HAL open science

Morphological segmentation of hyperspectral images

Guillaume Noyel, Jesus Angulo, Dominique Jeulin

To cite this version:

Guillaume Noyel, Jesus Angulo, Dominique Jeulin. Morphological segmentation of hyperspectral images. Image Analysis & Stereology, 2007, 26 (3), pp.101-109. 10.5566/ias.v26.p101-109 . hal-01220414v1

HAL Id: hal-01220414

https://hal.science/hal-01220414v1

Submitted on 18 Dec 2015 (v1), last revised 1 Oct 2020 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

CENTRE DE MORPHOLOGIE MATHEMATIQUE
35, rue Saint-Honoré, 77305 FONTAINEBLEAU (FRANCE)

Morphological segmentation of hyperspectral images

Guillaume Noyel, Jesus Angulo,
Dominique Jeulin

N-36/06/MM

Accepted in the journal Image Analysis and Stereology
Published in Proceedings of
International Congress of Stereology
ICS XII
Saint-Etienne – France - Aug. 30-Sept. 7, 2007

Décembre 2006

MINES PARIS

Submitted to Image Anal Stereol, 9 pages
Original Research Paper

MORPHOLOGICAL SEGMENTATION OF HYPERSPECTRAL IMAGES

GUILLAUME NOYEL, JESÚS ANGULO AND DOMINIQUE JEULIN

Centre de Morphologie Mathématique, Ecole des Mines de Paris, 35 rue Saint-Honoré, Fontainebleau, F - 77305, France
e-mail: {guillaume.noyel, jesus.angulo, dominique.jeulin}@ensmp.fr
(Submitted)

ABSTRACT

The present paper develops a general methodology for the morphological segmentation of hyperspectral images, i.e. with an important number of channels. This approach, based on watershed, is composed of a spectral classification to obtain the markers and a vectorial gradient which gives the spatial information. Several alternative gradients are adapted to the different hyperspectral functions. Data reduction is performed either by Factor Analysis or by model fitting. Image segmentation is done on different spaces: factor space, parameters space, etc. On all these spaces the spatial/spectral segmentation approach is applied, leading to relevant results on the image.

Keywords: factor analysis, hyperspectral imagery, mathematical morphology, watershed segmentation.

INTRODUCTION

Hyperspectral images are multivariate discrete functions with typically several tens or hundreds of spectral bands. In a formal way, each pixel of an hyperspectral image is a vector with values in wavelength, in time, or associated with any index j. To each wavelength, time or index corresponds an image in two dimensions called channel. In the sequel we use only the term of spectrum and spectral channel. The number of channels depends on the nature of the specific problem under study (satellite imaging, spectroscopic images, temporal series, etc.). Let

$$\mathbf{f}_\lambda : \left\{ \begin{array}{lcl} E & \rightarrow & \mathscr{T}^L \\ x & \rightarrow & \mathbf{f}_\lambda(x) = \left(f_{\lambda_1}(x), f_{\lambda_2}(x), \ldots, f_{\lambda_L}(x)\right) \end{array} \right. \quad (1)$$

be an hyperspectral image, where:

- $E \subset \mathbb{R}^2$, $\mathscr{T} \subset \mathbb{R}$ and $\mathscr{T}^L = \mathscr{T} \times \mathscr{T} \times \ldots \times \mathscr{T}$
- $x = x_i \setminus i \in \{1, 2, \ldots, P\}$ is the spatial coordinates of a vector pixel $\mathbf{f}_\lambda(x_i)$ (P is the pixels number of E)
- $f_{\lambda_j} \setminus j \in \{1, 2, \ldots, L\}$ is a channel (L is the channels number)
- $f_{\lambda_j}(x_i)$ is the value of vector pixel $\mathbf{f}_\lambda(x_i)$ on channel f_{λ_j}.

In this paper, we introduce a general methodology for hyperspectral image segmentation, using a watershed based morphological approach.

