Region-Based Spatial Preprocessing for Endmember Extraction and Spectral Unmixing

Gabriel Martín and Antonio Plaza, *Senior Member, IEEE*

Abstract—**Linear spectral unmixing is an important task in remotely sensed hyperspectral data exploitation. This approach first identifies a collection of spectrally pure constituent spectra, called *endmembers*, and then expresses the measured spectrum of each mixed pixel as a combination of endmembers weighted by fractions or *abundances* that indicate the proportion of each endmember in the pixel. Over the last decade, several algorithms have been developed for automatic extraction of spectral endmembers from hyperspectral data, with many of them relying exclusively on the spectral information. In this letter, we develop a novel unsupervised spatial preprocessing (SPP) module which adopts a region-based approach for the characterization of each endmember class prior to endmember identification using spectral information. The proposed approach can be combined with any spectral-based endmember extraction technique. Our method is validated using both synthetic scenes constructed using fractals and a real hyperspectral data set collected by NASA's Airborne Visible Infrared Imaging Spectrometer over the Cuprite Mining District in Nevada and further compared with previous efforts in the same direction such as the spatial–spectral endmember extraction, automatic morphological endmember extraction, or SPP methods.**

Index Terms—**Endmember extraction, hyperspectral imaging, spatial–spectral analysis, spectral unmixing.**

I. INTRODUCTION

THE availability of hyperspectral instruments with a number of spectral bands that exceed the number of spectral mixture components, such as NASA Jet Propulsion Laboratory's Airborne Visible Infrared Imaging Spectrometer (AVIRIS) [1], has allowed one to cast the unmixing problem in terms of an overdetermined system of equations in which, given a set of pure spectral signatures (called *endmembers*), the actual unmixing to determine *abundance fractions* can be defined in terms of a numerical inversion process [2].

Several algorithms have been developed over the last decade for automatic or semiautomatic extraction of spectral endmembers directly from the input scene. Classic techniques include the orthogonal subspace projection (OSP) [3], N-FINDR [4], or vertex component analysis (VCA) [5], among many others [6]–[8], but only a few techniques have included the spatial information. For instance, extended morphological operations have been used as a baseline to develop an automatic morphological endmember extraction algorithm [9] for spatial–spectral

endmember extraction (SSEE). Another spatial–spectral approach includes spatial averaging of spectrally similar endmember candidates found via singular value decomposition, called SSEE algorithm [10]. Recently, a spatial preprocessing (SPP) algorithm [8] has been proposed, which estimates, for each pixel vector in the scene, a spatially derived factor that is used to weigh the importance of the spectral information associated to each pixel in terms of its spatial context. The SPP is intended as a preprocessing module that can be used in combination with an existing spectral-based endmember extraction algorithm. All the aforementioned spatial–spectral approaches include, at some point, a spatial window (of fixed or variable size) which is translated over the spatial domain of the scene in order to analyze contextual information around each image pixel. This strategy is feasible when spectral variation is smooth and relatively constant over the image, but in order to better model natural scenes with step discontinuities and multiscale features and regions, techniques that are able to model such regions in a spatially adaptive fashion are needed [11].

In this letter, we develop a novel region-based SPP (RBSPP) strategy which uses spatial information as a *guide* to exploit spectral information more effectively by adequately exploiting spatial context in adaptive fashion. Our proposed approach first adaptively searches for the most spectrally pure regions (understood as groups of several contiguous pixel vectors with similar spectral content) by using a hybrid procedure that combines unsupervised clustering and OSP concepts. This process is then followed by a spectral-based endmember extraction process conducted using only the pixels located in such regions.

The remainder of this letter is organized as follows. Section II describes the proposed RBSPP approach. Section III describes our experiments with hyperspectral data simulated using fractals. Section IV describes our experiments with a real hyperspectral data set collected by AVIRIS over the Cuprite Mining District in Nevada. Section V concludes with some remarks and hints at plausible future research.

II. RBSPP

The proposed preprocessing method exploits the intuitive idea that transition areas between two or more different land-cover classes are more likely to contain mixed pixels. Thus, it is reasonable to assume that pure signatures are less likely to be found in transition areas and more likely to be present in well-defined spatially homogeneous regions. With the aforementioned design principle in mind, our preprocessing approach incorporates the following properties [12].

1) The inclusion of spatial information is performed in the form of an adaptive and fully automated preprocessing module which uses a region-based approach to reduce

the sensitivity of the method to different spatial window shapes and sizes [8]–[10].

