Automated Extraction of Image-Based Endmember Bundles for Improved Spectral Unmixing

Ben Somers, Maciel Zortea, Antonio Plaza, *Senior Member, IEEE*, and Gregory P. Asner

***Abstract*—Spectral unmixing is an important task in hyperspectral data exploitation. It amounts to estimating the abundance of pure spectral constituents (endmembers) in each (possibly mixed) observation collected by the imaging instrument. In recent years, several endmember extraction algorithms (EEAs) have been proposed for automated endmember extraction from hyperspectral data sets. Traditionally, EEAs extract/select only one single standard endmember spectrum for each of the presented endmember classes or scene components. The use of fixed endmember spectra, however, is a simplification since in many cases the conditions of the scene components are spatially and temporally variable. As a result, variation in endmember spectral signatures is not always accounted for and, hence, spectral unmixing can lead to poor accuracy of the estimated endmember fractions. Here, we address this issue by developing a simple strategy to adapt available EEAs to select multiple endmembers (or bundles) per scene component. We run the EEAs in randomly selected subsets of the original hyperspectral image, and group the extracted samples of pure materials in a bundle using a clustering technique. The output is a spectral library of pure materials, extracted automatically from the input scene. The proposed technique is applied to several common EEAs and combined with an endmember variability reduction technique for unmixing purposes. Experiments with both simulated and real hyperspectral data sets indicate that the proposed strategy can significantly improve fractional abundance estimations by accounting for endmember variability in the original hyperspectral data.**

***Index Terms*—Endmember extraction algorithms (EEAs), endmember variability, hyperspectral imaging, multiple endmember spectral mixture analysis (MESMA), spectral mixture analysis (SMA).**

I. Introduction

SPECTRAL MIXTURE ANALYSIS (SMA) is a popular tool to analyze remotely sensed hyperspectral images [1]. No matter the spatial resolution, the spectral signatures collected in natural environments are invariably a mixture of the signatures of the various materials found within the spatial extent of the ground instantaneous field of view (GIFOV) of the imaging instrument [2]. A classic technique for spectral unmixing is *linear* SMA [2], which involves two steps: (i) identifying the unique signatures of the pure ground components or endmembers, i.e., *endmember extraction* [3], [4]; and (ii) estimating the proportions of different endmembers for each pixel by model inversion, i.e., *abundance estimation* [5]. This model assumes that the collected spectra at the spectrometer can be expressed in the form of a linear combination of endmembers weighted by their corresponding abundances. Let us assume that the spectral signature collected by an imaging spectrometer with n bands is denoted by $\mathbf{X} = [x_1, x_2, \cdots, x_n] \in \Re^n$, where $\Re$ denotes the set of real numbers in which the spectral response x_k at sensor channels $k = 1, \ldots, n$ is included. Under the linear mixture model assumption, $\mathbf{X}$ can be modeled using the following expression:

$$\mathbf{X} = \sum_{z=1}^{p} \Phi_z \cdot \mathbf{E}_z + \mathbf{n} \tag{1}$$

where $\mathbf{E}_z$ denotes the spectral response of endmember z, Φ_z is a scalar value designating the fractional abundance of the endmember z at $\mathbf{X}$, p is the total number of endmembers, and $\mathbf{n}$ is a noise vector. Two physical constraints are generally imposed into the model described in (1), these are the abundance non-negativity constraint (ANC), i.e., $\Phi_z \geq 0$, and the abundance sum-to-one constraint (ASC), i.e., $\sum_{z=1}^{p} \Phi_z = 1$. As indicated in [5], a fully constrained estimate can be obtained in least-squares sense by solving the optimization problem in (1) with the ASC and ANC constraints. Such fully constrained linear spectral unmixing estimate is generally referred to in the literature by the acronym FCLSU [6]. Commonly used inversion techniques for solving (1) are Gramm-Schmidt orthogonalization [7], maximum-likelihood [8] or least squares regression analysis [6], [9]. The latter solution is adopted in this study.

Recent research efforts in the spectral unmixing community have strongly focused on the development and optimization of algorithms for (semi)automated endmember extraction. Endmembers can also be extracted from field or laboratory spectra [10], but these spectra are rarely acquired under the same conditions as the remotely sensed data. Therefore, many classic endmember extraction algorithms (EEAs) focus on the automated selection of endmembers directly from the image data. Techniques in this category include the pixel purity index (PPI) [11], N-FINDR [12], iterative error analysis (IEA) [13], unsupervised fully constrained least squares (UFCLS) [6], optical real-time adaptive spectral identification system (ORASIS) [14], convex

Manuscript received October 01, 2011; revised November 20, 2011; accepted December 13, 2011. Date of publication January 31, 2012; date of current version May 23, 2012. This work was supported by the Belgian Science Policy Office in the framework of the STEREO II programme—Project VEGEMIX (SR/67/146), by the European Community's Marie Curie Research Training Networks Programme under reference MRTN-CT-2006-035927, Hyperspectral Imaging Network (HYPER-I-NET), by the Spanish Ministry of Science and Innovation (HYPERCOMP/EODIX project, reference AYA2008-05965-C04-02), and by the Junta de Extremadura (local government) under project PRI09A110.

B. Somers is with the Flemish Institute for Technological Research (VITO), Centre for Remote Sensing and Earth Observation Processes (TAP), BE-2400 Mol, Belgium.

M. Zortea and A. Plaza are with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, E-10071 Cáceres, Spain.

G. P. Asner is with the Department of Global Ecology, Carnegie Institution for Science, Stanford, CA 94305 USA.

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

Digital Object Identifier 10.1109/JSTARS.2011.2181340

cone analysis (CCA) [15], vertex component analysis (VCA) [16], and an orthogonal subspace projection (OSP) technique in [17], among many others [18]–[34]. Other endmember determination techniques have been developed without assuming the presence of pure signatures in the input data. These methods aim at generating *virtual* endmembers [35] (not necessarily present in the set comprised by input data samples) by finding the simplex with minimum volume that encompasses all observations [36]–[41]. However, the generation of such virtual signatures may not be easily associated with physical components in the scene.

Image endmembers have the advantage of being collected at the same scale and under the same environmental conditions as the remotely sensed data and can, thus, be more easily associated with spectral components in the image scene [3]. However, a critical note on classic EEAs is that they usually fail to cope with the severe effects of intra- and inter-class endmember variability [42]. Traditionally, EEAs extract image-wide endmember spectra. This implies that only one single standard endmember spectrum is defined for each of the presented endmember classes or scene components. The standard endmember spectra are subsequently assigned to each image pixel, and sub-pixel cover fractions are calculated. The use of fixed endmember spectra, however, is a simplification since in many cases the conditions of the scene components are spatially and temporally variable. The resulting variation in endmember spectral signatures is not accounted for and, hence, unmixing can lead to poor accuracy of the estimated endmember fractions. This problem is referred to as the endmember variability problem and has been identified as one of the most profound sources of error in cover fraction estimates provided by linear SMA [42].

Solutions to overcome endmember variability have been presented in recent literature (see [42] and references therein). Multiple endmember spectral mixture analysis (MESMA) [43] and AutoMCU [44] are generally adopted endmember variability reduction techniques showing systematic and significant improvements in cover fraction estimates over traditional linear SMA techniques [45]. Yet, the premise of these endmember variability reduction techniques is the availability of a spectral library which should contain representative instances of all endmembers present within the scene. In other words, the library should allow modelling the spectral variability of the endmembers. Such extended spectral libraries are commonly built from field measurements, mainly because the aforementioned EEAs are bound to extract a fixed number of endmembers which is given by the use of a single endmember per pure class. In an operational setting, these libraries of field measurements may not be available, thus introducing limitations in the operational applicability of endmember variability reduction techniques.

To address this problem, in this work we propose a simple strategy for adapting available EEAs (currently providing a single spectrum per endmember class) in order to select multiple endmembers (or bundles) per scene component. This idea is not new, as previous work [46], [47] already highlighted this need more than a decade ago. Unfortunately, the message and the presented techniques have not been generally adapted, and the discussion on this dual endmember extraction-variability concept quickly lost attention. As a result, one of our goals in this paper is to direct the attention of the community to the advantages that can be gained by adding variability to already existing EEAs. To demonstrate the added value of mutually addressing endmember extraction and endmember variability in linear SMA, a conceptually straightforward technique to derive endmember bundles from an image scene is presented and discussed. This technique is applied to a set of widely used automated EEAs and combined with a common endmember variability reduction technique providing abundance estimates for each scene component, i.e., the well-known MESMA approach originally developed in [43]. The integration of these concepts gives a new unmixing chain for linear SMA addressing endmember variability, which is quantitatively evaluated in this work using both simulated and real hyperspectral data sets.

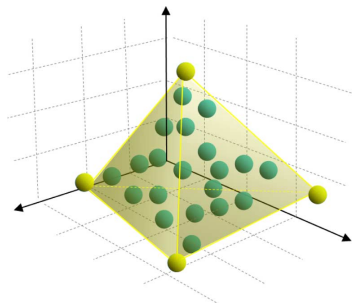

Fig. 1. Geometric interpretation of the mixture problem in a three-dimensional space.

