A Quantitative and Comparative Analysis of Endmember Extraction Algorithms From Hyperspectral Data

Antonio Plaza, Pablo Martínez, Rosa Pérez, and Javier Plaza

***Abstract*—Linear spectral unmixing is a commonly accepted approach to mixed-pixel classification in hyperspectral imagery. This approach involves two steps. First, to find spectrally unique signatures of pure ground components, usually known as endmembers, and, second, to express mixed pixels as linear combinations of endmember materials. Over the past years, several algorithms have been developed for autonomous and supervised endmember extraction from hyperspectral data. Due to a lack of commonly accepted data and quantitative approaches to substantiate new algorithms, available methods have not been rigorously compared by using a unified scheme. In this paper, we present a comparative study of standard endmember extraction algorithms using a custom-designed quantitative and comparative framework that involves both the spectral and spatial information. The algorithms considered in this study represent substantially different design choices. A database formed by simulated and real hyperspectral data collected by the Airborne Visible and Infrared Imaging Spectrometer (AVIRIS) is used to investigate the impact of noise, mixture complexity, and use of radiance/reflectance data on algorithm performance. The results obtained indicate that endmember selection and subsequent mixed-pixel interpretation by a linear mixture model are more successful when methods combining spatial and spectral information are applied.**

***Index Terms*—Comparative and quantitative framework, endmember extraction, spatial/spectral analysis, spectral mixture analysis.**

I. Introduction

THE ADVENT OF hyperspectral technology has introduced a completely new perspective in many remote sensing applications. One of the most important hyperspectral imagers in current use is the National Aeronautics and Space Administration (NASA)/Jet Propulsion Laboratory's 224-band Airborne Visible and Infrared Imaging Spectrometer (AVIRIS), which covers the wavelength region from 0.38–2.5 mm at a nominal spectral resolution of 10 nm [1]. Linear spectral unmixing [2] is a commonly accepted approach to analysis and classification of the massive volume of data provided by available imaging spectrometers. Such a method applies a linear mixture model to estimate the abundance fractions of spectral signatures within mixed pixels. These pixels involve a mixture of more than one distinct substance, and they exist for one of two reasons. First, if the spatial resolution of the sensor is not high enough to separate different materials, these can jointly occupy a single pixel, and the resulting spectral measurement is a composite of the individual spectra. Second, mixed pixels may also result when distinct materials are combined into homogeneous (intimate) mixtures. It has been reported that the reflectance spectrum of a mixed pixel is a systematic combination of the component spectra in the mixture [3]. The systematics tend to be linear if components of interest in a pixel appear in spatially segregated patterns. If, however, the components are in intimate association, light typically interacts with more than one component as it is multiply scattered, and the mixing systematics between the different components are essentially nonlinear. Although subpixel nonlinear mixing can be important for some types of analysis, the effects of multiple scattering in the majority of applications are assumed to be negligible if a linear model is used [4].

The linear mixture model can be defined as follows. Let $\boldsymbol{h}(x,y)$ be the hyperspectral signature collected by the sensor at the pixel with spatial coordinates (x,y). This signature can be considered an N-dimensional (N-D) vector, where N is the number of spectral bands; it can also be modeled as a linear combination of endmember vectors e_i, $i=1,\ldots,E$ using the following expression:

$$\boldsymbol{h}(x,y)=\sum_{i=1}^{E}\Phi_i(x,y)\cdot\boldsymbol{e}_i \tag{1}$$

where $\Phi_i(x,y)$ is a scalar value representing the fractional coverage of endmember vector e_i at pixel $\boldsymbol{h}(x,y)$. Two constraints are usually imposed in the previous equation [2]. These are the abundance nonnegativity constraint (ANC) and abundance sum-to-one constraint (ASC), respectively defined as

$$\Phi_i(x,y)\geq 0, \qquad \text{for all } 1\leq i\leq E \tag{2}$$

$$\sum_{i=1}^{E}\Phi_i(x,y)=1. \tag{3}$$

The key task in linear spectral mixture analysis is to find an appropriate suite of pure spectral signatures (endmembers), which are then used to model at-sensor pixel spectra through a linear combination of endmember signatures. The selection of endmembers can be performed in two ways: 1) by deriving them directly from the image (image endmembers) or 2) from field or

Manuscript received January 21, 2003; revised September 16, 2003. This work was supported in part by the Spanish Government under Grant TIC-2000-0739-C04-03 and in part by the Junta de Extremadura under Grant 2PR01A085.

The authors are with the Neural Networks and Signal Processing Group, Computer Science Department, Escuela Politécnica, University of Extremadura, 10071 Cáceres, Spain (e-mail: aplaza@unex.es; pablomar@unex.es; rosapere@unex.es; jplaza@unex.es).

Digital Object Identifier 10.1109/TGRS.2003.820314

laboratory spectra of known target materials (library endmembers); see [5] for a comparison between the two. The risk in using library endmembers is that these spectra are rarely acquired under the same conditions as the airborne data. Image endmembers have the advantage of being collected at the same scale as the data and can, thus, be more easily associated with features on the scene [2].

A number of algorithms have been developed over the past decade to accomplish the task of finding appropriate image endmembers for spectral mixture analysis. It should be taken into account that the presence of pure class pixels in the image data depends on available sensor spatial resolution. As a result, there may be cases where it is not possible for a certain algorithm to find such pure pixels in a scene. In those situations, the fractional components found for the mixed pixels are usually expressed in terms of other mixed pixels (the endmembers identified by the algorithm) and not in terms of pure classes. With the above statements in mind, and taking into account the increasing number of endmember identification methods readily available, the need for standardized strategies to evaluate the quality of selected endmembers has been identified as a desired goal by the scientific community devoted to hyperspectral data analysis. One of the simplest evaluation approaches has been the comparison of endmembers with available ground-truth spectra [6]. Spectral mixture analysis utilizes the high dimensionality of hyperspectral imagery to produce a suite of abundance fraction images for each endmember. Each fraction map shows a subpixel estimate of endmember relative abundance, as well as the spatial distribution of the endmember. Therefore, the quality of a suite of endmembers can also be evaluated by looking at the spatial distribution of fractional abundances, i.e., by comparing estimated abundance fractions to existing ground-truth abundance maps [7]. The comparative approaches mentioned above represent complementary strategies that can be applied to decide whether a method works correctly from two different points of view (spectral and spatial). However, these approaches are only possible when high-quality ground-truth information concerning the original scene is available. We must also consider that the generation of reliable ground-truth in real scenarios is difficult and expensive [8], a fact that has traditionally prevented the existence of comparative surveys using large databases of real images. In order to avoid this shortcoming, simulation of hyperspectral imagery has been suggested as a simple and intuitive way to perform a preliminary evaluation of analysis techniques [9]. The primary reason for the use of simulated imagery as a complement to real data analysis is that all details of the simulated images are known. These details can be efficiently investigated because they can be manipulated individually and precisely. As a result, algorithm performance can be examined in a controlled manner.

Experiments using simulated hyperspectral data have been previously reported in the literature, e.g., see [7] and [10]–[12]. Although the number of research studies involving experimentation with simulated imagery is relevant, there is a lack of commonly accepted data that can be used to evaluate individual endmember selection algorithms. In addition, no unified criterion has been accepted for rigorous and impartial comparison of these algorithms. The importance of this issue cannot be understated since, without such data and effective evaluation criteria, the performance of any new method cannot be substantiated [13]. In this paper, we take a first step by conducting a comparative study of performance analysis among several endmember extraction algorithms using both simulated and real hyperspectral data in radiance and reflectance units. The major contributions of this work are 1) the development of a framework for experimental comparison of endmember selection algorithms and 2) an assessment of the state of the art for endmember identification by drawing comparisons between substantially different approaches to the problem in rigorous fashion, so that each method is fairly compared with others on the same common ground.

The following section presents an overview of various endmember extraction algorithms, along with a brief description of those techniques that will be compared in this work. Section III describes the image database used in the present study and the quantitative and comparative framework used for evaluation of endmember selection and subsequent mixed-pixel classification accuracy. In Section IV, a comparative performance analysis for the algorithms described in Section II is presented and discussed. Section V points out main concluding statements derived from this paper.

II. Endmember Extraction Techniques

During the last decade, several algorithms have been proposed for the purpose of autonomous/supervised endmember selection from hyperspectral scenes. Table I shows an overview of the main characteristics of several available methods, including the manual endmember selection tool (MEST) [14], pixel purity index (PPI) [15], N-FINDR [6], iterative error analysis (IEA) [16], optical real-time adaptative spectral identification system (ORASIS) [17], convex cone analysis (CCA) [7], automated morphological endmember extraction (AMEE) [18], and simulated annealing algorithm (SAA) [19]. As deducible from Table I, the following observations are made: 1) only three of the methods listed have been validated by using both simulated and experimental data and 2) few available methods have been rigorously compared with other approaches. We provide next a brief overview of the methods that will be assessed in the present study.

A. PPI

The PPI algorithm is characteristic in its supervised nature; it consists of the following steps. First, a "noise-whitening" and dimensionality reduction step is performed by using the MNF transform [15]. Then, a pixel purity score is calculated for each point in the image cube by randomly generating L lines in the N-D space comprising the MNF-transformed data. All the points in that space are projected onto the lines, and the ones falling at the extremes of each line are counted. After many repeated projections to different random lines, those pixels that count above a certain cutoff threshold C are declared "pure." These potential endmember spectra are loaded into an interactive N-D visualization tool and rotated in real time until a desired number of endmembers E are visually identified as extreme pixels in the data cloud.

