SPATIAL-SPECTRAL ENDMEMBER EXTRACTION FROM REMOTELY SENSED HYPERSPECTRAL IMAGES USING THE WATERSHED TRANSFORM

Maciel Zortea

Department of Mathematics and Statistics
University of Tromsø
N-9037 Tromsø, Norway
E-mail: maciel.zortea@hyperinet.eu

Antonio Plaza

Department of Computer Science
University of Extremadura
E-10071 Cáceres, Spain
E-mail: aplaza@unex.es

1. INTRODUCTION

Spectral unmixing is an important task in remotely sensed hyperspectral data exploitation [1]. When the spatial resolution of the sensor is not fine enough to separate different spectral constituents, these can jointly occupy a single pixel and the resulting spectral measurement will be a *mixed* pixel, i.e., a composite of the individual pure spectra (called *endmembers* in the literature [2]). In order to define the mixture problem in mathematical terms, let us assume that a remotely sensed hyperspectral scene with n bands is denoted by $\mathbf{X}$, in which the pixel at the discrete, spatial coordinates (i, j) of the scene is represented by a feature vector given by $\mathbf{X}(i, j) = [x_1(i, j), x_2(i, j), \cdots, x_n(i, j)] \in \Re^n$, and $\Re$ denotes the set of real numbers corresponding to the pixel's spectral response $x_k(i, j)$ at sensor channels $k = 1, \ldots, n$. Under a linear mixture model assumption [1], 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 the z-th endmember, $\Phi_z(i, j)$ is a scalar value designating the fractional abundance of the z-th endmember at pixel $\mathbf{X}(i, j)$, p is the total number of endmembers, and $\mathbf{n}(i, j)$ is a noise vector. Two physical constrains are generally imposed into the model described in (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]. The solution of the linear spectral mixture problem described in (1) relies on two fundamental issues: the successful estimation of how many endmembers, p, are present in the input hyperspectral scene $\mathbf{X}$, and 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 $\mathbf{X}(i, j)$.

Over the last decade, several algorithms have been developed for extraction of spectral endmembers directly from the input hyperspectral data set [4]. However, most of them have been focused on analyzing the data without incorporating information on the spatially adjacent data, i.e., the hyperspectral data is not really treated as an image but as an unordered listing of spectral measurements with no spatial arrangement. In turn, one of the distinguishing properties of hyperspectral data is the multivariate information coupled with a two-dimensional (pictorial) representation amenable to image interpretation. Subsequently, the endmember extraction process could benefit from an integrated framework in which not only the spectral information, but also the spatial arrangement of pixel vectors were taken into account in the search for image-derived endmembers.

A few attempts exist in the literature aimed at including the spatial information in the process of extracting spectral endmembers. Approaches include morphological endmember extraction [5], which integrates spatial and spectral information, or spatial preprocessing using a sliding-window approach [6]. These approaches are sensitive to the shape and size of the morphological structuring element (or spatial window) used as an input parameter to define a spatial context around each pixel vector in the scene. Instead of relying on a structuring element or window of fixed size, the watershed transform [7] considers the image data as imaginary topographic relief; the brighter the intensity, the higher the corresponding elevation. Let us assume that a drop of water falls on such a topographic surface. The drop will flow down along the steepest slope path until it reaches a minimum. The set of points of the surface whose steepest slope path reaches a given minimum constitutes the *catchment basin* associated with that minimum, while the *watershed lines* are the zones dividing adjacent catchment basins.

