A Fast Sequential Endmember Extraction Algorithm Based on Unconstrained Linear Spectral Unmixing

Javier Plaza[a], Antonio Plaza[a], Gabriel Martín[a]

[a]Department of Technology of Computers and Communications
University of Extremadura, Avda. de la Universidad s/n
E-10071 Cáceres, Spain

ABSTRACT

Spectral unmixing is an important tool for interpreting remotely sensed hyperspectral scenes with sub-pixel precision. It relies on the identification of a set of spectrally pure components (called endmembers) and the estimation of the fractional abundance of each endmember in each pixel of the scene. Fractional abundance estimation is generally subject to two constraints: non-negativity of estimated fractions and sum-to-one for all abundance fractions of endmembers in each single pixel. Over the last decade, several algorithms have been proposed for simultaneous and sequential extraction of image endmembers from hyperspectral scenes. In this paper, we develop a new sequential algorithm that automatically extracts endmembers by using an unconstrained linear mixture model. Our assumption is that fractional abundance estimation using a set of properly selected image endmembers should naturally incorporate the constraints mentioned above, while imposing the constraints for an inadequate set of spectral endmembers may introduce errors in the model. Our proposed approach first applies an unconstrained linear mixture model and then uses a new metric for measuring the deviation of the unconstrained model with regards to the ideal, fully constrained model. This metric is used to derive a set of spectral endmembers which are then used to unmix the original scene. The proposed algorithm is experimentally compared to other algorithms using both synthetic and real hyperspectral scenes collected by NASA/JPL's Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS).

Keywords: Hyperspectral data, spectral mixture analysis, endmember extraction, linear spectral unmixing, fully constrained abundance estimation, unconstrained abundance estimation.

1. INTRODUCTION

Hyperspectral imaging instruments are capable of collecting hundreds of images, corresponding to different wavelength channels, for the same area on the surface of the Earth.[1] For instance, NASA is continuously gathering imagery data with instruments such as the Jet Propulsion Laboratory's Airborne Visible-Infrared Imaging Spectrometer (AVIRIS),[2] able to record the visible and near-infrared spectrum (wavelength region from 0.4 to 2.5 micrometers) of the reflected light of an area 2 to 12 kilometers wide and several kilometers long using 224 spectral bands. The resulting hyperspectral data cube[3] is a stack of images (see Fig. 1) in which each pixel (vector) is represented by a spectral signature or 'fingerprint' that uniquely characterizes the underlying objects, and the resulting multidimensional data volume typically comprises several GBs per flight.

Spectral mixture analysis (SMA) is a popular tool for characterizing mixed pixels in remotely sensed hyperspectral data sets.[4] In this model, it is assumed that the collected spectra are linearly mixed. As a result, a linear (macroscopic) mixture is obtained when the endmember substances are sitting side-by-side within the field of view of the imaging instrument as displayed in Fig. 2(a). The standard processing chain for linear SMA comprises two stages: 1) extraction of pure spectral signatures (endmembers), and 2) estimation of the fractional abundance of each endmember in each pixel of the scene. The use of spectral endmembers allows one to deal with the problem of mixed pixels. For instance, it is likely that the pixel labeled as 'vegetation' in Fig. 1 actually comprises a mixture of vegetation and soil. In this case, the measured spectrum can be decomposed into a linear

Send correspondence to Javier Plaza:
E-mail: jplaza@unex.es; Telephone: +34 927 257000 (Ext. 82578)

Image and Signal Processing for Remote Sensing XV, edited by Lorenzo Bruzzone, Claudia Notarnicola, Francesco Posa,
Proc. of SPIE Vol. 7477, 74770L · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.830734

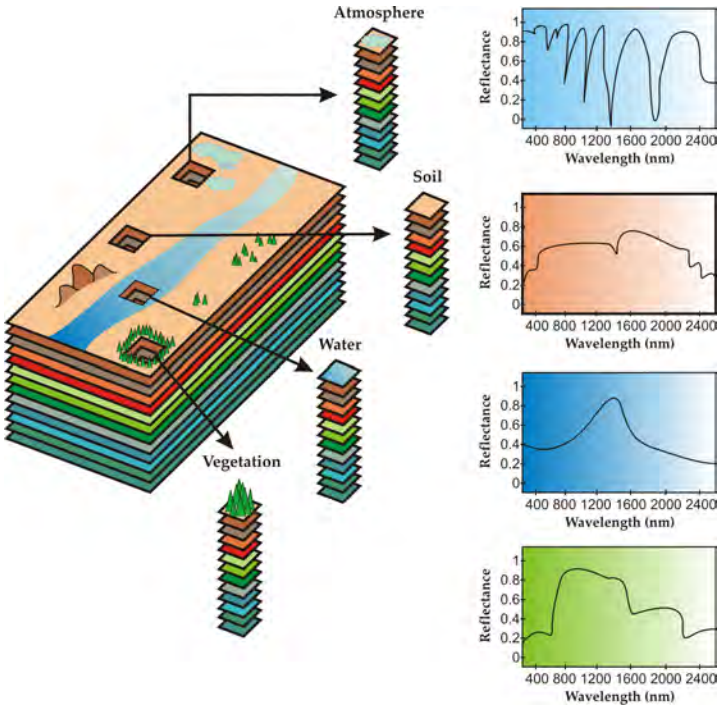


Figure 1. Concept of hyperspectral imaging.