2) At the preprocessing level, spatial information is used as a *guide* to exploit spectral information more effectively by directing the endmember searching process to image areas which are *both* spatially homogeneous and spectrally pure, as opposed to previous developments which mainly accounted for the spatially homogeneous nature of the regions instead of their spectral purity [8].

3) The integration of spatial and spectral information is performed at the preprocessing level, and hence, it does not require any modifications in the endmember extraction stage conducted afterward. This feature is available in the methodology proposed in [8] but not in other spatial–spectral developments [9], [10].

In the following sections, we describe in more detail the individual stages sequentially applied by the proposed method.

A. Estimation of the Number of Endmembers

Two different methods have been used in this letter to estimate the number of endmembers p in the original hyperspectral image: the HySime method [13] and the virtual dimensionality (VD) concept [14]. The former does not have any input parameters, while the latter is a Neyman–Pearson detector based on a prescribed parameter P_F (i.e., false-alarm probability). In spite of different criticisms to these methods, these are the two most widely used approaches for estimating the number of endmembers available in the literature.

B. Unsupervised Clustering and Segmentation

Clustering aims at grouping pixels in feature space, so that pixels belonging to the same cluster are spectrally similar. In this letter, we have used three different clustering algorithms.

1) The first one is the ISODATA algorithm [15]. It starts with a random initial partition of the set of available pixel vectors in the original hyperspectral image into c candidate clusters. It then iteratively optimizes this initial partition using least squares [15]. A relevant issue for this algorithm is how to set the number of clusters c in advance. Specifically, in order for the algorithm to automatically determine the value of c, we can define a minimum number of clusters $c_{\min}$ and a maximum number of clusters $c_{\max}$ based on the information about the considered image (i.e., how many groups of materials with similar spectra are present).

2) The second one is the k-means algorithm [16]. Its goal is to determine a set of c points, called centers, so as to minimize the mean squared distance from each pixel vector to its nearest center. The algorithm is based on the observation that the optimal placement of a center is at the centroid of the associated cluster. It starts with a random initial placement. At each stage, the algorithm moves every center point to the centroid of the set of pixel vectors for which the center is a nearest neighbor according to the spectral angle (SA) [17] and then updates the neighborhood by recomputing the SA from each pixel vector to its nearest center. These steps are repeated until the algorithm converges to a point that is a minimum for the distortion [16].

3) The third one is the hierarchical segmentation (HSEG) algorithm [18]. It is a segmentation technique based on an iterative hierarchical stepwise optimization region growing method. Furthermore, it allows merging of nonadjacent regions by means of spectral clustering. The algorithm starts by labeling each pixel as a separate region. It then calculates a dissimilarity criterion (in our case, the SA) between all pairs of spatially adjacent regions. Then, it finds the smallest distance value $dist\_val$ and sets $thres\_val$ equal to it, merging all pairs of spatially adjacent regions with $dist\_val = thres\_val$. If a parameter $spectral\_wght > 0.0$, the algorithm merges all pairs of spatially nonadjacent regions with $dist\_val \leq spectral\_wght \times thres\_val$ and iterates until c clusters are obtained [18]. In order to reduce computational demands, a recursive divide-and-conquer approximation of HSEG (RHSEG) has been used [19].

In all cases, the output is a set of spectral clusters, each made up of one or more spatially connected regions. This set can also be seen as a set of r connected regions $\{\mathbf{R}_l\}_{l=1}^{r}$.

C. Region Selection Using Orthogonal Projections

The main goal of this stage is to select, out of the set of r regions $\{\mathbf{R}_l\}_{l=1}^{r}$ resulting from the previous stage, a subset of $\{\mathbf{R}_j\}_{j=1}^{c}$ regions which are spectrally distinct, with $r \geq c$. For this purpose, we first select a representative spectral signature for each connected region (in this letter, the mean spectrum $\mathbf{M}_l$, with $l = 1, 2, \ldots, r$). Then, we apply the OSP algorithm to find a set of c spatially representative regions with associated spectra which are both spectrally pure and orthogonal between them. The OSP starts by identifying the first region as the one with the associated mean spectrum with higher intensity as follows:

$$\mathbf{R}_1 = \arg\left\{\max_l \sum_{l=1}^{r} \mathbf{M}_l \mathbf{M}_l^{\mathrm{T}}\right\}. \tag{1}$$