II. Methodology

A. Geometric Interpretation of the Mixture Problem

Equivalent to the mathematical description in (1), the linear mixed-pixel problem can also be approached from a geometrical point of view [11], [12]. The pixels of a (hyper)spectral image cube can be cast as a scattering of points in an n-dimensional Euclidean space, where n is the number of spectral bands. Each spectral band is one axis of the space and every pixel in the image has a point associated with it. The data cloud can be enclosed by a simplex whose vertices are the endmembers or most purest pixels in the image. Cover fractions are determined by the position of spectra within the simplex and can be considered relative coordinates in a new reference system determined by the endmembers. This geometrical interpretation of the mixed-pixel problem is demonstrated in Fig. 1 for a simple mixture model represented in a three-dimensional space, using four endmembers to create the simplex. Many of the endmember extraction algorithms described in the next subsection can be interpreted using this simple geometric representation of the mixture problem.

B. Automated Endmember Extraction

Five EEAs are considered in this work to automatically extract endmember spectra from hyperspectral imagery. All these well known methods have in common that only a single pixel (spectrum) is extracted to represent each pure material present

in the image, under the assumption of the presence of pure pixels in the scene:

1) N-FINDR [12]: This algorithm looks for the set of pixels with the largest possible volume by *inflating* a simplex inside the data. The procedure begins with a random initial selection of p pixels. Every pixel in the image must be evaluated in order to refine the estimate of endmembers, looking for the set of pixels that maximizes the volume of the simplex defined by selected endmembers. The corresponding volume is calculated for every pixel in each endmember position by replacing that endmember and finding the resulting volume. If the replacement results in an increase of volume, the pixel replaces the endmember. This procedure is repeated until there are no more endmember replacements.
2) OSP [17]: This algorithm starts by selecting the pixel vector with maximum length in the scene as the first endmember, then it looks for the pixel vector with the maximum absolute projection in the space orthogonal to the space linearly spanned by the initial pixel, and labels that pixel as the second endmember. A third endmember is found by applying an orthogonal subspace projector to the original image [17], where the signature that has the maximum orthogonal projection in the space orthogonal to the space linearly spanned by the first two endmembers. This procedure is repeated until the desired number of endmembers, p, is found.
3) VCA [16]: This algorithm also makes use of the concept of orthogonal subspace projections. However, as opposed to the OSP algorithm described above, the VCA exploits the fact that the endmembers are the vertices of a simplex, and that the affine transformation of a simplex is also a simplex [16]. As a result, VCA models the data using a positive cone, whose projection onto a properly chosen hyperplane is another simplex whose vertices are the final endmembers. After projecting the data onto the selected hyperplane, VCA projects all image pixels to a random direction and uses the pixel with the largest projection as the first endmember. The other endmembers are identified in sequence by iteratively projecting the data onto a direction orthogonal to the subspace spanned by the endmembers already determined. The new endmember is then selected as the pixel corresponding to the extreme projection, and the procedure is repeated until a set of p endmembers is found.
4) UFCLS [6]: This algorithm takes advantage of the FCLSU method in [5] to find a set of endmembers in an iterative fashion. It first finds the pixel vector with the maximum length in the hyperspectral image and selects it as the first endmember. It then assumes that all other pixels in the image scene are pure pixels made up of the initial endmember with 100% abundance. Of course, this is not generally true, and thus it subsequently finds a pixel that has the largest least squares error from the initial pixel and selects it as the second endmember. The same procedure is repeated until a set of p endmembers is found.
5) IEA [13]: As in the case of UFCLS, this algorithm performs a series of fully constrained unmixing operations, 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. FCLSU is then applied in using this vector as endmember, 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 p, a number of pixels t, and an angle value θ. The parameter t 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 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 p endmembers have been selected. In this work, we use a special case of the IEA algorithm in which the final spectra obtained by the algorithm are not derived as averaged values of a set of pixels but as real pixels in the data instead, i.e., we set $t = 1$ and $\theta = 0$. The reason is that all other EEAs derive image pixels as endmembers and we would like to have a fair inter-comparison of methods. Therefore, the only difference between the UFCLS algorithm and our version of the IEA algorithm is the starting value that each algorithm uses. While the UFCLS selects the pixel vector with the maximum length as the initial target pixel, the IEA algorithm selects the mean spectrum of the data to start its searching process.

C. Proposed Endmember Bundle Extraction

A straightforward approach to introduce endmember variability in EEAs is presented in Fig. 2, which diagrammatically summarizes the different steps involved in our proposed unmixing chain. The idea is to run the EEAs in randomly selected subsets of the original hyperspectral image. Sampling with or without replacement can be considered for the generation of subsets. Our main assumption for adopting this random strategy in the selection of pixel vectors is that one can approximate the statistics of the original image with a smaller percentage of image pixels [26]. This means that, if a reasonable number of pure pixels are available for each endmember in the scene, the pure pixels will also be present in the image subsets generated by random sampling. Obviously, the validity of this assumption depends upon the size of the image, the number of endmembers present in the scene, and how the subsets are generated. In principle, if they are indeed present, distinct pure pixels could be found in each subset of the image using an EEA such as those described in Section II-B. For each run, a unique set of endmembers is extracted by the considered EEA. Once all image subsets have been analyzed a spectral library of extracted endmembers is generated. The library is now a set of spectra from the different ground components, that needs to be ordered. By using a k-means clustering algorithm, with the Euclidean distance as similarity measure [48], the spectral library is divided into separate endmember bundles for each ground component. Once the endmember bundles are obtained, they are used as input to MESMA. Candidate endmembers are randomly selected from the extracted bundles and the FCLSU iterated each time using another combination of candidate endmembers. The minimum

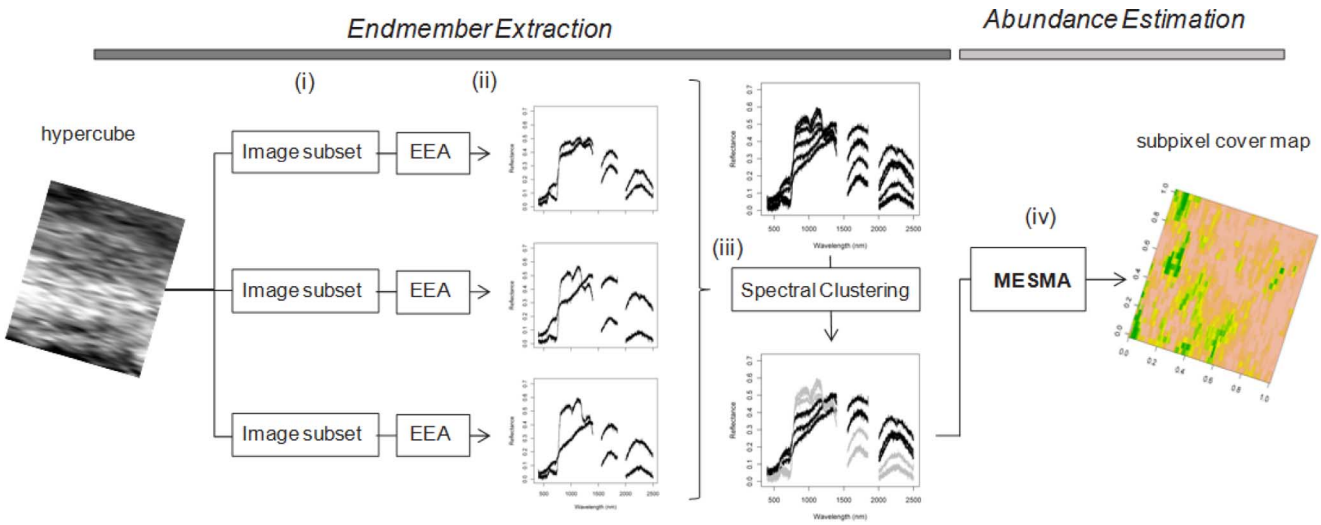

Fig. 2. Block diagram illustrating the proposed methodology for automatic extraction of endmember bundles and unmixing: the image is first subdivided in a number of subsets, whose pixels are randomly selected from the image (i) and the EEA is applied on each subset (ii); endmember bundles are constructed using a clustering algorithm (iii); the bundles are implemented in MESMA providing a sub-pixel cover distribution map (iv).

RMSE criterion is then used to provide the sub-pixel cover fraction estimates. It is worth noticing that, when sampling without replacement is used to generate the subsets (each pixel is assigned to a unique subset), the computational burden of the endmember extraction process is not increased.

D. Abundance Estimation

Two different strategies have been considered in this work to solve (1) in order to derive the endmember fractional abundances:

1) FCLSU [5]: This approach corresponds to the fully constrained, least-squares unmixing technique which imposes the ASC and ANC constraints for abundance estimation using fixed endmembers.
2) MESMA [43]: This approach allows endmembers to vary on a per-pixel basis and thereby allows to cope with the effects of intra- and inter-class endmember variability in spectral unmixing. The MESMA procedure can be described as follows (i) FCLSU is run in iterative fashion; (ii) Each run different endmember combinations, randomly selected from a spectral library, are used to decompose each pixel; (iii) The model with the best fit, i.e., with the lowest root mean square error (RMSE) [3] in the reconstruction of the original pixel, is adopted. MESMA requires as input for each ground component or endmember an extensive library of field, laboratory, and/or image spectra.

III. Simulated Image Experiments

In this section we evaluate the proposed unmixing chain based on automated extraction of image-based endmember bundles using simulated hyperspectral data. The simulated data provides perfect reference or ground-truth data, allowing a thorough validation of the proposed methodology in a fully controlled analysis scenario.

