

VALIDATION OF EXTRACTED ENDMEMBERS FROM HYPERSPECTRAL IMAGES

A. Sharifi^{1*}, M. Hosseingholizadeh²

¹ Department of Surveying engineering, Faculty of civil engineering, Shahid Rajaei Teachers Training University, Tehran, a_sharifi@sru.ac.ir

² Department of Surveying engineering, Faculty of civil engineering, Shahid Rajaei Teachers Training University, Tehran, m.hgzadeh@gmail.com

KEY WORDS: Endmember extraction, Field spectral, Hyperspectral image, Pixel Purity Index (PPI), Sequential Maximum Angle Convex Cone (SMACC), Spectral library, Spectral mixture analysis.

ABSTRACT:

An essential step in the characterization of surface materials using hyperspectral image analysis is image classification using endmembers. Spectral unmixing is the best method for hyperspectral image classification. This method assumes that the pixel-to-pixel variability in scene results from varying proportions of spectral endmembers. Spectral endmembers can be derived from the imagery or measurements in the laboratory or field. The primary objective of this paper was to assess the ability of extracted endmembers against some different solutions for extraction of endmembers for hyperspectral image classification. In this paper, we compared the Pixel Purity Index (PPI) and the Sequential Maximum Angle Convex Cone (SMACC) as two popular methods of endmember extraction with library and field spectral. We used spectral information divergence for detection desirable endmembers from field spectral. For accuracy assessment of spectral mixture analysis and production of endmember abundance images for each of methods, the linear spectral unmixing algorithm is used. After a comparison between the results of these methods, it has been verified that field spectral have a better classification result in comparing with other endmember extraction methods. Also, the PPI has reliable results as an automatic endmember extraction method in comparing.

1. INTRODUCTION

With advances of imaging sensors in these years, hyperspectral imagery is being applied increasingly more in different scientific fields (Chang, 2003) common problem associated with remote sensing images is the full existence of mixed pixels, within which more than one type of material is present. Usually, to utilize the remote sensing images in this situation, one has to decompose these mixed pixels into a set of simple ground cover spectral signatures (known as endmembers) and their corresponding proportions. This process is called spectral unmixing, which generally involves two procedures: the first step is to identify the endmembers which exist in the remote sensing images (endmember collection); and the second step is to acquire the proportion of each endmember for each pixel (abundance estimation) (Tao, 2007). Usually, endmember collection, the first step of spectral unmixing, is much more challenging. Endmember collection algorithms identify the endmembers based on the theory of convex geometry that under the linear spectral mixture analysis, the observations from a scene are in a simplex whose vertices correspond to the endmembers. To determine physical endmembers, several algorithms have been developed, such as the pixel purity index (PPI) (Boardman, 1995), and the sequential maximum angle convex cone (SMACC) (Gruninger, 2004). For an image cube with N spectral bands, these algorithms attempt to find the endmembers as spectral convexities in an N -dimensional space.

2. METHODOLOGY

2.1 Linear Spectral Unmixing

Over the past years, linear spectral mixture analysis (LSMM) has been widely used for mixed pixel decomposition. It assumes

that the spectral signature of an image pixel is linearly mixed by the spectral signatures of objects present in the image (Tao, 2007).

Define X as a hyperspectral vector of a single pixel in hyperspectral remote sensing images, and A as a reflectance characteristic matrix composed of reflectance of each object in each spectral band, and S as a vector composed of the percentage of each object. So we can obtain the equation:

$$X = A \times S \quad (1)$$

If the hyperspectral remote sensing images have n bands and m sorts of interesting objects, then X is $n \times 1$ vector, A is an $n \times m$ matrix, and S is a $m \times 1$ vector. In this model, the selection of matrix A is essential to the precision of unmixing results.

(a) The sum of percentages s_i of objects of interest in every single pixel should be 1, i.e.

$$\sum_{i=1}^m s_i = 1 \quad (2)$$

(b) The sum of percentages s_i of objects of interest in every single pixel should be 1, i.e.

$$0 \leq s_i \leq 1 \quad (i = 1, 2, \dots, m) \quad (3)$$

Also, because the spectral of ground objects is a sort of energy, it cannot be negative, so the elements of matrix A should be positive, i.e.

$$a_{ij} \geq 0 \quad (4)$$

2.2 Endmember Extraction Methods

According to the definition in (Chang, 2003), an endmember is an idealized pure signature for a class. Finding pure signatures

* Corresponding author

in hyperspectral imagery is considered to be an essential and crucial task in hyperspectral data exploitation. The best methods for endmember collection are mentioned in this section.