TABLE I
SUMMARY OF AVAILABLE ENDMEMBER EXTRACTION ALGORITHMS

Method	Nature of algorithm	Datasets used in evaluation	Ground truth used in evaluation	Algorithm assumptions	Computational complexity	Convergence property	Most probable applications	Use of spatial info	Compared against
MEST [14]	Supervised, totally interactive	None	None	PCA-based reduction in the data	Depends on manual supervision	Pre-defined number of endmembers	Land-cover and mineral mapping	No	—
PPI [15]	Supervised, partially interactive	None	None	MNF-based reduction in the data	High number of iterations required	Maximum number of iterations	Land-cover and mineral mapping	No	—
N-FINDR [6]	Fully automated, non-parameterized	Simulated and real data	Reference spectra, abundance fractions	None	Depends on initial random pixel selection	Simplex with the maximum volume found	Land-cover and mineral mapping	No	ORASIS, IEA
IEA [16]	Fully automated, parameterized	Real data	Reference spectra	None	Significant due to repeated un-mixing steps	Minimum error in unmixing	Land-cover and mineral mapping	No	ORASIS
AMEE [18]	Fully automated, parameterized	Simulated and real data	Reference spectra, abundance fractions	None	Moderate for medium-sized kernel sizes	Range of kernel sizes used	Land-cover classification, target detection	Yes	PPI, N-FINDR
ORASIS [17]	Fully automated, parameterized	Real data	None	None	Low-algorithm tuned for rapid execution	All exemplar spectra found in the data	Oceanography and target detection	No	PPI
CCA [7]	Fully automated, parameterized	Simulated and real data	Reference spectra, abundance fractions	None	Very high if many corners in convex cone	All corners of the convex cone found	Target and anomaly detection	No	—
SSA [19]	Fully automated, parameterized	Real data	None	None	Volume-based objective function	Simulated annealing procedure	Discrimination of vegetation Species	No	—

B. N-FINDR

The N-FINDR method finds the set of pixels that define the simplex with the maximum volume, potentially inscribed within the dataset. First, a dimensionality reduction of the original image is accomplished by using the MNF transform. Next, randomly selected pixels qualify as endmembers, and a trial volume is calculated as follows. Let $\mathbf{E}$ be defined as

$$\mathbf{E} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \boldsymbol{e}_1 & \boldsymbol{e}_2 & \dots & \boldsymbol{e}_E \end{bmatrix} \tag{4}$$

where e_i are endmember column vectors, and E is the number of endmembers used to calculate the simplex volume. The volume of the simplex formed by the endmembers is proportional to the determinant of $\mathbf{E}$

$$V(\mathbf{E}) = \frac{1}{(E-1)!}\,\text{abs}\left(|\mathbf{E}|\right). \tag{5}$$

In order to refine the initial volume estimate, a trial volume is calculated for every pixel in each endmember position by replacing that endmember and recalculating the volume. If the replacement results in a volume increase, the pixel replaces the endmember. This procedure, which does not require any input parameters, is repeated until there are no replacements of endmembers left [6].

C. IEA

In the IEA algorithm, a series of constrained unmixing operations is performed, each time selecting as endmembers the pixels that minimize the remaining error in the unmixed image. An initial vector (usually the mean spectrum of the data) is chosen to start the process. A constrained linear spectral unmixing in terms of this vector is performed, and the error image, formed by the errors remaining at each pixel after the unmixing operation, is calculated. The user then selects a desired number of endmembers E, a number of pixels R, and an angle value θ. R is the number of pixels with the largest number of errors, selected from the error image. The spectral vector corresponding to the pixel with the single largest error is found. A subset of R consisting of all those pixels that fall within an angle θ of the maximum error vector is then calculated, and these pixels are averaged to produce the new endmember vector. This process is continued until E endmembers have been selected [16].

D. CCA

This method is based on the fact that some physical quantities, such as radiance and reflectance, are nonnegative. The vectors formed by discrete radiance/reflectance spectra can be expressed as linear combinations of nonnegative components, which lie inside a nonnegative, convex region. The objective of CCA is to find the boundary points for that region. To implement this concept, the method finds the eigenvectors of the sample spectral correlation matrix of the image, and selects those eigenvectors corresponding to the E largest eigenvalues (where E is the *a priori* number of endmembers to model). The method then looks for the boundaries of the convex cone, where the linear combinations of these eigenvectors produce vectors that are strictly nonnegative, by using the following expression:

$$\boldsymbol{h}(x,y) = \boldsymbol{p}_1 + a_1\boldsymbol{p}_2 + \cdots + a_{E-1}\boldsymbol{p}_E \geq 0 \tag{6}$$

where $\boldsymbol{h}(x,y)$ is the hyperspectral radiance/reflectance signature at the pixel with spatial coordinates (x,y), the p_i are the eigenvectors corresponding to the largest eigenvalues, and 0 is the zero vector. As demonstrated in [20], it is possible to find sets of coefficients that produce a linear combination that contains $E-1$ elements of the linear combination that are exactly zero, with all of the other elements nonnegative. These points

represent the corners of the convex cone. In the implementation of CCA discussed in [7], (6) is rewritten as

$$\boldsymbol{h}(x,y) = [\boldsymbol{p}_1 \cdots \boldsymbol{p}_E] \begin{bmatrix} 1 \\ a_1 \\ \cdots \\ a_{E-1} \end{bmatrix} = \mathbf{P}\boldsymbol{a} \geq \mathbf{0} \tag{7}$$

where the $\boldsymbol{p}_i$ are N-dimensional column vectors. For $N > E$, $\mathbf{P}\boldsymbol{a} = 0$ is an overdetermined system of linear equations. If the elements of $\mathbf{P}$ are viewed as coefficients, and those of $\boldsymbol{a}$ as variables, then there are N equations of the form

$$\boldsymbol{p}_{j1} + a_1\boldsymbol{p}_{j2} + \cdots + a_{E-1}\boldsymbol{p}_{jE} = \mathbf{0}, \qquad \text{for } j = 1, \ldots, N \tag{8}$$

which define $(E-1)$-dimensional hyperplanes in E-dimensional space. Exact solutions can be found by taking $(E-1)$-tuples among the N equations. These solutions produce linear combinations of the eigenvectors that have at least $E-1$ zeros. The boundary of the convex cone is the set of all solution vectors that satisfy (6) or, equivalently, $\text{Min}[\boldsymbol{h}(x,y)] = 0$, where the minimum is taken over all $h_i(x,y) \in \boldsymbol{h}(x,y), i = 1, \cdots, N$. These vectors can be used as endmember spectra for unmixing.

E. ORASIS

This method has been under development at the Naval Research Laboratory (NRL) for over five years. The first step applied by the method is a preprocessing, which consists of flat fielding and various tests on the spectra [17]. Next, an *exemplar selection* process is run with the purpose of prescreening the data for unique spectra, thus creating a representative set of exemplar vectors that can be used as endmembers for spectral unmixing. This procedure rejects redundant spectra by calculating the spectral angle distance (SAD) between spectral vectors [2]. Any vector that is not separated by a certain threshold angle TA is removed from the data. The procedure then finds a basis set of much lower dimension than the original data by a modified Gram–Schmidt process. The exemplar spectra are then projected onto this basis subspace, and a simplex is found through a minimum volume transform. An improvement in the prescreener of ORASIS algorithm has been recently reported [21]. This module, whose goal is to reduce the volume of data with minimal loss of important information, can be seen as a single-pass learning vector quantization (LVQ) process. The exemplars are representative of the entire dataset, i.e., all spectra in the original data are within the TA of at least one exemplar (this is the quantization aspect of the prescreener). However, the first exemplar found to match a spectrum might not be the best fit in the entire set. Thus, more error than is necessary might be present if the method stops after finding the first exemplar within the TA for each spectrum in the scene. The current implementation of ORASIS continues the exemplar selection process even after a first fit is found, being therefore possible to find another exemplar that is closer to the spectrum under test (best fit).

F. AMEE

The input to the AMEE method is the full image data cube, with no previous dimensionality reduction. The method is based on two parameters: a minimum $S_{\min}$ and a maximum $S_{\max}$ spatial kernel size. First, a minimum kernel $K = S_{\min}$ is considered. This element is moved through all the pixels of the image, defining a spatial context around each hyperspectral pixel $\boldsymbol{h}(x,y)$. The spectrally purest $(\boldsymbol{p})$ and the spectrally most highly mixed $(\boldsymbol{m})$ spectral signatures are respectively obtained at the neighborhood of $\boldsymbol{h}(x,y)$ defined by K using the following extended morphological operations [18]:

$$\boldsymbol{p} = \arg\text{-Max}_{(s,t)\in K}\left\{\sum_s\sum_t \text{dist}\left(\boldsymbol{h}(x,y), \boldsymbol{h}(x-s, y-t)\right)\right\} \quad \forall (s,t) \in K \tag{9}$$

$$\boldsymbol{m} = \arg\text{-Max}_{(s,t)\in K}\left\{\sum_s\sum_t \text{dist}\left(\boldsymbol{h}(x,y), \boldsymbol{h}(x+s, y+t)\right)\right\} \quad \forall (s,t) \in K \tag{10}$$

where dist is the SAD distance. A morphological eccentricity index (MEI) [22] is then obtained by calculating the SAD distance between the two signatures above. This operation is repeated for all the pixels in the scene, using kernels of progressively increased size, and the resulting scores are used to evaluate each pixel in both spatial and spectral terms. The algorithm performs as many iterations as needed until $K = S_{\max}$. The associated MEI value of selected pixels at subsequent iterations is updated by means of newly obtained values, as a larger spatial context is considered, until a final MEI image is generated. Endmember selection is performed by a fully automated approach consisting of two steps: 1) autonomous segmentation of the MEI image, and 2) spatial/spectral growing of resulting regions [18].

The six endmember extraction algorithms described in this section represent substantially different design choices. PPI, N-FINDR, and CCA might be characterized as instances of the classic approach to endmember selection, based on the search for spectral convexities in N-D space [15]. While PPI is partially automated, both N-FINDR and CCA are fully automated. IEA is based on an iterative process in which those pixels that reduce the error obtained in constrained spectral unmixing operations are used as endmembers. On other hand, ORASIS performs endmember selection by using LVQ concepts. Contrary to the methods above, which rely on spectral properties of the data alone, AMEE uses a morphological approach where spatial and spectral information are equally employed to derive endmembers.

III. Experimental Design

The aim of this section is twofold: 1) to describe the hyperspectral data that will be used in experiments and 2) to describe an accuracy assessment framework to validate endmember extraction algorithms [23].