Watershed segmentation requires a gradient (i.e. a scalar function) and markers on the target structures to obtain a correct image segmentation (Beucher and Meyer, 1992). A gradient on a multivariate function can be obtained in different ways. One way is to calculate on each image channel a modulus of a gradient, and to take the sum or the supremum of the gradients (Meyer, 1992; Angulo and Serra, 2003). Another way is to use vectorial gradients based on distance between vector pixels (Angulo and Serra, 2003; Evans and Liu, 2006). We consider here various alternatives for hyperspectral images.

Moreover, when dealing with hyperspectral images, the large number of channels generates data redundancies. Consequently, it is necessary to reduce the amount of data, to extract pertinent information. To do this, two ways are explored: a linear factor analysis and a model approach.

Previously in the literature, several approaches to multispectral image segmentation were explored. Flouzat *et al.* (1998) use a spatial and a spectral segmentation based on the filtering of the image adjacency graph. Paclík *et al.* (2003) obtain, with statistical classifiers, the pixels probabilities of membership to clusters for spectral domain and the pixels probabilities of membership to clusters for spatial domain. The pixels probabilities of membership to a cluster are obtained by multiplying both probabilities, because they assume independence between spatial and spectral information. The pixels are classified and the process is repeated until convergence. Li and Xiao (2004) compute on each channel a morphological multiscale gradient by summation of morphological gradients with increasing size structuring elements. As watershed requires a scalar gradient (i.e. one channel), the channels gradients are summed with a weight equal to one. After morphological filtering on the gradient, the

local minima are used as markers for watershed segmentation. Scheunders (2001) computes a gradient by summation of channels gradients followed by a watershed segmentation. Soille (1996) combines spectral classification on histograms and spatial segmentation. The multidimensional histogram is segmented, using the watershed algorithm, to obtain a classified image. On the classified image the minima of the gradient of the hyperspectral image are imposed. With the gradient and the markers he applies the watershed segmentation.

The methodology and the different alternatives studied in the current paper are illustrated by means of a 60 channels image acquired by active thermography on plastic lids. The size of the image is 256×256 pixels $\times$ 60 channels. The aim is to segment glue occlusions within plastic lids. This sequence of images comes from Laboratoire Le2i, Le Creusot, France. Legrand *et al.* (2002) explain how the image was acquired. They also present a segmentation that was done on a channel, using difference of lid images with and without glue, thresholding and filtering. Their segmentation is used as a reference for comparison (fig. 1). In our case, only the sequence with glue occlusions is available. That's why we cannot use image difference or any kind of calibration with a reference image.

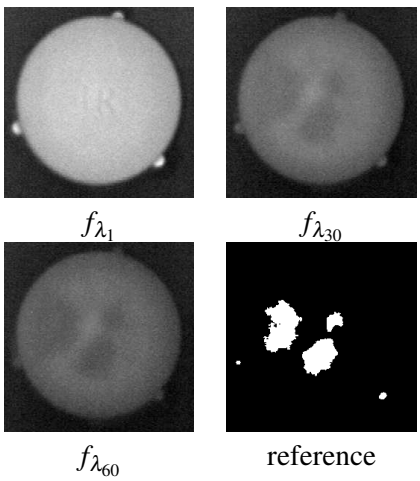


Fig. 1. *Three channels of the original image and the reference obtained by Legrand et al. (2002) method.*

FRAMEWORK FOR MORPHOLOGICAL SEGMENTATION ON MULTIVARIATE DATA

The watershed transformation is one of the most powerful tools for segmenting images. According to a flooding paradigm, the watershed lines associate a catchment basin to each minimum of the function (Beucher and Meyer, 1992). Typically, the function to flood is a gradient function which defines the transitions between the regions. Using the watershed on a scalar image without any preparation leads to a strong over-segmentation (large number of minima). There are two alternatives in order to get rid of the over-segmentation. The first one consists in initially determining markers for each region of interest: using the homotopy modification, the local minima of the gradient function are only the region markers. A difficult issue is to determine the markers, especially for generic images. The second alternative involves hierarchical approaches based on non-parametric merging of catchment basins (waterfall algorithm) or based on the selection of the most significant minima according to different criteria (dynamics, area or volume extinction values) (Meyer, 2001).

In this paper, we focus on markers based segmentation. In fact, we consider that the hyperspectral images have enough information to define markers from a spectral classification.

