-one constraint (ASC) and abundance nonnegativity constraint (ANC) [7], [30]–[32]. FCLS is a solution mathematically optimized for expression of mixed pixels in terms of *a priori* known pure signatures (endmembers), which is normally not the real case [33]. In real cases, the spectra of each endmember is often assumed to be the average spectra of the related representative samples directly extracted from the scene [34]. Therefore, despite enjoying a robust mathematical form (a Lagrangian multiplier), FCLS only adheres to the mean value of the training sample sets and ignores the other statistical properties of the extracted endmember representatives.

Recently, an interesting technique using extended support vector machine (SVM) which exploits the statistics of the training samples in inversion step is proposed by Wang and Jia [35]. This method adopts the concept of class cores introduced by Brown *et al.* [36] in which the location of a class core is estimated by considering the complete set of training samples as (relative) pure pixels of the defined class. The support vectors are the pixel vectors that are used to form the boundary between pure and mixed regions so as the two regions above or below the boundary lines are pure pixels (class core) and the region between the lines is associated with the mixed pixels. According to LSMA, the SVM decision boundary (in the middle of this region) becomes a 50% mixing line, and the linear decision rules for estimating the contribution of each endmember (in a two-class problem) are developed accordingly. In multiclass cases, the one-against-all approach can be adopted (Section II-A). The method is applicable using linear SVM (u-LSVM) for linearly separable data, as well as the more general form, nonlinear SVM (u_NLSVM), for linearly inseparable data which is often the real case.

In this paper, we propose a different approach for abundance quantification of hyperspectral images based on the idea of conversion of the inversion task into a classification problem with the help of SVM. The premise of this approach is simulating different class labels through constructing synthetic mixtures of endmembers using the available training samples. This approach exploits the statistical information embedded in the extracted training samples, endmember representatives, instead of simply taking their average spectra into account. As a result, it provides higher unmixing accuracies with a reasonable computational cost. The proposed method is referred to as unmixing-to-classification conversion model with SVM and noted as uccm_SVM in the rest of this paper.

The remainder of this paper is organized as follows. In Section II, SVM and the adopted design for this work is briefly described. This section discusses the theoretical and structural basis of uccm_SVM as well. Section III presents the experiments on simulated and real hyperspectral images for performance analysis and provides a discussion on the obtained results. Finally, the conclusion is drawn in Section IV.

II. Proposed Unmixing Technique (uccm_SVM)

A. Support Vector Machine (SVM)

SVM is an approved approach for classification of multispectral and hyperspectral images [37]–[44]. SVM is intrinsically a two-class classifier but some techniques for application of SVMs in multiclass classification problems are addressed in the literature [45], [46]. For a two-class problem with training set $\{x_i, y_i\}_{i=1}^{N}$ consisting of N vectors in the d-dimensional feature space $x_i \in R^d (i = 1, 2, \ldots, N)$ with associated targets $y_i \in \{-1, 1\}$, the discriminant function of the separating hyperplane for linearly separable classes is defined as

$$f(x) = w \cdot \Phi(x) + b \tag{1}$$

where $w \in R^d$ is the vector normal to the hyperplane, $b \in R$ is the bias, and the classification decision rule is according to sign $[f(x)]$. The aim is finding the optimal hyperplane which minimizes the cost function consisted of two criteria, namely margin maximization and error minimization. Cost function is defined as

$$\Psi(w, \xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \tag{2}$$

subjected to the following constraints:

$$y_i \times (w \cdot \Phi(x_i) + b) \geq 1 - \xi_i,\ \xi_i \geq 0,\ i = 1, 2, \ldots, N \quad (3)$$

where C is a regularization constant to control the shape of the decision boundary and ξ_i is a slack variable, both introduced to account for inseparable data. This constrained optimization problem can be translated (using a Lagrangian formulation) into a dual optimization leading to a quadratic programming problem [45]. In real applications, the data are seldom linearly separable in the input space. To further generalize the above method to linearly inseparable information, the data can be mapped into a higher dimensional feature space, i.e., $\Phi(x) \in R^{d'} (d' \ll d)$, where the separation is easier through an optimal hyperplane defined by a weight factor $w^* \in R^{d'}$ and a bias $b^* \in R$. The final result is a discriminant function expressed in the original dimensional feature space

$$f(x) = \sum_{i \in S} a_i y_i K(x_i, x) + b^* \quad (4)$$

where $K(\cdot, \cdot)$ is a kernel function and S is a subset of the indices $\{1, 2, \ldots, N\}$ corresponding to the nonzero Lagrange multipliers a_i's. The training samples associated to nonzero weights are called *support vectors (SVs)*.

B. Key Premise of uccm_SVM

The central idea of uccm_SVM is converting the abundance quantification problem (inversion step) into a classification problem through assigning a set of class labels to different possible abundances (0%–100%) of an endmember in a pixel. Such a modeling enables us to exploit the statistical properties of the training samples, i.e., extracted endmember representatives. The well-known one-against-all strategy, the most common design used in multiclass classification applications based on binary SVM classifiers, is adopted in uccm_SVM. For simplicity, let us assume we have two endmember classes ω_A and ω_B in a region of interest ROI : $\{x_i\}_{i=1}^N \in R^d$ of a hyperspectral image. Using the available training sample sets $A = \{a_i\}_{i=1}^m \in \omega_A$ and $B = \{b_i\}_{i=1}^n \in \omega_B$ (assuming $m \leq n$), we simulate artificial classes by producing synthetic training vector matrices $\{C_j \in \omega_{C_j}\}_{j=1}^{n_C}$, $n_C \in \{1, 2, \ldots\}$ with different compositions of ω_A and ω_B in two steps. First, we equalize the size of the training sets A and B simply by enlarging the smaller one through repeating some of its samples, and/or shrinking the larger one through trimming it, to obtain two $q \times d$ training vector matrices $A' = \{a_i\}_{i=1}^q$ and $B' = \{b_i\}_{i=1}^q$, $m \leq q \leq n$ of the same size. Then, $\{C_j\}_{j=1}^{n_C}$ are generated by adding A' and B' matrices in n_C different constrained-to-unity ratios as follows.