Once an initial region has been identified, the preprocessing algorithm assigns $\mathbf{U}_1 = [\mathbf{M}_1]$ and applies an orthogonal subspace projector [3] to the mean spectra of all r regions, thus calculating the second region as the one with the associated mean spectrum with maximum projection value as follows:

$$\mathbf{R}_2 = \arg\left\{\max_l \left[\left(P_{\mathbf{U}_1}^{\perp} \mathbf{M}_l\right)^{\mathrm{T}} \left(P_{\mathbf{U}_1}^{\perp} \mathbf{M}_l\right)\right]\right\},$$
$$\text{with} \quad P_{\mathbf{U}_1}^{\perp} = \mathbf{I} - \mathbf{U}_1 \left(\mathbf{U}_1^{\mathrm{T}} \mathbf{U}_1\right)^{-1} \mathbf{U}_1^{\mathrm{T}} \tag{2}$$

where $\mathbf{I}$ is the identity matrix and $l = 1, 2, \ldots, r$. The algorithm then assigns $\mathbf{U}_2 = [\mathbf{M}_1 \mathbf{M}_2]$ and repeats the procedure iteratively so that a new connected region is obtained at each iteration as follows:

$$\mathbf{R}_j = \arg\left\{\max_l \left[\left(P_{\mathbf{U}_{j-1}}^{\perp} \mathbf{M}_l\right)^{\mathrm{T}} \left(P_{\mathbf{U}_{j-1}}^{\perp} \mathbf{M}_l\right)\right]\right\},$$
$$\text{with} \ P_{\mathbf{U}_{j-1}}^{\perp} = \mathbf{I} - \mathbf{U}_{j-1} \left(\mathbf{U}_{j-1}^{\mathrm{T}} \mathbf{U}_{j-1}\right)^{-1} \mathbf{U}_{j-1}^{\mathrm{T}}. \tag{3}$$

The procedure is terminated when $j = c$, i.e., once a predefined number of c spatially connected regions $\{\mathbf{R}_j\}_{j=1}^{c}$ have been identified from the initial set of l regions $\{\mathbf{R}_l\}_{l=1}^{r}$.
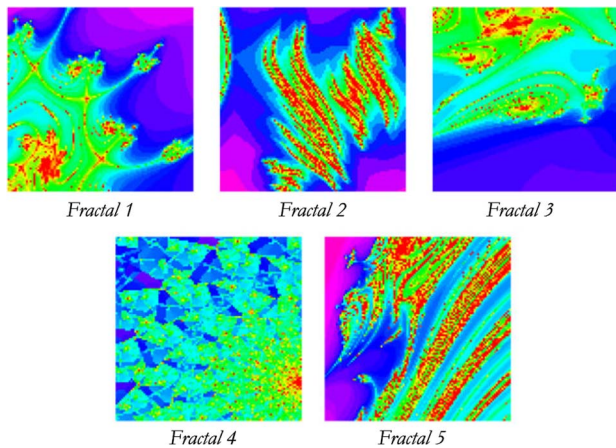
Fig. 1. Synthetic images used in experiments, where spatial patterns were generated using fractals.

D. Automatic Endmember Extraction

Once the SPP has been completed, a spectral-based endmember extraction algorithm is applied to the set of pixels associated to the c spatially connected regions $\{\mathbf{R}_j\}_{j=1}^{c}$ retained after the preprocessing in order to select p endmembers. In this letter, three well-known algorithms have been considered to perform such automatic endmember extraction: OSP [3], N-FINDR [4], and VCA [5]. The reasons for our selection are as follows: 1) These algorithms are representative of a class of convex geometry-based techniques which have been quite successful in endmember extraction literature [6]–[8]; 2) they are fully automated; and 3) they do not require any input parameters other than the desired number of endmembers to be extracted p.