Multivariate gradients

A gradient image, in fact the norm, is a scalar function with values in the reduced interval $[0,1]$, i.e. $\nabla : E \rightarrow [0,1]$. This normalization is always applied to multivariate gradients given below in order to have homogeneous gradient functions. To define a gradient, four approaches are considered: a morphological gradient on one channel, a metric-based gradient on all channels, a gradient defined as the supremum, or as the sum, of morphological gradients on each channel. The morphological gradient is a marginal gradient (i.e. it can only be applied on scalar images) defined as the difference between channel dilation and erosion with a unit structuring element B:

$$g(f_{\lambda_j}(x)) = \delta_B(f_{\lambda_j}(x)) - \varepsilon_B(f_{\lambda_j}(x)) \qquad (2)$$

The gradient supremum of morphological gradients on each channel is a vectorial gradient defined as:

$$\nabla_{\vee}\mathbf{f}_{\lambda}(x) = \vee[g(f_{\lambda_j}(x)), j \in \{1,\ldots,L\}] \qquad (3)$$

Each morphological gradient must be normalized between $[0,1]$ before taking the supremum.

The gradient weighted sum of morphological gradients is given by:

$$\nabla_{+}\mathbf{f}_{\lambda}(x) = \sum_{j=1}^{L} w_{\lambda_j} g(f_{\lambda_j}(x)) \qquad (4)$$

where w_{λ_j} denotes the weight of the gradient of channel f_{λ_j}.

The metric-based gradient is a vectorial gradient defined as the difference between the supremum and the infimum of a defined distance on a unit neighbourhood $B(x)$:

$$\nabla_d \mathbf{f}_\lambda(x) = \vee[d(\mathbf{f}_\lambda(x),\mathbf{f}_\lambda(y)), y \in B(x)] - \wedge[d(\mathbf{f}_\lambda(x),\mathbf{f}_\lambda(y)), y \in B(x)] \quad . \tag{5}$$

Various metric distances are available for this gradient such as:

•Euclidean distance:

$$d_E(\mathbf{f}_\lambda(x),\mathbf{f}_\lambda(y)) = \sqrt{\sum_{j=1}^{L}(f_{\lambda_j}(x) - f_{\lambda_j}(y))^2} \tag{6}$$

•Mahalanobis distance:

$$d_M(\mathbf{f}_\lambda(x),\mathbf{f}_\lambda(y)) = \sqrt{(\mathbf{f}_\lambda(x)-\mathbf{f}_\lambda(y))^t \Sigma^{-1}(\mathbf{f}_\lambda(x)-\mathbf{f}_\lambda(y))} \tag{7}$$

where Σ is the covariance matrix between variables (channels) of $\mathbf{f}_\lambda$. If channels are uncorrelated, the covariance matrix is diagonal. The diagonal values are equal to channels variance $\sigma^2_{\lambda_j} \setminus j \in \{1,2,\ldots,L\}$. Therefore, the Mahalanobis distance becomes:

$$d_M(\mathbf{f}_\lambda(x),\mathbf{f}_\lambda(y)) = \sqrt{\sum_{j=1}^{L}\left(\frac{f_{\lambda_j}(x)-f_{\lambda_j}(y)}{\sigma_{\lambda_j}}\right)^2} \tag{8}$$

•chi-squared distance:

$$d_{\chi^2}(\mathbf{f}_\lambda(x_i),\mathbf{f}_\lambda(x_{i'})) = \sqrt{\sum_{j=1}^{L}\frac{N}{f_{.\lambda_j}}\left(\frac{f_{\lambda_j}(x_i)}{f_{x_i.}} - \frac{f_{\lambda_j}(x_{i'})}{f_{x_{i'}.}}\right)^2} \tag{9}$$

with $f_{.\lambda_j} = \sum_{i=1}^{P} f_{\lambda_j}(x_i)$, $f_{x_i.} = \sum_{j=1}^{J} f_{\lambda_j}(x_i)$ and $N = \sum_{j=1}^{L}\sum_{i=1}^{P} f_{\lambda_j}(x_i)$.