$$C_j = \Gamma_j A' + (1 - \Gamma_j) B',\ \Gamma_j = \frac{(n_C - j)}{(n_c)},$$
$$j = 1, \ldots, n_C - 1 \quad (5)$$
$$T = \left\{A', B', \{C_j\}_{j=1}^{n_C}\right\},\ C_j = \{c_{ji}\}_{i=1}^q,$$
$$c_{ji} = \Gamma_j a_i + (1 - \Gamma_j) b_i. \quad (6)$$

Now, we have $n_c + 2$ classes including two original endmember classes ω_A and ω_B and n_C artificial classes $\{\omega_{C_j}\}_{j=1}^{n_C}$ with known compositions of endmembers ω_A and ω_B. Therefore, each component of T is a training vector matrix which represents a predefined abundance of ω_A and ω_B in pixels. It is worth mentioning that this process is possible with mixed training sample sets too, i.e., if we have training sample sets $E = \{e_i\}_{i=1}^m$ and $F = \{f_i\}_{i=1}^n$ with known compositions of ω_A and ω_B, T can be constructed using the same method. After constructing T, the multiclass SVM classifier is trained with T's training vector matrices for classification of the pixels according to the abundance-related classes. Then, every pixel of the scene is analyzed by the trained SVM to define its belonging to one of the information classes $\{\omega_A, \omega_B, \omega_{c_j}\}$, $j \in \{1, 2, \ldots, n_c\}$, and finally, the obtained classification results are converted to abundance information for endmembers ω_A and ω_B.

uccm_SVM is a digitized quantification algorithm, i.e., it defines the abundances in prefixed proportion values. Therefore, the expected resolution is a key factor to define the number of required artificial classes n_c. For a resolution of $r\%$ for unmixing, n_c is calculated as follows:

$$n_c = \left\| \left(\frac{100}{r} - 1 \right) \right\|. \quad (7)$$

For instance, resolutions of 5% and 10% correspond to 19 and 9 artificial classes, respectively.

In a two-endmember problem as the above example, abundance fractions for both endmembers $\beta_{x_i}(\omega_A)$, $\beta_{x_i}(\omega_B)$, $i = 1, \ldots, N$ intrinsically satisfying sum-to-one and nonnegativity constraints are obtainable in one run. In a multiendmember problem, e.g., with three endmembers ω_A, ω_B, and ω_C with corresponding training sample sets A, B, and C, the algorithm should be applied three times in a one-against-rest fashion to estimate the abundances of all the endmembers. In each run for this case, "one" is the endmember (ω_{one}) to be quantified, and "rest" stands for the other endmembers. In this three-endmember example, when it is run for ω_C as ω_{one}, ω_A and ω_B are considered as ω_{rest_C} (or ω_{rest} in general) for which appropriate training sample set R_{rest_C} should be firstly prepared using A and B. The complete unmixing process for this three-endmember example can be described as follows:

1) Integrating A and B into $R_{\mathrm{A,B}} = \{A, B\}$ and then batching all the components in the form of a single vector matrix, R_{rest_C} is obtained;
2) Using C for C and R_{rest_C} for ω_{rest_C}, the multiendmember problem is transformed into a two-endmember problem;
3) After equalizing the size of C and R_{rest_C} (size of R_{rest_C} is likely to be larger, thus, can be taken as the reference), T_C (T for C) is constructed using C' and R'_{rest_C} according to (6);
4) SVM is trained using the training vector matrices of T_C, and classification is carried out to quantify the relative abundance of ω_C in all pixels $\beta^r_{x_i}(\omega_C)$;
5) Steps one to four are repeated for A and B to quantify their relative abundances $\beta^r_{x_i}(\omega_A)$, $\beta^r_{x_i}(\omega_B)$;