A. Data

A simulated image scene with spectral range of 400–2500 nm, spectral resolution of 1 nm and image size of 500 pixels was compiled from 150 *in situ* measurements of bare soil (sandy texture, gravimetric moisture content ranging between 0 and 15%), citrus trees (*Citrus sinensis*, 9-year old with tree height of 3 m and canopy width of 3 m) and weed canopy (*Echium* sp. L.) spectra. The spectra were measured in a citrus orchard field near Wellington, South Africa, using a full-range FieldSpec JR spectroradiometer with a 25° foreoptic (Analytic Spectral Devices, Boulder, CO, USA). More details on the field measurements can be found in [49]. Using these measurements, mixed spectra were generated according to (1). Endmember combinations and subpixel cover fractions were randomly assigned to each pixel. A flat 1% reflectance spectrum was used as a surrogate for shading and shadowing effects. Random noise was added to simulate contributions from ambient and instrumental sources. White Gaussian noise was added using a standard normal distribution of randomly generated numbers, i.e.

$$\hat{x}_k = x_k + \left(A \times \frac{e(0,1)}{\mathrm{SNR}} \right) \tag{2}$$

where x_k is the reflectance at the k-th wavelength or spectral band, $e(0,1)$ is a randomly generated number from a normal distribution with a mean of zero and a standard deviation of one [3], and A is the assumed reflectance for the spectrum. For the simulations, we consider a 50:1 SNR for each band as the ratio of the 50% (i.e., A) signal level to the standard deviation of the noise. This results in a noise standard deviation that is roughly proportional to the average signal, a phenomenon often observed in radiometric data [50]. For each endmember, except for shadow, approximately 10% pure pixels (i.e., 100% fraction cover) were added to the scene. No additional pure shadow pixels were included because we believe that, given the relatively coarse spatial resolution of current operational hyperspectral sensors (i.e.,

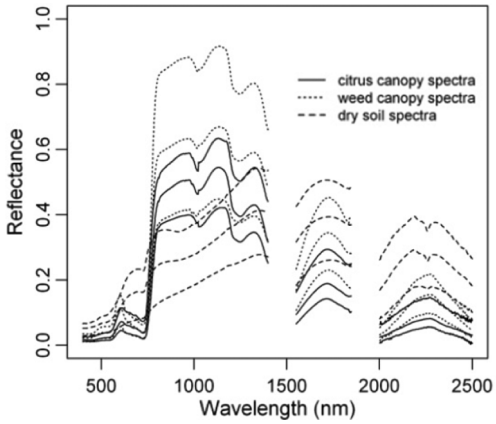

Fig. 3. The minimum, mean and maximum spectra of the reference endmember bundles used as input for generating the simulated orchard scene.

in the order of magnitude of at least a few meters) and the structural design of a fruit orchard, pure shadow pixels are sparse in an operational setting [51].

B. Experiments

The accuracy of the proposed MESMA-based unmixing chain in Fig. 2 was compared to the traditional FCLSU approach (i.e., no endmember bundles) using the simulated data described in the previous subsection. Both unmixing chains are further cross-referenced to a 'best case solution' in which the endmember spectra are directly extracted from the spectral libraries of field measurements used as input in the image simulation (i.e., reference or ground truth). The minimum, mean and maximum spectra of the reference endmember bundles is shown in Fig. 3. The RMSE between the original and the reconstructed hyperspectral scene, the abundance error (Δf [21]), the coefficient of determination (R^2 [45]), the slope, and the intercept of the estimated versus real (i.e., ground truth) cover fractions in a scatterplot were used to evaluate the performance of the considered unmixing chains.

Before presenting the unmixing results, we first briefly discuss on how to deal with shadow effects when implementing EEAs. In our experiment, the endmembers of interest are: (i) citrus canopy, (ii) weed canopy, (iii) bare soil and (iv) shadow. Accordingly, the EEAs should be run to detect $p = 4$ endmembers. However, when doing so we observed that the OSP and VCA approaches were not able to extract a shadow endmember. Instead, an additional soil spectrum was selected as the fourth endmember candidate. This situation is graphically illustrated in Fig. 4, which shows the $p = 4$ endmembers extracted by the different EEAs with the fourth endmember marked in bold. The reason why OSP and VCA did not extract the shadow endmember as the fourth endmember (as it was the case with the other tested methods) is likely due to the fact that no pure shadow pixels were present in the scene and the endmember searching strategy adopted by these algorithms did not properly separate the shadow endmembers from the previously identified endmembers, particularly, due to the high presence of noise that can be observed in the simulated data, as indicated by the spectral plots in Fig. 4. In order to address this issue, we decided to run both OSP and VCA to detect $p = 3$ endmembers and manually added a shadow endmember. This is traditionally done by modeling the shadow as a flat 1% reflectance spectrum [52], [53].

The performance of the different unmixing chains as applied on the simulated citrus orchard scene is reported in Table I. The reference chain shows a 'best case scenario' in which the endmembers are directly extracted from the spectral libraries used as input in the image simulation (see Fig. 3). In the FCLSU approach only the average spectrum of each endmember bundle is used, while the MESMA approach iteratively searches through the entire library as such, accounting for endmember variability. Results clearly demonstrate the relevance of addressing endmember variability. MESMA visibly outperforms the traditional FCLSU ($\Delta\mathrm{RMSE} = 0.005, \Delta f = 0.03, \Delta R^2 = 0.11$). Overall, the best accuracy among EEAs was observed for the OSP ($\mathrm{RMSE} = 0.012, \Delta f = 0.08, R^2 = 0.85$). The most significant improvements were obtained for the two vegetation endmembers (i.e., weed and crop; $\Delta_{\Delta f} = 0.02$, $R^2 = 0.06$). As indicated by Table I, by introducing endmember bundles through MESMA, the high spectral similarity between the weed and citrus canopies can be modeled more effectively, resulting in improved abundance estimates.It should be noted that in our analysis we used EEAs which provided different endmembers for different runs (as it is the case with N-FINDR and VCA, which start from a random initial condition), and also EEAs which start from a fixed initial condition (i.e., the brightest pixel in the image in the case of OSP and UFCLS, or the median spectrum of the data in the case of IEA). For those algorithms with a random initial condition (N-FINDR and VCA) we experimentally observed that the variations in RMSE for different runs are very small as the algorithms provide endmembers which are spectrally very similar across different runs and, hence, the variations in RMSE due to different initialization conditions are negligible.

Although the accuracy of the EEA-driven unmixing chains did not equal that of the reference case, on average the differences were very small ($\Delta\mathrm{RMSE} = 10^{-4}, \Delta f = 0.02, \Delta R^2 = 0.04$). This again illustrates the potential of using automated EEAs for unmixing purposes. The residual difference in unmixing accuracy can be partly explained by the reduced SNR associated with the image-based endmembers [3], but in our context it is mainly due to the fact that the extracted endmembers deviate from the real endmember means used in the reference scenario [53]. This is illustrated in Fig. 5, in which the endmembers automatically extracted by different EEAs are cross-referenced to the ground-truth endmembers. In this scenario, IEA produces a higher average RMSE compared to the other techniques which can explain its lower overall unmixing accuracy in Table I.

At this point, it is important to emphasize that the properties of the extracted endmember bundles were not only dependent upon the EEA applied but also by the size and number of image subsets used in the analysis. To analyze this relevant aspect, the average deviation (in percentage) between the (mean and standard deviation of the) reference and extracted endmember bundles, plotted against the size (i.e., expressed as the percentage of the total image size) and the number of image subsets is shown in Fig. 6. For the citrus orchard scene, the optimum is obtained

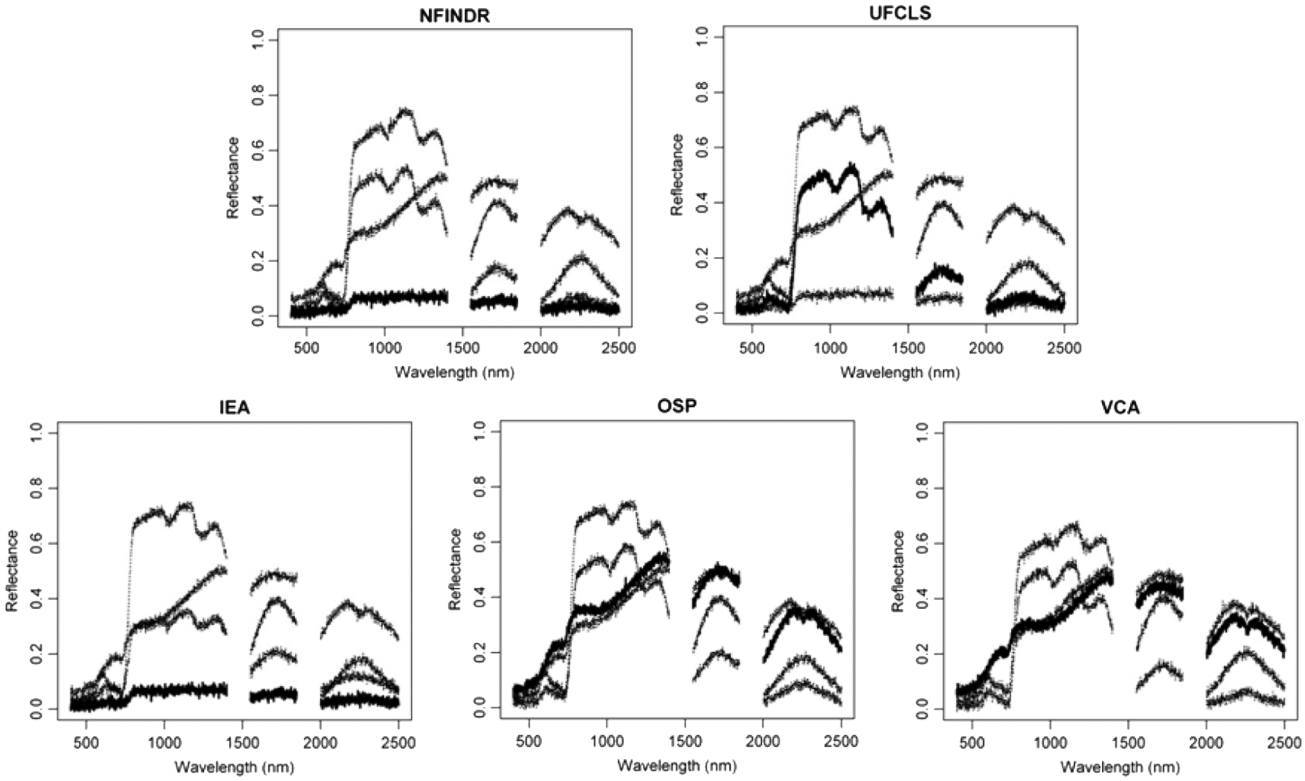

Fig. 4. Endmembers extracted by the five considered EEAs in the simulated image experiment. High noise can be appreciated in the derived spectral signatures. The fourth endmember extracted by the different methods is marked in bold.

