

# The Endmembers Selection and Spectral Unmixing Based on the Optimal Combination of the Endmembers Extracted by N-FINDR Algorithm and SSWA Algorithm

Jun Xu<sup>1</sup>

School of Information Engineering,  
East China Jiaotong University,  
Nanchang, China  
e-mail:xjsdcq@163.com

Fuhong Xu<sup>2</sup>

Division of Planning and Finance,  
East China Jiaotong University,  
Nanchang, China  
e-mail:249771538@qq.com

**Abstract**—Based on the convex geometry, a lot of endmember extraction algorithms have been proposed, in which N-FINDR algorithm use sample points in data cloud to construct simplex and maximize its volume, it suitable for the images which contain “pure pixels”, we can randomly select pixels as the endmembers to form the simplex and calculate its volume, and then use other pixels to replace one of the endmembers, calculate the simplex volume again. If the replacement volume increase, then the replacement is accepted, otherwise the replacement is given up, until all the endmembers are founded. SSWA algorithm constructs a simplex to surround the data cloud and shrink it continuously, that is to say, the extracted endmembers may have no corresponding “pure pixels” in the original hyperspectral images, it firstly finds a simplex to contain all the sample points, and then shrinks the simplex volume according to the gradient descent rule with a penalty function. This paper combines the advantages of these two algorithms, selects the optimal combination within the endmembers extracted by these two algorithms, finally obtains the best spectral unmixing accuracy.

**Keywords**-hyperspectral; endmember; simplex; spectral unmixing

## I. INTRODUCTION

As we all know, hyperspectral images has high spectral resolution, but its spatial resolution is generally low, which results in hyperspectral image pixels can not only contain a certain kind of spectral signature, but the mixture of several materials spectral signature. For these mixed pixels, we can not determine their material property based on the extracted spectral vectors, nor simply classify them into to a certain type of material in the classification process<sup>[1-2]</sup>. The existence of a large number of mixed pixels in hyperspectral images makes the accuracy of the traditional pixel-level image classification and segmentation difficult to improve, which is a serious obstacle to the quantitative hyperspectral remote sensing. To improve the accuracy of hyperspectral remote sensing applications and overcome this obstacle, it is necessary to unmix the mixed pixels into

different “basic unit” (endmembers) and obtain the proportion of these “basic unit” (abundance) within each pixel simultaneously<sup>[3]</sup>.

For the mixed pixel, two mixing models have been established: linear spectral mixture model (LSMM) and nonlinear spectral mixture model (NSMM)<sup>[4-5]</sup>. NSMM is more complex, and LSMM is relatively simple and now widely used, in most cases the spectral unmixing accuracy of LSMM can meet application requirements. Before spectral unmixing, endmembers extraction is necessary. The early method of acquiring endmembers is based on the means of human-computer interaction, which depends on the experience and knowledge of the operator and is unfavourable for the fast and accurate processing of the image. In recent years, a series of automatic endmembers extraction algorithms have been proposed and applied by domestic and foreign scholars. N-FINDR<sup>[6]</sup> proposed by Winter in 1999 searches for the maximum volume simplex in the spectral feature space, whose vertices are identified as endmembers. In the same year, Fuhrmann proposed SSWA algorithm<sup>[7]</sup>, which assumes that there is no longer the pure pixels containing a single spectral signature, searches for a minimum volume simplex in the spectral feature space to surround all samples, similarly, the vertices of the simplex are identified as endmembers. For some endmembers, there are the corresponding pure pixels in hyperspectral images; for other endmembers, its corresponding pixels do not exist. Because of the simplex vertices are the actual existing sample points in N-FINDR algorithm, so for those endmembers which have not corresponding “pure pixels” in the image, N-FINDR is unable to accurately extract them. On the contrary, in SSWA algorithm, the vertices of the simplex which surrounds data cloud are not necessarily the real sample points, therefore, for those endmembers who have corresponding “pure pixels” in the image, SSWA algorithm may not be able to accurately extract them. In view of this situation, this paper proposes an endmembers extraction method combining N-FINDR algorithm and

SSWA algorithm, makes best use of the advantages and bypass the disadvantages of these two algorithms, regardless of the endmembers that has the corresponding "pure pixels" in the image or those endmembers have not "pure pixels", through screening and comparison, all the endmembers can be accurately extracted.