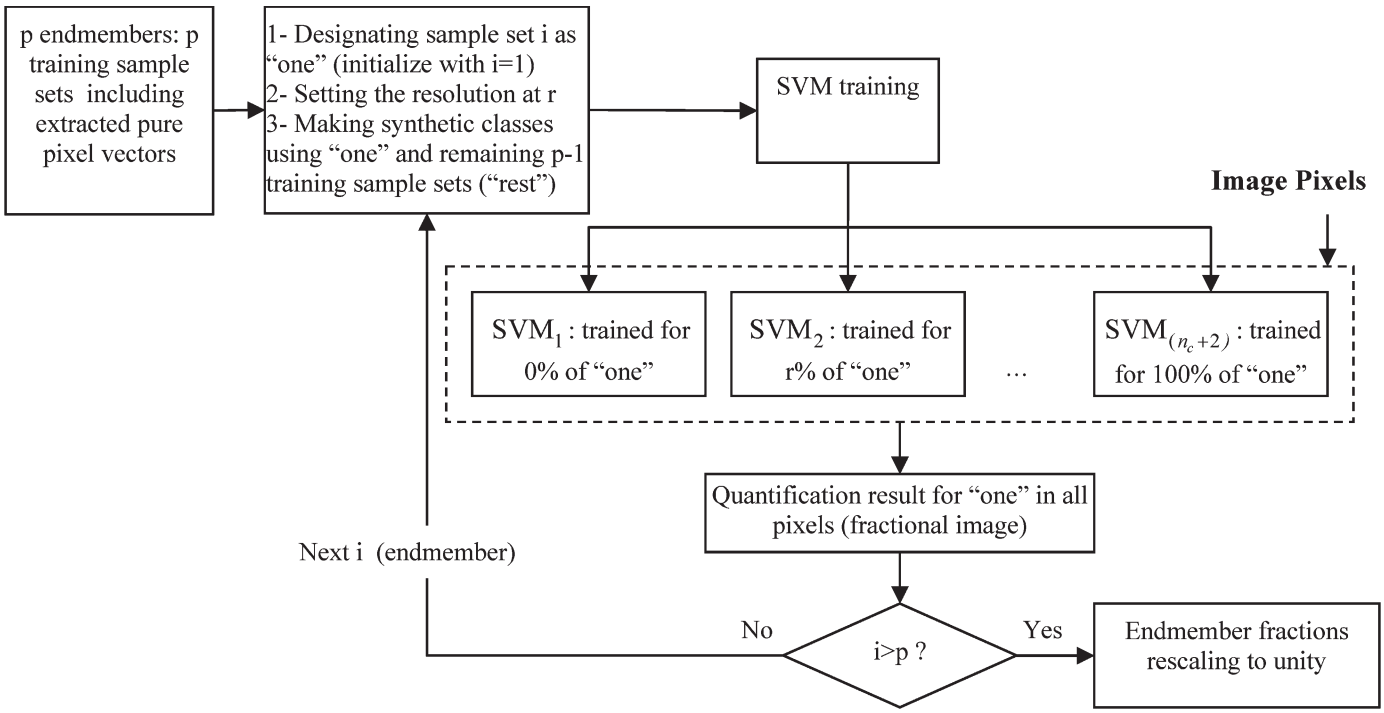

Fig. 1. Layout and process flow of uccm_SVM designed for unmixing with a resolution of r%.

6) The obtained results are normalized through a sum-to-one rescaling function, e.g., for $\beta^r_{x_i}(\omega_C)$ it is $\beta_{x_i}(\omega_C) = \beta^r_{x_i}(\omega_C)/(\beta^r_{x_i}(\omega_A) + \beta^r_{x_i}(\omega_B) + \beta^r_{x_i}(\omega_C))$.

The idea is similar in dealing with problems containing more than three endmembers. As the number of endmembers increases, R_{rest} gets bigger leading to more computational time for each run. However, in real scenarios where the number of training samples for each endmember is low and/or for reasonable practical values of resolution like $r = 10$ (which was validated in our experiments), the proposed method is noncomplex. The layout and process flow of uccm_SVM is shown in Fig. 1.

Assume that we have p endmembers in an ROI and we plan to synthesize necessary artificial classes for quantification of each endmember according to (5) and (6). The general equation for making the training sample set R_{rest} for abundance quantification of endmember $\omega_t \in \{\omega_i\}_{i=1}^p$ can be represented as follows:

$$R_{rest_t} = \{A_i : i \in \{1, 2, \ldots, p\}, i \neq t\} \tag{8}$$

where A_i are the training sample sets corresponding to ω_i. Therefore, for quantification of each endmember ω_t with $|R_{rest_t}|$ training samples, a total of n_{ts} labeled training samples are required to be generated in each run.

$$n_{ts} = |R_{rest_t}| \times (n_c + 2). \tag{9}$$

III. Experimental Design and Results

To validate the performance of the proposed algorithm, the experiments are designed in 3 parts: 1) with a simulated image constructed using ground truth information from a real hyperspectral image; 2) with a real hyperspectral image of an urbanized area; and 3) with a real hyperspectral image mainly consisting natural landcovers. Two very efficient techniques are used to benchmark the results: 1) FCLS algorithm as one of the most widely used inversion techniques in remote sensing [9], [11], [31], [32], [47], [48]; and 2) the recently proposed method by Wang and Jia, u_NLSVM [35].

For the experiments of this work, a Gaussian radial basis function kernel which has the following mathematical form was adopted for SVM:

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right). \tag{10}$$

Gamma (γ) is the inner product coefficient which tunes the smoothing of the discriminant function. The other important factor is the regularization parameter (C) in SVM. To reliably define the optimum values for γ and C, a five-fold cross-validation experimental framework was adopted in all the experiments with SVM (with the used class information). Estimated through a grid search using the sets 0.1, 0.2, ..., 3} for γ and $\{2^1, 2^2, \ldots, 2^{15}\}$ for C, γ and C were set to 1.1 and 2^9, 1.3 and 2^6, and 1.1 and 2^6, respectively for the experiments Sections III-B–D. The resolution for uccm-SVM was set to 10% in all the experiments though $r = 4\%$ or 5% results slightly better with about twice computational time. All the algorithms were programmed with MATLAB. The used computer was a laptop with an Intel Centrino Duo 1.83 GHz processor and 4 GB RAM.

For quantitative performance evaluation of the different algorithms, the root mean square error (RMSE) and correlation coefficient (CC) of obtained fractional images are calculated using the true fractional images. However, due to the almost similar expressions presented by these two measures over all the obtained results and for the sake of conciseness, only the average CC values besides the RMSE values are presented

(a) (b)

Fig. 2. ROSIS data sets: (a) Center of Pavia. (b) The University of Pavia.

Fig. 3. AVIRIS data set: Kennedy Space Center.

TABLE I
CENTER OF PAVIA DATA SET: INFORMATION CLASSES

Class Label	Class Name	# Training Samples	# Test Samples
1	Water	824	65278
2	Trees	820	6507
3	Meadows	824	2891
4	Self-Blocking Bricks	808	2140
5	Soil	820	6549
6	Asphalt	816	7525
7	Bitumen	808	7287
8	Tiles	1260	3122
9	Shadows	476	2165
Total		7456	103464

in this paper. Furthermore, owing to the plenty of fractional images in experiments with each data set, only the fractional images obtained in experiment Section III-C are presented for visual evaluation.

A. Data Sets

Three data sets used for the experiments include two hyperspectral images acquired by ROSIS-03 (Reflective Optics System Imaging Spectrometer) optical sensor and a hyperspectral image collected by the NASA Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) instrument. One of the ROSIS-03 data sets is a 1096 × 490 image acquired over the center of Pavia city in Italy. The other data set by this sensor is a 610 × 340 image acquired over the University of Pavia. The number of bands of ROSIS-03 sensor is 115 with a spectral coverage ranging from 0.43 to 0.86 μm and a spatial resolution of 1.3 meter per pixel. After removing water absorption and low SNR bands, 102 and 103 bands remain for the center of Pavia and the University of Pavia data sets, respectively. The third data set acquired by AVIRIS over the Kennedy Space Center (KSC), Florida, is a 512 × 614 image. It consists of 224 bands over the wavelength range from 0.4 to 2.5 μm and a spatial resolution of 18 m per pixel. After