Markers by spectral clustering

The markers defining the targets are obtained with an unsupervised classification based on clustering. We have used in this study the clustering algorithm "Clara" (Kaufman and Rousseeuw, 1990). This is a similar way, but more robust than the "kmeans" classification, in order to cluster large numbers of points. Then, the pertinent clusters are selected to be the markers. To perform clustering, each channel is considered as a variable. Therefore each pixel has its own value on each variable.

DATA REDUCTION AND FILTERING USING FACTOR ANALYSIS

Due to the redundancy of channels, a data reduction is performed using Factor Correspondence Analysis (FCA) (Benzécri, 1973). We prefer an FCA in place of a Principal Component Analysis because image values are positive and the spectral channels can be considered as probability distributions. The metric used in FCA is the chi-squared normalized by channels weight. This metric is adapted to probability laws. FCA can be seen as a transformation from image space to factorial space (eq. 10). In factorial space the coordinates of the pixels vector on each factorial axis are called pixels factors. The pixels factors can be considered as an hyperspectral image whose channels correspond to factorial axes:

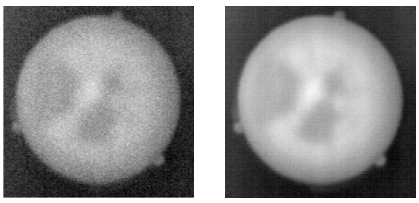
$$\zeta : \begin{cases} \mathscr{T}^L & \rightarrow \quad \mathscr{T}^K \,/\, K < L \\ \mathbf{f}_\lambda(x) & \rightarrow \quad \mathbf{c}^{\mathbf{f}}_\alpha(x) = \left(c^{\mathbf{f}}_{\alpha_1}(x),\ldots,c^{\mathbf{f}}_{\alpha_K}(x)\right) \end{cases} \tag{10}$$

A limited number K of factorial axes is usually chosen. Therefore FCA can be seen as a projection of initial pixels in a factor space with a smaller dimension.

The pseudo-inverse transform consists in reconstructing the images from factors (eq. 11). This is an approximation of the original image if one keeps a part of factorial axes for the reconstruction:

$$\widehat{\zeta}^{-1} : \begin{cases} \mathscr{T}^K & \rightarrow \quad \mathscr{T}^L \,/\, K < L \\ \mathbf{c}^{\mathbf{f}}_\alpha(x) & \rightarrow \quad \widehat{\mathbf{f}}_\lambda(x) = \left(\widehat{f}_{\lambda_1}(x),\ldots,\widehat{f}_{\lambda_L}(x)\right) \end{cases} \tag{11}$$

Besides, hyperspectral images usually contain noise due to the acquisition device, compression, etc. In this case it is possible to use FCA to filter images. As shown by Green *et al.* (1988), the noise is rejected on the last factorial axes whereas the signal remains on the first axes. By keeping only axes with sufficient signal and reconstructing the image, all channels are filtered. In our example, the image is very noisy ; consequently better results are obtained by filtering the whole sequence with a FCA, keeping the two first factorial axes (fig. 2). The eigenvalues associated with these two factors are $\mu_1 = 13.9\,10^{-4}$ and $\mu_2 = 3.15\,10^{-4}$. In other words, the first axis represent 13.8% of inertia and the second one 3.1%. Therefore, the inertia of the two first axes is equal to 16.9% which is small. In fact the next axes contain a lot of noise. This explains why their share in inertia is high. Therefore, they must be removed to get information without noise.

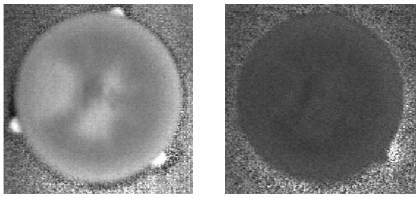


$f_{\lambda_{50}}$ $\widehat{f}_{\lambda_{50}}$

Fig. 2. *$f_{\lambda_{50}}$ and $\widehat{f}_{\lambda_{50}}$ are the channel 50 before and after filtering by FCA.*

SEGMENTATION OF $\widehat{\mathbf{f}}_\lambda(x)$, $\mathbf{c}^{\mathbf{f}}_\alpha(x)$

It is possible to generate a segmentation on the hyperspectral image composed of factorial axes $\mathbf{c}^{\mathbf{f}}_\alpha(x)$ (fig. 3) or on the filtered image $\widehat{\mathbf{f}}_\lambda(x)$.