## II. ENDMEMBERS EXTRACTION AND SELECTION SCHEME COMBING N-FINDR ALGORITHM AND SSWA ALGORITHM

### A. N-FINDR algorithm

The distribution of hyperspectral images pixels in the spectral feature space looks like a cluster of point cloud, the data cloud is surrounded in a convex simplex<sup>[8,9]</sup>, and the vertices of the simplex are endmembers, therefore the endmembers extraction is transformed into how to find optimal simplex vertices. Domestic and foreign scholars have conducted massive research on endmembers extraction algorithm based on convex geometry.

The N-FINDR algorithm proposed by Winter<sup>[6]</sup>, it selects the convex simplex volume as the optimization object, firstly a group of sample points are selected as the vertices to constructed simplex and calculates the simplex volume, and then the other sample points in point cloud take place of the simplex vertices one by one, if the replaced simplex vertices can make the simplex volume be larger than before, then the simplex vertices are retained, otherwise they are abandoned. Repeating this process until all the sample points are in spectral feature space are calculated, pixels in the image, a convex simplex with the maximum volume is obtained finally, whose vertices are just the endmembers of the hyperspectral images.

The geometry of any dimension can calculate its volume, for example, the one-dimensional space volume  $R^1$  is the length, 2-dimensional space volume  $R^2$  is the area. The simplex volume constructed by  $n+1$  different sample points in  $n$ -dimensional space volume can be calculated by the following formula:

$$V(M) = \frac{1}{n!} \text{abs}(|W|) \quad (1)$$

In formula (1),

$$W = \begin{bmatrix} 1 & 1 & 1 & L & L & 1 \\ w_1 & w_2 & w_3 & L & L & w_{k+1} \end{bmatrix} \quad (2)$$

Where  $\text{abs}(\cdot)$  is the absolute value operator,  $|\cdot|$  is the determinant operator.

According to the simplex volume definition, in the ideal state, the simplex volume formed by  $m$  endmembers is necessarily the biggest among all the simplexes formed by the equal number of pixels. Therefore, we can randomly select  $m$  pixels as the endmembers to calculate simplex volume, and then use other pixel replace one endmember, calculate the simplex volume again, if the replacement volume increase, then the replacement is accepted, otherwise the replacement is given up, and then

use the new pixel replace the endmember, until every endmembers is replaced by all the pixels. Obviously, in this process, the simplex volume increases continually, and finally the simplex with the maximal volume is obtained, at this time, the vertexes of the simplex are just the endmembers. The flow chart of the endmember extraction process is shown in Fig .1.

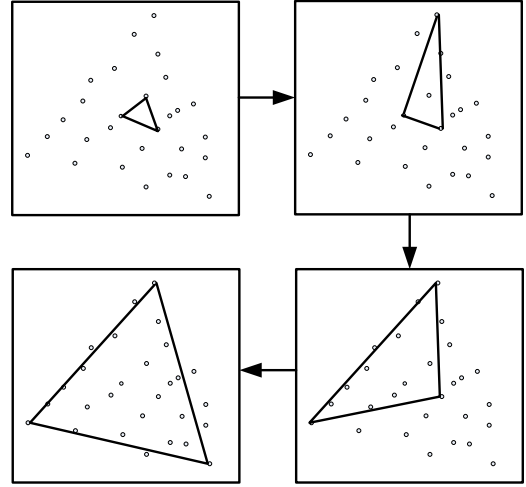


Figure 1. Flow chart of the endmember extraction process of N-FINDR algorithm

N-FINDR algorithm assumes that there are "pure pixels" in the image, but the real image probably contain no "pure pixels", in this case, this running results of N-FINDR algorithm will extract the most slightly mixed pixels as endmembers, which is one of the limitations of this algorithm.