TABLE II
UNIVERSITY OF PAVIA DATA SET: INFORMATION CLASSES

Class Label	Class Name	# Training Samples	# Test Samples
1	Asphalt	548	6641
2	Meadows	540	18649
3	Gravel	392	2099
4	Trees	524	3064
5	(Painted) Metal Sheets	265	1345
6	Soil	532	5029
7	Bitumen	375	1330
8	Self-Blocking Bricks	514	3682
9	Shadows	231	947
Total		3921	42776

TABLE III
KENNEDY SPACE CENTER DATA SET: INFORMATION CLASSES AND CLASSIFIED RESULTS

Class Label	Class Name	# Training Samples	# Pixels in Classified Image
1	Scrub	761	20674
2	Willow	243	4737
3	CP Hammok	256	11177
4	CP/Oak	252	12812
5	Slash pine	161	5128
6	Oak/Broad leaf	229	42071
7	Harwood swamp	105	1481
8	Graminoid marsh	431	21611
9	Spartina marsh	520	33128
10	Gattail marsh	404	11769
11	Salt marsh	419	24685
12	Mud flats	503	27334
13	Water	927	97761
Total		5211	314368

1	2	3	4	5	6	7	8	9
9	1	2	3	4	5	6	7	8
8	9	1	2	3	4	5	6	7
7	8	9	1	2	3	4	5	6
6	7	8	9	1	2	3	4	5
5	6	7	8	9	1	2	3	4
4	5	6	7	8	9	1	2	3
3	4	5	6	7	8	9	1	2
2	3	4	5	6	7	8	9	1

Fig. 4. Simulated image made from nine classes of center of Pavia data set.

TABLE IV
COMPARISON OF DIFFERENT ALGORITHMS IN TERMS OF RMSE FOR UNMIXING OF THE SIMULATED IMAGE

# Train Samp.	Method	Single Class and Average RMSE (%)										Ave. CC (%)	# SVs	n_{ts}	Time (s)	
		1	2	3	4	5	6	7	8	9	Ave.				Train	Test
2	FCLS	**4.17**	22.60	24.46	20.62	**20.19**	14.80	22.25	12.71	**7.44**	16.58	80.01	-	-	-	9.1
	u-NLSVM	9.49	28.38	29.32	23.22	22.46	**14.58**	**16.40**	20.50	9.82	19.35	73.68	47	-	0.03	0.12
	uccm-SVM	9.25	**18.79**	**14.01**	**19.58**	20.78	16.22	21.83	**10.89**	10.83	**15.80**	**84.13**	2533	352	3.3	2.2
5	FCLS	**4.71**	24.65	27.77	23.13	22.13	12.72	22.60	12.49	**5.99**	17.36	77.30	-	-	-	10.2
	u-NLSVM	9.39	**12.35**	14.54	22.19	18.62	14.41	**19.19**	16.72	8.73	15.13	85.59	48	-	0.03	0.13
	uccm-SVM	8.63	15.56	**13.27**	**17.65**	**17.35**	**11.57**	21.43	**9.70**	9.66	**13.87**	**87.58**	3312	440	3.8	2.6
10	FCLS	**4.83**	24.44	27.24	25.29	24.73	12.26	22.36	13.79	**5.79**	17.86	76.37	-	-	-	10.6
	u-NLSVM	9.33	**10.35**	13.06	22.75	19.45	13.41	**18.88**	16.34	8.99	14.73	86.29	59	-	0.03	0.14
	uccm-SVM	8.09	13.77	**11.41**	**16.01**	**16.28**	**10.35**	21.59	**9.49**	9.74	**12.97**	**89.02**	6129	880	9.3	6.5
30	FCLS	**4.86**	22.36	24.29	18.16	**13.16**	12.24	21.98	8.92	**6.03**	14.67	84.20	-	-	-	10.6
	u-NLSVM	9.35	19.43	19.86	25.30	19.09	13.22	**20.50**	**8.43**	7.87	15.89	83.94	61	-	0.05	0.15
	uccm-SVM	7.64	**14.91**	**13.89**	**14.72**	14.56	**10.24**	21.59	9.60	8.94	**12.90**	**88.96**	15761	2640	43.2	16.4
50	FCLS	**5.66**	22.26	23.91	16.95	**12.81**	12.96	20.87	8.78	**6.15**	14.48	84.92	-	-	-	10.8
	u-NLSVM	9.34	18.52	19.01	16.55	17.11	11.90	15.67	**7.78**	10.39	14.03	88.16	80	-	0.07	0.16
	uccm-SVM	7.66	**16.91**	**15.86**	**12.21**	14.09	**9.08**	**12.19**	10.24	8.82	**11.90**	**91.45**	26030	4400	134.2	51.9
100	FCLS	**6.98**	9.99	10.39	13.70	12.80	14.14	16.57	9.69	**5.94**	11.13	92.10	-	-	-	12.1
	u-NLSVM	9.44	**7.30**	**8.40**	12.78	14.43	12.25	16.33	**7.63**	10.71	11.03	92.89	115	-	0.23	0.16
	uccm-SVM	7.30	9.78	8.83	**8.88**	**9.02**	**9.20**	**10.87**	9.14	8.29	**9.04**	**95.52**	52987	8800	465.5	52.6