A. Data Description

Two AVIRIS imaging spectrometer datasets of the Jasper Ridge Biological Preserve (JRBP) in California have been selected for experiments. The datasets are available (from http://aviris.jpl.nasa.gov) in radiance and reflectance units (from now on, we will respectively refer to the data as AVJRBP_RAD and AVJRBP_REF). The datatets, acquired on April 1998, consist of 512×614 pixels and 224 spectral bands,

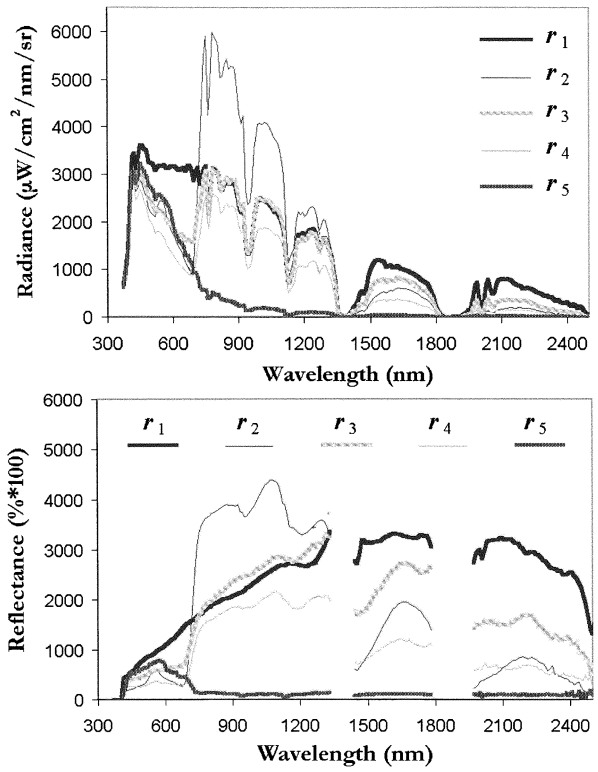

Fig. 1. Selected spectral signatures from (top) AVJRBP_RAD and (bottom) AVJRBP_REF.

with a nominal ground resolution of 20 m, spectral resolution of 10 nm, and 16-bit radiometric resolution. In a previous study of surface materials over JRBP, image endmembers were derived from the scenes above based on extensive ground knowledge [24]. Fig. 1 plots spectral signatures in radiance and reflectance units associated to the main constituent materials at JRBP. These signatures, denoted as $\boldsymbol{r}_1$ (soil), $\boldsymbol{r}_2$ (evergreen forest), $\boldsymbol{r}_3$ (dry grass), $\boldsymbol{r}_4$ (chaparral vegetation), and $\boldsymbol{r}_5$ (shade), were obtained from the image scene by using a hybrid method combining visual inspection and prior information about the scene. Specifically, ground knowledge was used to identify homogeneous vegetation, shadow and soil areas in the scene. Inside those areas, representative pixels were selected as ground-truth spectra by comparing them to a spectral library of field data, used to represent landscape components at JRBP. In this process, we ensured that library spectra matched the phenology at the time of the image, and that there was little miscalibration between field spectra and image spectra. To simplify notation, we hereby use $\{\boldsymbol{r}_j\}_{j=1}^{5}$ to represent spectral signatures in the form of either radiance or reflectance. The above signatures were artificially mixed in computer simulations to create two simulated scenes.

1) Scene CS1: This scene, with a size of 100×100 pixels, is formed by 100 regions, $R_1, \ldots, R_{100}$, of one-pixel width, representing linear mixtures between $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. Abundance fractions of $\boldsymbol{r}_1$ at region R_i are assigned by (i/100), while abundance fractions of $\boldsymbol{r}_2$ at R_i are assigned by $1 - (i/100)$, as depicted in Fig. 2(a) and (b). The scene represents a subtle mixing scenario where $\boldsymbol{r}_1$ progressively infiltrates into $\boldsymbol{r}_2$ and vice-versa. As Fig. 2 shows, the scene contains binary mixtures only.

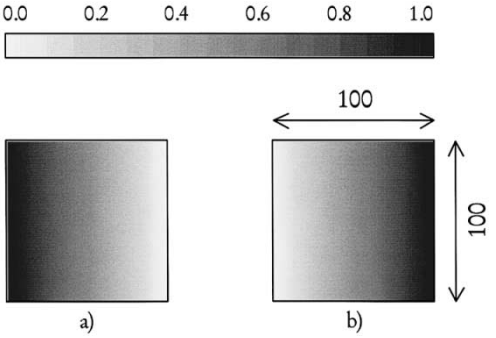

Fig. 2. Computer-simulated scene CS1. (a) Abundance fractions for $\boldsymbol{r}_1$. (b) Abundance fractions for $\boldsymbol{r}_2$.

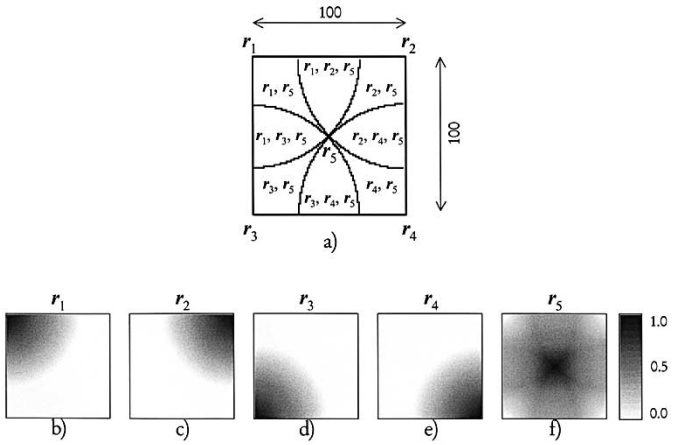

Fig. 3. Computer-simulated scene CS2. Mixture regions created in (a) the simulation. Abundance fractions for (b) $\boldsymbol{r}_1$, (c) $\boldsymbol{r}_2$, (d) $\boldsymbol{r}_3$, (e) $\boldsymbol{r}_4$, and (f) $\boldsymbol{r}_5$.

2) Scene CS2: The second simulated scene was created using five spectral signatures: $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, $\boldsymbol{r}_3$, $\boldsymbol{r}_4$, arranged at the vertices of the image, and $\boldsymbol{r}_5$, located at the central pixel. Signature abundance decreases linearly away from the specified points, as shown in Fig. 3(b)–(f). Abundance fractions have been assigned so that the contributions of the components add to one for every pixel. The resulting image, with a size of 100 $\times$ 100 pixels, contains pure pixels, but also binary, ternary and quaternary mixtures. Fig. 3(a) illustrates the endmembers that participate in each resulting mixture region after the simulation.

Random noise was added to the two scenes above to simulate contributions from ambient (clutter) and instrumental sources. White gaussian noise was created by using numbers with a standard normal distribution obtained from a pseudorandom number generator and added to each pixel. For the simulations, we consider the SNR for each band as the ratio of the 50% signal level to the standard deviation of the noise, hence following the definition given in [25]. This results in noise standard deviation that is roughly proportional to the average signal, a phenomenon often observed in radiometric data [7]. Thus, the simulated hyperspectral data are created, based on a simple linear mixture model, by the following expression:

$$\boldsymbol{s}(x,y) = \left(\frac{\text{SNR}}{2} + \boldsymbol{n}(x,y)\right) \cdot \sum_{j=1}^{R} \alpha_j(x,y) \cdot \boldsymbol{r}_j \qquad (11)$$

where $s(x,y)$ denotes a vector containing the simulated discrete spectrum at the pixel with spatial coordinates (x,y) of the simulated image, R is the total number of reference spectral signatures used to simulate the scene, $\alpha_j(x,y)$ is the assigned fractional abundance of spectral signature $\boldsymbol{r}_j$ at the pixel, and $\boldsymbol{n}(x,y)$ is the noise factor. The ANC and ASC constraints have been imposed in the expression in order to provide the simulation with adequate physical meaning [26]. Each simulated scene was generated by using radiance and reflectance spectra (we name the data according to the acronyms: CS1_RAD, CS1_REF, CS2_RAD, and CS2_REF). Six different SNR values, i.e., 10 : 1, 30 : 1, 50 : 1, 70 : 1, 90 : 1, and 110 : 1, have been considered in the generation of each simulated hyperspectral dataset. As a result, a database formed by 24 simulated scenes and two real AVIRIS images will be used in the experiments.

B. Accuracy Assessment Framework

The primary objective of this section is to describe a quantitative framework to assess and compare the effectiveness of endmember extraction algorithms from real and simulated imagery. When real hyperspectral data are used, selected image endmembers are applied as ground-truth spectra for quantitative assessment. In contrast, the proposed approach to assess endmember extraction accuracy when simulated data are used is outlined in Fig. 4. Due to the controlled environment used in the generation of simulated imagery, ground-truth is available in two forms (see Fig. 4): 1) a (spectral) library of image endmembers, used to create the scenes, and 2) a series of (spatial) fractional maps representing the pixel by pixel abundance of selected image endmembers. In situation 1), extracted endmembers after algorithm application can be directly compared with ground-truth spectra. In situation 2), estimated fractional abundances of extracted signatures can be compared with ground-truth fractional maps. We next provide a description of the comparative strategies used to evaluate endmember quality in each case.

1) Comparison With a Ground-Truth Spectral Library: In this comparative approach, the quality of a suite of extracted endmembers is evaluated by comparing them with ground-truth spectral signatures using spectral similarity criteria. Let $\boldsymbol{e}_i$ be a spectral signature selected from the simulated scene by an endmember extraction algorithm, and $\boldsymbol{r}_j$ be a ground-truth spectral signature. According to the previous notations, we define $\Gamma_E = \{\boldsymbol{e}_i\}_{i=1}^{E}$ as a set containing E endmembers extracted by an algorithm, and $\Gamma_R = \{\boldsymbol{r}_j\}_{j=1}^{R}$ as a set containing R ground-truth spectral signatures. Two distance metrics are considered in this work to compare the similarity of $(\boldsymbol{e}_i, \boldsymbol{r}_j)$ pairs: the spectral angle distance (SAD) and the spectral information divergence (SID). SAD is given by the cosine of the spectral angle formed by N-D vectors. As a result, this measurement is invariant in the multiplication of $\boldsymbol{e}_i$ and $\boldsymbol{r}_j$ by constants and, consequently, is invariant before unknown multiplicative scalings that may arise due to differences in illumination and angular orientation [2]. On the other hand, SID is based on the concept of divergence, and measures the discrepancy of probabilistic behaviors between two spectral signatures. SID offers a new look at the spectral similarity between two spectral signatures by making use of relative entropy to account for the spectral information provided by each signature [27]. A spectral similarity matching algorithm (SSMA) is used in this work to match extracted endmembers to available reference signatures, according to a SAD-based similarity criterion. The criterion is based on the use of a similarity threshold value T, so that the algorithm allows a matching only in the case that the SAD score between the endmember and the reference is as low as possible and is strictly under T. Table II addresses a pseudocode of SSMA. As shown in Table II, the inputs to SSMA are Γ_E, Γ_R, and T. The output of SSMA is a new set Γ_M, which contains those endmembers in Γ_E that have been matched to available reference signatures according to a spectral similarity criterion.