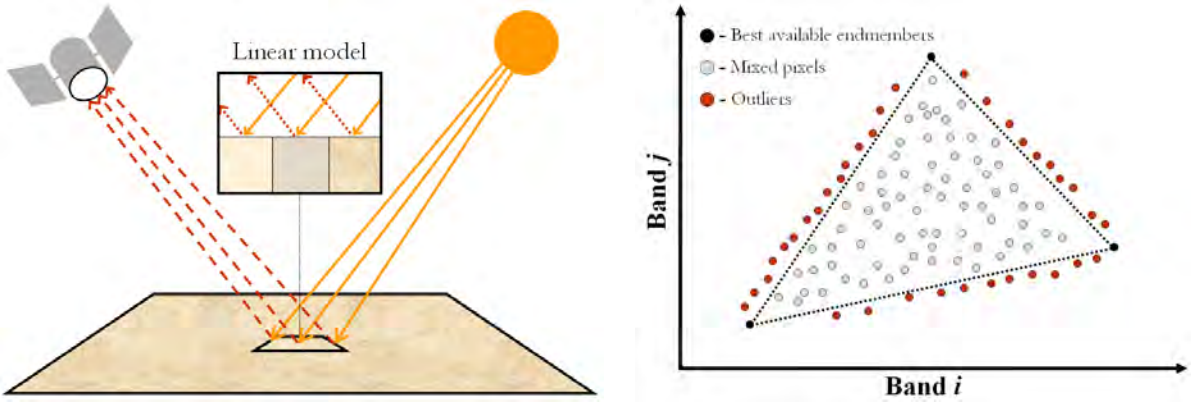
combination of pure spectral endmembers of soil and vegetation, weighted by abundance fractions that indicate the proportion of each endmember in the mixed pixel.

Let us assume that a remotely sensed hyperspectral scene with n bands is denoted by $\mathbf{F}$, in which the pixel at the discrete spatial coordinates (i, j) of the scene is represented by a vector $\mathbf{X}(i, j) = [x_1(i, j), x_2(i, j), \cdots, x_n(i, j)] \in \Re^n$, where $\Re$ denotes the set of real numbers in which the pixel's spectral response $x_k(i, j)$ at sensor channels $k = 1, \ldots, n$ is included. Under the linear mixture model assumption, each pixel vector in the original scene can be modeled using the following expression:

$$\mathbf{X}(i, j) = \sum_{z=1}^{p} \Phi_z(i, j) \cdot \mathbf{E}_z + \mathbf{n}(i, j), \tag{1}$$

where $\mathbf{E}_z$ denotes the spectral response of endmember z, $\Phi_z(i, j)$ is a scalar value designating the fractional abundance of the endmember z at the pixel $\mathbf{X}(i, j)$, p is the total number of endmembers, and $\mathbf{n}(i, j)$ is a noise vector. The solution of the linear spectral mixture problem described in (1) relies on a successful estimation of how many endmembers, p, are present in the input hyperspectral scene $\mathbf{F}$, and also on the correct determination of a set $\{\mathbf{E}_z\}_{z=1}^{p}$ of endmembers and their correspondent abundance fractions $\{\Phi_z(i, j)\}_{z=1}^{p}$ at each pixel $\mathbf{X}(i, j)$. Two physical constrains are generally imposed into the model described in equation (1), these are the abundance non-negativity constraint (ANC), i.e., $\Phi_z(i, j) \geq 0$, and the abundance sum-to-one constraint (ASC), i.e., $\sum_{z=1}^{p} \Phi_z(i, j) = 1$.[3,5] However, if the endmembers are not properly selected or there are not any available pure signatures in the scene (resulting from the available spatial resolution of the sensor or other issues[6]), it may be difficult to form a simplex using the spectrally purest pixels available in the data so that the simplex encloses all other (mixed) pixels in the scene, as it would be naturally desirable in a linear mixture modeling scenario.[7] In this situation, illustrated in Fig. 2(b), the outliers correspond to pixels which cannot be explained using the ASC or ANC constraints, since their endmember abundances may be negative or summing less or more than one. In this case, imposing the ASC and ANC constraints in fractional abundance estimation may be harmful for linear mixture modeling of the data.