III. EXPERIMENTS WITH SYNTHETIC DATA

A. Construction of Synthetic Scenes

A database of five 100×100-pixel synthetic hyperspectral scenes has been created using fractals to generate distinct spatial patterns often found in nature. In this letter, we used fractals to simulate linear mixtures of a set of endmember signatures randomly selected from a spectral library compiled by the U.S. Geological Survey (USGS)[1] and made up of a total of 420 signatures. Fig. 1 shows the five fractal images used in the simulations. These images are further divided into a number of clusters using the k-means algorithm [16], where the number of clusters extracted from the five fractal images was always larger than the number of endmember signatures, fixed in our experiments to $p = 9$, and the abundance proportions in the regions associated to each cluster have been set so that the pixels that are closer to the border of the region are more heavily mixed while the pixels located at the center of the region are more spectrally pure in nature. For this purpose, a Gaussian filter is applied, where the width of the Gaussian is carefully adjusted according to the width of each region. Zero-mean Gaussian noise was added to the synthetic scenes in different signal-to-noise ratios (SNRs)—from 30:1 to 110:1—to simulate contributions from ambient and instrumental sources, following the procedure described in [3]. Fig. 2 shows the spectra of the USGS signatures used in the simulation of one of
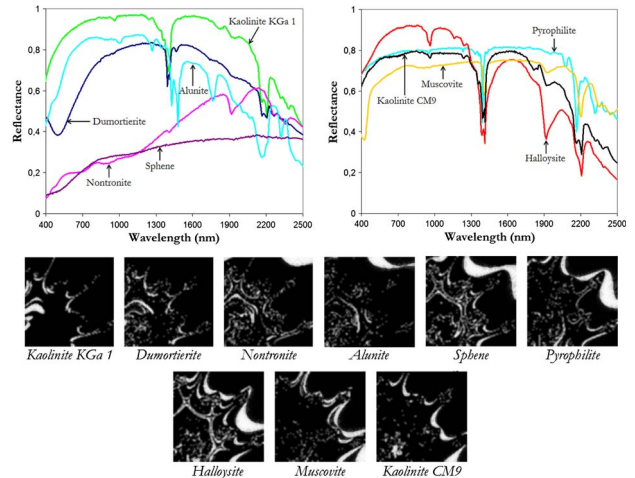

Fig. 2. (Top) USGS library signatures and (bottom) fractional abundance distributions considered for generating the simulated hyperspectral scene labeled as "Fractal 1" in experiments.

the synthetic scenes[2] (labeled as "Fractal 1" in Fig. 1). The fully constrained [2] abundance maps associated to each reference USGS signature in the construction of the aforementioned synthetic scene are also shown in Fig. 2.

B. Performance Analysis

Before describing our experiments, it is first important to address the parameter values used for the different clustering methods considered. The parameter $c_{\min}$ of ISODATA was set empirically to p, which is the number of endmembers estimated by a consensus between VD and HySime, while the parameter $c_{\max}$ was set to $2p$, which is a value that generates an explicit oversegmentation in order to let the algorithm automatically decide about the number of classes. In the case of k-means and RHSEG, we explicitly set $c = p$. For RHSEG, we also set $spectral\_wght = 0.2$, which results in a good tradeoff between spatial and spectral information [18].

Two different metrics have been used to compare the performances of endmember extraction algorithms in the synthetic fractal scenes. The first metric is the SA [17] between each extracted endmember and the set of available USGS ground-truth spectral signatures. The second metric used in this work is the root-mean-square error (rmse) between the original and a reconstructed version of the hyperspectral scene [8]. Table I shows the average SA scores (in degrees) between the reference USGS mineral spectra and their corresponding endmember pixels produced by several endmember extraction algorithms, across the five synthetic scenes in Fig. 1. As a result, each value reported in Table I corresponds to the average SA obtained after processing the five considered scenes with the same SNR (five different SNR values, ranging from 30:1 to 110:1, are reported in the table). In our experiments, three different clustering methods (ISODATA, k-means, and RHSEG) are used to implement the clustering stage in RBSPP. As shown in Table I, the combination of the proposed RBSPP method with the spectral-based algorithms does not always provide the best results in terms of SA. The SA is probably not the best tool for measuring spectral similarities when the differences are

[1]http://speclab.cr.usgs.gov/spectral.lib06

[2]http://www.umbc.edu/rssipl/people/aplaza/fractals.zip

TABLE I
AVERAGE SPECTRAL SIMILARITY SCORES (IN DEGREES) BETWEEN THE
USGS MINERAL SPECTRA AND THEIR CORRESPONDING ENDMEMBER
PIXELS PRODUCED BY SEVERAL ENDMEMBER EXTRACTION
ALGORITHMS ACROSS THE FIVE SYNTHETIC SCENES IN FIG. 1