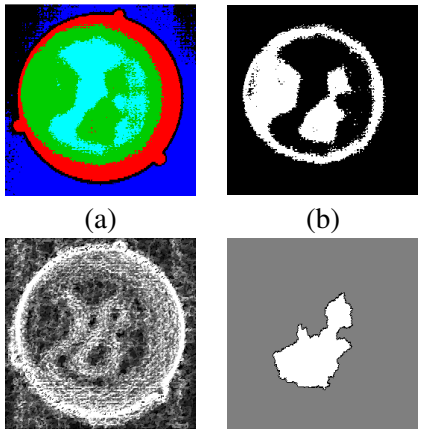


$c^{\mathbf{f}}_{\alpha_1}(x)$ $c^{\mathbf{f}}_{\alpha_2}(x)$

Fig. 3. *FCA factors on axes* 1 *and* 2 *of original image. For visualization, factors are scaled by a factor.*

In the factorial space made up of the two first axes, a classification by "Clara" is processed. Then the green cluster corresponding to glue is selected defining the markers (fig. 5). In order to regularize these markers, an opening with an hexagonal structuring element of size 5 is applied. As differences between glue and lid are small, a gaussian filter of size 11 followed by a morphological leveling are applied on each channel, to enhance the contours and to obtain a better gradient. The levelings are a subclass of symmetric connected filters (or filters by reconstruction) that suppress details but preserve the contours of the remaining objects (Meyer, 2004). The levelings need an image marker, a rough simplification of the reference image, to determine the structures to be leveled.

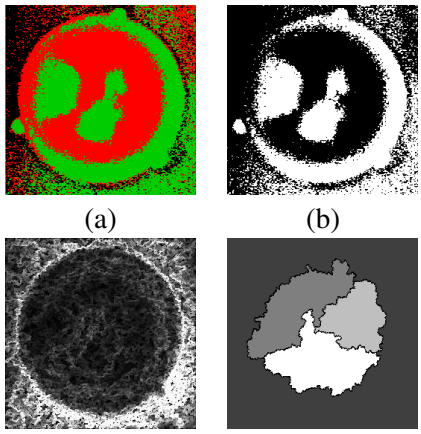
The Euclidean distance in FCA factorial space is equivalent to the chi-squared distance in image space (Benzécri, 1973). Therefore a chi-squared distance based gradient $\nabla_{\chi^2}\widehat{\mathbf{f}}_\lambda(x)$ is performed on the filtered image $\widehat{\mathbf{f}}_\lambda(x)$ and an Euclidean distance based gradient $\nabla_E\mathbf{c}^{\mathbf{f}}_\alpha(x)$ on the FCA factors $\mathbf{c}^{\mathbf{f}}_\alpha(x)$. Then the watershed segmentation is computed (fig. 4 and 5). Both segmentations, compared to the reference model (fig. 1), are not satisfactory for the present image.



(a) (b)

(c) (d)

Fig. 4. *Processes are computed on the filtered image $\widehat{\mathbf{f}}_\lambda(x)$. (a) Clara segmentation, (b) Green cluster corresponding to glue used as markers, (c) Chi-squared gradient, (d) Watershed segmentation. For visualization, gradients are scaled by a factor.*



(a) (b)

(c) (d)

Fig. 5. *Processes are computed on the factor pixels $\mathbf{c}^{\mathbf{f}}_\alpha(x)$. (a) Clara segmentation, (b) Green cluster corresponding to glue used as markers, (c) Euclidean gradient, (d) Watershed segmentation. For visualization, gradients are scaled by a factor.*

DATA REDUCTION USING MODEL APPROACH

Another way to reduce data is to fit a model on the spectrum of each pixels vector. The parameters of the model fitted in each pixel are seen as parametrical cartographies or maps. The whole maps generate an hyperspectral image $\mathbf{p}(x) = (p_1(x), \ldots, p_M(x))$. This approach is advantageous in the way it takes the order of channels into account. In fact, in FCA channels the order is without importance.