In this paper, we investigate the use of the watershed transform for integrating spatial and spectral information in the process of endmember extraction. The proposed approach is presented as a preprocessing module which can be used in combination

with available endmember extraction algorithms. The method takes advantage of a particular characteristic of remotely sensed scenes, in which spatially adjacent pixel vectors are expected to share similar information, from a spectral point of view. Thus, a possible approach to take advantage of such knowledge when searching for image endmembers is to exploit the spatial similarity between adjacent pixels by defining a criterion which is sensitive to the nature of both *homogenous* and *transition* areas between different land-cover classes. Intuitively, the transition areas between two or more different land-cover types would likely contain some mixed pixels. Thus, it would be reasonable to assume that pure pixels are less likely to be found in such transition areas. Our assumption in this work is that the watershed transform can be used to intelligently guide the endmember searching process to spatially homogeneous areas located at the local minima of the *catchment basins* and far away from *watershed lines* that define the transition areas between different regions, expected to contain mixed pixels. In a similar fashion, local maxima are also included in the analysis by conveniently manipulating the component being analyzed by the watershed transform.

2. USE OF WATERSHED TRANSFORM FOR ENDMEMBER CANDIDATE SELECTION

The proposed approach for spatial preprocessing prior to endmember extraction consists of the following steps:

1. *Feature reduction*. Apply a dimensionality reduction transformation such as principal component analysis (PCA) [8], singular value decomposition (SVD) [9] or a recently developed Hysime algorithm [10] to reduce the dimensionality of the data from n to $p-1$. In the case of using PCA or SVD for feature reduction, the value of p is estimated using the virtual dimensionality (VD) concept [11]. When Hysime is used, the value of p is automatically set by this method.

2. *Noise reduction*. Prior to the application of the watershed transform, which is sensitive to noise and subject to over-segmentation, we perform a kernel-based smoothing of the transformed image after feature reduction by using a Gaussian kernel in which the width of the Gaussian, σ, is used as an additional parameter to our proposed framework.

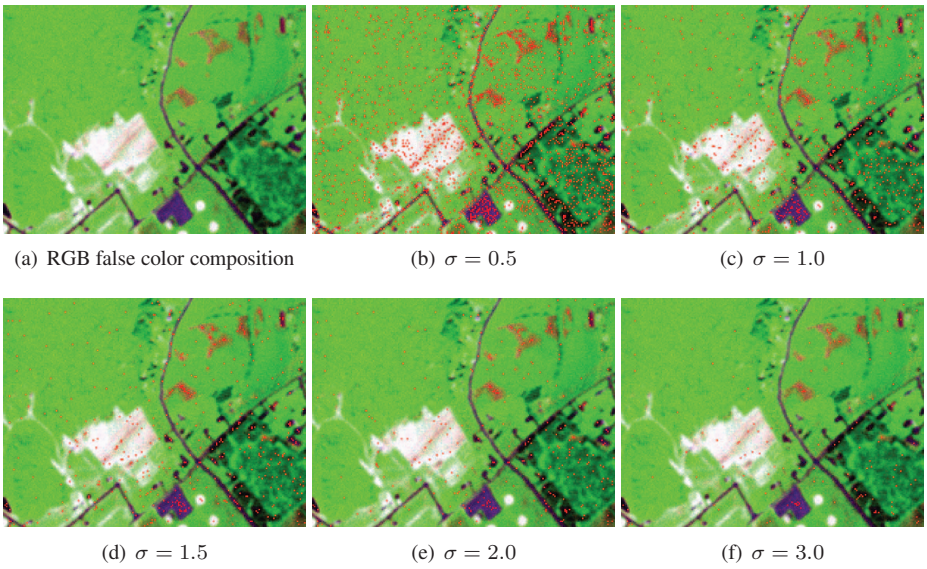
3. *Watershed transform preprocessing*. Assuming the presence of pure pixels in the image, each of the p components retained after the feature reduction transformation in the previous step is now treated as a topographic surface. The watershed transform is used to segment each component in *catchment basins* and *watershed lines*. In the context of our endmember extraction problem, pixels that are spatially purer than its neighbors are expected to be located at the minima of the catchment basins. To avoid losing those points corresponding to a local maxima in each component, also expected to contain purer pixels than the neighborhood, each single component under analysis is multiplied by minus one and the watershed is applied again. Thus, original local maxima are conveniently transformed to local minima, and the candidates can be conveniently located again using the catchment basins provided by the watershed transform. To avoid including many redundant candidate pixels (local minima or maxima), only a small percentile of the candidates from each component are considered. In our current experiments, the threshold is set to 5% of the value corresponding to the global minima or maxima of the component being analyzed. Therefore, a relevant subset of pixels located at the catchment basins at each of the p components after the feature reduction transformation are included in the candidate endmember set and fed to the next stage.