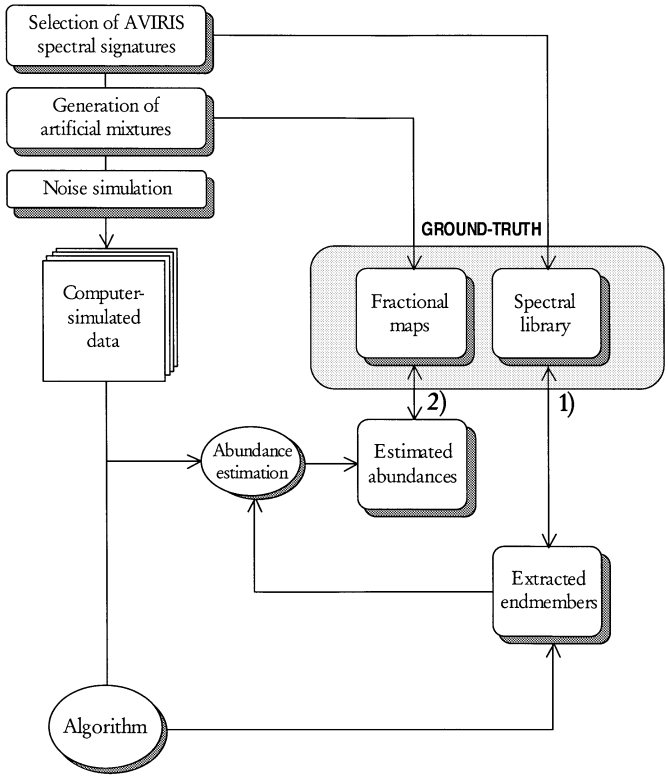

Fig. 4. Schematical description of the approach used to assess endmember extraction accuracy with simulated data.

2) Comparison With Ground-Truth Fractional Abundance Maps: In case of abundance ground-truth availability, the quality of endmembers can be assessed by estimating their abundance in the scene and comparing the obtained values with reference fractions. This approach allows an evaluation of endmembers from a spatial viewpoint, and is particularly suitable for our database of simulated scenes, where high-quality ground-truth values are given by the artificially assigned abundance fractions used in the generation of data. In this work, fractions of extracted endmembers are estimated by a fully constrained linear spectral unmixing (FCLSU) approach [26], which incorporates the ANC and ASC constraints used in the image simulations. Let N and M be the corresponding total number of samples and lines in a simulated hyperspectral image. Also, let $\boldsymbol{m}_j$ be an extracted endmember that has been matched to an available reference signature $\boldsymbol{r}_j$ by SSMA, i.e., $\boldsymbol{m}_j \in \Gamma_M$. Similarly, let $\hat{\alpha}_j(x,y)$ denote the FCLSU-estimated abundance of $\boldsymbol{m}_j$ at the pixel with spatial coordinates (x,y). A

TABLE II
PSEUDOCODE OF SPECTRAL SIMILARITY MATCHING ALGORITHM

```
Inputs: Γ_E , Γ_R , T
Output: Γ_M
Let Γ_M = ∅
For i = 1 to E label e_i as "not matched"
For j = 1 to R do begin
    Min = ∞
    k = 0
    For i = 1 to E do begin
        If e_i is labeled as "not matched" then begin
            If SAD(e_i, r_j) < T and SAD(e_i, r_j) < Min begin
                Min = SAD(e_i, r_j)
                k = i
            Endif
        Endif
    Endfor
    If k > 0 then begin
        Label e_k as "matched"
        m_j = e_k
        Γ_M = Γ_M ∪ m_j
    Endif
Endfor
```

simple statistical measurement to evaluate the similarity of true versus estimated values is provided by the root mean square error (RMSE), given by

$$\mathrm{RMSE}(\boldsymbol{m}_j, \boldsymbol{r}_j) = \left(\frac{1}{N \cdot M} \cdot \sum_{x=1}^{N} \sum_{x=1}^{M} [\alpha_j(x,y) - \hat{\alpha}_j(x,y)]^2 \right)^{\frac{1}{2}}. \tag{12}$$

In order to compare a specified ground-truth with the results provided by endmember extraction algorithms, we have created a software toolkit that incorporates the metrics addressed in this section [28]. The toolkit was used to generate the simulated data described in Section III-A, as well as the results that will be presented and discussed in the following section.

IV. COMPARATIVE PERFORMANCE ANALYSIS

This section describes a series of experiments that use simulated and real hyperspectral data to conduct a comprehensive comparison among standard endmember extraction algorithms. In order to ensure the fairest possible comparison, the best performance must be obtained from each alternative method. Then, prior to a full examination and discussion of results, it is important to outline parameter values used for PPI, IEA, AMEE, and ORASIS algorithms, bearing in mind that the N-FINDR and CCA methods do not require any input parameters. We also remark that some of the methods (IEA, CCA, and ORASIS) include an optional choice of setting a fixed number of endmembers to be identified by the algorithm, or allowing the algorithm to decide how many endmembers are present in the scene. In this work, the number of endmembers to be found by the above methods was fixed in experiments involving simulated data, where the number of available endmember constituents is known *a priori*, but was not previously set in the case of processing real imagery. In addition, although in practice it is difficult to fully optimize every method, we have used our experience with the methods to select parameters that are reasonably close to optimal for the test data. Specifically, it was observed that PPI produced the same final set of endmembers for the two real scenes (AVJRBP_RAD and AVJRBP_REF) when the value of L was above 3000 iterations (values of $L = 10^4$, $L = 10^5$, and $L = 10^6$ iterations were tested). For the simulated data, only 1000 iterations were required to produce a final suite of endmembers that were essentially the same as those found with $L = 10^6$ iterations. Based on the above simple experiment, the cutoff threshold parameter C was set to the mean of PPI scores obtained after $L = 1000$ iterations for the simulated data, and $L = 3000$ iterations for the real data. These parameter values are in agreement with those used before in the literature [29]. Pixels were then grouped into smaller subsets based on their clustering in the N-D space. Finally, resulting groups of extreme pixels were linked to the original image, and the mean spectrum of each group was used as a candidate endmember for unmixing [15]. In order to determine appropriate parameters for IEA, we tested the algorithm with our simulated/real datasets using parameter values in the following intervals: $R = [10\,100]$ pixels and $\theta = [0.5, 1.4]$ rad. In general terms, we observed that the algorithm was not very sensitive to the setting of those parameters. In fact, even if the parameters R and θ were changed, within rather liberal limits, essentially the same endmembers were obtained. The only caution that we exercised, in accordance with suggestions made by the authors of IEA [16], was to avoid values for parameters R and θ that were so small that single pixel endmembers were returned. After analyzing the SSMA similarity scores between selected and ground-truth endmembers in both real and simulated imagery, we set R and θ to their corresponding 50 pixels and 1.2 rad. In similar terms, a previous performance study of our AMEE algorithm revealed that satisfactory results in most situations can be found by setting $S_{\min}$ and $S_{\max}$ parameters to 3×3 pixels and 15×15 pixels, respectively [18]. These values were also found appropriate for the simulated/real datasets used in this work, mainly due to the observed width in pixels of patterns of interest in the data. Finally, the developers of the ORASIS method at NRL optimized each algorithm run by carefully setting the TA parameter to an appropriate value. This was done by taking into account the SNR of each tested dataset. Using the empirically selected parameter values described above, we have performed three comparative experiments, summarized in Table III.

A. Experiment 1

In this subsection, we conduct an experiment-based comparison among PPI, N-FINDR, IEA, AMEE, CCA, and ORASIS algorithms by using simulated scenes CS1_RAD, CS1_REF, CS2_RAD, and CS2_REF with different SNR values. For clarity, we emphasize that CS1_RAD and CS1_REF represent a simple mixing scenario formed by binary mixtures only, while CS2_RAD and CS2_REF contain ternary and quaternary mixtures in addition to binary mixtures. We focus on analyzing

TABLE III
SUMMARY OF EXPERIMENTS

Experiment	Datasets used	Objective
1	CS1_RAD, CS1_REF, CS2_RAD, CS2_REF.	To compare algorithm performance in the task of extracting endmembers from simulated hyperspectral data.
2	CS1_RAD, CS1_REF, CS2_RAD, CS2_REF.	To compare algorithm performance in the task of estimating endmember abundance fractions in simulated hyperspectral data.
3	AVJRBP_RAD, AVJRBP_REF.	To compare algorithm performance in the task of extracting endmembers from real hyperspectral data.