TABLE I
THE RMSE BETWEEN THE ORIGINAL AND THE RECONSTRUCTED HYPERSPECTRAL SCENE, THE ABUNDANCE ERROR (Δf), AND THE R^2, SLOPE AND INTERCEPT OF THE ESTIMATED VERSUS REAL (I.E., GROUND-TRUTH) COVER FRACTIONS IN A SCATTERPLOT FOR THE TRADITIONAL FCLSU CHAIN, I.E., NO ENDMEMBER BUNDLES, AND THE MESMA UNMIXING CHAIN INCLUDING THE ENDMEMBER BUNDLE APPROACH IMPLEMENTED USING DIFFERENT EEAS. THE REFERENCE IS A 'BEST CASE SCENARIO' IN WHICH THE ENDMEMBERS ARE DIRECTLY EXTRACTED FROM THE SPECTRAL LIBRARIES USED AS INPUT IN THE IMAGE SIMULATION (SEE FIG. 3)

Score	RMSE		Δf						R^2						Slope						Intercept					
Unmixing	FCLSU	MESMA	FCLSU			MESMA			FCLSU			MESMA			FCLSU			MESMA			FCLSU			MESMA		
Endmembers	all	all	tree	weed	soil	tree	weed	soil	tree	weed	soil	tree	weed	soil	tree	weed	soil	tree	weed	soil	tree	weed	soil	tree	weed	soil
Reference	0.015	0.010	0.09	0.10	0.06	0.05	0.05	0.05	0.80	0.72	0.92	0.94	0.90	0.93	0.92	0.77	0.99	1.08	1.00	1.03	0.05	0.04	0.02	0.01	0.00	0.01
N-FINDR	0.015	0.012	0.10	0.10	0.10	0.08	0.09	0.07	0.76	0.67	0.88	0.84	0.77	0.89	1.01	1.00	1.20	1.05	0.99	1.09	0.03	0.01	0.07	0.02	0.00	0.06
OSP	0.015	0.012	0.09	0.10	0.09	0.07	0.08	0.07	0.79	0.70	0.90	0.86	0.78	0.90	1.10	1.01	1.08	1.17	1.02	1.09	0.04	0.07	0.02	0.01	0.03	0.05
VCA	0.014	0.011	0.09	0.11	0.12	0.08	0.08	0.08	0.79	0.68	0.86	0.79	0.74	0.91	0.96	0.89	1.23	0.90	0.93	1.08	0.03	0.02	0.07	0.01	0.01	0.05
IEA	0.017	0.012	0.12	0.12	0.11	0.09	0.10	0.09	0.74	0.65	0.86	0.82	0.69	0.88	0.93	1.01	1.17	1.17	0.93	1.08	0.05	0.07	0.08	0.02	0.01	0.06
UFCLS	0.015	0.012	0.12	0.11	0.09	0.09	0.09	0.09	0.79	0.71	0.88	0.80	0.74	0.89	1.13	0.93	1.09	1.09	1.01	1.24	0.02	0.01	0.05	0.05	0.03	0.05

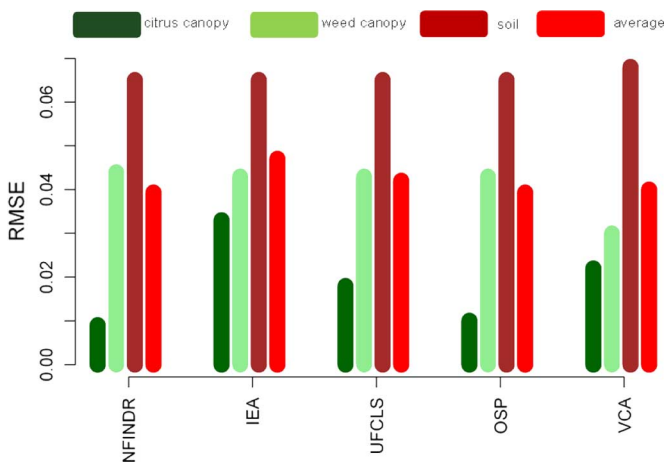

Fig. 5. RMSE between the reference endmembers and those extracted automatically by different EEAs.

with ten subsets each 10% of the total image size. As shown by Fig. 6, the use of too few subsets or subsets that are too big in size (e.g., 50% of the total image size) results in an incomplete modeling of endmember variability. Quite opposite, the use of too many subsets or subsets that are too small (e.g., 5% of total image size) reduces the correlation between extracted and reference endmember bundles. The EEAs select the 'purest' pixel(s) for each subset. As a result, the smaller the subset, the higher the risk that no pure pixel is present, and that mixed pixels are selected as endmembers instead. This is demonstrated in Fig. 7, which shows the endmember bundles extracted by OSP for different subset sizes. Results for the 5% subset size reveal the selection of endmembers that are not in the reference endmember bundle. This is particularly true for the weed and soil endmembers. Results for the 50% subset size, on the other hand, show endmember bundles that do not incorporate enough spectral variability. The best agreement between extracted and refer-

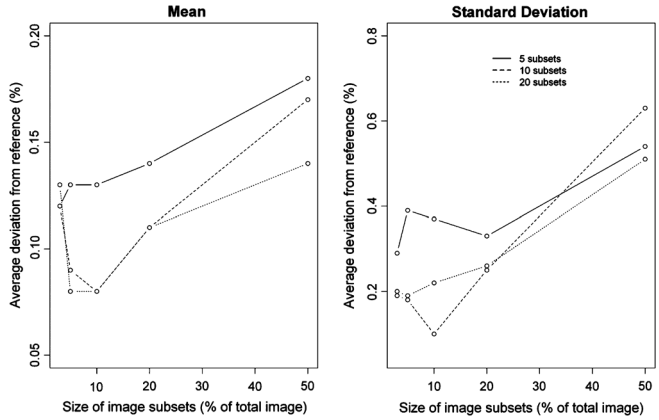

Fig. 6. The average deviation (in percentage) between the mean (left) and standard deviation (right) of the reference endmember bundles in Fig. 3 and the endmember bundles extracted by the EEAs, plotted against the size of the image subsets. Results are shown for a different number of image subsets.

ence bundles in our experiments was obtained for a 10% subset size.

Finally, the trend in accuracy of the OSP-MESMA unmixing chain for ten subsets of increasing size is provided in Fig. 8. Notwithstanding the fact that result verifies the assumption that the highest accuracies were obtained for a subset size of 10%, improvements of the MESMA approach over the traditional FCLSU approach were also observed for the other subset sizes (i.e., 3, 5, 20 and 50). Similar conclusions could be drawn for all EEAs. These results suggest the relevance of introducing endmember bundles in automated endmember extraction.

IV. Real Image Experiments

While the simulations conducted in the previous section are ideally suited to test the proposed methodological concept, real hyperspectral datasets are used in this section in order to allow for a demonstration of the operational potential of the proposed approach. We consider two different case studies showing the potential of the proposed methodology in two different analysis scenarios: (a) invasive species mapping in Hawaiian rainforests, and (b) mineral mapping in the Cuprite mining district, Nevada.

A. Invasive Species Mapping in Hawaiian Rainforests

In this case study, we evaluated the ability of the proposed unmixing chain to monitor the spread of invasive plant species in a montane rainforest area of the Hawaii Volcanoes National Park, Island of Hawaii. The forests are mainly characterized by a patchy mosaic of two common overstory species, i.e., native Hawaiian *Metrosideros polymorpha* and the invasive, N-fixing *Myrica faya*. An Earth Observing-1 Hyperion satellite image collected on August 15, 2005 was used for this analysis, and a set of field observations of species location (average uncertainty of $\approx$2 meters) was available for the scene. Based on the field observations, spectrally 'pure' image pixels could be identified and used to reconstruct reference endmember bundles for each tree species (see Fig. 10). Further analyses focused on a 30×90-pixel subset. The field observations contained within this region of interest were not used to construct the endmember bundles, allowing them to act as an independent validation set.

