An unsupervised segmentation approach for hyperspectral images

N. Gorretta,[a] J.M. Roger,[a] G. Rabatel,[a] C. Fiorio,[b] C. Lelong[c] and V. Bellon-Maurel[a]

[a]*Cemagref, UMR ITAP 361, avenue J-F. Breton 34033 Montpellier Cedex 5, France*
[b]*LIRMM, UMR 5506, 161 rue Ada 34392 Montpellier Cedex 5, France*
[c]*CIRAD, UMR TETIS 500, rue Jean-François Breton 34093 Montpellier Cedex 5, France*

Introduction

Most of the methods devoted to hyperspectral imaging processing conduct data analysis without taking into account spatial information. Pixels are processed individually, as an array of spectral data without any spatial structure. An intensive literature is available on pixel-level processing approaches, i.e. classifications techniques, where each pixel is assigned to one class, based on its spectral value[1] However, with the development of high resolution sensors, coupling spectral and spatial information when processing complex images appears to be a relevant approach. The purpose of this article is to propose a generic processing method that strikes a balance in the treatment of both kinds of information. This method, called Butterfly, aims at performing an iterative cross analysis of data in the spectral and the spatial domains, leading to segmentation of a hyperspectral image.

The outline of this paper is as follows. First, theorical aspects of the butterfly approach are presented. Second, an implementation scheme of this generic method using specific notions of topology, i.e. connectivity and adjancy is described. Experimental results are then presented. Finally, conclusions are drawn.

The butterfly approach

Generic scheme

In computer vision, segmentation refers to the process of partitioning an image into multiple segments, or sets of pixels sharing some properties (homogeneity with respect to some criterion of interest e.g. intensity or texture) and corresponding to individual surfaces or objects. Indeed, given a hyperspectral image **Z**, the purpose of a segmentation approach is to obtain a partition P including k regions Ri, with $i = 1...k$. However, due to the Hugues phenomena and properties of high dimensional space, it is not possible to achieve a segmentation directly on a hyperspectral image **Z** by using classical image analysis tools. To face these difficulties, various strategies have been attempted.[2–4] [Benediktsson, Palmason *et al*. (2005); Fauvel (2007); Plaza, Benediktsson *et* *al*. (2009)]. However, these methods tend to favour only one description (spectral or spatial) either directly or indirectly.

The Butterfly approach proposes to strike a balance in the treatment of both kinds of information by identifying a spatial and a spectral structure in the spatial and the spectral space respectively and by using an iterative process to provide collaboration between each space via extracted structures. Thus, one "butterfly" round is made up of two steps:

1. Extraction of a spatial structure (topology) based on a spectral structure,
2. Extraction of a spectral structure (latent variables) based on a spatial structure.

The first step deals with the use of commonly-used image processing tools (segmentation algorithms) on a limited number of score images. To carry out the second step, chemometric tools are employed to reveal a subspace which enables us to characterise the data according to topological notions used.

Specialisation and implementation

The scheme proposed above is a generic one. To apply it, we have to focus on specific spatial and spectral structures and to define a cooperation scheme between them. Here, we will focus solely on particular notions of topology for the spatial space i.e. the notions of connectivity and adjacency and the building of latent variables for the spectral space. To ensure collaboration, concepts of regions and false colours or score images (spatial space) are transferred to concepts of classes and latent variables (spectral space) respectively. The above is summarised in Figure 1.

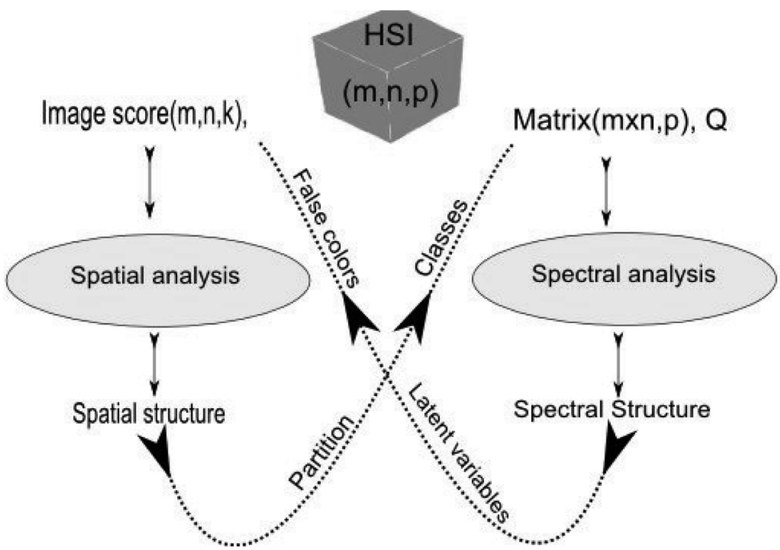


Figure 1. The butterfly approach: an unsupervised segmentation strategy.

To apply the Butterfly approach using these particular notions of topology, we propose to use a top down segmentation scheme, i.e. a split strategy named Normalised Cut approach (Shi and Malik 2000).[5] According to the theoretical aspect of the butterfly approach and to deal with this segmentation scheme, chemometric tools must reveal latent variables which are able to describe the within inertia (**W**) of the data. A simple and efficient way to do this is to diagonalise the within variance matrix of the data and to use the first eigen vectors as latent variables.