TABLE IV
TOTAL NUMBER OF EXTRACTED ENDMEMBERS (E) BY PPI, N-FINDR, IEA, AMEE, CCA, AND ORASIS METHODS, ALONG WITH SAD-BASED (BOLD TYPEFACE) AND SID-BASED SIMILARITY SCORES BETWEEN SSMA-MATCHED ENDMEMBERS AND CORRESPONDING REFERENCE SPECTRAL SIGNATURES FOR CS1_RAD AND CS1_REF WITH DIFFERENT SNR VALUES

		SNR																	
		10:1			30:1			50:1			70:1			90:1			110:1		
	Method	E	m_1	m_2	E	m_1	m_2	E	m_1	m_2	E	m_1	m_2	E	m_1	m_2	E	m_1	m_2
CS1_RAD	PPI	2	N/M	N/M	2	**0.062** 0.004	**0.058** 0.003	2	**0.040** 0.001	**0.040** 0.001	2	**0.028** 0.000	**0.027** 0.000	2	**0.022** 0.000	**0.021** 0.000	2	**0.019** 0.000	**0.017** 0.000
	N-FINDR	7	N/M	N/M	5	**0.077** 0.005	**0.064** 0.004	2	**0.044** 0.001	**0.050** 0.002	2	**0.032** 0.001	**0.028** 0.000	2	**0.025** 0.000	**0.022** 0.000	2	**0.020** 0.000	**0.019** 0.000
	IEA	2	**0.089** 0.007	**0.091** 0.008	2	**0.014** 0.000	**0.029** 0.001	2	**0.022** 0.000	**0.021** 0.001	2	**0.020** 0.000	**0.016** 0.000	2	**0.013** 0.000	**0.015** 0.000	2	**0.008** 0.000	**0.009** 0.000
	AMEE	2	**0.082** 0.007	**0.086** 0.007	2	**0.017** 0.000	**0.014** 0.000	2	**0.011** 0.000	**0.010** 0.000	2	**0.008** 0.000	**0.009** 0.000	2	**0.007** 0.000	**0.008** 0.000	2	**0.006** 0.000	**0.008** 0.000
	CCA	2	N/M	N/M	2	**0.057** 0.003	**0.062** 0.004	2	**0.041** 0.001	**0.045** 0.002	2	**0.028** 0.001	**0.032** 0.001	2	**0.024** 0.000	**0.026** 0.001	2	**0.020** 0.000	**0.022** 0.000
	ORASIS	2	N/M	N/M	2	**0.077** 0.006	**0.066** 0.004	2	**0.046** 0.002	**0.039** 0.001	2	**0.032** 0.001	**0.025** 0.000	2	**0.024** 0.000	**0.019** 0.000	2	**0.020** 0.000	**0.015** 0.000
CS1_REF	PPI	2	N/M	N/M	2	**0.061** 0.003	**0.066** 0.004	2	**0.039** 0.001	**0.035** 0.001	2	**0.025** 0.000	**0.029** 0.000	2	**0.020** 0.000	**0.022** 0.000	2	**0.017** 0.000	**0.019** 0.000
	N-FINDR	6	N/M	N/M	4	**0.068** 0.004	**0.064** 0.004	2	**0.040** 0.001	**0.039** 0.001	2	**0.026** 0.000	**0.029** 0.001	2	**0.022** 0.000	**0.024** 0.000	2	**0.016** 0.000	**0.019** 0.000
	IEA	2	**0.083** 0.007	**0.087** 0.007	2	**0.028** 0.000	**0.034** 0.005	2	**0.034** 0.005	**0.032** 0.004	2	**0.023** 0.000	**0.035** 0.005	2	**0.017** 0.000	**0.031** 0.005	2	**0.017** 0.000	**0.012** 0.001
	AMEE	2	**0.081** 0.007	**0.075** 0.005	2	**0.015** 0.000	**0.021** 0.001	2	**0.010** 0.000	**0.015** 0.000	2	**0.008** 0.000	**0.014** 0.000	2	**0.007** 0.000	**0.013** 0.000	2	**0.006** 0.000	**0.012** 0.000
	CCA	2	N/M	N/M	2	**0.064** 0.004	**0.068** 0.005	2	**0.041** 0.001	**0.044** 0.002	2	**0.028** 0.001	**0.029** 0.001	2	**0.020** 0.000	**0.024** 0.000	2	**0.019** 0.000	**0.021** 0.000
	ORASIS	2	N/M	N/M	2	**0.067** 0.004	**0.076** 0.005	2	**0.040** 0.001	**0.045** 0.002	2	**0.027** 0.000	**0.032** 0.001	2	**0.021** 0.000	**0.025** 0.000	2	**0.017** 0.000	**0.020** 0.000

the performance of the algorithms by deriving appropriate endmembers from simulated data.

Table IV tabulates the spectral similarity scores produced by SAD and SID metrics for extracted endmembers from CS1_RAD and CS1_REF scenes with SNR values ranging from 10 : 1 to 110 : 1. Results in Table IV show that PPI, N-FINDR, CCA, and ORASIS produce endmembers that are not matched (N/M) by SSMA to available ground-truth signatures when the SNR equals 10 : 1 (SID scores are only shown for SSMA-matched endmembers). It should be noted that the threshold value used to implement the SAD-based similarity criterion of SSMA in experiments was $T = 0.1$, a reasonable limit of tolerance for this metric. This value was selected by observing the confusion matrix of SAD similarity values among the five original spectral signatures used to create the simulated scenes [27]. The previously addressed result indicates that the above four methods, based on convex geometry concepts (PPI, N-FINDR, CCA) and LVQ (ORASIS), are more sensitive to noise than IEA and AMEE in a simple binary mixing scenario. Overall, IEA and AMEE produce the best SAD and SID similarity scores for CS1_RAD and CS1_REF, with AMEE scores being slightly better than those obtained by IEA. In contrast, we can observe that the results for all the tested methods are similar for radiance and reflectance endmembers. Even though the best results for all methods are obtained when the SNR equals 110 : 1, there is a significant performance increase in most of the approaches examined when the SNR is increased from 10 : 1 to 30 : 1.

Table V shows the statistics produced by SAD and SID for extracted endmembers using CS2_RAD and CS2_REF with SNR values ranging from 30 : 1 to 110 : 1. Results for SNR of 10 : 1 have been omitted, since most methods produced a majority of N/M endmembers. This fact seems to indicate that the presence of ternary and quaternary mixtures has a strong influence on algorithm performance. Although the desired number of endmembers to be detected by PPI, IEA, CCA, and ORASIS was set to five—the real number of endmembers in the scene—these methods produced N/M endmembers for SNR values of 70 : 1 and below. In general terms, the six methods find difficulties to provide adequate correlations for $\boldsymbol{r}_5$, i.e., the shade endmember, especially at low SNR values. In addition, with the single exception of ORASIS, most methods produced low SAD- and

TABLE V
TOTAL NUMBER OF EXTRACTED ENDMEMBERS (E) BY PPI, N-FINDR, IEA, AMEE, CCA, AND ORASIS METHODS, ALONG WITH SAD-BASED (BOLD TYPEFACE) AND SID-BASED SIMILARITY SCORES BETWEEN SSMA-MATCHED ENDMEMBERS AND CORRESPONDING REFERENCE SPECTRAL SIGNATURES FOR CS2_RAD AND CS2_REF WITH DIFFERENT SNR VALUES

		SNR																	
		30:1						70:1						110:1					
	Method	E	m_1	m_2	m_3	m_4	m_5	E	m_1	m_2	m_3	m_4	m_5	E	m_1	m_2	m_3	m_4	m_5
CS1_RAD	PPI	5	**0.069** 0.005	**0.073** 0.005	**0.064** 0.006	N/M	N/M	5	**0.029** 0.000	**0.034** 0.001	**0.031** 0.000	**0.048** 0.002	N/M	5	**0.018** 0.000	**0.017** 0.000	**0.018** 0.000	**0.018** 0.000	**0.078** 0.006
	N-FINDR	8	**0.076** 0.004	**0.086** 0.008	**0.093** 0.009	N/M	N/M	8	**0.029** 0.000	**0.031** 0.001	**0.028** 0.001	**0.095** 0.010	N/M	5	**0.018** 0.000	**0.022** 0.000	**0.031** 0.002	**0.090** 0.010	**0.081** 0.007
	IEA	5	**0.066** 0.004	**0.081** 0.007	**0.069** 0.004	N/M	N/M	5	**0.026** 0.000	**0.029** 0.000	**0.042** 0.001	**0.055** 0.005	N/M	5	**0.018** 0.000	**0.018** 0.000	**0.022** 0.000	**0.028** 0.001	**0.078** 0.006
	AMEE	6	**0.035** 0.001	**0.036** 0.001	**0.033** 0.001	N/M	N/M	5	**0.025** 0.000	**0.031** 0.001	**0.024** 0.000	**0.020** 0.000	N/M	5	**0.012** 0.000	**0.013** 0.000	**0.013** 0.000	**0.021** 0.000	**0.069** 0.005
	CCA	5	**0.076** 0.004	**0.081** 0.006	**0.074** 0.004	N/M	N/M	5	**0.028** 0.001	**0.033** 0.001	**0.031** 0.001	**0.079** 0.008	N/M	5	**0.018** 0.000	**0.021** 0.000	**0.024** 0.000	**0.068** 0.005	**0.076** 0.006
	ORASIS	5	N/M	**0.081** 0.006	**0.082** 0.007	**0.083** 0.007	N/M	5	**0.069** 0.004	**0.038** 0.002	**0.035** 0.002	**0.019** 0.000	N/M	5	**0.031** 0.001	**0.023** 0.000	**0.019** 0.000	**0.018** 0.000	**0.079** 0.006
CS1_REF	PPI	5	**0.072** 0.004	**0.067** 0.004	**0.074** 0.004	N/M	N/M	5	**0.027** 0.000	**0.029** 0.000	**0.032** 0.000	**0.042** 0.001	N/M	5	**0.019** 0.000	**0.017** 0.000	**0.027** 0.000	**0.034** 0.001	**0.082** 0.007
	N-FINDR	9	**0.064** 0.004	**0.060** 0.004	**0.067** 0.004	N/M	N/M	7	**0.025** 0.000	**0.029** 0.001	**0.041** 0.002	**0.091** 0.009	N/M	5	**0.016** 0.000	**0.016** 0.000	**0.027** 0.000	**0.036** 0.001	**0.092** 0.009
	IEA	5	**0.069** 0.004	**0.071** 0.004	**0.067** 0.004	N/M	N/M	5	**0.012** 0.000	**0.025** 0.000	**0.033** 0.000	**0.041** 0.001	N/M	5	**0.008** 0.000	**0.007** 0.000	**0.027** 0.000	**0.033** 0.000	**0.076** 0.006
	AMEE	7	**0.031** 0.001	**0.025** 0.000	**0.037** 0.001	N/M	N/M	6	**0.009** 0.000	**0.012** 0.000	**0.021** 0.000	**0.032** 0.001	N/M	5	**0.006** 0.000	**0.007** 0.000	**0.020** 0.000	**0.031** 0.001	**0.065** 0.005
	CCA	5	**0.073** 0.004	**0.075** 0.004	**0.074** 0.004	N/M	N/M	5	**0.031** 0.001	**0.037** 0.001	**0.039** 0.002	**0.083** 0.008	N/M	5	**0.022** 0.000	**0.020** 0.000	**0.025** 0.000	**0.071** 0.005	**0.079** 0.006
	ORASIS	5	N/M	**0.073** 0.004	**0.064** 0.004	**0.073** 0.004	N/M	5	**0.057** 0.003	**0.033** 0.001	**0.036** 0.001	**0.031** 0.001	N/M	5	**0.028** 0.001	**0.018** 0.000	**0.024** 0.000	**0.030** 0.001	**0.081** 0.007