TABLE V
COMPARISON OF DIFFERENT ALGORITHMS IN TERMS OF RMSE FOR UNMIXING OF THE UNIVERSITY OF PAVIA DATA SET

# Train Samp.	Method	Single Class and Average RMSE (%)										Ave. CC (%)	# SVs	n_{ts}	Time (s)	
		1	2	3	4	5	6	7	8	9	Ave.				Train	Test
2	FCLS	32.30	**29.17**	**20.11**	26.33	10.82	**24.52**	14.73	17.62	27.83	22.60	54.52	-	-	-	223.4
	u-NLSVM	30.28	43.08	22.19	29.03	**9.53**	32.60	**12.94**	16.29	**5.20**	22.35	52.11	61	-	0.03	1.5
	uccm-SVM	**25.79**	36.45	21.09	**26.01**	9.77	24.79	14.66	**15.66**	8.31	**20.28**	**56.88**	1448	176	0.45	15.5
5	FCLS	28.47	27.02	17.96	**21.42**	8.84	**25.05**	13.79	16.70	21.19	20.05	62.27	-	-	-	218.5
	u-NLSVM	23.36	37.10	14.89	27.02	**8.04**	31.94	15.41	15.87	**4.89**	19.84	64.16	95	-	0.03	1.8
	uccm-SVM	**22.67**	**26.41**	**14.16**	23.77	9.89	25.80	**13.26**	**13.77**	9.63	**17.71**	**66.53**	3635	440	2.4	38.9
10	FCLS	29.67	28.70	18.09	22.66	9.55	25.36	15.54	16.75	21.78	20.90	60.23	-	-	-	239.8
	u-NLSVM	26.21	27.52	15.56	**21.74**	**8.64**	26.91	15.21	18.22	**4.65**	18.30	69.91	134	-	0.04	2.3
	uccm-SVM	**22.69**	**25.69**	**12.70**	24.64	9.07	**22.35**	**13.57**	**13.73**	11.03	**17.27**	**70.07**	7234	880	6.9	94.3
30	FCLS	28.61	28.56	16.69	22.26	**9.49**	25.38	14.07	16.33	19.00	20.04	62.28	-	-	-	183.2
	u-NLSVM	**19.82**	24.60	**11.51**	**19.32**	11.89	23.21	**11.33**	17.24	**5.71**	16.07	**75.99**	219	-	0.15	3.1
	uccm-SVM	21.70	**20.87**	11.81	22.66	9.54	**20.90**	11.95	**13.56**	10.98	**16.00**	73.57	21224	2640	48.6	318.5
50	FCLS	27.35	32.14	15.53	23.56	9.40	26.73	14.00	15.92	18.69	20.37	61.60	-	-	-	187.1
	u-NLSVM	**19.53**	22.96	**10.07**	**19.81**	**8.34**	23.68	13.01	17.66	**5.45**	15.61	**77.43**	280	-	0.21	3.8
	uccm-SVM	20.37	**19.46**	10.26	21.19	8.73	**18.84**	**11.00**	**12.22**	10.99	**14.78**	77.08	35369	4400	152.2	486
100	FCLS	29.11	28.69	15.53	22.38	9.31	25.95	14.92	16.67	18.22	20.09	62.52	-	-	-	186.3
	u-NLSVM	**18.89**	21.62	**10.90**	**17.00**	10.76	22.51	11.53	14.94	**6.02**	14.91	**79.23**	412	-	0.5	4.8
	uccm-SVM	19.72	**19.44**	11.11	20.31	**9.92**	**18.57**	**10.75**	**11.99**	10.44	**14.70**	76.91	69340	8800	511.1	985.2

removing water absorption and low SNR bands, 176 bands were used for the analysis. Figs. 2 and 3 show an RGB representation of the ROSIS and KSC data sets, and Tables I–III show the related information classes.

B. Experiments With the Simulated Hyperspectral Image Constructed With Real Observed Data

In this part, we evaluated the performance of the proposed algorithm using the ground truth information of a real hyperspectral image. Using 1764 labeled test samples of every class from the center of Pavia data set, we made 81 14 $\times$ 14 pixel images, i.e., nine 14 $\times$ 14 pixel images for every class. These small images were put together according to Fig. 4 to obtain a 126 $\times$ 126 simulated image of nine landcovers. The simulated image was then down sampled with a 3 $\times$ 3 mean filter to result a 42 $\times$ 42 image of pure and mixed pixels with known abundances of endmembers in every pixel. Then, using some separate pure pixels of the downsampled image, the different algorithms were trained, and the image was unmixed. The experiments were performed from very low to medium number of training samples, i.e., 2 to 100. Table IV depicts the single-class RMSE, class-average RMSE, and class-average CC values for different training set sizes.

The obtained results approve the competency of uccm-SVM on most of classes over all sizes of training set. More importantly, it always presents the lowest class-average RMSE and CC. In terms of computational complexity, the u-NLSVM is the fastest owing to its shortcut approach to the abundance quantification task. uccm-SVM is faster than FCLS for very low number of training samples and becomes more complex as the number of training samples increases. However,