2.2.1 Pixel Purity Index (PPI)

The pixel purity index (PPI) has been widely used in hyperspectral image analysis for endmember collection due to its publicity and availability in the Environment for Visualizing Images (ENVI) software. This algorithm was developed by Boardman et al. (Boardman, 1995). It is an unsupervised algorithm as opposed to the PPI, which requires human intervention to manually select a final set of endmembers (Envi, 2006). The PPI algorithm procedure (Chang, 2006) is as follows:

Initialization: Apply a maximum noise fraction (MNF) transform (Green, 1988) to reduce the dimensionality of the dataset, and generate a set of k unit vectors called “skewers,” $\{skewer_j\}_j^k = 1$ randomly, where k is sufficiently a large positive integer.

PPI Calculation: For each $skewer_j$, all the data sample vectors are projected onto the skewer to find sample vectors at its extreme positions and form the last set for this particular $skewer_j$, denoted by $S_{extrema}(skewer_j)$. Even though a different $skewer_j$ generates a different set $S_{extrema}(skewer_j)$, it is very likely that some sample vectors may appear in more than one set. An indicator functions of a set $S, I_S(r)$ by:

$$I_S(r) = \begin{cases} 1, & \text{if } r \in S \\ 0, & \text{if } r \notin S \end{cases} \quad (5)$$

and

$$N_{PPI}(r) = \sum_j I_{S_{extrema}(skewer_j^{(k)})}(r) \quad (6)$$

where $N_{PPI}(r)$ is defined as the PPI score of sample vector?

Candidate Selection: Find the PPI scores $N_{PPI}(r)$ for all the sample vectors defined in initialization section.

Endmember Collection: Let be a threshold value set for the PPI score. Collect all the sample vectors with $N_{PPI}(r) \geq t$.

The pixels resulting from the algorithm above are usually input to ENVI’s “L-dimensional visualization tool” which is an additional tool of the PPI algorithm that allows manual selection of a final set of endmembers (Envi, 2006).

2.2.2 SMACC

The sequential maximum angle convex cone (SMACC) is one of the newest methods for finding spectral endmembers and their abundances throughout hyperspectral images. This algorithm was developed by Gruninger et al. (Gruninger, 2004). SMACC uses a convex cone model (known as Residual Minimization) with these constraints to identify image endmember spectrally. Extreme points are used to determine a convex cone, which defines the first endmember. A constrained oblique projection is then applied to the existing cone to derive the next endmember (Envi, 2006). The cone is increased to include the new endmember. The process is repeated until a projection derives an endmember that already exists within the convex cone (to a specified tolerance) or until the specified

number of endmembers is found. In other words, SMACC first finds the brightest pixel in the image; then it finds the pixel most different from the brightest. Then, it finds the pixel most different from the first two.

The process is repeated until SMACC finds a pixel already accounted for in the group of the previously found pixels, or until it finds a specified number of endmembers. The spectral of pixels that SMACC finds become the endmembers of the resulting spectral library. Endmembers derived from SMACC are unique, a one-to-one correspondence does not exist between the number of materials in an image and the number of endmembers. SMACC derives endmembers from pixels in an image. Each pixel may contain only one material, or it may contain a high percentage of a single material with unique combinations of other materials. Each material identified in an image is described by a subset spanning its spectral variability. The SMACC provides an endmember basis that defines each of these material subsets and provides abundance images for determining the fractions of the total spectrally integrated radiance, or reflectance of a pixel contributed by each resulting endmember. Mathematically, SMACC uses the following convex cone expansion for each pixel spectrum (endmember), defined as:

$$H(c, i) = \sum_k^N R(c, k)A(k, j) \quad (7)$$

where: i is the pixel index, j and k are the endmember indices from 1 to the expansion length (N), R is a matrix that contains the endmember spectral as columns, c is the spectral channel index, A is a matrix that contains the fractional abundance of each endmember j in each endmember k for each pixel (Gruninger, 2004).

The 2D matrix representation of a spectral image is factored into a convex 2D basis (a span of a vector space) times a matrix of positive coefficients. In the image matrix (R), the row elements represent individual pixels, and each column represents the spectrum of that pixel. The coefficients in A are the fractional contributions or abundances of the essential members of the original matrix. The basis forms an n -D convex cone within its subset. The convex cone of the data is the set of all favorable linear combinations of the data vectors, while the convex hull is the set of all weighted averages of the data. The factor matrices are then determined sequentially. At each step, a new convex cone is formed by adding the selected vector from the original matrix that lies furthest from the cone defined by the existing basis (Envi, 2006).