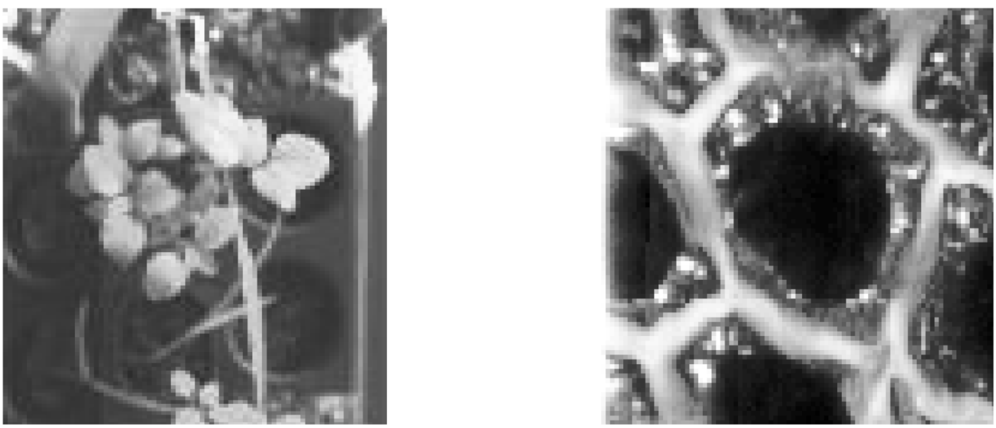
Results and discussion

Hyperspectral image data set

Two different sets were used for the experiments, with different contexts and characteristics [Figures 2(a) and (b)]:

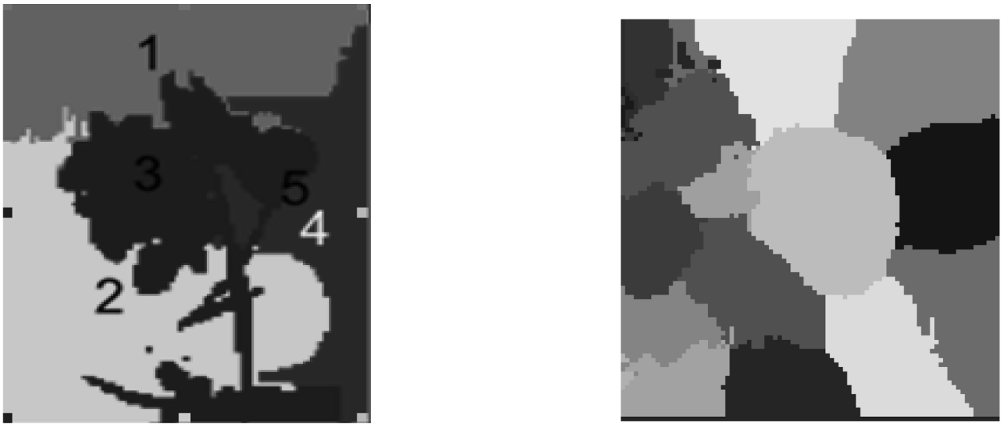
- The first one is a proxi-detection image of a vegetation scene that was acquired in field with a Hyspex VNIR-1600 hyperspectra system (Norsk Elecktro Optikk). The image is 120×150 pixels with a spatial resolution of 0.020 mm^2 per pixel. The number of data channels is 160 (spectral band 400–980 nm) with a spectral resolution of 3.6 nm. The image contains two species of weed [referenced 1 and 2 on Figure 2(b)].
- The second one is an IR-Raman hyperspectral image of a wheat grain endosperm. The image is 89×89 pixels with a spatial resolution of few micrometers per pixel. The number of data channels is 460 (spectral band 860–1748 nm) with a spectral resolution of 2 nm.

Segmentation of these images was performed with the butterfly approach using a split strategy made with 5 and 20 rounds for the Hyspex and the Raman image respectively. Each split round



(a) Hyspex VNIR-1600 image (b) Raman Image

Figure 2. Hyperspectral images data sets–gray scale representation (equalised image).



(a) Hyspex VNIR-1600 image (b) Raman Image

Figure 3. Segmentation results.

was conducted using adapted parameters for the Normalised Cut algorithm and one latent variable (first eigenvector of the within variance matrix of the data).

Figure 3(a) and 3(b) show the unsupervised segmentation map obtained for the Hyspex and the Raman image respectively.

In these figures, different gray colors correspond to different regions. Indeed, based on a visual inspection, in both images, the main spatial structures are well determined: on the Hyspex image, *background and weed* species are well segmented; In the same manner, on the Raman image, butterfly segmentation allows the characterisation of cell organisation in wheat endosperm.

Conclusion

A generic approach to segment hyperspectral images has been proposed. This method, called butterfly, is based on an iterative process and a cross analysis of spectral and spatial structures. To demonstrate the relevance of this layout, we have focused on specific notions of topology (spatial structures). Then, an implementation using a top down approach (split), i.e. the Normalised cut algorithm, has been tested on two hyperspectral data images. Results obtained are promising. However, they must be confirmed by additional testing on other hyperspectral images and with ground truth for a quantitative evaluation.

Acknowledgement

The authors would like to thank Cécile Barron from INRA for providing the Raman hyperspectral data set.

References

1. Y.A. Tarabalka, A. Benediktsson and J. Chanussot, *IEEE Trans. Geosci. Remote Sens.* **47,** 2973 (2009).
2. J.A. Benediktsson, J.A. Palmason and J.R. Sveinsson, *IEEE Trans. Geosci. Remote Sens.* **43,** 480 (2005).
3. M. Fauvel, *Spectral and Spatial Methods for the Classification of Urban Remote Sensing Data*. Grenoble Institute of Technology and Faculty of Engineering, University of Iceland, p. 180 (2007).
4. A. Plaza, A. Benediktsson, J.W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J.C. Tilton and G. Trianni "Recent advances in techniques for hyperspectral image processing", *Remote Sens. Environ.*, in Press, Corrected Proof (2009).
5. J. Shi and J. Malik, *IEEE Trans. Pattern Anal.* **22,** 888 (2000).