### B. SSWA algorithm

The simplex shrink-wrap algorithm(SSWA) still utilizes the characteristic that all the mixed pixels lie in the simplex which formed by the endmember vertices to extract the endmembers. It no longer assumes the endmembers in the image and extracted from the existing pixels, so its biggest advantage is to ensure all the pixels lie in the simplex formed by the endmembers, but its drawback is that the extracted endmembers may not be with in the image, therefore we can only identify the corresponding objects of the endmembers by the spectral information and can not use the spatial information.

Unlike N-FINDR algorithm which continually "expands" the simplex, the implementation way of SSWA is to "shrink" the simplex volume, it firstly finds a simplex to contain all the sample points, and then shrinks the simplex volume according to the gradient descent rule with a penalty function.

The initial simplex can be selected according to two times the maximum value of each band of the image pixels, obviously, the simplex formed by these pixels and the origin must contain all of the pixels in the image. Due to the volume calculation, therefore this algorithm needs to reduce the dimensionality of data.

If all the pixels in the image  $p_i$  ( $i=1,2,... M+1$ ) in the feature space are surrounded by the simplex constructed

by endmembers  $w_i$  ( $i=1,2,\dots,N+1$ ), then the following objective function can be established:

$$H(W, P) = V(W) + \alpha F(W, P) \quad (3)$$

In equation(3), the definition of  $V(W)$  is the same to  $V(W)$  in equation(1).  $\alpha$  is an adjustment factor,  $F(W, P)$  is the penalty function.

Penalty function is used to ensure all the sample points to be within a single simplex, its role is to keep from the surface of the simplex closing to the data cloud. When the simplex is very large and its surfaces are far away from data cloud, the function value is small; when each surface of the simplex is very close to the data cloud, the function value is large.

The adjustment coefficient  $\alpha$  is used to balance the effect produced by punishment function and volume shrinkage, when the simplex shrink, the "extrapolation" effect of the penalty will gradually increase, which requires the relationship of the simplex volume  $V(W)$  with adjustment coefficient  $\alpha$  is positively related.  $\alpha$  increases with the increment of  $V(W)$  and decrease with the decrement of  $V(W)$ .

Calculating the gradient of the target function  $H(W, P)$ :

$$\nabla H(W, P) = \nabla V(W) + \alpha \nabla F(W, P) \quad (4)$$

The gradient  $\nabla H(W, P)$  obtained from Equation(4) is used as the descent value in each iteration, the following iterative formula is obtained:

$$W^{(k+1)} = W^{(k)} - \mu \nabla H^{(k)}[W^{(k)}, P] \quad (5)$$

$$\nabla H^{(k)}[W^{(k)}, P] = \nabla V[W^{(k)}] + \alpha(k) F[W^{(k)}, P] \quad (6)$$

$\mu$  is the adjustment coefficient of step length, when the times of iterations reach the maximum or the residual error is small enough, the algorithm terminate.

Due to the simplex constructed by SSWA algorithm surround the whole data cloud, the simplex vertices do not belong to the data cloud, that is to say, the extracted endmembers have no corresponding "pure pixels" in the original hyperspectral images. This algorithm can obtain good results for those hyperspectral images which do not contain "pure pixels". If the image does exist the "pure pixel" block, then it is not in accordance with the assumption of SSWA algorithm, then the extracted endmember is virtual here, therefore does not meet the actual situation.

### C. Endmembers selection scheme

If only some endmembers in hyperspectral images have corresponding "pure pixels", the other endmembers are severely mixed and have no corresponding "pure pixels", according to the above analysis, different endmember extraction results will be obtained by N-FINDR algorithm and SSWA algorithm. Because N-FINDR algorithm can only find the actual pixels as the simplex vertices, so the endmembers have corresponding "pure pixels" in the image can be extracted accurately, however, for those endmembers which have no

corresponding "pure pixels" in the image, N-FINDR algorithm will extracted the most severely mixed pixels as the endmembers, obviously it is unreasonable. SSWA algorithm uses the method of outsourcing simplex contraction, all the sample points are contained in a simplex, therefore the endmembers which have no corresponding "pure pixels" in the image those in the image can be extracted accurately, but those endmembers which have corresponding "pure pixels" in the image will be extracted inaccurately. If we extracted two groups of endmembers by N-FINDR algorithm and SSWA algorithm respectively, then filter and optimize the endmembers combination, the most accurate endmembers combination can be found out to unmix the spectral mixture pixel and realize the quantitative hyperspectral remote sensing.