4. *Endmember extraction*. Finally, a spectral-based endmember extraction algorithm is applied to the set of endmember candidates obtained after watershed preprocessing to select a final set of p endmembers $\{\mathbf{E}_z\}_{z=1}^{p}$. Possible algorithms include the N-FINDR algorithm in [12], the VCA algorithm in [13], or the OSP algorithm in [14], among others [4].

3. EXPERIMENTAL RESULTS

We have conducted a preliminary validation of the proposed approach using a database consisting of 28 simulated hyperspectral data sets (all with 156×207 pixels in size), obtained through manipulation of real hyperspectral images acquired by the Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) over a mixed scenario including agricultural, vegetation, and urban areas[1]. This scene has spatial resolution of about 4 meters per pixel and 224 spectral bands. Our simulated scenes are created by initially applying an averaging filter of size 3 pixels to reduce noise and increase the mixing of the pixels in the original input AVIRIS scene, simulating a new but similar scene acquired with lower spatial resolution. In the sequence, the number of endmembers are estimated and the endmembers are extracted from the manipulated scene using a spectral-based algorithm. Pairs of endmembers with spectral angle lower than a pre-defined threshold were merged to increase the representativeness of the retained final set of endmembers. In the sequence, the ground truth abundance fractions where generated by applying the fully constrained linear

[1] Available online: http://aviris.jpl.nasa.gov/html/aviris.freedata.html, the scene labeled as f960705t01p02_r05



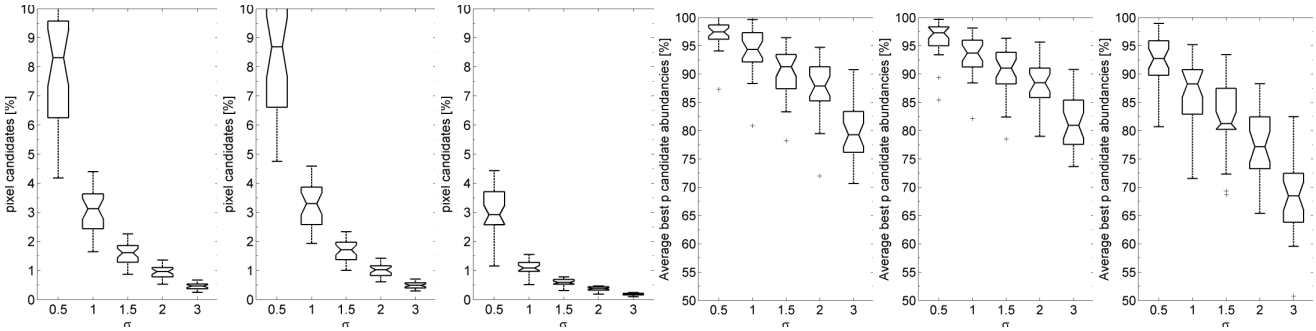
(a) RGB false color composition (b) $\sigma = 0.5$ (c) $\sigma = 1.0$

(d) $\sigma = 1.5$ (e) $\sigma = 2.0$ (f) $\sigma = 3.0$

Fig. 1. (a) False color composition of one of the simulated scenes. (b–f) Candidate endmembers (marked as red dots) extracted by our proposed method for different values of σ.

spectral unmixing (FCLSU) algorithm in [3] to the manipulated image and the retained set of endmembers (ASC and ANC constraints are imposed). Then, we use the linear mixture model in Eq. (1) to construct the final image from the retained endmembers and the ground-truth fractional abundances. In the reconstructed images, created by assuming that at least one pure pixel per source is available in the image data, the endmembers and fractional abundances are known and can be used as ground-truth to evaluate the performance of our proposed algorithm. Random noise was added (in different proportions) to the scenes above to simulate contributions from ambient (clutter) and instrumental sources.