SID-based similarity scores for r_4, i.e., the chaparral vegetation endmember. The shade component may be more difficult to model because it is present in most mixtures in the scene, including ternary and quaternary cases [see Fig. 3(f)]. In contrast, the spectral similarity of r_4 with respect to r_2 and r_3 could be the main cause for the lower performance of methods when modeling the chaparral component in both radiance and reflectance data. The improved scores produced by ORASIS for the SSMA-matched endmember m_4 (see Table V) seem to indicate that the LVQ module makes the algorithm robust in the presence of spectrally similar samples. Overall, the performance of all tested methods is good when the SNR equals 110 : 1, with AMEE providing high similarity scores in many cases.

B. Experiment 2

In this subsection, CS1_RAD, CS1_REF, CS2_RAD, and CS2_REF are used for experiments to evaluate the performance of PPI, N-FINDR, IEA, AMEE, CCA, and ORASIS in the task of estimating abundances from extracted endmembers. Fig. 5(a)–(d) shows graphs of the RMSE error scores obtained in abundance estimation of SSMA-matched endmembers m_1 and m_2, provided by the six compared methods, on CS1_RAD and CS1_REF using different SNR values. In general terms, the methods work slightly better with reflectance data. This fact could be explained by the higher spectral similarity between radiance endmembers compared to reflectance endmembers, a fact that was observed in the respective confusion matrices of SAD and SID similarity values. Overall, both AMEE and IEA produce the lowest error scores when the SNR is high or moderate, and N-FINDR produces the highest error values. In particular, when moderate-to-high SNR data are used, IEA produces results that are much better than those found by the same algorithm with lower image quality. The performance of CCA and ORASIS is also significantly increased when high SNR data are used. Contrary to the approaches above, AMEE is characterized by a more stable behavior with respect to SNR. This result provides some objective confirmation of our introspection: that the incorporation of spatial information reduces algorithm sensitivity to noise and outliers. However, it must be noted that, because of the simple nature of mixtures in CS1_RAD and CS1_REF, this observation is not conclusive.

In order to test the above statement in a more complex mixing scenario, Fig. 6 shows the RMSE error scores obtained after estimating the abundance of SSMA-matched endmembers m_1, m_2, m_3, m_4 and m_5, provided by the six methods, on CS2_RAD and CS2_REF using different SNR values. In the case of m_4 and m_5, performance comparison is only available for SNR values superior to 50 : 1 and 90 : 1, respectively (remaining SNR values result in N/M endmembers). As a general comment, it may be deduced from Fig. 6 that the performance of the six compared methods is slightly better when reflectance data are used. In addition, convex geometry-based methods (N-FINDR, PPI, CCA) seem more sensitive to noise and complexity of spectral mixtures than IEA and AMEE, with AMEE scores being slightly better than those found by IEA in most cases. With the single exception of m_1, the LVQ-based endmember selection module of ORASIS produces better results than those found by methods based on convex geometry concepts. In addition, this algorithm produces the best abundance estimation results for m_4 [see Fig. 6(g) and (h)]. This fact confirms the improved performance of ORASIS when modeling the chaparral vegetation endmember, a result that was also observed in experiment 2. On other hand, it should be noted that the performance of IEA decreases significantly when the SNR is reduced from 50 : 1 to

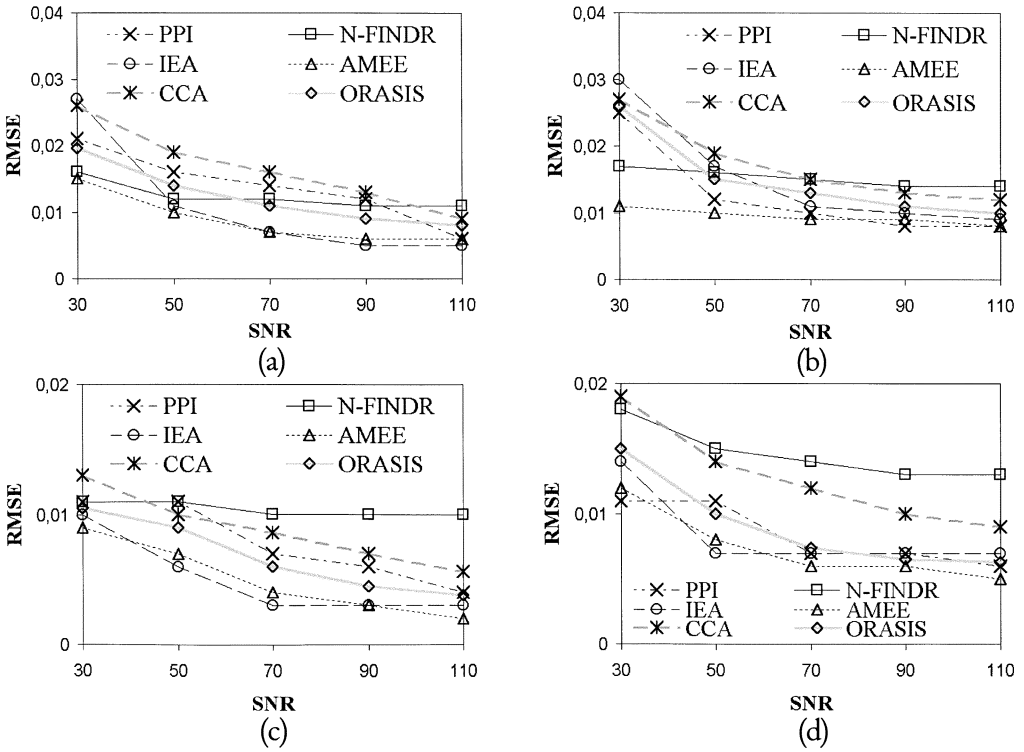

Fig. 5. RMSE scores for SSMA-matched endmembers by PPI, N-FINDR, IEA, AMEE, CCA, and ORASIS in CS1 at different SNR values. (a) $\boldsymbol{m}_1$ in CS1_RAD. (b) $\boldsymbol{m}_2$ in CS1_RAD. (c) $\boldsymbol{m}_1$ in CS1_REF. (d) $\boldsymbol{m}_2$ in CS1_REF.

30 : 1 [see Fig. 6(a)–(f)]. This last item seems to indicate that the iterative minimization of constrained unmixing errors made by IEA is not as effective as the integrated spatial/spectral treatment of data performed by AMEE, in particular when lower quality data are used.

C. Experiment 3

A comparative study of PPI, N-FINDR, IEA, AMEE, CCA, and ORASIS in the task of extracting endmembers from real hyperspectral data is conducted. Contrary to computer-simulated data, where noise and spectral mixture complexity can be easily investigated, real imagery is conditioned by a great deal of additional circumstances. These factors, which include variations in illumination through the scene, observation angle, and nonlinear mixture effects caused by multiple scattering, have been reported to influence the endmember selection process [2]. However, it is difficult to substantiate the individual effect of those strongly interrelated parameters on algorithm performance. In this experiment, we show a general performance comparison of algorithms on real data (AVJRBP_RAD and AVJRBP_REF), using selected image endmembers in Fig. 1 as ground-truth information. Before addressing the obtained results, we should point out that SNR in the AVIRIS sensor at the time of data acquisition was considerably higher than the values used in computer simulations: the highest SNR levels calculated in 1998 for a 50% reflectance target at sea level and 23.5° zenith angle corresponds in the A spectrometer (0.7 μm) to levels of 1000 : 1, and the lowest were measured in the SWIR region (2.2 μm) with levels of 400 : 1 [30].

Table VI shows the SAD and SID similarity scores obtained by the six compared methods in AVJRBP_RAD and AVJRBP_REF. Results in Table VI reveal that all tested methods produce SSMA-matched endmembers for the five materials studied. With the exception of ORASIS runs, where some vegetation features are better modeled in the radiance scene, the methods produce results that are slightly better for reflectance data than for radiance data. When radiance data are used, AMEE provides the best similarity scores for the soil $(\boldsymbol{m}_1)$, evergreen forest $(\boldsymbol{m}_2)$, and dry grass $(\boldsymbol{m}_3)$ constituents, while ORASIS produces the best result for chaparral vegetation $(\boldsymbol{m}_4)$. When reflectance data are used, AMEE endmembers provide the best similarity scores in four out of five materials. These results apply for the two similarity metrics (SAD and SID) used in the comparison. It should be noted that the shade constituent $(\boldsymbol{m}_5)$ is better characterized by IEA, in both radiance and reflectance data. One explanation for this result might be due to the fact that the spatial trend of the shade component in the real scenes follows a very irregular pattern, not necessarily correlated as observed in the spatial distribution of the other studied endmembers [24]. Another reason might be that the shade ground-truth spectrum was modeled by using a water spectral signature from a lake outside JRBP scenes, chosen because its albedo was the lowest in the entire scene [24]. The results above demonstrate the importance of considering both spatial and spectral information in the selection of endmembers for linear spectral unmixing. It is important to emphasize that the linear mixture model is not flexible enough to accommodate the full range of natural vegetation variability throughout the landscape. In order to accurately characterize