For a hyperspectral image, all the pixels can be mapped to the spectral feature space. We first use N-FINDR algorithm to extract the endmembers. Because the endmembers in the image does not necessarily have the corresponding "pure pixel", therefore these extracted endmembers may be partially accurate. Then we use SSWA algorithm extract the endmembers again in the spectral feature space. At this time, these extracted endmembers are not necessarily the real sample points, but they may be the pure pixels corresponding to the endmembers. Among all the endmembers extracted by these two algorithms, we can iterate through all the combinations to unmix the hyperspectral image. The endmembers with the highest spectral unmixing accuracy will be identified as the hyperspectral image endmember set of this hyperspectral image. The optimal spectral unmixing results are also obtained synchronously.

## III. EXPERIMENT AND DISCUSSION

Three spectral signatures are chosen from the USGS mineral spectral library in the ENVI USGS as the endmembers<sup>[10]</sup>, their spectral curves are shown in Fig .2, based on which the simulated  $200 \times 200$  hyperspectral images are constructed. In the three endmembers, only endmember① has the corresponding "pure pixels" in the simulated images, the endmember②, and endmember③ have no corresponding "pure pixels".

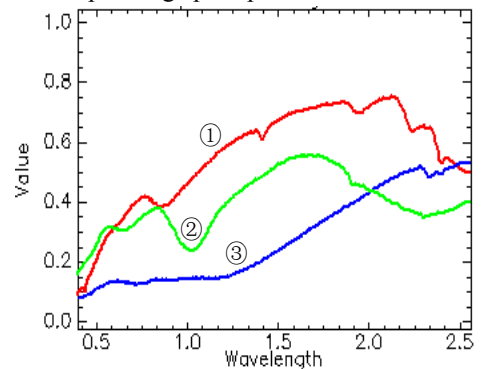


Figure 2. Three pure spectral signatures from USGS

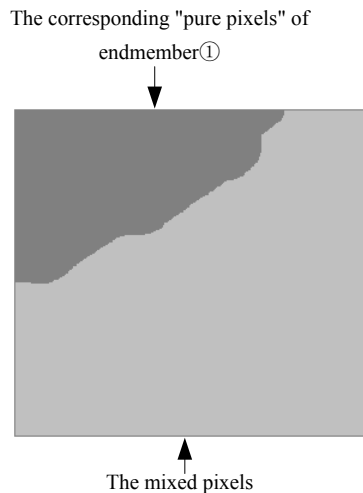


Figure 3. The simulated image

Fig .4 is the 2-dimensional scatter diagram drew by two bands of spectral images. In Fig .4, the red points are the endmembers extracted by F-FINDR algorithm in the simulated images, the blue points are the endmembers extracted by SSWA algorithm. All the endmembers combination abc, abC, aBc, aBC, ABC, ABc, Abc are used to unmix the simulated hyperspectral images respectively based on the unconstrained linear spectral mixture model, then we compare the RMSE of the reconstructed images with the one of the original images, found that the endmembers combination aBC gain the minimal RMSE, therefore, the final extracted endmembers are determined as a, B, C. By comparison, the endmembers a, B, C extracted in the simulated images are highly similar to the spectral curves in Fig .2. For the real hyperspectral data, obviously, the optimal endmembers combination can also selected in the same way.

