

ENDMEMBER EXTRACTION FROM HYPERSPECTRAL IMAGERY USING A PARALLEL ENSEMBLE APPROACH WITH CONSENSUS ANALYSIS

F. Ayuso, J. Setoain, M. Prieto, C. Tenllado, F. Tirado

J. Plaza, A. Plaza

Dept. Computer Architecture
Complutense Univ. of Madrid, Spain

Dept. Computer Science
Univ. Extremadura, Cáceres, Spain

ABSTRACT

We have explored in this paper a framework to test in a quantitative manner the stability of different endmember extraction and spectral unmixing algorithms based on the concept of Consensus Clustering. The idea is to investigate if the sensibility of those algorithms to the number of endmembers can be used to estimate this parameter itself. Preliminary results on synthetic data reveal that the proposed scheme, which can be implemented efficiently in parallel, can compete with state-of-the-art schemes.

1. INTRODUCTION

A general approach to unmix hyperspectral data consists of two main steps:

1. *Endmember extraction.* This step identifies a collection of hidden pure spectra (the endmembers) from the hyperspectral scene. Many proposals exist in the literature. Among the most popular ones we highlight N-FINDR, the Simplex Growing Algorithm (SGA), the Pixel Purity Index (PPI), the Vertex Component Analysis (VCA) or the Automatic Target Generation Process (ATGP) [1].
2. *Estimation of fractional abundances.* This step finds the fractional abundances that indicate the proportion of each endmember material in the mixed pixels. Essentially, there are two kind of methods: linear techniques such as Linear Spectral Unmixing (LSU) [2] and non-linear ones such as neural-based approaches including Multi-layer Perceptron (MLP) [3].

Endmember extraction is one of the challenging steps in this process and has received considerable attention in recent years. From a viewpoint of algorithm design, three major issues determine its performance [2]:

1. The **learning rules** that drive the searching process. Most algorithms implicitly or explicitly assume that pure pixels are present in the data to speed up this process, and use different criteria when searching for

those pure pixels. Two major criteria are multidimensional geometry-based simplex volume maximization and pixel spectral signature similarity [4]. A promising scheme based on the Non-Negative Matrix Factorization without the pure-pixel assumption was recently proposed in [5].

2. The **stopping criteria** of the process. Most popular endmember extraction algorithms, such as the aforementioned methods, use as stopping criteria the number of endmembers to be searched. This predefined parameter must be estimated beforehand and becomes critical for performance. If it becomes too low, then not all desired endmembers will be extracted—specifically, those being *weak* endmembers—, or, conversely if it becomes too high, some extracted endmembers may turn out to be mixed signatures [2]. A popular approach is to resort to the Virtual Dimensionality (VD) proposed in [6].
3. The **initial conditions** use by the algorithm. Many algorithms use a random restart but a proper initial selection of the endmembers can be very beneficial [2].

Research efforts have mainly focused on the first issue [7, 5, 8, 4] but both the stopping criteria and the initial conditions also have a significant impact on performance [2]. In this research, we introduce a new framework for the estimation of the number of endmembers that integrates itself with the unmixing step in a closed loop. Our approach is based on *Consensus Clustering* [9], a methodology proposed in the context of gene expression data analysis to represent the consensus across multiple runs of a clustering algorithm. It has been use effectively to determine the number of clusters in the data and to assess the stability of the discovered clusters [10].

For the sake of introducing this methodology in the hyperspectral processing chain, we have focused on a single endmember extraction algorithm (N-FINDR) and one unmixing algorithm (MLP). The effectiveness of this combination is promising but we have to admit that its computational cost is high. However, the methodology is embarrassingly parallel and its execution cost can be drastically reduced on any high-performance computing platforms such as clusters, multicore processors, graphics processing units (GPUs) or even Grids.

The rest of the paper is organized as follows. Section 2 briefly summarizes N-FINDR and MLP. The proposed methodology is described in Section 3. We conclude with some experimental results and remarks in Sections 4 and 5, respectively.

2. PROCESSING CHAIN

This section describes the processing chain that will be used in this work to illustrate the proposed consensus-based framework. It comprises two steps: 1) endmember extraction using the N-FINDR algorithm, and 2) abundance estimation using an MLP neural network.