Fig. 1 shows one of the simulated data sets and the candidate endmembers (marked in red color) obtained after applying steps 1–3 of the methodology described in section 2, using different values of σ and the PCA for feature reduction. The signal-to-noise ratio (SNR) in the simulated scene is 30 db. As shown by Fig. 1, the spatial representativeness of endmember candidates increases with the value of σ since the endmembers are located at the minima of progressively larger *catchment basins*, but there is also a chance that good endmember candidates may be lost as the value of σ is increased. In order to substantiate this issue, Figs. 2(a-c) show the box and whisker plots summarizing the percentage of endmember candidates retained for different values of σ for different feature reduction methods (PCA, SVD and Hysime) applied to the entire set of 28 simulated scenes with SNR $= 30$dB. For values of $\sigma \geq 1$, the percentage of endmember candidates retained was below 4%, which represents a substantial reduction of the number of pixels to be analyzed by a given endmember extraction algorithm. However, despite the fact that reducing the number of candidate endmembers may be perceived as a beneficial feature in terms of computational complexity and spatial representativeness of such endmember candidates, it is also very important to substantiate if the candidates retained after watershed transform can be still successfully used for spectral unmixing purposes (i.e., to analyze the degree of spectral purity of each endmember candidate). In order to evaluate this important aspect, we analyzed the spectral purity of retained candidates by measuring the fractional abundance of the predominant endmmeber in each candidate. This analysis is possible in the controlled scenario given by simulated hyperspectral scenes that we have adopted for experiments in this paper. For this purpose, Figs. 2(d-f) show the box and whisker plots summarizing the average abundance of the predominant endmember in the candidate pixels retained after the watershed preprocessing, as a function of the value of σ, for the 28 simulated hyperspectral scenes with SNR $= 30$dB. As with the percentage of selected candidates, we can observe in Figs. 2(d-f) that the value of σ plays a crucial role in the spectral purity of retained endmember candidates. On the other hand, the performance of the three considered dimensionality reduction techniques (PCA, SVD and Hysime) is similar in terms of the availability of spectrally pure candidates after watershed preprocessing. However, Figs. 2(d-f) also reveal that the spectral purity of the candidates decreases with increasing values of σ, although the spatial relevance of such endmembers may increase. In order to substantiate these remarks, further experiments on endmember extraction and fractional abundance estimation using the retained endmember candidates will be presented with the final version of this paper. Specifically, we will resort to the N-FINDR, VCA and OSP algorithms for endmember extraction, and to the FCLSU algorithm for abundance



(a) PCA (b) SVD (c) Hysime (d) PCA (e) SVD (f) Hysime

Fig. 2. (a-c) Percentage of endmember candidates retained by the proposed methodology. (d-f) Average abundance of the predominant endmember in the candidates retained by the proposed methodology. In both cases, different feature extraction techniques (PCA, SVD and Hysime) are used.

estimation, relating the results with the ground-truth information available for the simulated scenes. Experiments using real hyperspectral scenes will be also discussed in the final version of the paper.

4. REFERENCES

[1] 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*, vol. 91, pp. 8098–8112, 1986.

[2] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.

[3] 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*, vol. 39, pp. 529–545, 2000.

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

[5] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 9, pp. 2025–2041, 2002.

[6] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, pp. 2679–2693, 2009.

[7] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer-Verlag, Secaucus, NJ, 2003.

[8] R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing, 2nd ed.*, Academic Press: NY, 1997.

[9] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, Springer, 2006.

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

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

[12] 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.

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

[14] J. C. Harsanyi and C.-I Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 4, pp. 779–785.