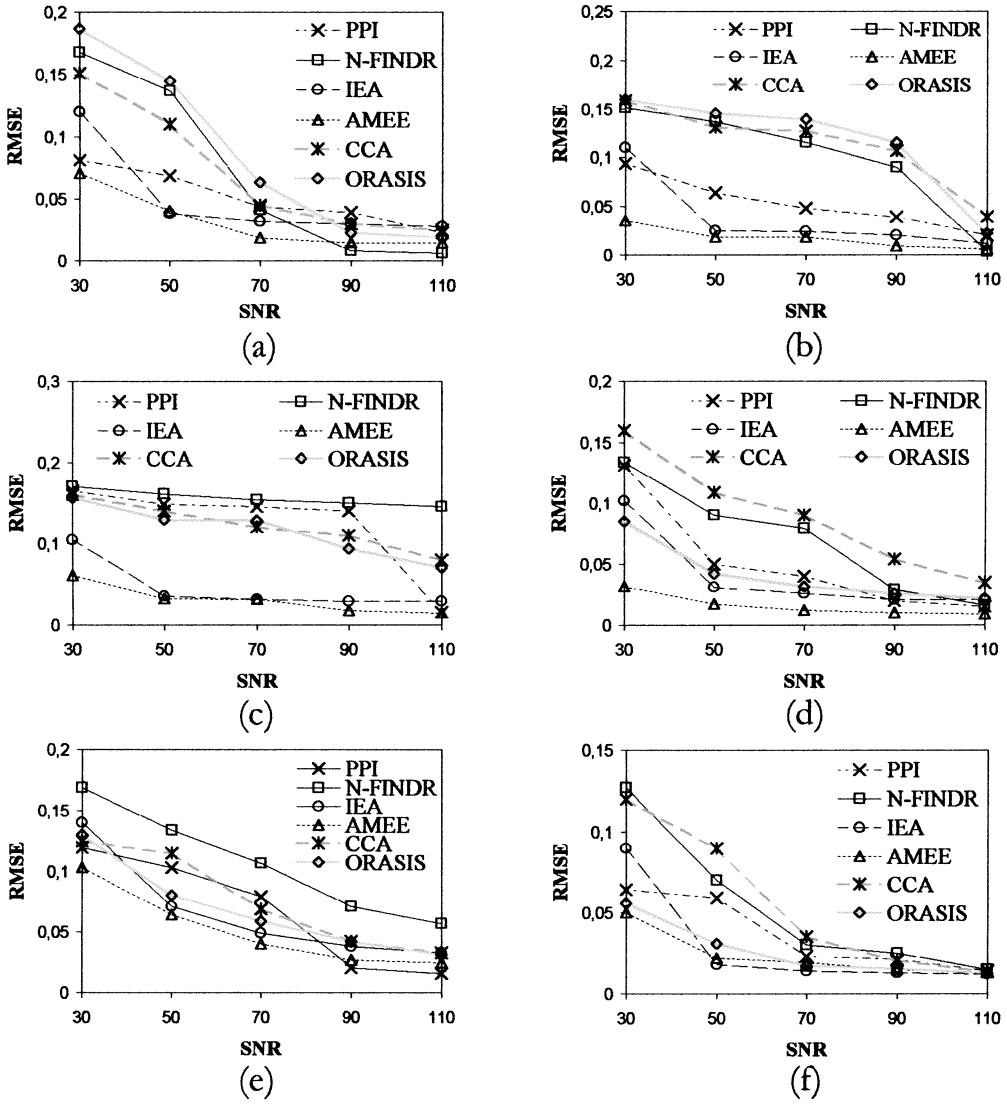

Fig. 6. RMSE scores for SSMA-matched endmembers by PPI, N-FINDR, IEA, AMEE, CCA, and ORASIS in CS2 at different SNR values. (a) $\boldsymbol{m}_1$ in CS2_RAD. (b) $\boldsymbol{m}_1$ in CS2_REF. (c) $\boldsymbol{m}_2$ in CS2_RAD. (d) $\boldsymbol{m}_2$ in CS2_REF. (e) $\boldsymbol{m}_3$ in CS2_RAD. (f) $\boldsymbol{m}_3$ in CS2_REF.

the JRBP ecosystem structure, it might be necessary to account for changes in bidirectional reflectance (BRDF) in the subpixel components, since multiple scattering overestimates the results from a linear scattering model. In order to calibrate the impact of the linear spectral unmixing algorithm used on the task of estimating subpixel fractional abundances in the real datasets, we have performed a qualitative comparison between two different algorithms: fully constrained (FCLSU) and unconstrained (ULSU) linear spectral unmixing. In the qualitative experiments, it was found that the correspondent endmember fractional abundance maps, derived by using FCLSU and ULSU, were in visual agreement for all the methods tested. In addition, negative and/or unrealistic ULSU-derived abundance fractions, which usually indicate a bad fit of the model and reveal inappropriate endmember selections, were very rarely found, in particular when the AMEE algorithm was applied. Having those circumstances in mind, the results obtained in this experiment indicate that the linear mixture model, improved by the integration of spatial and spectral information in the task of selecting endmembers, is able to provide a relatively good characterization of general landscape conditions.

V. Conclusion and Future Lines

Several endmember extraction algorithms have been proposed in the literature over the past decade. Comparing these approaches has been a challenging task due to a lack of rigorous criteria to substantiate any new algorithm. Another difficulty arises from the fact that there are no commonly used data to perform quantitative comparisons. In this paper, we have conducted an evaluation of several endmember extraction algorithms for linear spectral unmixing, using a widely available database of simulated and real AVIRIS images. The comparison of methods has been carried out from two different scopes. First, the issue was undertaken under the assumption that reliable ground-truth spectral signatures are available. A further attempt from the perspective of a mixed-pixel classification problem was also conducted by comparing estimated

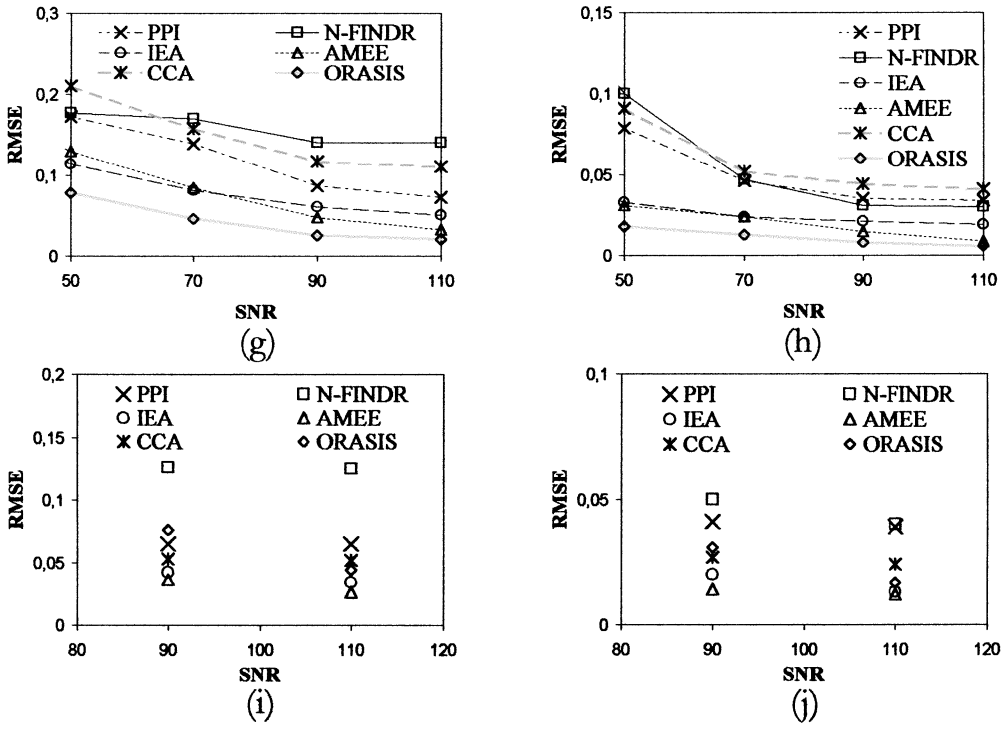

Fig. 6. *(Continued.)* RMSE scores for SSMA-matched endmembers by PPI, N-FINDR, IEA, AMEE, CCA, and ORASIS in CS2 at different SNR values. (g) $\boldsymbol{m}_4$ in CS2_RAD. (h) $\boldsymbol{m}_4$ in CS2_REF. (i) $\boldsymbol{m}_5$ in CS2_RAD. (j) $\boldsymbol{m}_5$ in CS2_REF.

TABLE VI
SAD-BASED (BOLD TYPEFACE) AND SID-BASED SIMILARITY SCORES BETWEEN SSMA-MATCHED ENDMEMBERS BY PPI, N-FINDR, IEA, AMEE, CCA, AND ORASIS METHODS AND CORRESPONDING REFERENCE SPECTRAL SIGNATURES FOR AVJRBP_RAD AND AVJRBP_REF