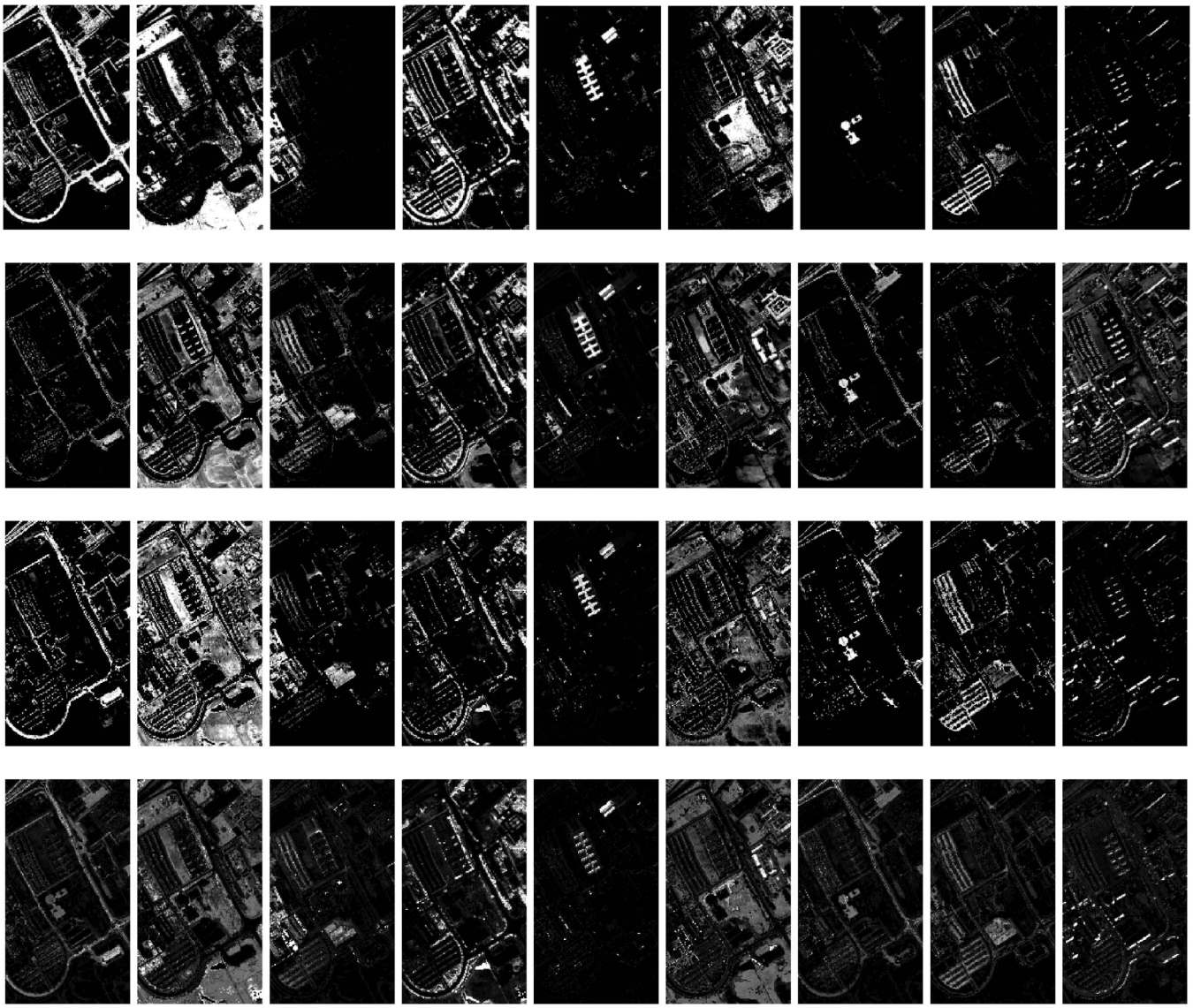

Fig. 5. Fractional images of the University of Pavia: from top to bottom: ground truth, results by FCLS, results by u-NLSVM, and results by uccm-SVM. From left to right classes 1 to 9.

appreciating the fact that only very limited training samples are often available in real scenarios, the proposed technique is a desirable choice for practical problems.

C. Experiments With Real Hyperspectral Images: Unmixing the University of Pavia Data Set

This data set includes nine classes. To obtain the ground truth map the image was classified using a ML classifier with the help of all available labeled training and test samples. To insure the high accuracy of the process, first, we divided the test samples into two groups of the same size. After training the ML classifier with one of these groups, the image was classified, and the performance of the algorithm was evaluated using the other group. We compared the results with the results obtained with a recently proposed SVM-based technique [49]. A quantitative evaluation in terms of classification accuracy followed by a spectral analysis on the few areas which were classified differently by these two techniques validated the higher accuracy of the ML algorithm (with half of the test samples used to train it). Finally, the image was classified using the ML algorithm with all the available training and test samples as the training samples to obtain the highest possible classification accuracy. The classified data was subsequently used as the ground truth map of the University of Pavia data set. To obtain an image with *a priori* known contribution of its landcovers in every pixel, we downsampled the original image (truncating the last row and column) using a 3×3 mean filter. Then, using 2 to 100 pure pixels (as separate as possible) of every class from the obtained 203×113 image, the algorithms were trained and applied for unmixing of it. The results are tabulated in Table V.

As is shown in Table V, the proposed technique performs better than FCLS and u-NLSVM in majority of single-class RMSEs and also in the class-average RMSE for all sizes of the training set (particularly for low numbers where it is the best in terms of CC too). As the number of training samples increases, FCLS only shows a little improvement while uccm-SVM and n-LSVM improve considerably. N-LSVM

TABLE VI
COMPARISON OF DIFFERENT ALGORITHMS IN TERMS OF RMSE FOR UNMIXING OF THE KENNEDY SPACE CENTER DATA SET

# Train Samp.	Method	Single Class and Average RMSE (%)														Ave. CC (%)	# SVs	n_{ts}	Time (s)	
		1	2	3	4	5	6	7	8	9	10	11	12	13	Ave.				Train	Test
2	FCLS	**16.97**	9.51	11.48	14.41	**7.42**	20.73	9.47	19.29	**19.22**	14.47	19.85	17.35	14.63	14.99	49.76	-	-	-	435.8
	u-NLSVM	17.59	8.98	11.29	12.01	8.42	25.48	7.45	23.33	21.93	20.25	**18.31**	20.17	13.12	16.02	44.55	316	-	0.03	10.1
	uccm-SVM	17.45	**7.51**	**11.25**	**11.58**	16.17	**20.19**	**6.62**	**17.72**	19.95	**12.93**	19.49	**16.65**	**12.36**	**14.61**	**51.01**	5770	720	3.7	182.3
5	FCLS	17.78	9.12	14.52	14.08	10.31	18.07	9.43	14.89	16.11	**11.17**	18.94	21.93	17.48	14.91	56.18	-	-	-	542.8
	u-NLSVM	16.15	6.89	12.12	**11.93**	**8.39**	23.46	**8.01**	18.16	16.30	16.82	**17.16**	21.11	16.91	14.88	54.73	647	-	0.8	19.4
	uccm-SVM	**17.01**	**6.74**	**11.20**	12.08	6.89	**18.98**	8.70	**14.57**	**14.95**	12.55	18.70	**18.63**	**15.04**	**13.54**	**58.32**	7485	900	5.1	256.3
10	FCLS	18.66	11.83	14.54	14.10	**7.12**	**20.14**	10.39	16.85	16.13	12.40	18.87	21.03	15.53	15.20	53.24	-	-	-	541.7
	u-NLSVM	**15.31**	6.01	**12.00**	11.22	11.34	21.66	**6.16**	18.87	19.39	13.20	**16.45**	19.12	**11.01**	13.98	60.27	1084	-	0.15	31.5
	uccm-SVM	16.28	**5.97**	12.01	**10.87**	9.37	20.47	6.90	**13.59**	**15.43**	**11.66**	19.94	**17.70**	13.99	**13.40**	**61.63**	14753	1800	19.2	503.3
30	FCLS	17.86	10.49	14.88	11.89	14.01	**17.05**	10.20	16.36	16.54	13.95	**15.95**	19.44	17.79	15.11	57.00	-	-	-	544.4
	u-NLSVM	14.75	6.58	10.37	10.59	10.34	19.36	**5.89**	16.85	16.29	14.22	16.29	16.01	**10.80**	12.95	67.20	2075	-	1.1	53.2
	uccm-SVM	**14.34**	**5.90**	**10.18**	**10.21**	**8.73**	17.76	6.23	**12.21**	**14.26**	**11.67**	16.62	**15.33**	12.55	**12.00**	**69.09**	41887	5400	220.3	1486
45	FCLS	17.90	10.09	15.13	11.32	14.11	**16.50**	9.72	16.97	17.25	12.46	**15.21**	21.04	20.63	15.26	56.92	-	-	-	546.6
	u-NLSVM	**13.68**	**5.48**	11.15	11.07	9.06	17.81	**6.44**	16.64	15.75	12.57	16.09	17.15	13.14	12.77	69.24	2656	-	2.2	77.8
	uccm-SVM	14.25	5.56	**10.77**	**10.31**	**8.11**	17.20	6.60	**11.70**	**15.13**	**10.76**	15.34	**15.07**	**12.11**	**11.76**	**70.05**	61659	8100	458.9	1816
90	FCLS	18.14	11.30	15.50	11.57	**8.09**	**15.83**	10.06	16.75	17.94	12.18	15.24	20.50	19.46	14.81	57.64	-	-	-	457.4
	u-NLSVM	13.55	**5.11**	10.08	10.57	10.28	17.51	**5.94**	14.59	15.57	12.61	16.27	16.30	12.14	12.35	72.88	4222	-	5.5	128
	uccm-SVM	**13.37**	5.51	**9.82**	**10.12**	8.95	17.42	6.44	**11.72**	**13.82**	**10.63**	**14.60**	**13.66**	**10.93**	**11.31**	**73.45**	117434	10800	1720	3987