Over the last decade, several algorithms have been developed for extracting spectral endmembers directly from an input hyperspectral data set.[8,9] Available algorithms may be generally categorized into two classes



(a) (b)

Figure 2. (a) Linear mixture model. (b) Graphical interpretation of the concept of spectral endmember.

of algorithms,[10] namely, simultaneous endmember extraction algorithms (SMEEAs) and sequential endmember extraction algorithms (SQEEAs):

- On the one hand, SMEEAs generally assume that the number of endmembers (p) is known in advance and produce a set of endmembers simultaneously. So, for a different value of p, an SMEEA generally produces a different set of endmembers. In other words, for any given number of endmembers, p, an EEA must recalculate all the endmembers and cannot take advantage of a set of $p-1$ endmembers previously generated by the same algorithm. In addition, these $p-1$ endmembers do not necessarily constitute a subset of the set of p endmembers generated subsequently, which may have implications in computational performance when various values of p need to be tested. Among a wide set of available SMEEAs, we have selected the N-FINDR method[11] as a representative of this class of algorithms, among several other choices.[12–15]
- In contrast, SQEEAs produce a set of endmembers in sequential order. In other words, a set of p endmembers generated by an SQEEA always includes the set of previously generated $p-1$ endmembers. This feature is highly desirable for an endmember extraction algorithm because it can save a great deal of computational time when the value of p has to be adjusted, as it is often the case in real applications. Among a wide set of available SMEEAs, we have selected a method inspired by the concept of orthogonal subspace projection (OSP)[16] as a representative algorithm[17] of this class of methods, among several other choices.[18–21]

In this paper, we develop a new SQEEA that automatically extracts endmembers from hyperspectral scenes by using an unconstrained linear mixture model. Our assumption is that the ANC and ASC constraints should be naturally satisfied if the set of endmembers has been properly extracted, thus providing an adequate physical description of the data. However, this is not always the case and, in many cases, the constraints are imposed regardless of whether the extracted endmembers are spectrally pure or not. As a result, these constraints are not satisfied *per se*. In order to address this issue, our proposed approach first applies an unconstrained linear mixture model and then uses a metric for measuring the deviation of the unconstrained model with regards to the fully constrained one. To do so, we extract as endmembers those extreme pixel vectors that minimize the number of negative abundance estimations (i.e., those falling outside the estimated simplex), thus searching for the optimal combination of available image pixels which results in the smallest number of abundance estimates that do not satisfy the ASC and ANC constraints. The proposed algorithm is experimentally compared to other representative SQEEAs (OSP algorithm) and SMEEAs (N-FINDR algorithm) in order to substantiate its performance, using both synthetic and real hyperspectral scenes collected by NASA/JPL's Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS).

The remainder of the paper is organized as follows. Section 2 describes the OSP and N-FINDR algorithms used as representative SMEEAs and SQEEAs, respectively. Section 3 introduces the proposed method. Section

4 describes the synthetic and real hyperspectral data sets used for experiments in this work. Section 5 conducts a quantitative and comparative assessment of the proposed method with regards to available approaches. Finally, section 6 concludes with some remarks and hints at plausible future research lines.

2. CLASSIC ENDMEMBER EXTRACTION ALGORITHMS

2.1 N-FINDR

The original N-FINDR algorithm developed by Winter[11] is a representative example of a SMEEA, and can be summarized by the following steps:

1. *Feature reduction.* Apply a dimensionality reduction transformation such as the minimum noise fraction (MNF) or the principal component analysis (PCA)[22, 23] to reduce the dimensionality of the data from n to $p-1$, where p is an input parameter to the algorithm (number of endmembers to be extracted).

2. *Initialization.* Let $\{\mathbf{E}_1^{(0)}, \mathbf{E}_2^{(0)}, \cdots, \mathbf{E}_p^{(0)}\}$ be a set of endmembers randomly extracted from the input data.

3. *Volume calculation.* At iteration $k \geq 0$, calculate the volume defined by the current set of endmembers as:

$$V(\mathbf{E}_1^{(k)}, \mathbf{E}_2^{(k)}, \cdots, \mathbf{E}_p^{(k)}) = \frac{\left| det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mathbf{E}_1^{(k)} & \mathbf{E}_2^{(k)} & \cdots & \mathbf{E}_p^{(k)} \end{bmatrix} \right|}{(p-1)!} \tag{2}$$

4. *Replacement.* For each pixel vector $\mathbf{X}(i,j)$ in the input hyperspectral data set, recalculate the volume by testing the pixel in all p endmember positions, i.e., first calculate $V(\mathbf{X}(i,j), \mathbf{E}_2^{(k)}, \cdots, \mathbf{E}_p^{(k)})$, then $V(\mathbf{E}_1^{(k)}, \mathbf{X}(i,j), \cdots, \mathbf{E}_p^{(k)})$, and so on until $V(\mathbf{E}_1^{(k)}, \mathbf{E}_2^{(k)}, \cdots, \mathbf{X}(i,j))$. If none of the p recalculated volumes is greater than $V(\mathbf{E}_1^{(k)}, \mathbf{E}_2^{(k)}, \cdots, \mathbf{E}_p^{(k)})$, then no endmember is replaced. Otherwise, the combination with maximum volume is retained. Let us assume that the endmember absent in the combination resulting in the maximum volume is denoted by $\mathbf{E}_j^{(k+1)}$. In this case, a new set of endmembers is produced by letting $\mathbf{E}_j^{(k+1)} = \mathbf{X}(i,j)$ and $\mathbf{E}_i^{(k+1)} = \mathbf{E}_i^{(k)}$ for $i \neq j$. The replacement step is repeated for all the pixel vectors in the input data until all the pixels have been exhausted.