3. EXPERIMENTAL RESULT

The image data to be used for experiments is the Cuprite AVIRIS image scene, which was collected over the Cuprite mining district, Nevada, in 1997. The data cube consists of 50 spectral bands in the visible, near-IR, and shortwave-IR regions of the spectrum with 20 meter GSD. For this study, we use 50 SWIR bands over a 56 square kilometer area, corresponding to a 400x350x50 data cube. This scene is well understood mineralogically, where field spectral provide the precise spatial locations of bright pixels that correspond to the four minerals, alunite, buddingtonite, calcite, and kaolinite. Their spectral signatures are shown in Figure 1. We used the laboratory

reflectance spectral of the known minerals are publicly available at USGS spectral library and a reference image for comparing results in our experiments.

3.1 Accuracy Assessment of Endmember Extraction Methods

In this experiment, the PPI and SMACC were implemented for the image scene in the Environment for Visualizing Images (ENVI) 4.3 software system developed by Research Systems, Inc. We set the number of iterations in PPI parameters to 10000, and the threshold factor to 1. We select individual pixels falling into the corners of the data cloud and use those pixel signatures as the final endmember set. For collecting endmembers using SMACC algorithm, we set coalesce redundant endmember to 0.1 and unmixing constraint to “positive only. Also, we use the mean spectral of field spectral for comparison of results.

We found about 20 endmembers for PPI and SMACC, but we could not assess the accuracy of all endmember because we had only four fields spectral. We measured the spectral similarity between detected endmember spectral and reference spectral. Then we selected four endmembers as the most similar spectral signature to reference endmember spectral. In this case, we observed that PPI results are better than SMACC result because of the more significant spectral angle between detected endmembers and reference endmembers for SMACC. On the other hand, the smaller spectral angle in two methods for field spectral in comparison with spectral library means the better accuracy for field spectral as ground truth.

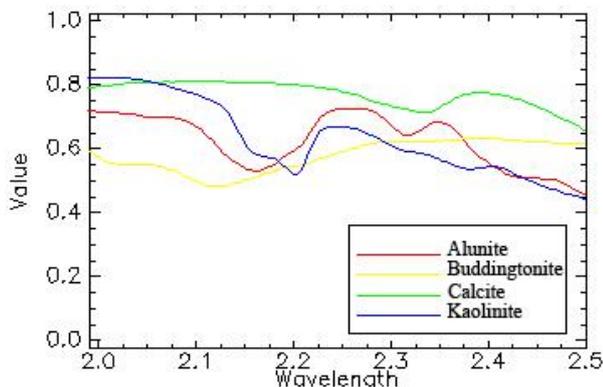


Figure 1. Extracted endmember Comparison using PPI and SMACC methods by library and field spectral.

3.2 Accuracy Assessment of Spectral Unmixing

In this experiment, a comparative study is conducted to see how many differences are between abundance images for each of endmember collection methods and their accuracy for solving spectral mixture. Abundance images for all methods are shown in Figure 2.