The Hyperion image (30-meter spatial resolution; 220 spectral bands covering the spectral range 400–2500 nm) was atmospherically corrected using the ACORN-4 model (Imspec, Palmdale, CA, USA). A de-striping algorithm was applied to correct for miscalibration between cross track detectors [54]. Due to its low SNR, the SWIR2 (2000–2500 nm) spectral region, as well as the major water-vapor absorption band centered around 1400 nm were excluded [53], leaving 118 spectral bands available for further analysis. After atmospheric correction and image georegistration, the different FCLSU and MESMA unmixing chains were applied as to provide invasive species maps. The endmember bundles extracted using the different EEAs were compared to the reference endmember bundles. Unmixing accuracy was further evaluated using the RMSE between the original and the reconstructed hyperspectral scene, while field observations of tree locations within the region of interest were used to evaluate the ability of the different approaches in successfully detecting patches of invasive species. The left panel of Fig. 9 shows the average deviation (in percentage) between the extracted and the reference bundles, plotted against the number of image subsets. The center and right panel respectively show the EEA specific trends for the mean and standard deviation. Overall, the optimum was reached for 30 subsets (as opposed to 10 for the simulated scene, as displayed in Fig. 6) and an image subset size of 10%. As shown by Fig. 9, VCA systematically showed higher accuracies as compared to the other techniques, while the opposite was true for IEA. The optimal performance of VCA in extracting proper endmember bundles could further be illustrated by Fig. 10. This figure shows the 95% confidence interval of the reference endmember bundles (dotted lines) and the endmember bundles as extracted by VCA (left panels) and N-FINDR (right panels) for 30 subsets of a 10% image size (solid lines). Whereas the endmember bundles extracted by N-FINDR slightly over- and underestimate the SWIR reflectance of *Myrica Faya* and *Metrosideros Polymorpha* respectively, an overall better accordance between the extracted and the reference bundles could be observed for VCA.

Subsequently, the endmember bundles extracted by VCA were used as input for MESMA to conform the proposed unmixing chain in this particular case study. The benefits of introducing endmember variability in unmixing are quantitatively demonstrated in Fig. 11. In this figure, the unmixing accuracy of (i) the 'best case scenario' with the reference endmember bundles as input in MESMA (left panels), (ii) the traditional VCA approach with FCLSU (middle panels), and (iii) the MESMA approach with the VCA extracted endmember bundles (30 subsets, subset size of 10%; right panels) are compared. The top panels show a significant decrease in RMSE for the MESMA approaches as compared to the traditional VCA-FCLSU approach. The bottom panels demonstrate the improved efficiency of the MESMA approach to detect patches of invasive species. The pixels with a cover fraction value of $\geq$75% for the *Myrica faya* endmember are highlighted and cross-referenced to field observations of the invasive species (black crosses in Fig. 11). The increased accuracy of the endmember bundle approach translates itself in a significant reduction in false negative (FN) observations (estimated as 20% for MESMA with reference spectra, as 33% for MESMA

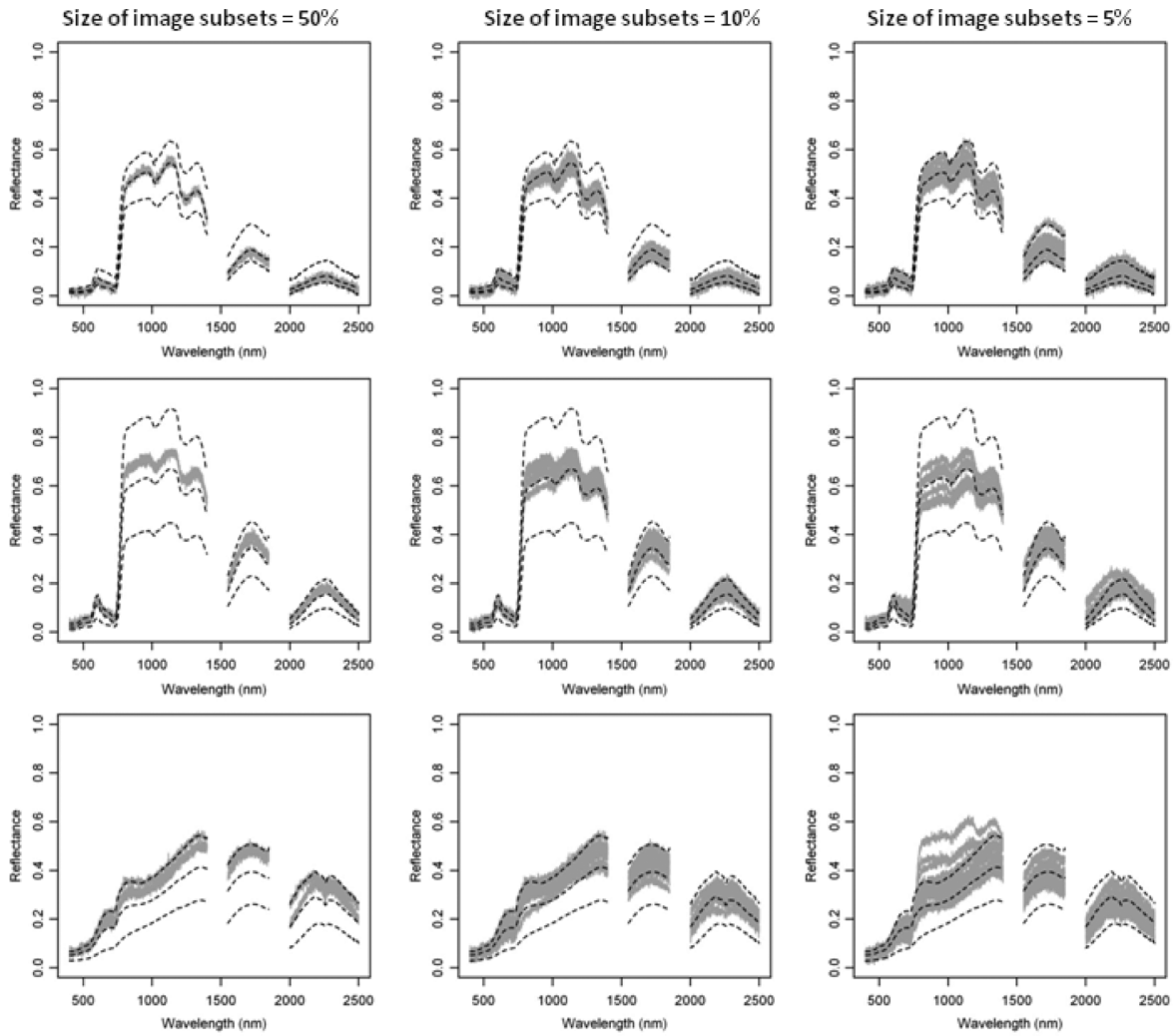

Fig. 7. Mean and 95% confidence intervals of the endmember bundles extracted by OSP (grey) and the reference endmember bundles (dotted lines). Results are shown for ten subsets of different sizes (expressed as the percentage of the total image size). The top row corresponds to the *citrus* canopy, the middle row corresponds to the *weed* canopy, and the bottom row corresponds to the *bare soil*.

with VCA endmembers, and as 60% for FCLSU with VCA endmembers). Overall, it was observed that traditional EEAs result in a significant underestimation of *Myrica faya* cover. As such, the invasive species patch, highlighted with a black circle in Fig. 11 could not be detected by the traditional VCA approach with FCLSU. When endmember variability was introduced in the VCA approach through image subsetting and MESMA, the invasive species patch could be detected successfully (see bottom right panel in Fig. 11).

B. Mineral Mapping in the Cuprite Mining District, Nevada

To evaluate the effectiveness of the endmember bundle extraction technique in a different analysis scenario with a high number of endmembers, a 1997 Airborne Visible-Nearinfrared Imaging Spectrometer (AVIRIS) image of Cuprite, Nevada USA (224 spectral bands between 0.4–2.5 μm, full width at half maximum of 10 nm, spatial resolution of 20 meters) was used. This scene has been widely used to validate the performance of endmember extraction algorithms [55]. After exclusion of the major atmospheric water vapor absorption regions, 192 spectral bands remained available for analysis. As in previous studies [56], the accuracy of the different FCLSU and MESMA unmixing chains was evaluated based on the RMSE between the original and the reconstructed hyperspectral scene. In all cases, the number of endmembers to be extracted was set to $p = 14$ after the consensus reached between two of the most popular methods for estimating the number of endmembers in hyperspectral data: virtual dimensionality [57] and HySime [58]. In this experiment, no reference endmember bundles were available for the Cuprite data. Fig. 12 and Fig. 13 show the 14 endmember bundles as extracted by, respectively, the VCA and N-FINDR subsetting approach with ten subsets of a 10% image size (i.e., with identical computational complexity). As shown by both figures, nicely delineated and unique spectral bundles can be identified which strongly suggests the potential of endmember bundle extraction in complex scenes. Similar results (not reported here for space considerations) were observed for the other EEAs. Implementing these endmember bundles in MESMA resulted in a significant increase in modeling accuracy, expressed using RMSE as the baseline metric,