2.2 OSP

The orthogonal subspace projection (OSP) approach[16] has been combined with an automatic target generation process[17] to automatically derive spectrally distinct signatures in typical SQEEA fashion. Let $\mathbf{E}_1$ be an initial signature (i.e., the pixel vector with maximum length). The OSP relies on an orthogonal subspace projector operator specified by the following expression:

$$P_{\mathbf{U}}^{\perp} = \mathbf{I} - \mathbf{U}(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T, \tag{3}$$

which is applied to all pixel vectors $\mathbf{X}(i,j)$ in the input hyperspectral data set, with $\mathbf{U} = \{\mathbf{E}_1\}$. It then finds a second signature, denoted by $\mathbf{E}_2$, with the maximum projection in $< \mathbf{E}_1 >^{\perp}$, which is the orthogonal complement space linearly spanned by $\mathbf{E}_1$. A third signature $\mathbf{E}_3$ can then be found by applying another orthogonal subspace projector $P_{\mathbf{U}}^{\perp}$ with $\mathbf{U} = \{\mathbf{E}_1, \mathbf{E}_2\}$ to the original image, where the signature that has the maximum orthogonal projection in $< \mathbf{E}_1, \mathbf{E}_2 >^{\perp}$ is selected as $\mathbf{E}_3$. The above procedure is repeated until a set of endmembers $\{\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_p\}$ is extracted, where p is an input parameter to the algorithm.

3. PROPOSED ENDMEMBER EXTRACTION APPROACH

In the following, we describe our proposed SQEEA, which has been designed to account both for the extremity of endmembers in the data cloud and for the capacity of such endmembers to properly unmix a hyperspectral scene (satisfying both the ANC and ASC constraints) when the scene is unmixed using an unconstrained linear model. The algorithm can be summarized by the following steps:

1. *Feature reduction.* Apply a dimensionality reduction transformation such as the minimum noise fraction (MNF) or the principal component analysis (PCA)[22,23] to reduce the dimensionality of the data from n to $p-1$, where p is an input parameter to the algorithm (number of endmembers to be extracted).

2. *Initialization.* Calculate the centroid of the data cloud as $\overline{\mathbf{X}} = \frac{\sum_{i=1,j=1}^{r,c} \mathbf{X}(i,j)}{r \times c}$. Here, r is the number of rows and c is the number of columns of the original hyperspectral scene. The algorithm now assumes that all other pixels in the image scene are pure pixels made up of $\overline{\mathbf{X}}$ with 100% abundance and performs an unconstrained linear unmixing of the scene using $\overline{\mathbf{X}}$ as the only endmember. Of course, the resulting abundance estimation is not accurate. However, it can be used to identify the pixel with highest fractional abundance error. This pixel is selected as the first endmember $\mathbf{E}_1$.

3. *Iterative endmember extraction.* Once the initial endmember $\mathbf{E}_1$ has been found, the following endmembers are extracted (in iterative fashion) by evaluating all pixels at the different spatial locations (i,j) in the original image and selecting the pixel which maximizes the following expression:

$$\epsilon_\Phi = \mathrm{SAD}(\overline{\mathbf{X}}, \mathbf{X}(i,j)) + \frac{1}{r \times c \times p} \sum_{i=1}^{r} \sum_{j=1}^{c} \Big| (1 - \sum_{z=1}^{p} |\Phi_z(i,j)|) \Big|, \tag{4}$$

   where the first term accounts for the spectral angle distance (SAD)[3] of the considered pixel with regards to the centroid of the data cloud (thus providing an idea about the extremity or *convexity* of the pixel in spectral angle sense) and the second term accounts for the number of negative fractions, or sets of fractions summing less than or above one for the considered pixel, when the pixel is unmixed using an unconstrained linear mixture model. Here, p refers to the number of endmembers extracted so far. As a result, the second term in equation (4) evaluates to what extent the fractional abundance estimates provided by the unconstrained model satisfy the ANC and ASC constraints, which should naturally hold if the endmembers are properly selected. The idea of selecting as a new endmember the pixel that maximizes equation (4) is intended to minimize the number of fractional abundance estimates that do not satisfy the ANC and ASC constraints when the unconstrained linear mixture model is applied.