	AVJRBP_RAD					AVJRBP_REF				
Method	$\boldsymbol{m}_1$	$\boldsymbol{m}_2$	$\boldsymbol{m}_3$	$\boldsymbol{m}_4$	$\boldsymbol{m}_5$	$\boldsymbol{m}_1$	$\boldsymbol{m}_2$	$\boldsymbol{m}_3$	$\boldsymbol{m}_4$	$\boldsymbol{m}_5$
PPI	**0.062**	**0.058**	**0.040**	**0.040**	**0.028**	**0.027**	**0.022**	**0.021**	**0.019**	**0.017**
	0.004	0.003	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
N-FINDR	**0.077**	**0.064**	**0.044**	**0.050**	**0.032**	**0.028**	**0.025**	**0.022**	**0.020**	**0.019**
	0.005	0.004	0.001	0.002	0.001	0.000	0.000	0.000	0.000	0.000
IEA	**0.014**	**0.029**	**0.022**	**0.021**	**0.020**	**0.016**	**0.013**	**0.015**	**0.008**	**0.009**
	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000
AMEE	**0.017**	**0.014**	**0.011**	**0.010**	**0.008**	**0.009**	**0.007**	**0.008**	**0.006**	**0.008**
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CCA	**0.057**	**0.062**	**0.041**	**0.045**	**0.028**	**0.032**	**0.024**	**0.026**	**0.020**	**0.022**
	0.003	0.004	0.001	0.002	0.001	0.001	0.000	0.001	0.000	0.000
ORASIS	**0.067**	**0.076**	**0.040**	**0.045**	**0.027**	**0.032**	**0.021**	**0.025**	**0.017**	**0.020**
	0.004	0.005	0.001	0.002	0.000	0.001	0.000	0.000	0.000	0.000

abundance planes to ground-truth reference maps representing the spatial distribution of endmember constituents in simulated imagery. Experiments with simulated and real data indicate that the combination of spatial and spectral information produces results that are superior to those found by using the spectral information alone. Low SNR data and complex spectral mixtures are modeled better when a spatial/spectral model is incorporated to endmember selection. Despite our effort to conduct a comprehensive, impartial, and rigorous comparative analysis of various algorithms, completion is not claimed. In particular, the number of algorithms compared in this work is limited to six methods, selected because they represent very different design alternatives. In contrast, the application of the proposed comparative framework to a larger number of images with high-quality ground-truth data is required in order to extrapolate the main conclusions drawn from the present study. Further research on nonlinear mixture effects, observation angle, and illumination/shadow effects is also needed. Finally, another topic of interest is the development of new comparative strategies to substantiate endmember extraction algorithms, in particular when no ground-truth information is available. This last item may help data analysts to decide whether any particular comparative approach could generally prove more suitable than others for evaluating endmember quality. From the perspective of the experimentation carried out in this study, the answer to this question has been posed from the viewpoint of the particular use of the endmember concept made in the application, either as a spectral feature that can be used to identify materials by direct comparison to available reference spectra or as a tool to estimate subpixel abundance fractions in the scene.

ACKNOWLEDGMENT

The authors gratefully acknowledge J. Bowles and D. Gillis for their significant effort in the task of applying Naval Research Laboratory's ORASIS algorithm to the simulated/real database of images discussed in this paper, and also for many helpful discussions and suggestions. They thank M. Winter for providing an evaluation version of N-FINDR, which was used to generate the results shown in the paper, and also for his assistance in the use of the software and implementation of other methods. Contributions by R. O. Green, in the form of valuable comments to improve the comparative effort, are also gratefully acknowledged. Last but not least, the authors would like to thank the anonymous reviewers for their helpful suggestions, as well as A. Curado for his linguistic review of this paper.

REFERENCES

[1] R. O. Green *et al.*, "Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, pp. 227–248, 1998.

[2] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Mag.*, vol. 19, pp. 44–57, Jan. 2002.

[3] J. B. Adams, M. O. Smith, and A. R. Gillespie, "Imaging spectroscopy: Interpretation based on spectral mixture analysis," in *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, C. M. Pieters and P. Englert, Eds. Cambridge, U.K.: Cambridge Univ. Press, 1993, pp. 145–166.

[4] P. M. Mather, "Land cover classification revisited," in *Advances in Remote Sensing and GIS Analysis*, P. M. Atkinson and N. J. Tate, Eds. Chichester, U.K.: Wiley, 1999, pp. 7–16.

[5] D. A. Roberts, G. T. Batista, J. L. G. Pereira, E. K. Waller, and B. W. Nelson, "Change identification using multitemporal spectral mixture analysis: Applications in eastern Amazonia," in *Remote Sensing Change Detection: Environmental Monitoring Applications and Methods*, R. S. Lunetta and C. D. Elvidge, Eds. Ann Arbor, MI: Ann Arbor Press, 1998, pp. 137–161.

[6] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266–275, 1999.

[7] A. Ifarraguerri and C.-I. Chang, "Multispectral and hyperspectral image analysis with convex cones," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 756–770, Mar. 1999.

[8] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Mag.*, vol. 19, pp. 17–28, Jan. 2002.

[9] P. L. Vora, J. E. Farrell, J. D. Tietz, and D. H. Brainard, "Image capture: Simulation of sensor responses from hyperspectral images," *IEEE Trans. Image Processing*, vol. 10, pp. 307–316, Feb. 2001.

[10] G. D. Robinson, H. H. Gross, and J. R. Schott, "Evaluation of two applications of spectral mixing models to image fusion," *Remote Sens. Environ.*, vol. 71, pp. 272–281, 2000.

[11] E. F. Collins, D. A. Roberts, and C. C. Borel, "Spectral mixture analysis of simulated thermal infrared spectrometry data: An initial temperature estimate bounded TESSMA search approach," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 1435–1446, July 2001.

[12] C.-I. Chang, S.-S. Chiang, J. A. Smith, and I. W. Ginsberg, "Linear spectral random mixture analysis for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, pp. 375–392, Feb. 2002.

[13] C.-I. Chang and H. Ren, "An experiment-based quantitative and comparative analysis of target detection and image classification algorithms for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1044–1063, Mar. 2000.

[14] C. A. Bateson and B. Curtiss, "A method for manual endmember selection and spectral unmixing," *Remote Sens. Environ.*, vol. 55, pp. 229–243, 1996.

[15] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Summaries of the VI JPL Airborne Earth Science Workshop*. Pasadena, CA, 1995.

[16] R. A. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, "Automatic endmember extraction from hyperspectral data for mineral exploration," in *Proc. 21st Can. Symp. Remote Sensing*, Ottawa, ON, Canada, 1999.

[17] J. Bowles, P. J. Palmadesso, J. A. Antoniades, M. M. Baumback, and L. J. Rickard, "Use of filter vectors in hyperspectral data analysis," *Proc. SPIE*, vol. 2553, pp. 148–157, 1995.

[18] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, pp. 2025–2041, Sept. 2002.

[19] C. A. Bateson, G. P. Asner, and C. A. Wessman, "Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 1083–1094, Mar. 2000.

[20] M. L. Mavrovouniotis, A. M. Harper, and A. Ifarraguerri, "Classification of pyrolysis mass spectra of biological agents using convex cones," *J. Chemometrics*, vol. 8, pp. 305–333, 1994.

[21] J. Bowles, D. Gillis, and P. Palmadesso, "New improvements in the ORASIS algorithm," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, 2000.

[22] A. Plaza, "Proposal, validation and testing of a new morphological method for the analysis of hyperspectral data which combines spatial and spectral information," Ph.D. dissertation, Comput. Sci. Dept., Univ. Extremadura, Cáceres, Spain, 2002.

[23] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A comparative analysis of endmember extraction algorithms using AVIRIS hyperspectral imagery," in *Summaries of the 11th JPL Airborne Earth Science Workshop*. Pasadena, CA, 2002.

[24] M. Garcia and S. L. Ustin, "Detection of interannual vegetation responses to climatic variability using AVIRIS data in a coastal savanna in California," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 1480–1490, July 2001.

[25] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 779–785, July 1994.

[26] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, pp. 529–545, Mar. 2000.

[27] C.-I. Chang, "An information theoretic-based approach to spectral variability, similarity and discriminability for hyperspectral image analysis," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1927–1932, Aug. 2000.

[28] J. Plaza, A. Plaza, P. Martinez, and R. Perez, "H-COMP: A tool for quantitative and comparative analysis of endmember extraction algorithms," in *Proc. IGARSS*, Toulouse, France, 2003.

[29] T. Rashed, J. Weeks, M. Gadalla, and A. Hill, "Revealing the anatomy of cities through spectral mixture analysis of multispectral satellite imagery: A case study of the Greater Cairo region, Egypt," *Geocarto Int.*, vol. 16, pp. 5–16, 2001.

[30] R. O. Green, B. Pavri, J. Faust, and O. Williams, "AVIRIS radiometric laboratory calibration, inflight validation and a focused sensitivity analysis in 1998," in *Summaries of the 8th JPL Airborne Earth Science Workshop*. Pasadena, CA, 1999.

Antonio Plaza received the M.S. and Ph.D. degrees in computer science from the University of Extremadura, Caceres, Spain, in 1997 and 2002, respectively.

He is currently an Associate Professor with the Computer Science Department, University of Extremadura, since 2000, where he was a Research Associate and Assistant Professor from 1997 to 1999. He was also a Visiting Researcher at Applied Information Sciences Branch, Goddard Space Flight Center, Greenbelt, MD, and the AVIRIS Data Facility, Jet Propulsion Laboratory, Pasadena, CA. His main research interests span computer vision, image processing, pattern recognition, and development and efficient implementation of hyperspectral image analysis algorithms on massively parallel computing facilities and VLSI hardware-based computer architectures.

Dr. Plaza is currently serving as a reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and IEEE TRANSACTIONS ON IMAGE PROCESSING.

Pablo Martínez received the Ph.D. degree in physics from the University of Granada, Granada, Spain, in 1992.

He has been a Professor of computer science with the University of Extremadura, Cáceres, Spain, since 1985. He is currently the Head Scientist of the Neural Networks and Signal Processing Group (GRNPS). He has held Visiting Researcher positions at the Applied Information Sciences Branch, Goddard Space Flight Center, Greenbelt, MD, and the Department of Electrical Engineering, University of Maryland, College Park. His main research interests include remote sensing, digital image analysis, hardware-based architectures, operating systems management and configuration, and neural network-based pattern recognition.

Rosa Pérez received the M.S. degree in mathematics from the University of Extremadura, Cáceres, Spain, in 1985, and the Ph.D. degree in computer science from the Polytechnic University of Madrid, Madrid, Spain, in 1995.

She is currently a Professor of computer science in the Computer Science Department, University of Extremadura, Cáceres, Spain, where she has been since 1985. Her main research interests include neural networks, systolic array- and FPGA-based design, pattern recognition, and signal processing.

Javier Plaza received the M.Sc. degree in computer science from the University of Extremadura, Cáceres, Spain, in 2002, where he is currently pursuing the Ph.D. degree.

He is currently an Assistant Professor in the Computer Science Department, University of Extremadura. His current research work is focused on the development of efficient implementations of nonlinear mixture model-based algorithms for abundance estimation of materials in hyperspectral scenes. He is also involved in the design and configuration of commodity cluster computing architectures for high-performance hyperspectral analysis. Other major research interests include development of quantitative and comparative applications for remote sensing, and the design of web-based applications for hyperspectral image processing.