Algorithm	SNR=30:1	SNR=50:1	SNR=70:1	SNR=90:1	SNR=110:1
N-FINDR	2.089	0.463	0.383	0.388	0.361
OSP	2.118	0.452	0.349	0.361	0.345
VCA	2.187	0.520	0.367	0.433	0.435
SPP+N-FINDR	2.293	0.778	0.701	0.694	0.693
SPP+OSP	2.342	0.622	0.536	0.529	0.529
SPP+VCA	2.271	0.455	0.327	0.319	0.347
RBSPP*+N-FINDR	2.174	0.874	0.676	0.786	0.804
RBSPP*+OSP	2.226	0.764	1.020	1.088	1.034
RBSPP*+VCA	1.033	0.685	0.678	0.809	0.714
RBSPP†+N-FINDR	2.220	0.648	0.708	0.618	0.690
RBSPP†+OSP	2.275	0.588	0.724	0.605	0.567
RBSPP†+VCA	1.307	0.973	0.890	0.761	0.512
RBSPP§+N-FINDR	2.272	0.982	0.845	0.677	0.895
RBSPP§+OSP	3.332	0.910	0.860	0.809	0.941
RBSPP§+VCA	1.301	0.818	0.936	0.686	1.083
AMEE	2.670	1.260	0.969	1.193	1.252
SSEE	2.124	1.077	0.576	0.722	0.645

*Implemented with ISODATA. † Implemented with k-means. § Implemented with RHSEG.

TABLE II
AVERAGE ERROR SCORES AFTER RECONSTRUCTING THE FIVE
SYNTHETIC SCENES IN FIG. 1 USING THE ENDMEMBERS
EXTRACTED BY SEVERAL METHODS

Algorithm	SNR=30:1	SNR=50:1	SNR=70:1	SNR=90:1	SNR=110:1
N-FINDR	0.356	0.039	0.009	0.007	0.008
OSP	0.359	0.039	0.010	0.008	0.008
VCA	0.369	0.044	0.017	0.021	0.018
SPP+N-FINDR	0.359	0.045	0.017	0.016	0.016
SPP+OSP	0.368	0.048	0.016	0.015	0.015
SPP+VCA	0.368	0.040	0.011	0.009	0.009
RBSPP*+N-FINDR	0.358	0.049	0.028	0.026	0.028
RBSPP*+OSP	0.359	0.055	0.072	0.054	0.054
RBSPP*+VCA	0.308	0.048	0.033	0.043	0.028
RBSPP†+N-FINDR	0.360	0.042	0.018	0.018	0.022
RBSPP†+OSP	0.361	0.042	0.019	0.023	0.017
RBSPP†+VCA	0.320	0.070	0.034	0.070	0.029
RBSPP§+N-FINDR	0.372	0.078	0.023	0.023	0.024
RBSPP§+OSP	0.388	0.087	0.025	0.029	0.027
RBSPP§+VCA	0.340	0.066	0.040	0.035	0.046
AMEE	0.627	0.484	0.468	0.473	0.484
SSEE	0.358	0.135	0.035	0.066	0.026

*Implemented with ISODATA. † Implemented with k-means. § Implemented with RHSEG.

TABLE III
SPECTRAL SIMILARITY SCORES (IN DEGREES) BETWEEN USGS
MINERAL SPECTRA AND THEIR CORRESPONDING ENDMEMBERS
EXTRACTED BY SEVERAL ALGORITHMS FROM
THE AVIRIS CUPRITE SCENE

Algorithm	Alunite GDS84	Budding. GDS85	Calcite WS272	Kaolinite KGa-1	Muscovite GDS107	Mean
N-FINDR	4.81	4.29	7.60	9.92	5.05	6.33
OSP	4.81	4.16	9.52	10.76	5.29	6.91
VCA	6.91	5.38	9.53	9.65	6.47	7.59
SPP+N-FINDR	7.72	4.27	9.34	11.26	5.69	7.66
SPP+OSP	6.06	4.27	8.43	12.28	4.64	7.14
SPP+VCA	14.11	8.49	11.94	13.86	5.61	10.80
RBSPP+N-FINDR	7.22	5.71	5.59	10.43	5.08	6.81
RBSPP+OSP	6.86	4.16	10.05	11.14	5.70	7.58
RBSPP+VCA	6.86	4.34	8.31	10.19	5.42	7.02
AMEE	4.81	4.17	5.87	8.74	4.61	5.64
SSEE	4.81	4.16	8.48	10.73	4.63	6.57

to implement RBSPP provided very similar results in terms of performance and computational complexity, in the following, we use only the ISODATA algorithm for simplicity.