4. *Stopping rule.* The aforementioned procedure is repeated in order to find a second endmember $\mathbf{E}_2$, a third endmember $\mathbf{E}_3$, and so on, until a set of p endmembers $\{\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_p\}$ are extracted, where p is the only input parameter to the algorithm.

4. HYPERSPECTRAL DATA SETS

4.1 Synthetic hyperspectral scenes

A database of five synthetic hyperspectral scenes has been created using fractals to generate distinct spatial patterns, which are then used to simulate linear mixtures of reflectance signatures selected from a spectral library compiled by the U.S. Geological Survey (USGS)*. Fig. 3 shows the five scenes considered in experiments, and 4 shows the nine fractional abundance maps used in the generation of one of such scenes. In Fig. 4, black color indicates 0% abundance of the corresponding mineral, white color indicates 100% abundance of the mineral, and fractional abundances in each pixel of the scene sum to unity. Zero-mean Gaussian noise was added to the scenes in different signal to noise ratios (SNRs) –from 30:1 to 110:1– to simulate contributions from ambient (clutter) and instrumental sources, following a previously described procedure.[16]

*http://speclab.cr.usgs.gov/spectral-lib.htm

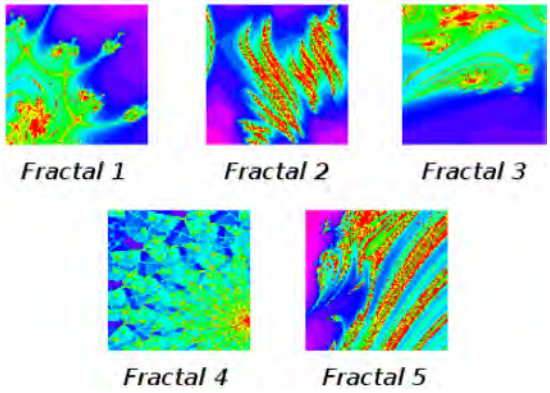


Figure 3. Synthetic images used in experiments, where spatial patterns where generated using fractals.

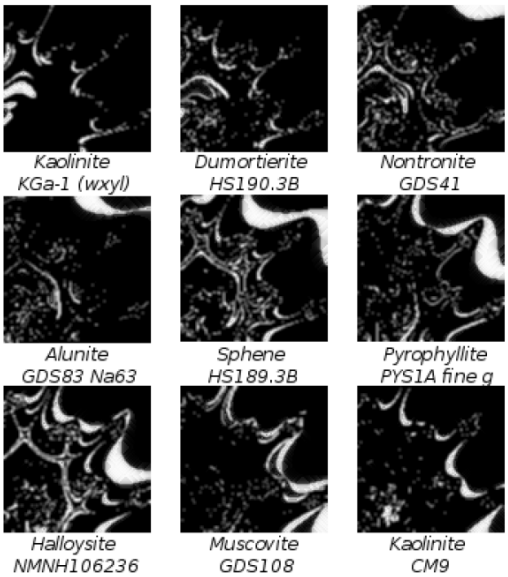
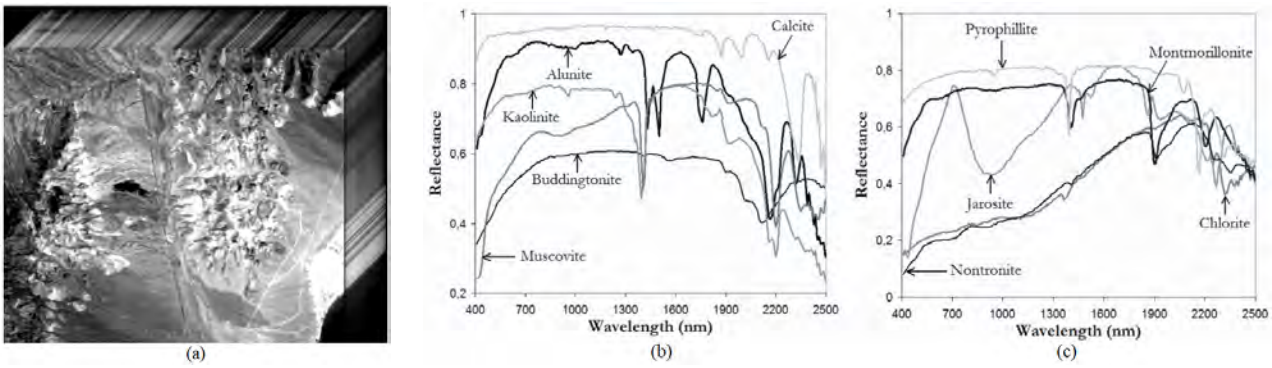


Figure 4. Fractional abundance distributions used in the generation of one of the synthetic images.