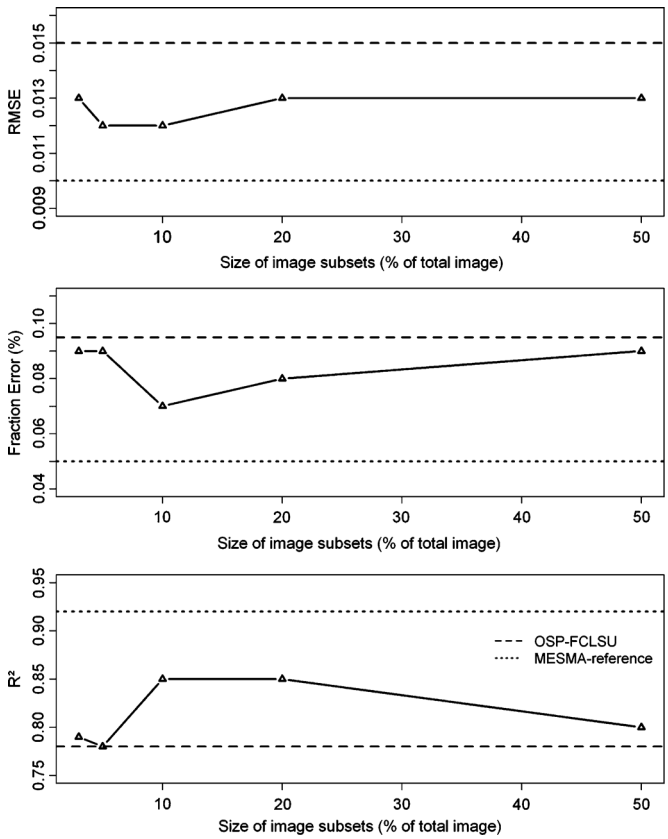

Fig. 8. RMSE, Δf and R^2 for the OSP-MESMA approach plotted against different image subset sizes. The indicated benchmarks are (i) the MESMA approach with the reference endmember bundles (Fig. 3) as input (MESMA-reference) and (ii) the traditional OSP approach, not incorporating endmember bundles (OSP-FCLSU).

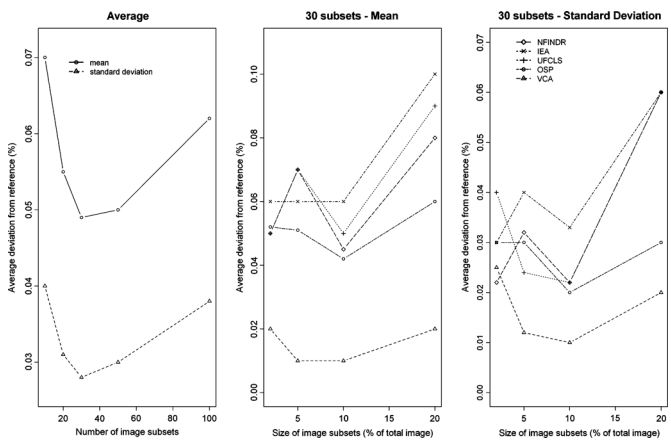

Fig. 9. The average deviation (in percentage) between the extracted and the reference bundles, plotted against the number of image subsets (left); the EEA specific trends in deviation for the mean (center), and standard deviation (right).

as compared to the traditional FCLSU approach. In Fig. 14 this aspect is demonstrated using VCA but again we emphasize that similar results were obtained for the other tested EEAs.

To conclude this section, we emphasize that the computational complexity for the EEAs considered in this study can be found in the literature [17], [12], [16]. For the proposed endmember bundles approach, the computational complexity is approximately given by the product of the computational com-

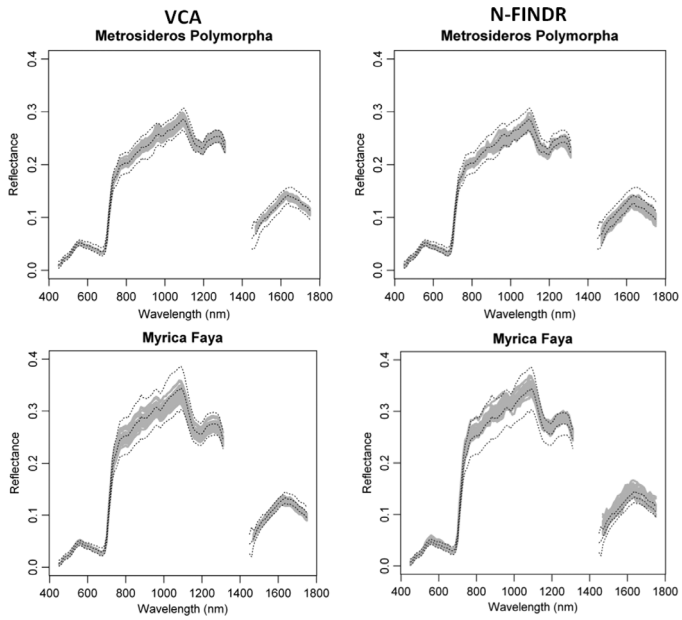

Fig. 10. Mean and 95% confidence intervals of the *Metrosideros polymorpha* (top) and *Myrica faya* (bottom) endmember bundles as extracted (grey) by VCA (left) and N-FINDR (right) and the corresponding reference endmember bundles (dotted lines).

plexity of each base EEA algorithm, times the number of image subsets, times the size of each subset. Note, for instance, that by using 10 image subsets, each of size 10% of the image, no virtual increase in the computational load is expected. The processing time for random sampling of the image, and the k-means clustering applied to relatively small bundle sets (when compared to the size of the scene) is of little practical concern.

V. Conclusions and Future Research Lines

The value of sub-pixel cover maps provided by linear SMA has been widely acknowledged and, together with others, we are convinced that spectral unmixing has the potential of becoming a mainstream image processing protocol in many operational remote sensing applications. However, the pertinent lack of synergy between existing solutions for the different steps of the unmixing chain (i.e., endmember extraction and abundance estimation) currently still prevents fully exploring the operational potential of SMA:

- On the one hand, currently available automated EEAs fail to cope with the effect of intra- and inter-class endmember variability. As opposed to endmembers taken from field or laboratory spectra, image endmembers can be used without any calibration. However, by defining image-wide endmember spectra the EEAs do not account for spatial or temporal variability in spectral endmembers. Hence, unmixing can lead to suboptimal accuracies when estimating endmember fractions.
- On the other hand, a number of solutions are already available to effectively account for endmember variability and thus significantly improve the accuracy of the sub-pixel abundance estimations. These techniques, however, require the availability of large spectral libraries of endmember constituents accounting for each plausible

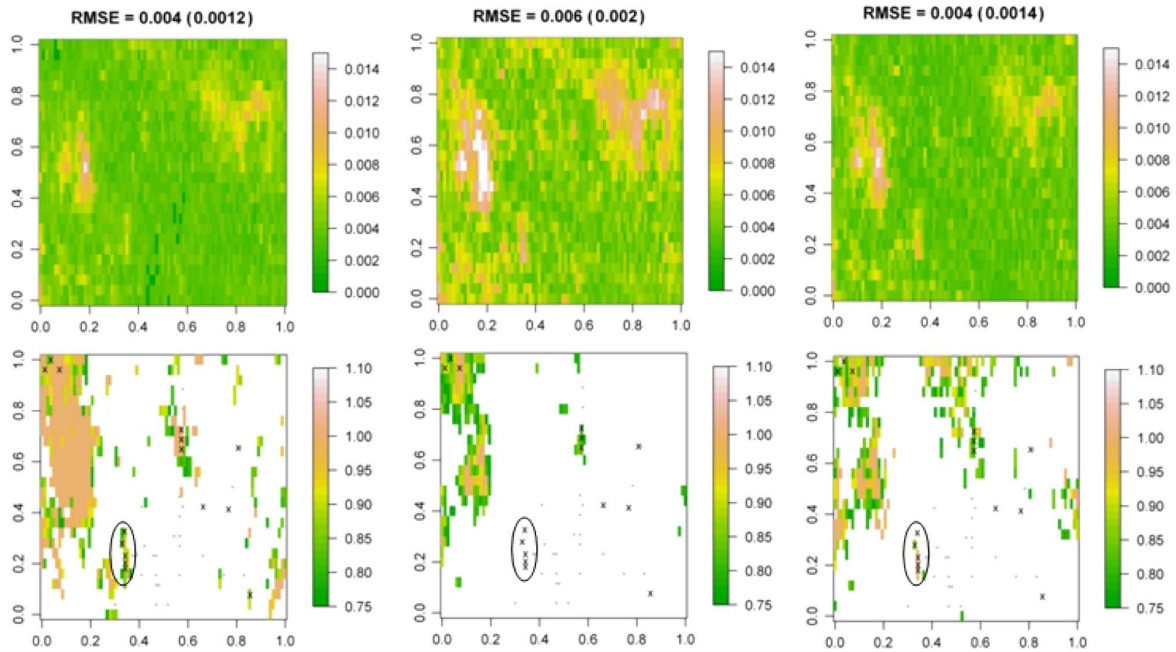

Fig. 11. (Top row) RMSE map obtained after reconstructing the Hyperion Hawaii scene for (i) the MESMA approach with the reference endmember bundles used as input (leftmost column), (ii) the traditional VCA approach with FCLSU, not incorporating endmember bundles (middle column), and the VCA approach incorporating endmember variability through subsetting and MESMA (rightmost column). The mean RMSE as well as the standard deviation (in the parenthesis) are indicated on top of the image. (Bottom row) *Myrica faya* cover fraction map as provided by the different unmixing chains. Only the pixels indicating a significant cover of the invasive species, i.e., $\geq$75%, are highlighted. The field observations of *Myrica faya* are indicated as black crosses, while the observations of *Metrosideros polymorpha* are visible as grey dots. The black circle in the bottom panels highlights an invasive species patch not detected by the traditional VCA approach (with FCLSU) in the middle panel, yet successfully detected by the MESMA approaches.