IV. EXPERIMENTS WITH REAL DATA

In this experiment, we use the well-known AVIRIS Cuprite data set, available online in reflectance units[3] after atmospheric correction. This scene has been widely used to validate the performance of endmember extraction algorithms. The portion used in experiments corresponds to a 350×350-pixel subset of the sector labeled as f970619t01p02_r02_sc03.a.rfl in the online data. The scene comprises 224 spectral bands between 0.4 and 2.5 μm, with a full width at half maximum of 10 nm and a spatial resolution of 20 m per pixel. Prior to the analysis, several bands were removed due to water absorption and low SNR in those bands, leaving a total of 192 reflectance channels to be used in the experiments. The Cuprite site is well understood mineralogically [20]. A few selected spectra from the USGS library, corresponding to highly representative minerals in the Cuprite Mining District, are used in this work to substantiate endmember purity.

Table III tabulates the SA scores (in degrees) obtained after comparing the USGS library spectra of *alunite*, *buddingtonite*, *calcite*, *kaolinite*, and *muscovite*, with the corresponding endmembers extracted by different algorithms from the AVIRIS Cuprite scene. In all cases, the input parameters of the different endmember extraction methods tested have been carefully optimized so that the best performance for each method is reported. It should be noted that Table III only displays the smallest SA scores of all endmembers with respect to each USGS signature for each algorithm. For reference, the mean SA values across all five USGS signatures are also reported. The number of endmembers to be extracted was set to $p = 19$ in all experiments after the consensus reached between HySime [13] and the VD concept [14], implemented using $P_F = 10^{-3}$ as the input false-alarm probability. As shown in Table III, the use of RBSPP often improved the results achieved by other preprocessing methods such as SPP.

Fig. 3 shows the error maps obtained after reconstructing the AVIRIS Cuprite scene using $p = 19$ endmembers extracted by different methods. Fig. 3 shows that the application of RBSPP as preprocessing generally results in similar error maps across different endmember extraction algorithms, while the

just in a few—but diagnostic—wavelengths (e.g., the spectra of Muscovite and Pyrophilite in Fig. 2). However, the results obtained by RBSPP are always comparable to those provided by other methods (regardless of the clustering algorithm used). The best overall results in Table I are found by RBSPP + VCA. This is because VCA is the only considered endmember extraction algorithm that performs noise characterization in the data before conducting the endmember search. Since RBSPP discards those pixels which are contained in mixed regions, the noise characterization process conducted by VCA is more efficient for endmember extraction purposes because it is performed only with those pixels located in spectrally pure regions.

On the other hand, Table II shows the rmse scores (measured across the five synthetic scenes in Fig. 1) after reconstructing the scenes using the endmembers extracted by several methods. Again, each value reported in the table corresponds to the rmse obtained after reconstructing five different scenes but simulated with the same SNR. Table II indicates that spectral-based methods generally provide the best overall results (lower rmse values), while RBSPP can produce comparable results when the SNR decreases. Since the three clustering methods used
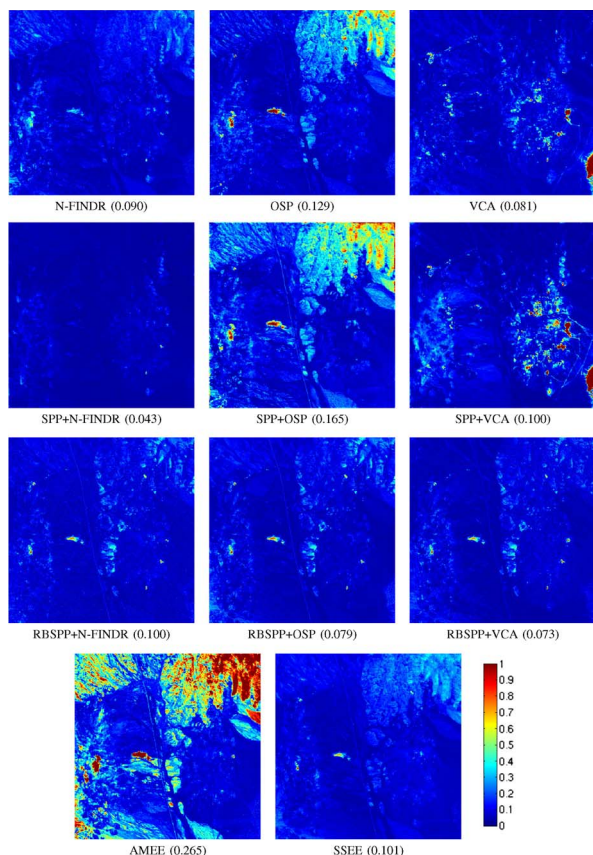
Fig. 3. Errors measured for various endmember extraction algorithms after reconstructing the AVIRIS Cuprite scene.