2.1. Endmember extraction

The original N-FINDR algorithm developed by Winter [11] 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) [12] to reduce the dimensionality of the data from N to $p-1$ bands, where p is an input parameter to the algorithm (number of endmembers to be extracted).
2. *Initialization.* Let $\{\mathbf{E}_1^{(0)}, \mathbf{E}_2^{(0)}, \dots, \mathbf{E}_p^{(0)}\}$ 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 follows:

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

4. *Replacement.* For each pixel vector \mathbf{r} in the input hyperspectral data, recalculate the volume by testing the pixel in all p endmember positions, i.e., first calculate $V(\mathbf{r}, \mathbf{E}_2^{(k)}, \dots, \mathbf{E}_p^{(k)})$, then calculate $V(\mathbf{E}_1^{(k)}, \mathbf{r}, \dots, \mathbf{E}_p^{(k)})$, and so on until calculating the volume $V(\mathbf{E}_1^{(k)}, \mathbf{E}_2^{(k)}, \dots, \mathbf{r})$. If none of the p recalculated volumes is greater than $V(\mathbf{E}_1^{(k)}, \mathbf{E}_2^{(k)}, \dots, \mathbf{E}_p^{(k)})$, 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 $\mathbf{E}_j^{(k+1)} = \mathbf{r}$ 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.

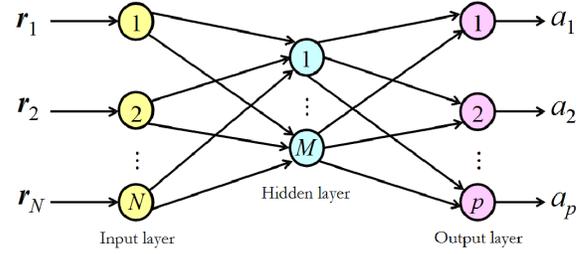


Fig. 1. MLP neural network architecture

2.2. Abundance estimation

The neural architecture used in this work is shown in Fig. 1, where each column of nodes is a layer and the leftmost layer is the input layer. The second layer is the hidden layer, and the third layer is the output layer. The neuron count at the input layer, N , equals the number of spectral bands. The input patterns are pixel vectors directly obtained from the input data. The number of neurons at the output layer, p , equals the number of spectral endmembers. It should be noted that M , the number of hidden neurons, is generally fine-tuned depending on the problem under consideration. Based on previous results in the literature and our own experimentation [3], the considered architecture is based on one hidden layer only, with the number of neurons empirically set to the square root of the product of the number of input features and information classes. The sigmoid function is used as the nonlinear activation function. The MLP is trained using the well-known back-propagation algorithm, where the endmembers obtained from the endmember extraction stage are used as training patterns with known fractional abundances.

3. CONSENSUS-BASED FRAMEWORK

Figure 2 graphically illustrates the proposed methodology. We initially set a target range for the p parameter, which denotes the putative number of endmembers in the image. For each value in this range, we perform different runs of the endmember extraction method and the abundance estimation algorithm. For each run, the pixel vectors of the image are clustered into different classes after applying a winner takes all criterion. Since N-FINDR exhibits a non-deterministic behavior – for each run, a set of endmembers randomly extracted from the input data is used as initial condition –, these classifications vary across multiple runs. For representing and quantifying the agreement among them we use a consensus matrix as introduced in [9].

A consensus matrix is an $(T \times T)$ matrix –where T is the total number of pixels in the hyperspectral scene– that stores, for each pair of pixels, the proportion of classification runs in which both pixels are clustered together. This matrix is obtained by taking the average over the connectivity matrices

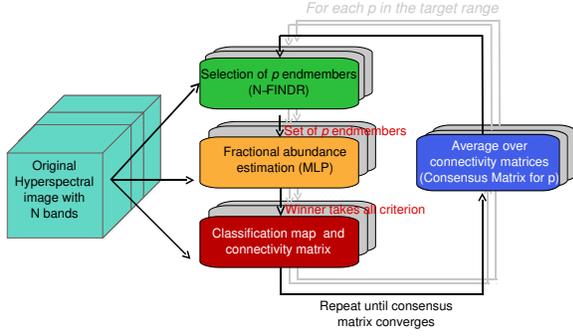


Fig. 2. Consensus-based framework. The p parameter denotes the putative number of endmembers in the image. We search across a given range of this parameter, looking for the actual number of endmembers that achieves the most stable classification.

of every run. The entries of a connectivity matrix are defined as follows:

$$M(i, j) = \begin{cases} 1, & \text{if both pixels belong to the same class} \\ 0, & \text{otherwise} \end{cases}$$

Perfect consensus corresponds to a consensus matrix C with all the entries equal to either 0 or 1, whereas any deviation from this optimal scenario can be interpreted as a lack of stability of the putative classes. With this abstraction, the problem of finding the number of endmembers in the image translates into finding the number of endmembers that yield the *cleanest* consensus matrix – ideally, a matrix containing 0’s and 1’s only –.