4.2 Real hyperspectral data

A well-known hyperspectral data set has been selected for the purpose of illustrating the spectral unmixing algorithm described in this work. The data set was collected by the AVIRIS sensor over the Cuprite mining district in Nevada [see Fig. 5(a)], and is available online in both radiance and reflectance units[†]. In our experiments, we use reflectance data in order to relate our results to the reference USGS spectral library. The scene selected for experiments is the one labeled as f970619t01p02_r02_sc03.a.rfl. This scene comprises a relatively large area (350×350 pixels and 20-meter pixels) and 224 spectral bands between 0.4 and 2.5 μm. Bands 1-3, 105-115 and 150-170 were removed prior to the analysis due to water absorption and low SNR in those bands. The site is well understood mineralogically, and has several exposed minerals of interest. Fig. 5(b-c) shows reference ground signatures of the above minerals. These signatures will be used to assess endmember signature purity in this work.

[†]http://aviris.jpl.nasa.gov/html/aviris.freedata.html



(a) (b) (c)

Figure 5. AVIRIS Cuprite scene (a) and reference spectra for the scene (b-c).

5. EXPERIMENTAL RESULTS

5.1 Experiments with synthetic hyperspectral data

We have first conducted an analysis of the proposed technique in the task of extracting pure spectral signatures from the five synthetic scenes in Fig. 3 with different signal to noise ratio (SNR) proportions, and compared the obtained results to those provided by the OSP and N-FINDR algorithms. In all cases, the number of endmembers to be extracted was set to $p = 9$, the number of random USGS spectral signatures used for the construction of those scenes.

Table 1 shows the mean spectral angle distance (SAD) of the extracted endmembers that were matched (using a previously developed spectral similarity matching algorithm[8]) to their corresponding reference USGS spectral signatures, for the proposed algorithm and for the OSP and N-FINDR algorithms when those algorithms were applied to the five synthetic scenes without noise and with SNRs ranging from 30:1 to 110:1. In the table, lower SAD indicates higher spectral similarity. As shown by Table 1, the proposed algorithm outperforms OSP and N-FINDR in three of the considered scenes (Fractal2, Fractal4 and Fractal5) although the average SAD scores reported for the method decrease when the SNR is very high (30:1). On the other hand, both the OSP and N-FINDR perform similarly, thus indicating that the design of an endmember extraction algorithm in the form of a SMEEA or a SQUEEA does not have a significant impact on algorithm performance. In this case, the design in the form of a SQUEEA has the computational advantage that it is not necessary to recalculate all the endmembers if a set of p endmembers has already been obtained and the number of endmembers needs to be increased to q with $q > p$. In this case, a SQUEEA such as the proposed method (and OSP) can take advantage of a set of $p - 1$ endmembers previously generated by the same algorithm. It should be noted that, in our experiments with synthetic scenes, the determination of the value of p is straightforward since the number of endmembers available in the simulated scenes is known in advance. However, in real data experiments several tests often need to be performed before setting the value of p to a desirable value.

5.2 Experiments with real hyperspectral data

To illustrate the performance of the proposed method with a real hyperspectral scene, we have conducted an experimental assessment of endmember extraction algorithms using the well-known AVIRIS Cuprite data set. In this work, we estimate the number of endmembers in the AVIRIS Cuprite data using the virtual dimensionality (VD) concept, which has been shown in previous work to be an effective approach for this purpose.[24] The VD estimated a total of $p = 18$ endmembers, which is consistent with the available ground-truth for the scene.

An experiment-based cross-examination of algorithm endmember extraction accuracy is presented in Table 2, which tabulates the SAD scores obtained after comparing some selected USGS library spectra with the corresponding endmembers extracted by the three considered algorithms (OSP, N-FINDR and the proposed method). The smaller the SAD values across the ten minerals considered, the better the results. It should be noted that Table 2 only displays the smallest SAD scores of all endmembers with respect to each USGS signature for each algorithm. As shown in the table, all tested methods produced endmembers which were similar, spectrally, to the

Table 1. Average SAD score (cosine of the spectral angle) of matched endmembers with regards to USGS reference spectral signatures for each set of $p = 9$ endmembers extracted by OSP, N-FINDR and the proposed algorithm from the synthetic scenes without noise and with SNRs from 30:1 to 110:1.