TABLE IV
PROCESSING TIMES (IN SECONDS) MEASURED IN A DESKTOP PC WITH INTEL CORE i7 920 CPU AT 2.67 GHz WITH 4 GB OF RAM

Algorithm	Spatial preprocessing	Endmember extraction	Total processing time
N-FINDR	-	466.08	466.08
OSP	-	136.09	136.09
VCA	-	31.12	31.12
SPP+N-FINDR	50.06	466.08	516.14
SPP+OSP	50.06	136.09	186.15
SPP+VCA	50.06	31.12	81.18
RBSPP+N-FINDR	96.51	17.72	114.23
RBSPP+OSP	96.51	5.71	102.22
RBSPP+VCA	96.51	5.89	102.40
AMEE	-	76.06	76.06
SSEE	-	1051.23	1051.23

use of SPP results in more variations in terms of reconstruction accuracy, depending on the endmember selection method. For example, the results from OSP and SPP + OSP show regions with large errors in the upper right corner of the image, which are related to an incorrect or missing endmember. Using RBSPP + OSP, these errors can be avoided.

To conclude this section, we note that RBSPP is more computationally expensive than SPP. However, since the RBSPP acts as a pixel selection module that discards those areas of the scene which do not exhibit spatial correlation, it has the potential to significantly reduce the time needed by spectral-based algorithms to conduct the endmember searching process (see Table IV).

V. CONCLUSION AND FUTURE RESEARCH

In this letter, we have presented a new RBSPP technique which can be combined with classic endmember extraction and spectral unmixing algorithms. Our experimental results have been discussed (using both simulated and real hyperspectral data) in terms of three criteria: 1) the spectral similarity to pure spectra; 2) the capacity to produce an accurate reconstruction of the original hyperspectral image; and 3) the computational complexity of the considered algorithms. After analyzing these criteria, we conclude that the proposed region-based preprocessing can assist in the selection of spectral endmembers which are more relevant in spatial sense without increasing the computational complexity. In the future, we will work toward its parallel implementation.

REFERENCES

[1] R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, M. R. Olah, and O. Williams, "Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, Sep. 1998.

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

[3] 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, Jul. 1994.

[4] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266–277, 1999.

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

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

[7] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "End-member extraction for hyperspectral image analysis," *Appl. Opt.*, vol. 47, no. 28, pp. 77–84, 2008.

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

[9] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 2025–2041, Sep. 2002.

[10] D. M. Rogge, B. Rivard, J. Zhang, A. Sanchez, J. Harris, and J. Feng, "Integration of spatial–spectral information for improved extraction of endmembers," *Remote Sens. Environ.*, vol. 110, no. 3, pp. 287–303, Oct. 2007.

[11] D. R. Thompson, R. Castano, and M. S. Gilmore, "Sparse superpixel unmixing for exploratory analysis of CRISM hyperspectral images," in *Proc. IEEE WHISPERS*, 2009, pp. 1–4.

[12] A. Plaza, G. Martin, and M. Zortea, "On the incorporation of spatial information to endmember extraction: Survey and algorithm comparison," in *Proc. IEEE WHISPERS*, 2009, pp. 1–4.

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

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

[15] G. Ball and D. Hall, "ISODATA, a novel method of data analysis and classification," Stanford Univ., Stanford, CA, Tech. Rep. AD-699616, 1965.

[16] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *J. Roy. Stat. Soc., Ser. C (Appl. Stat.)*, vol. 28, no. 1, pp. 100–108, 1979.

[17] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

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

[19] J. C. Tilton, HSEG/RHSEG, HSEGViewer and HSEGReader User's Manual (Version 1.40)2008. [Online]. Available: http://opensource.gsfc.nasa.gov/projects/HSEG/index.php

[20] R. N. Clark, G. A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S. J. Sutley, "USGS Digital Spectral Library splib06a," Digital Data Series 2312007. [Online]. Available: http://speclab.cr.usgs.gov/spectral.lib06