Different measures that try to quantify this *cleanliness* have been introduced [9, 10]. A simple but effective approach is to plot the histogram of consensus matrix indices: perfect consensus would translate into two clear bins centered at 0 and 1 respectively, whereas unstable classifications would lead to an increase in the number of fractional entries in the consensus matrix and would introduce secondary bins. Similar insights can be derived by inspection of the corresponding *empirical cumulative distribution* (CDF) [9].

4. EXPERIMENTAL RESULTS

4.1. Hyperspectral data

We have constructed a synthetic hyperspectral data set to illustrate the performance of the proposed consensus-based approach. A major advantage of using synthetic imagery is that all the details of the simulated images are known under a completely controllable environment because they can be manipulated individually and precisely. As a result, algorithm performance can be examined objectively and impartially. The reflectance spectra of nine U.S. Geological Survey (USGS) ground-truth mineral spectra¹: alunite, buddingtonite, calcite,

¹Available online: <http://speclab.cr.usgs.gov/spectral-lib.html>

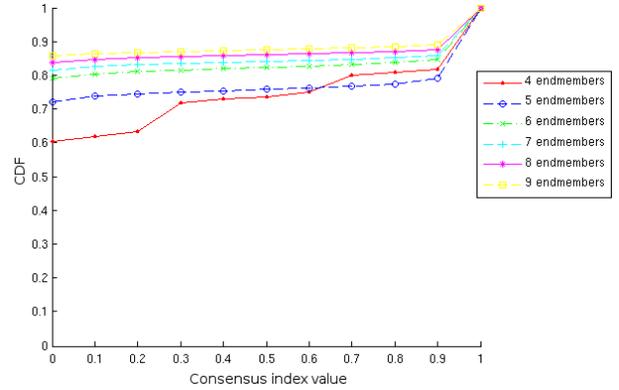


Fig. 3. Empirical CDFs corresponding to the entries of consensus matrices $C(p)$ for $p = 4..9$

kaolinite, chlorite, jarosite, montmorillonite and pyrophyllite, have been managed for computer simulations. These signatures have been used to simulate a square synthetic image scene with a size of 100×100 pixels. The four corner pixels of the image were simulated by the pure spectral signatures of alunite, buddingtonite, calcite, and kaolinite, with the center pixel simulated by the pure signature of muscovite. The four other pure signatures (chlorite, jarosite, montmorillonite, pyrophyllite) were placed halfway between a pair of any two corner pixels. The signature abundance decreased linearly from the pure pixels. As a result, only nine pure pixels (endmembers) are present in the simulated scene. Gaussian noise with a 10:1 signal-to-noise ratio (SNR), as defined in [6], was added to the synthetic scene to simulate contributions from ambient (clutter) and instrumental sources.

4.2. Consensus results

Once consensus matrices have been computed, our estimation of the number of endmembers proceeds by inspection of (1) the empirical CDFs’ shape and (2) their progression as the number of endmembers increases.

Figure 3 displays the empirical CDFs for values of $p = 4 \dots 9$. The predominance of 0’s and 1’s in those consensus matrices affects the shape of these curves. They exhibit a step around 0 –whose magnitude relates to the proportion of 0’s–, followed by a flat line reaching across the 0–1 range, and finally a second step around 1. For 9 endmembers, the shape of the curve closely approaches the ideal step function. However, as we increase the number of endmembers past 9 (Figure 4), the CDF curves display a different shape, with a gradual climb of values between 0 and 1, reflecting lack of stability. Essentially, as we introduce additional classes, there is an increase in the number of fractional entries in the consensus matrix since they are inherently unstable.