Scene	SNR	OSP	N-FINDR	Proposed
Fractal1	No noise	0.0031	0.0030	0.0041
	110:1	0.0031	0.0030	0.0041
	90:1	0.0029	0.0033	0.0041
	70:1	0.0030	0.0029	0.0042
	50:1	0.0048	0.0046	0.0073
	30:1	0.0349	0.0324	0.0352
Fractal2	No noise	0.0011	0.0012	0.0011
	110:1	0.0011	0.0012	0.0011
	90:1	0.0011	0.0012	0.0011
	70:1	0.0014	0.0014	0.0013
	50:1	0.0039	0.0040	0.0039
	30:1	0.0354	0.0339	0.0347
Fractal3	No noise	0.0077	0.0077	0.0175
	110:1	0.0077	0.0077	0.0175
	90:1	0.0077	0.0077	0.0175
	70:1	0.0078	0.0077	0.0176
	50:1	0.0092	0.0093	0.0192
	30:1	0.0384	0.0364	0.0381
Fractal4	No noise	0.0077	0.0058	0.0052
	110:1	0.0060	0.0058	0.0052
	90:1	0.0069	0.0072	0.0053
	70:1	0.0055	0.0068	0.0054
	50:1	0.0074	0.0088	0.0073
	30:1	0.0359	0.0363	0.0405
Fractal5	No noise	0.0121	0.0138	0.0121
	110:1	0.0121	0.0138	0.0121
	90:1	0.0129	0.0146	0.0121
	70:1	0.0129	0.0146	0.0121
	50:1	0.0140	0.0139	0.0133
	30:1	0.0401	0.0435	0.0557

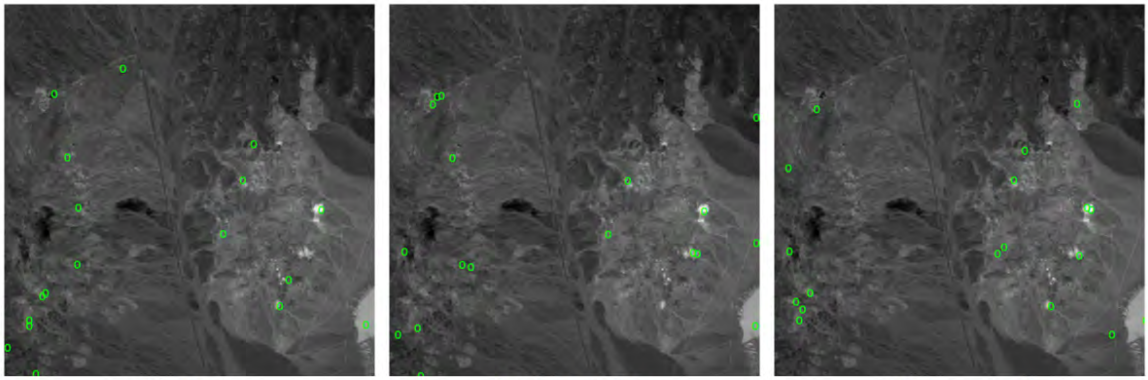
USGS reference signatures. Interestingly, the SAD spectral similarity scores obtained for the proposed algorithm were similar and, in some cases, superior to those reported by the OSP and N-FINDR algorithms.

For illustrative purposes, Fig. 6 maps the spatial locations of the $p = 18$ endmembers extracted by the OSP, N-FINDR and the proposed method. As shown by Fig. 6, the three considered methods provide several endmembers which are overlapped in terms of their spatial locations in the scene. On the other hand, Table 2 indicates that the endmembers which are not overlapped are similar in spectral terms, since the SAD scores reported in the table (comparing the best matched endmember for each method with regards to each USGS reference spectral signature) are very similar, thus indicating that the three methods perform similarly in the considered real hyperspectral image analysis scenario. Future work should comprise an evaluation of the considered methods under different analysis scenarios, using hyperspectral data sets collected by different instruments and in the context of different application domains.

To conclude our experimental results section, we provide an idea about the computational complexity of the proposed method in the task of processing the AVIRIS Cuprite hyperspectral scene in a personal computer with Intel Core 2 Duo processor at 1.83 GHz and 3 GB of RAM memory. In this platform, our Matlab implementation of the OSP algorithm extracted the $p = 18$ endmembers in 149.73 seconds, while our Matlab implementation of N-FINDR needed 780.67 seconds. Finally, our Matlab implementation of the proposed approach selected all

Table 2. SAD-based spectral similarity scores among selected USGS mineral spectra in Fig. 5(b-c) and the endmembers produced by different algorithms.

Mineral	OSP	N-FINDR	Proposed
Alunite	0.1002	0.1134	0.1002
Buddingtonite	0.0929	0.0929	0.0985
Calcite	0.1849	0.1216	0.1216
Kaolinite	0.1969	0.2185	0.2250
Chlorite	0.2695	0.2834	0.2695
Jarosite	0.1371	0.1715	0.1441
Montmorillonite	0.1045	0.1050	0.0903
Pyrophilite	0.1735	0.1434	0.1397
Muscovite	0.1077	0.1104	0.0946
Nontronite	0.2045	0.2185	0.2297



(a) OSP (b) N-FINDR (c) Proposed

Figure 6. Spatial locations of the $p = 18$ endmembers extracted by the three considered extraction methods.

$p = 18$ endmembers in 155.4 seconds. It should be noted that the proposed algorithm was implemented with PCA-based dimensionality reduction as its initial step, which consumed about 20 seconds of processing time (already included in the total processing time reported).