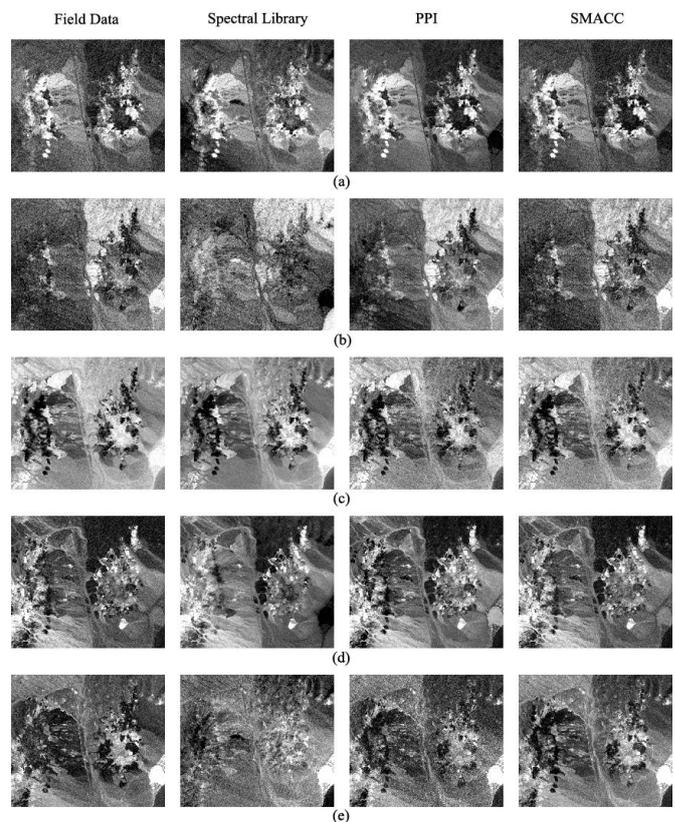


Figure 2. Abundance images for each of methods (a) alunite, (b) buddingtonite, (c) calcite, (d) kaolinite, and (e) error.

Also, table 1 tabulates the accuracy of classification for each of the methods. In this case, we observed that Field Spectral and SMACC have the best result for the spectral mixture. Also, they have the smallest commission (Com.) and omission (Omi.) errors for all classes. On the other hand, for all methods, calcite has the smallest error that means the better solution for spectral mixture analysis with this endmember because of its spectral signature that is different in comparison with other endmember spectral signatures as it is shown in Figure 1. Also, buddingtonite has the most significant error because of its spectral signature similarity to alunite and kaolinite spectral signatures. Also, kaolinite and alunite have the most mixed pixel in each other because of their spectral signature similarity.

Table 1. The overall accuracy of classification

	Field Spectral	Spectral library	PPI	SMACC
Overall accuracy	97.8	96.0	93.1	97.1

4. CONCLUSION

This paper presents comparative results for linear spectral unmixing via four endmember collection methods. Two standard endmember extraction algorithms: PPI and SMACC are compared. They are alternative methods to find endmember spectral signature that do not require any assumptions about the data or knowledge of the number of endmembers. For comparing results, Field Spectral and library Spectral are used. The two above methods with spectral library and Field Spectral are used to collect endmembers and have been applied to cuprite AVIRIS data, which has become a test image for

endmember extraction assessment due to available ground truth. Quantitative results obtained by these methods on four minerals that can be found prominently and in pure form in the test site. Linear spectral unmixing has implemented for achieving abundance images for each of methods and calculating accuracy of spectral mixture analysis. Collected endmembers are used for analyzing spectral mixture and compared abundance images for computing spectral mixture accuracy in each of methods. Experiment 1 presents the best result for Field Spectral as ground truth in comparing with spectral library and PPI as an endmember extraction method. Also, experiment 2 present that spectral library endmembers have a reliable result for unmixing of hyperspectral images with prior knowledge about the area. However, endmember spectral signatures similarity is a limitation for achieving better results.

REFERENCES

- Boardman, J. W., Kruse, F. A., and Green, R. O., "Mapping target signatures via partial unmixing of AVIRIS data," in Summaries of JPL Airborne Earth Science Workshop, Pasadena, CA, 1995.
- Chang, C. I., "Hyperspectral Imaging: Techniques for Spectral Detection and Classification," Amsterdam, The Netherlands, Kluwer, 2003.
- Chang, C. I., and Plaza, A., "A Fast Iterative Algorithm for Implementation of Pixel Purity Index," IEEE Geoscience and Remote Sensing Letters, vol 3, no 1, p.p. 98-106, 2006.
- ENVI User's Guide, Boulder, CO: Research Systems, Inc., 2006.
- Green, A. A., Berman, M., Switzer, P., and Craig, M. D., "A transformation for ordering multispectral data regarding image quality with implications for noise removal," IEEE Transaction on Geoscience and Remote Sensing, vol 26, no 1, p.p. 65–74, 1988.
- Gruninger, J., Ratkowski, A. J., and Hoke, M. L., "The Sequential Maximum Angle Convex Cone (SMACC) Endmember Model," Proceedings SPIE, vol 54, p.p. 25-33, Orlando FL, 2004.
- Tao, X., Wang, B., Zhang, L., and Zhang, J. Q., "A New Endmember Extraction Algorithm Based on Orthogonal Bases of Subspace Formed by Endmembers," IEEE Geoscience and Remote Sensing Symposium, IGARSS, p.p. 2006-2009, 2007.
- Tao, X., Wang, B., Zhang, L., and Zhang, J. Q., "A New Scheme for Decomposition of Mixed Pixels Based on Nonnegative Matrix Factorization," IEEE Geoscience and Remote Sensing Symposium, IGARSS, p.p. 1759-1762, 2007.