The inspection of these curves properly reveals 9 endmembers in the target image, whereas VD only finds 8 (when

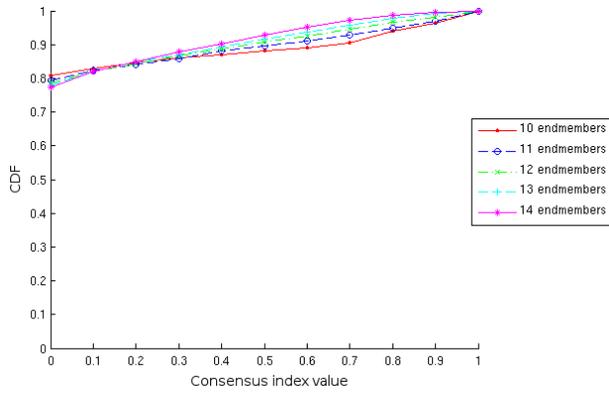


Fig. 4. Empirical CDFs corresponding to the entries of consensus matrices $C(p)$ for $p = 10..14$

PF was set to 0.1) or 7 endmembers (when the probability of false alarm was set to 0.01 or smaller). Similar results have been found with the same synthetic image introducing higher or smaller signal-to-noise ratios. VD underestimated the actual number of endmembers in these images and required noise to be present in the scenes in order to perform the estimation. As opposed to the VD concept, the proposed approach not only estimates the number of endmembers (even in synthetic scenes without any noise) but it also provides the actual endmembers along with their fractional abundance estimations in the scene. Although the results are encouraging, further experiments with additional synthetic and real scenes are required to substantiate the proposed approach in different application domains

5. CONCLUSIONS

The ultimate goal of this research is to devise a feasible and robust method that can set the stopping criteria – the number of endmembers – from the data itself without a priori knowledge about the underlying data distributions. While the popular VD method is quite efficient, it also suffers from several issues. First, it depends on an input parameter – false alarm probability – which has to be carefully set in advance in order to obtain satisfactory results. Furthermore, VD may not identify weak endmembers which correspond to anomalous endmembers.

In order to address these issues, we have developed a framework that exploits the sensibility of endmember extraction and spectral unmixing algorithms to the number of endmembers to estimate this parameter itself. Preliminary results on synthetic data reveal that the proposed scheme can compete with state-of-the-art schemes but its computational cost is very high. Fortunately, this analysis is inherently parallel and we could take advantage of the recent progress in commodity high-performance hardware to mitigate this problem. We plan to investigate the effectiveness of the

methodology in large studies using real hyperspectral data.

6. ACKNOWLEDGEMENT

This work has been supported by the European Community’s Marie Curie Research Training Networks Programme under reference MRTN-CT-2006-035927 and the Spanish government through the research contracts TIN2008-005089, Ingenio 2010 Consolider ESP00C-07-20811 and AYA2008-05965-C04-02.

7. REFERENCES

- [1] A. Plaza, P. Martinez, R. Perez, and J. Plaza, “A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data,” *IEEE TGRS*, vol. 42, no. 3, pp. 650–663, 2004.
- [2] A. Plaza and C.-I. Chang, “Impact of initialization on design of endmember extraction algorithms,” *IEEE TGRS*, vol. 44, no. 11, pp. 3397–3407, Nov. 2006.
- [3] J. Plaza, A. Plaza, R. Perez, and P. Martinez, “On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images,” *Pattern Recognition*, in press, 2009.
- [4] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, “End-member extraction for hyperspectral image analysis,” *Applied Optics*, vol. 47, no. 28, pp. F77–F84, 2008.
- [5] L. Miao and H. Qi, “Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization,” *IEEE TGRS*, vol. 45, no. 3, pp. 765–777, March 2007.
- [6] C.-I Chang and Q. Du, “Estimation of number of spectrally distinct signal sources in hyperspectral imagery,” *IEEE TGRS*, vol. 42, no. 3, pp. 608–619, March 2004.
- [7] C.-C. Wu and C.-I Chang, “Does an endmember set really yield maximum simplex volume?,” *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2007.*, pp. 3814–3816, July 2007.
- [8] J.M.P. Nascimento and J.M.B. Dias, “Vertex component analysis: a fast algorithm to unmix hyperspectral data,” *IEEE TGRS*, vol. 43, no. 4, pp. 898–910, April 2005.
- [9] S. Monti, P. Tamayo, J. Mesirov, and T. Golub, “Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data,” *Machine Learning*, vol. 52(1-2), 2003.
- [10] J. P. Brunet, P. Tamayo, T. R. Golub, and J. P. Mesirov, “Meta-genes and molecular pattern discovery using matrix factorization,” *Proceedings National Academic Science*, vol. 101, no. 12, pp. 4164–4169, March 2004.
- [11] M. E. Winter, “N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data,” *Proc. SPIE Image Spectrometry V*, vol. 3753, pp. 266–277, 2003.
- [12] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, Springer, 2006.