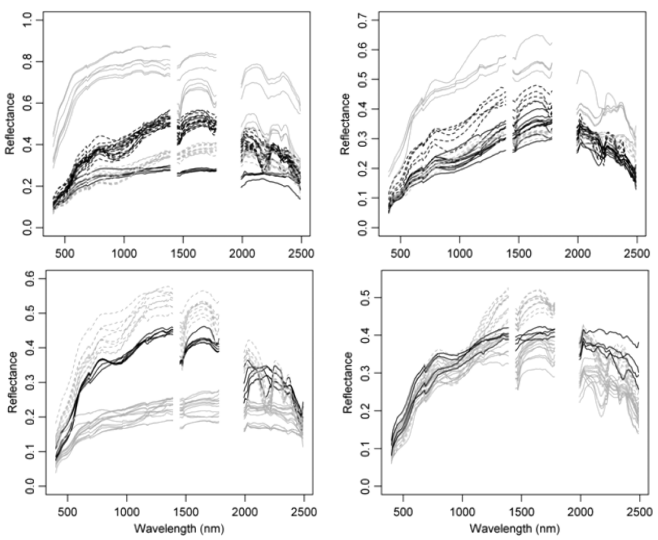

Fig. 12. The $p = 14$ endmember bundles extracted from the AVIRIS Cuprite image by the VCA subsetting approach with 10 subsets of a 10% image size. The bundles were spread over four images not to overload the figure.

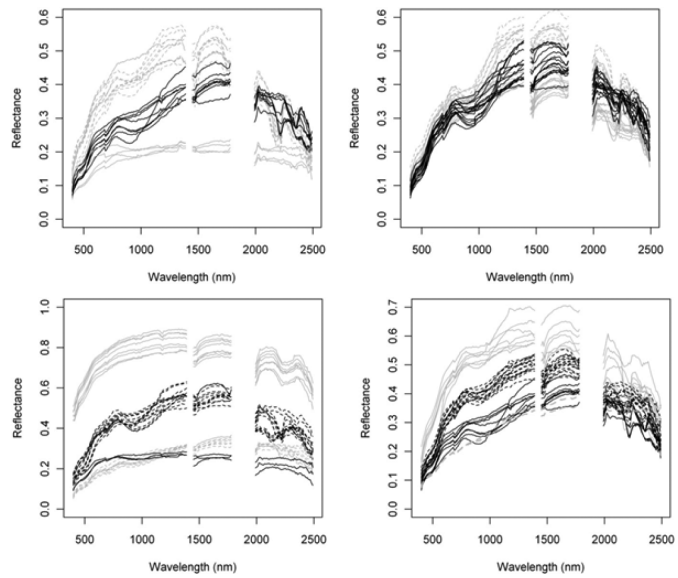

Fig. 13. The $p = 14$ endmember bundles extracted from AVIRIS Cuprite image by the N-FINDR subsetting approach with 10 subsets of a 10% image size. The bundles were spread over four images not to overload the figure.

endmember condition present in the scene. In an operational setting, these libraries are seldom available.

Here we have addressed these issues and developed an integrated solution that intends to bridge the existing gap which currently prevents available EEAs to perform appropriately in the presence of endmember variability. The proposed technique, which simply extracts endmembers from image subsets in an iterative fashion, has shown great promise when combined with MESMA, i.e., the most widely applied endmember variability reduction technique. The extracted endmember bundles showed significant resemblance to the reference endmember bundles, while systematically improving abundance estimates by incorporating endmember variability in the unmixing process. These observations have been verified in this work using both simulated and real hyperspectral datasets.

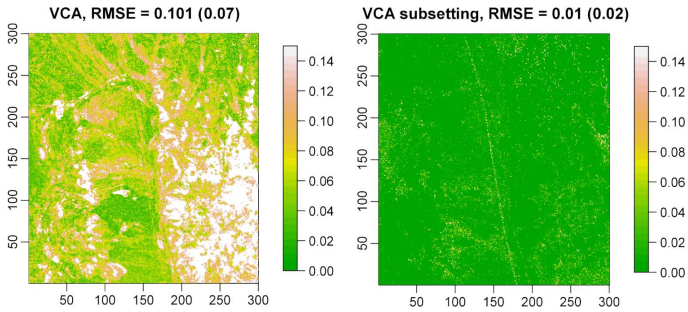

Fig. 14. The RMSE map obtained after reconstructing the AVIRIS Cuprite scene using the endmembers identified by (left) the traditional VCA approach with FCLSU—not incorporating endmember bundles—and (right) the VCA approach incorporating endmember variability through subsetting and MESMA. The mean RMSE as well as the standard deviation (in the parenthesis) are indicated on top of the images.

Despite the encouraging results obtained, we acknowledge the need for further research. In this study endmember clusters were derived from the spectral library of candidate endmembers that were automatically selected from the image data by an EEA. To define the clusters a k-means clustering was performed. As it concerned a proof of concept study, the popular Euclidean distance was used as the similarity measure. Although the current results clearly demonstrate the added value of combining traditional EEAs and clustering to implement endmember variability into spectral unmixing, additional research is required to evaluate whether endmember bundle reconstruction can benefit from implementing other similarity measures in the k-means algorithm (e.g., spectral angle mapper, associated with changes in the spectral shape rather than changes in the amplitude). We further assumed that each subset contained pure pixels of each endmember class. The validity of this assumption not only depends upon the fact that a reasonable number of pure pixels should be available for each endmember in the scene, but also on the size and number of image subsets used in the endmember bundles extraction algorithm. Additional research is needed to identify which and how these parameters influence the equilibrium between image subsetting and the properties of the extracted endmember bundles. A possible refinement of the proposed methodology would be to first estimate the number of endmembers present in each specific subset, and then apply the endmember extraction accordingly. For this purpose, techniques such as virtual dimensionality [57] and HySime [58] can be used. Moreover, if pure pixels are not available in the image subsets, algorithms that do not assume the presence of pure pixels [36]–[41] could be also considered in this framework. Further, no consistency was observed with regard to the most optimal EEA. In other words, the optimal EEA changed depending upon the considered scenario. Also in this area additional research is required. Future research should also focus on improving the computational efficiency of the proposed unmixing chain. This can be done by taking advantage of the parallel nature of the proposed framework, which can be efficiently implemented in latest-generation high performance computing platforms such as graphical processing units. As far as it concerns the endmember variability reduction techniques, MESMA was used as a prototype, yet, alternative techniques to address endmember variability, such as AutoMCU [44], stable zone unmixing [53] or a recently developed Fisher discriminant null space approach [59], should be tested in combination with the endmember bundle extraction approach presented in this study.

ACKNOWLEDGMENT

M. Zortea was a fellow of the Spanish "Juan de la Cierva" programme cofinanced by the European Social Fund. The Carnegie Airborne Observatory is supported by the W. M. Keck Foundation, Gordon and Betty Moore Foundation, and William Hearst III. The scientific input of Roberta Martin, Laurent Tits, David Knapp, and Prof. Pol Coppin is gratefully acknowledged. The authors acknowledge the two anonymous reviewers for their outstanding comments and suggestions, which greatly helped to improve the technical content and presentation of the manuscript.

REFERENCES

[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, 2002.
[2] J. B. Adams, M. O. Smith, and P. E. Johnson, "Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site," *J. Geophys. Res.*, vol. 91, pp. 8098–8112, 1986.
[3] 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, 2004.
[4] Q. Du, N. Raksuntorn, N. Younan, and R. King, "End-member extraction for hyperspectral image analysis," *Appl. Opt.*, vol. 47, no. 28, pp. 77–84, 2008.
[5] 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, pp. 529–545, 2000.
[6] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. New York: Kluwer Academic/Plenum Publishers, 2003.
[7] J. B. Adams, D. Sabol, V. Kapos, R. A. Filho, D. A. Roberts, and A. R. Gillespie, "Classification of multispectral images based on fraction endmembers, application to land cover change in the Brazilian Amazon," *Remote Sens. Environ.*, vol. 52, pp. 137–154, 1995.
[8] J. Settle, "On the effect of variable endmember spectra in the linear mixture model," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, pp. 389–396, 2006.
[9] A. Barducci and A. Mecocci, "Theoretical and experimental assessment of noise effects on least-squares spectral unmixing of hyperspectral images," *Opt. Eng.*, vol. 44, pp. 100–108, 2005.
[10] M. D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, 2011.
[11] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of aviris data," in *Proc. JPL Airborne Earth Sci. Workshop*, 1995, pp. 23–26.
[12] M. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proc. SPIE*, 1999, vol. 3753, pp. 266–270.
[13] R. A. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, "Automatic endmember extraction from hyperspectral data for mineral exploration," in *Proc. 21st Canadian Symp. Remote Sensing*, 1999, pp. 21–24.
[14] J. H. Bowles, P. J. Palmadesso, J. A. Antoniades, M. M. Baumback, and L. J. Rickard, "Use of filter vectors in hyperspectral data analysis," in *Proc. SPIE Infrared Spaceborne Remote Sensing III*, 1995, vol. 2553, pp. 148–157.
[15] A. Ifarraguerri and C.-I Chang, "Multispectral and hyperspectral image analysis with convex cones," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 756–770, 1999.

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

[17] J. C. Harsanyi and C.-I Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, 1994.

[18] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "ICE: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2085–2095, 2004.

[19] A. Plaza and C.-I Chang, "Impact of initialization on design of endmember extraction algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3397–3407, 2006.

[20] C.-I Chang and A. Plaza, "A fast iterative algorithm for implementation of pixel purity index," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 63–67, 2006.