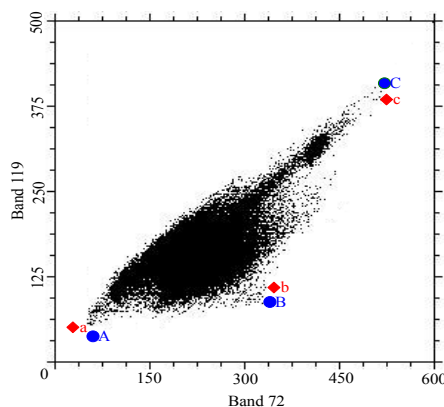


Figure 4. 2-dimensional scatter diagram drew by two bands of spectral images

#### IV. CONCLUSION

Spectral unmixing and object classification are two important goals of remote imaging spectroscopy. The idea of using endmembers derived from the data for classification and unmixing has been considered before.

A prominent problem of the hyperspectral image

quantitative analysis is mixed pixel phenomenon. Because of the spatial resolution limitation of sensors and the variety and complexity of the natural objects, especially many objects are small in size, the pixels are rarely formed by a single surface cover type, but a mixture of several objects. If a instantaneous field of view of the anisotropy is formed within a pixel, then this pixel is called a mixed pixel. Spectral unmixing can be divided into two steps: endmember extraction and abundance estimation. The endmember extraction refers to choose pure spectra of the objects from the image. The abundance estimate refers to calculate the mixing ratios of endmembers. At present, the most studied and the most widely used endmember extraction methods are based on the convex simplex. Since J.W. Boardman firstly used the convex simplex theory to extract the endmember in 1993, the endmember extraction method based on the convex simplex has attracted wide attention of many scholars in the world.

N-FINDR algorithm and SSWA algorithm are two different endmember extraction methods based on the convex simplex. They have their own advantages and shortcomings. N-FINDR algorithm is suitable to look for the endmembers in the hyperspectral images without "pure pixels"; SSWA algorithm is suitable to find the endmembers in the hyperspectral images with "pure pixels". This paper combines these two algorithms to find the optimal endmembers group which is favorable to obtain the optimal spectral unmixing precision. The results of this work are very encouraging. Results with simulated data reveal that the method can accurately extract the endmember spectral signatures.

#### ACKNOWLEDGMENT

This work was financially supported by Humanities and Social Science Foundation of the Ministry of Education of China(Grant No.14YJCZH172), and the Science and Technology Research Projects of Jiangxi Provincial Education Department of China(Grant No.GJJ14368)

#### REFERENCES

- [1]. Gillespie A R. Spectral mixture analysis of multispectral thermal infrared images[J]. Remote Sensing of Environment, 1992, 42(2): 137-145.
- [2]. Foody G M, Cox D P. Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions[J]. Remote sensing, 1994, 15(3): 619-631.
- [3]. Keshava N, Mustard J F. Spectral Unmixing[J]. IEEE Signal Processing Magazine, 2002,19(1):44-57.
- [4]. Manolakis D, Siracusa C, Shaw G. Hyperspectral subpixel target detection using the linear mixing model[J]. IEEE Transactions on Geoscience and Remote Sensing, 2001, 39(7): 1392-1409.
- [5]. Ichoku C, Karnieli A. A review of mixture modeling techniques for sub - pixel land cover estimation[J]. Remote Sensing Reviews, 1996, 13(3-4): 161-186.
- [6]. Winter M E. N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data[C]//SPIE's International Symposium on Optical Science, Engineering, and Instrumentation. International Society for Optics and Photonics, 1999: 266-275.

- [7]. Fuhrmann D R. Simplex shrink-wrap algorithm[C]//AeroSense'99. International Society for Optics and Photonics, 1999: 501-511.
- [8]. Boardman J W. Automating spectral unmixing of AVIRIS data using convex geometry concepts[C] //Summaries 4th Annu. JPL Airborne Geoscience Workshop. JPL Publication 93-26, 1993, 1: 11-14.
- [9]. Bioucas-Dias J M, Plaza A. An overview on hyperspectral unmixing: geometrical, statistical, and sparse regression based approaches[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International. IEEE, 2011: 1135-1138.
- [10]. <http://speclab.cr.usgs.gov/spectral.lib06b>, U.S Geological Survey Library, splib06b, September, 2007