presents the best average CC in higher number of training samples (more than 10) though always remains behind uccm-SVM in terms of average RMSE. Same as the experiments with simulated data the complexity of uccm_SVM is less than FCLS in very low and low number of training samples and gradually increases with the size of training set. The obtained fractional images by the three techniques and the ground truth maps are shown in Fig. 5 for visual comparison.

D. Experiments With Real Hyperspectral Images: Unmixing Kennedy Space Center Data Set

The third series of experiments were carried out using KSC Data set. Discrimination of landcovers for this environment is difficult due to the similarity of spectral signatures for certain vegetation types. As the ground truth information, the classification results provided by the University of Texas at Austin were used [50]. To obtain an image with *a priori* known contribution of its landcovers in every pixel, we downsampled the original image (truncating the last two rows and columns) by a 3×3 mean filter. Then, using 2 to 90 maximally separate pure pixels of every class of the obtained 170×204 image, the algorithms were trained and applied for the unmixing task. Since the number of pure pixels of classes 2, 5, and 7 in the downsampled image are 48, 92, and 97, respectively, the training set sizes 50 and 100 were replaced with 45 and 90 in this part. In the case of 90, the training set size for class 2 is 45.

As is tabulated in Table VI, the proposed technique performs better than FCLS and u-NLSVM over the whole range of training set sizes in majority of the classes and in all class averages. It is worth mentioning that despite the sensitivity of the SVM-based algorithms to the Hughes phenomenon, our technique works very well even with very low number of training samples. This is owing to the plenty of artificially made training samples. In terms of the computational time, uccm_SVM is better than FCLS in low number of training samples and gradually becomes more time consuming as the size of training set increases. As the previous experiments, n-LSVM's performance is often between FCLS and uccm-SVM; however, it is the fastest.

IV. DISCUSSIONS AND CONCLUSION

A novel learning-based unmixing technique is proposed in this paper. An unmixing-to-classification conversion model which transforms the unmixing task into a classification problem is the main premise of this method. Abundance quantification for each endmember is performed in a one-against-rest framework using a number of artificial classes. The artificial classes are constructed using different mixtures of the extracted training samples according to the model parameters and are labeled corresponding to the *a priori* known compositions of the endmembers in each class. The key advantage of the proposed method is exploiting the statistical nature of the extracted pure pixels for hyperspectral unmixing. This is in contrast to the traditional approaches which often only use the average spectra of the extracted pure samples. To realize this approach, a multiclass SVM classifier is successfully adopted into the proposed model. Experiments on simulated and real hyperspectral images validate the higher performance of the proposed technique compared with the efficient traditional mean-based unmixing techniques and newly proposed extended SVM algorithm.

ACKNOWLEDGMENT

Authors would like to highly appreciate Prof. X. Jia from the University of New South Wales, Prof. J. Richards from The Australian National University, and the anonymous reviewers for their helpful comments. The authors also would like to thank Prof. P. Gamba of the University of Pavia, Italy, for providing reference data.

REFERENCES

[1] J. Wang and C.-I Chang, "Applications of independent component analysis in endmember extraction and abundance qualification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2601–2616, Sep. 2006.

[2] R. B. Singer and T. B. McCord, "Mars: Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance," in *Proc. 10th Lunar Planet. Sci. Conf.*, 1979, pp. 1835–1848.

[3] B. Hapke, "Bidirection reflectance spectroscopy. I. Theory," *J. Geophys. Res.*, vol. 86, no. B4, pp. 3039–3054, 1981.

[4] P. Johnson, M. Smith, S. Taylor-George, and J. Adams, "A semiempirical method for analysis of the reflectance spectra of binary mineral mixtures," *J. Geophys. Res.*, vol. 88, no. B4, pp. 3557–3561, 1983.

[5] C.-I Chang, X. Zhao, M. L. G. Althouse, and J.-J. Pan, "Least squares subspace projection approach to mixed pixel classification in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 898–912, May 1998.

[6] M. Brown, H. G. Lewis, and S. R. Gunn, "Linear spectral mixture models and support vector machines for remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2346–2360, Sep. 2000.

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

[8] C.-I Chang and D. Heinz, "Constrained subpixel detection for remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1144–1159, May 2000.