[21] D. M. Rogge, B. Rivard, J. Zhang, and J. Feng, "Iterative spectral unmixing for optimizing per-pixel endmember sets," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3725–3736, 2006.

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

[23] C.-I Chang, C.-C. Wu, W. Liu, and Y.-C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2804–2819, 2006.

[24] A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 446–450, 2007.

[25] A. Zare and P. Gader, "Hyperspectral band selection and endmember detection using sparsity promoting priors," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 256–260, 2008.

[26] M. Zortea and A. Plaza, "A quantitative and comparative analysis of different implementations of n-findr: A fast endmember extraction algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, pp. 787–791, 2009.

[27] X. Tao, B. Wang, and L. Zhang, "Orthogonal bases approach for the decomposition of mixed pixels in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, pp. 219–223, 2009.

[28] A. Zare and P. Gader, "PCE: Piecewise convex endmember detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2620–2632, 2010.

[29] C.-I Chang, C.-C. Wu, C.-S. Lo, and M.-L. Chang, "Real-time simplex growing algorithms for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1834–1850, 2010.

[30] F. Schmidt, A. Schmidt, E. Treandguier, M. Guiheneuf, S. Moussaoui, and N. Dobigeon, "Implementation strategies for hyperspectral unmixing using Bayesian source separation," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4003–4013, 2010.

[31] O. Duran and M. Petrou, "Robust endmember extraction in the presence of anomalies," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 1986–1996, 2011.

[32] B. Zhang, X. Sun, L. Gao, and L. Yang, "Endmember extraction of hyperspectral remote sensing images based on the ant colony optimization (ACO) algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 7, pp. 2635–2646, 2011.

[33] M. Shoshany, F. Kizel, N. Netanyahu, N. Goldshlager, T. Jarmer, and G. Even-Tzur, "An iterative search in end-member fraction space for spectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 706–709, 2011.

[34] C.-I Chang, C.-C. Wu, and H.-M. Chen, "Random pixel purity index," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 324–328, 2010.

[35] X. Chen, J. Chen, X. Jia, B. Somers, J. Wu, and P. Coppin, "A quantitative analysis of virtual endmembers' increased impact on the collinearity effect in spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2945–2956, 2011.

[36] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, pp. 542–552, 1994.

[37] 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, 2007.

[38] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *Proc. IEEE Int. Geoscience and Remote Sensing Symp.*, 2008, vol. 3, pp. 250–253.

[39] J. Bioucas-Dias, "A variable splitting augmented lagragian approach to linear spectral unmixing," in *Proc. 1st IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Grenoble, France, Aug. 26–28, 2009, pp. 1–4.

[40] 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, pp. 4418–4432, 2009.

[41] J. M. Bioucas-Dias and J. Nascimento, "Hyperspectral unmixing based on mixtures of Dirichlet components," *IEEE Trans. Geosci. Remote Sens.*, 2011, in press.

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

[43] D. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. Green, "Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 267–279, 1998.

[44] G. Asner and D. Lobell, "A biogeophysical approach for automated SWIR unmixing of soils and vegetation," *Remote Sens. Environ.*, vol. 74, no. 1, pp. 99–112, 2000.

[45] B. Somers, S. Delalieux, J. Stuckens, W. W. Verstraeten, and P. Coppin, "A weighted linear spectral mixture analysis approach to address endmember variability in agricultural production systems," *Int. J. Remote Sens.*, vol. 30, no. 1, pp. 139–147, 2009.

[46] C. A. Bateson and B. Curtiss, "A method for manual endmember selection and spectral unmixing," *Remote Sens. Environ.*, vol. 55, no. 3, pp. 229–243, 1996.

[47] 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 Sens.*, vol. 38, no. 2, pp. 1083–1094, 2000.

[48] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *J. Royal Statistical Society, Series C (Applied Statistics)*, vol. 28, pp. 100–108, 1979.

[49] B. Somers, S. Delalieux, W. W. Verstraeten, and P. Coppin, "A conceptual framework for the simultaneous extraction of sub-pixel spatial extent and spectral characteristics of crops," *Photogramm. Eng. Remote Sens.*, vol. 75, pp. 57–68, 2009.

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

[51] J. Stuckens, B. Somers, S. Dzikiti, W. W. Verstraeten, G. Albrigo, R. Wennen, and P. Coppin, "Off-nadir viewing and sensor specifications for reducing spectral mixture issues in citrus orchards," *Photogramm. Eng. Remote Sens.*, vol. 76, pp. 1261–1274, 2010.

[52] C. C. D. Lelong, P. C. Pinet, and H. Poilve, "Hyperspectral imaging and stress mapping in agriculture: A case study on wheat in Beauce (France)," *Remote Sens. Environ.*, vol. 66, pp. 179–191, 1998.

[53] B. Somers, S. Delalieux, W. Verstraeten, J. Van Aardt, G. Albrigo, and P. Coppin, "An automated waveband selection technique for optimized hyperspectral mixture analysis," *Int. J. Remote Sens.*, vol. 31, no. 20, pp. 5549–5568, 2010.

[54] G. P. Asner and K. B. Heidebrecht, "Spectral unmixing of vegetation, soil, and dry carbon cover in arid regions: Comparing multispectral and hyperspectral observations," *Int. J. Remote Sens.*, vol. 23, pp. 3939–3958, 2002.

[55] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetracorder and expert systems," *J. Geophys. Res.*, vol. 108, pp. 1–44, 2003.

[56] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, pp. 2679–2693, 2009.

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

[58] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, 2008.

[59] J. Jin, B. Wang, and L. Zhang, "A novel approach based on fisher discriminant null space for decomposition of mixed pixels in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 699–703, 2010.

Ben Somers received the M. Sc. Degree and PhD Degree in bioscience engineering (land and forest management) from the Katholieke Universiteit Leuven (K.U.Leuven), Belgium, in 2005 and 2009 respectively. In 2010 he was research associate at the Geomatics Engineering group of the K.U.Leuven, Belgium. Since 2011 he is researcher at the Flemish Institute for Technological Research (VITO), Belgium. His research interests are the design of processing tools for hyperspectral remote sensing with a specific focus on Spectral Mixture Analysis and its application in precision farming and forest ecology.

Maciel Zortea received the Civil Engineer degree and the M.Sc. degree in remote sensing from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil, in 2002 and 2004, respectively, and the Ph.D. degree in Information and Communication Science and Technologies, curriculum in Space Science and Engineering from the University of Genoa, Genoa, Italy, in 2007.

He was a Marie Curie Experienced Researcher within the Hyperspectral Imaging Network, and a Postdoctoral Researcher at the Department of Technology of Computers and Communications, University of Extremadura, Caceres, Spain, and at the Department of Mathematics and Statistics, University of Tromsø, Norway. His current research interests include image analysis and pattern recognition for remote-sensing and medical imaging applications.

Dr. Zortea has served as a Reviewer for several journals, including the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

Antonio Plaza (M'05–SM'07) received the M.S. and Ph.D. degrees in computer engineering from the University of Extremadura, Caceres, Spain. He was a Visiting Researcher with the Remote Sensing Signal and Image Processing Laboratory, University of Maryland Baltimore County, Baltimore, with the Applied Information Sciences Branch, Goddard Space Flight Center, Greenbelt, MD, and with the AVIRIS Data Facility, Jet Propulsion Laboratory, Pasadena, CA. He is currently an Associate Professor with the Department of Technology of Computers and Communications, University of Extremadura, Caceres, Spain, where he is the Head of the Hyperspectral Computing Laboratory (HyperComp). He was the Coordinator of the Hyperspectral Imaging Network (Hyper-I-Net), a European project designed to build an interdisciplinary research community focused on hyperspectral imaging activities. He has been a Proposal Reviewer with the European Commission, the European Space Agency, and the Spanish Government. He is the author or coauthor of around 300 publications on remotely sensed hyperspectral imaging, including more than 60 Journal Citation Report papers, 20 book chapters, and over 200 conference proceeding papers. His research interests include remotely sensed hyperspectral imaging, pattern recognition, signal and image processing, and efficient implementation of large-scale scientific problems on parallel and distributed computer architectures. Dr. Plaza has coedited a book on high-performance computing in remote sensing and guest edited seven special issues on remotely sensed hyperspectral imaging for different journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (for which he serves as Associate Editor on hyperspectral image analysis and signal processing since 2007), the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (for which he serves as a member of the steering committee since 2011), the International Journal of High Performance Computing Applications, and the Journal of Real-Time Image Processing. He is also serving as an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER. He has served as a reviewer for more than 280 manuscripts submitted to more than 50 different journals, including more than 140 manuscripts reviewed for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He has served as a Chair for the IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing in 2011. He has also been serving as a Chair for the SPIE Conference on Satellite Data Compression, Communications, and Processing since 2009, and for the SPIE Remote Sensing Europe Conference on High Performance Computing in Remote Sensing since 2011. Dr. Plaza is a recipient of the recognition of Best Reviewers of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2009 and a recipient of the recognition of Best Reviewers of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING in 2010. He is currently serving as Director of Education activities and member of the Administrative Committee of the IEEE Geoscience and Remote Sensing Society.

Gregory P. Asner is on the faculty of the Department of Global Ecology, Carnegie Institution for Science and the Environmental Earth System Science Department at Stanford University. His research interests include ecosystem ecology, biogeochemistry and remote sensing. He develops new technologies for science-based conservation assessments of ecosystems including carbon emissions, hydrological function, and biological diversity.