6. CONCLUSIONS AND FUTURE RESEARCH LINES

In this paper, we have described a new algorithm for automatic endmember extraction which sequentially finds a pre-determined number of endmembers by considering both the extremity of sample pixel vectors (tested iteratively as potential endmember candidates) and the capacity of endmember candidates to satisfy the abundance non-negativity and sum-to-one constraints (on a per-pixel basis) after using an unconstrained linear spectral unmixing model. Our assumption in this work is that the aforementioned constraints should naturally hold in the fractional abundance estimations if the endmembers are properly selected, and thus make use of this assumption to intelligently guide the image endmember searching process to a combination of available extreme image pixels that minimize the number of fractional abundance estimations that do not satisfy such constraints. Our experimental results, obtained using both synthetic and real hyperspectral scenes, indicate the importance of considering not only the convexity or extremity of spectral endmembers (accounted for in this work in spectral angle sense), but also their capacity to produce fractional abundance estimations which can naturally fulfill the ASC and ANC constraints when applying an unconstrained linear spectral unmixing model in the process of automatically extracting such endmembers, i.e. without imposing such constraints during the endmember searching process. As future work, additional criteria for measuring the convexity of extracted endmembers (e.g., volume of the simplex defined by the endmembers or their orthogonality) in combination with the proposed error term (measuring the deviation of unconstrained versus constrained linear unmixing) should be considered. Further experiments should also be conducted in order to substantiate if the proposed method can also provide adequate

(either unconstrained or constrained) fractional abundance estimations when the set of derived endmembers is used to unmix additional synthetic and real hyperspectral scenes.

7. ACKNOWLEDGEMENT

This work has been supported by the European Community's Marie Curie Research Training Networks Programme under reference MRTN-CT-2006-035927 (HYPER-I-NET). Funding from the Spanish Ministry of Science and Innovation (HYPERCOMP/EODIX project, reference AYA2008-05965-C04-02) is gratefully acknowledged.

REFERENCES

[1] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for Earth remote sensing," *Science* **228**, pp. 1147–1153, 1985.

[2] 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, *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment* **65**(3), pp. 227–248, 1998.

[3] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic/Plenum Publishers: New York, 2003.

[4] 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," *Journal of Geophysical Research* **91**, pp. 8098–8112, 1986.

[5] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing* **39**, pp. 529–545, 2000.

[6] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine* **19**(1), pp. 44–57, 2002.

[7] 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 Transactions on Geoscience and Remote Sensing* **42**(10), pp. 2085–2095, 2004.

[8] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing* **42**(3), pp. 650–663, 2004.

[9] A. Plaza and C.-I. Chang, "Impact of initialization on design of endmember extraction algorithms," *IEEE Transactions on Geoscience and Remote Sensing* **44**(11), pp. 3397–3407, 2006.

[10] C.-I. Chang, *Hyperspectral Data Exploitation: Theory and Applications*, John Wiley & Sons: New York, 2007.

[11] M. E. Winter, "N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE Image Spectrometry V* **3753**, pp. 266–277, 2003.

[12] J. H. Bowles, P. J. Palmadesso, J. A. Antoniades, M. M. Baumback, and L. J. Rickard, "Use of filter vectors in hyperspectral data analysis," *Proc. SPIE Infrared Spaceborne Remote Sensing III* **2553**, pp. 148–157, 1995.

[13] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping Target Signatures Via Partial Unmixing of Aviris Data," *Proc. JPL Airborne Earth Sci. Workshop* , pp. 23–26, 1995.

[14] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Transactions on Geoscience and Remote Sensing* **40**(9), pp. 2025–2041, 2002.

[15] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing* **45**(3), pp. 765–777, 2007.

[16] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection," *IEEE Transactions on Geoscience and Remote Sensing* **32**(4), pp. 779–785.

[17] H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Transactions on Aerospace and Electronic Systems* **39**(4), pp. 1232–1249, 2003.

[18] R. A. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, "Automatic endmember extraction from hyperspectral data for mineral exploration," *Proc. 21st Canadian Symp. Remote Sens.* , pp. 21–24, 1999.
[19] A. Ifarraguerri and C.-I. Chang, "Multispectral and hyperspectral image analysis with convex cones," *IEEE Transactions on Geoscience and Remote Sensing* **37**(2), pp. 756–770, 1999.
[20] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex Component Analysis: A Fast Algorithm to Unmix Hyperspectral Data," *IEEE Transactions on Geoscience and Remote Sensing* **43**(4), pp. 898–910, 2005.
[21] C.-I. Chang and A. Plaza, "A fast iterative algorithm for implementation of pixel purity index," *IEEE Geoscience and Remote Sensing Letters* **3**(1), pp. 63–67, 2006.
[22] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing, 2nd ed.*, Academic Press: New York, 1997.
[23] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, Springer, 2006.
[24] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on* **42**(3), pp. 608–619, 2004.