[9] Y. F. Gu, Y. Zhang, and J. P. Zhang, "Integration of spatial-spectral information for resolution enhancement in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1347–1358, May 2008.

[10] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 542–552, May 1994.

[11] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.

[12] J. M. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," in *Proc. 1st IEEE GRSS WHISPERS*, Grenoble, France, 2009, pp. 1–4.

[13] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. Chichester, U.K.: Wiley, 2002.

[14] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001.

[15] J. M. P. Nascimento and J. M. Bioucas-Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[16] S. Jia and Y. Qian, "Spectral and spatial complexity-based hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3867–3879, Dec. 2007.

[17] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear Algebra Appl.*, vol. 416, no. 1, pp. 29–47, Jul. 2006.

[18] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Advances in Neural Information Processing Systems*, vol. 13. Cambridge, MA: MIT Press, 2000, pp. 556–562.

[19] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[20] M. Chu, F. Diele, R. Plemmons, and S. Ragni, Optimality, Computations and Interpretation of Nonnegative Matrix Factorizations, Oct. 2004. [Online]. Available: http://www.wfu.edu/~plemmons

[21] D. Gu, A. R. Gillespie, A. B. Kahle, and F. D. Palluconi, "Autonomous atmospheric compensation (AAC) of high resolution hyperspectral thermal infrared remote-sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 6, pp. 2557–2570, Nov. 2000.

[22] L. O. Jimenez, J. L. Rivera-Medina, E. Rodriguez-Diaz, E. Arzuaga-Cruz, and M. Ramirez-Velez, "Integration of spatial and spectral information by means of unsupervised extraction and classification for homogenous objects applied to multispectral and hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 844–851, Apr. 2005.

[23] S. Jia and Y. Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 161–173, Jan. 2009.

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

[25] 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.

[26] A. Plaza, P. Martinez, R. M. Perez, and J. Plaza, "A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles," *Pattern Recognit.*, vol. 37, no. 6, pp. 1097–1116, Jun. 2004.

[27] A. Villa, J. Chanussot, J. A. Benediktsson, and C. Jutten, "Spectral unmixing for the classification of hyperspectral images at a finer spatial resolution," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 521–533, May 2011.

[28] J. Settle, "On the relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, pp. 1045–1056, Jul. 1996.

[29] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2684–2695, Jul. 2008.

[30] C.-I Chang, H. Ren, C.-C. Chang, J. O. Jensen, and F. D'Amico, "Estimation of subpixel target size for remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1309–1320, Jun. 2004.

[31] C.-I Chang and B. Ji, "Weighted abundance-constrained linear spectral mixture analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 378–388, Feb. 2006.

[32] F. A. Mianji, Y. Gu, Y. Zhang, and J. Zhang, "Enhanced superresolution mapping technique for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 671–675, Jul. 2011.

[33] O. Duran and M. Petrou, "Spectral unmixing with negative and superunity abundances for subpixel anomaly detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 1, pp. 152–156, Jan. 2009.

[34] S. A. Robila and L. G. Maciak, "Considerations on parallelizing nonnegative matrix factorization for hyperspectral data unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 1, pp. 57–61, Jan. 2009.

[35] L. Wang and X. Jia, "Integration of soft and hard classifications using extended support vector machines," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 543–547, Jul. 2009.

[36] M. Brown, H. Lewis, and S. Gunn, "Linear spectral mixture models and support vector machine for remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2346–2360, Sep. 2000.

[37] A. Mathur and G. M. Foody, "Crop classification by a SVM with intelligently selected training data for an operational application," *Int. J. Remote Sens.*, vol. 29, no. 8, pp. 2227–2240, Apr. 2008.

[38] J. Muñoz-Marí, L. Bruzzone, and G. Camps-Valls, "A support vector domain description approach to supervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 8, pp. 2683–2692, Aug. 2007.

[39] B. Demir and S. Erturk, "Clustering-based extraction of border training patterns for accurate SVM classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 840–844, Oct. 2009.

[40] L. Bruzzone and M. Marconcini, "Toward the automatic updating of land-cover maps by a domain-adaptation SVM classifier and a circular validation strategy," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1108–1122, Apr. 2009.

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

[42] A. Plaza, J. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sens. Environ.*, vol. 113, pp. s110–s122, 2009.

[43] B.-C. Kuo, C.-H. Li, and J.-M. Yang, "Kernel nonparametric weighted feature extraction for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1139–1155, Apr. 2009.

[44] D. Tuia, G. Camps-Valls, G. Matasci, and M. Kanevski, "Learning relevant image features with multiple-kernel classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3780–3791, Oct. 2010.

[45] 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.

[46] F. A. Mianji and Y. Zhang, "Robust hyperspectral classification using relevance vector machine," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2100–2112, Jun. 2011.

[47] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.

[48] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.
[49] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.
[50] Y. Chen, Center for Space Research, The University of Texas at Austin. [Online]. Available: http://www.csr.utexas.edu/

Fereidoun A. Mianji (S'07–M'10) received the B.S. degree in electronics engineering from Guilan University, Rasht, Iran, the M.S. degree in radiomedical engineering from the Sharif University of Technology, Tehran, Iran, and Ph.D. degree in information and communication engineering from the Harbin Institute of Technology, Harbin, China, in 1992, 1996, and 2010, respectively.

He currently is with the National Radiation Protection Department of Iran. His research interests focus on digital image analysis, hyperspectral image processing and application, and medical imaging.

Zhang Ye (M'09) received the B.S. degree in communication engineering, and the M.S. and Ph.D. degrees, both in communication and electronic system, from the Harbin Institute of Technology (HIT), Harbin, China, in 1982, 1985, and 1996, respectively.

Since 1985, he joined HIT as a teacher. Between 1998 and 1999, he was a Visiting Scholar in the University of Texas at San Antonio. His research interests are remote sensing hyperspectral image analysis and processing, image/video compression and transmission, as well as multisource information collaboration processing and applications. Currently, he is a Professor and doctoral supervisor in information and communication engineering. He is the Director of Institute of Image and Information Technology in the school of Electronic and Information Engineering, HIT.