Moreover, with the hyperspectral image of parameters, a segmentation can be computed. It is also possible to make parameters orthogonal with a Principal Component Analysis and then generating a segmentation.

Principal Component Analysis (PCA) is used here instead of FCA, because it is possible to compute it on hyperspectral images with negative values, which can be the case for some parameters. PCA gives factorial axes with factors building another hyperspectral image: $\mathbf{c}^{\mathbf{p}}_{\beta}(x) = (c^{\mathbf{p}}_{\beta_1}(x), \ldots, c^{\mathbf{p}}_{\beta_M}(x))$.

For this thermographic sequence, a linear model is fitted on the image after removing the 10 first channels which correspond to a transitory phenomenon (fig. 6). The linear model $y = ax + b$ has two parameters, the slope a and the intercept b. On the first 10 channels we have defined another parameter, called rise m, as the maximum amplitude on these 10 channels:

$$m(x) = \max_{j\in[1:10]}(\widehat{f}_{\lambda_j}(x)) - \min_{j\in[1:10]}(\widehat{f}_{\lambda_j}(x))). \qquad (12)$$

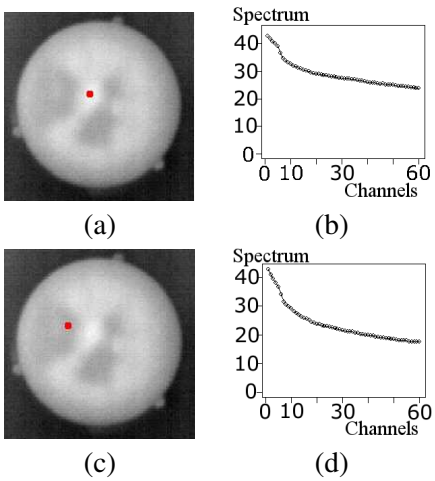


Fig. 6. *(a) Point outside glue where is measured spectrum (b), (c) point inside glue and corresponding spectrum (d).*

Then with the linear model we obtain images as maps of the parameters, which can be orthogonalised by PCA (fig 7). In this case the different axes represent the following inertia ratios 97.24%, 2.10% and 0.66%.

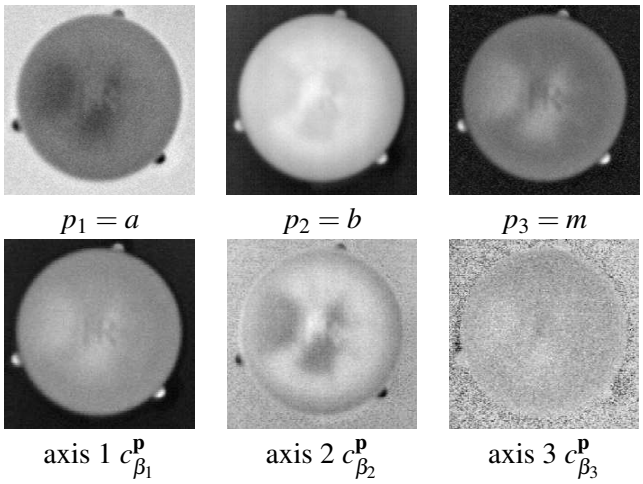


Fig. 7. Maps of the parameters of the linear model, PCA factors of the parameters on axis 1, 2 and 3..

SEGMENTATION OF $\mathbf{p}(x)$ OR $\mathbf{c}^{\mathbf{p}}_{\beta}(x)$

Different approaches of segmentation are tested on parameters $\mathbf{p}(x)$ or on parameters factors $\mathbf{c}^{p}_{\beta}(x)$. Markers are computed again by the cluster method "Clara" on the parameters and the image of factors . First, a clustering is processed on both images. The cluster corresponding to the lid center is selected because the glue is on the lid center. A second clustering "Clara" is made on the selected cluster. Non selected pixels are aggregated to the cluster of the largest size, because cluster corresponding to glue has the smallest size (fig. 8).

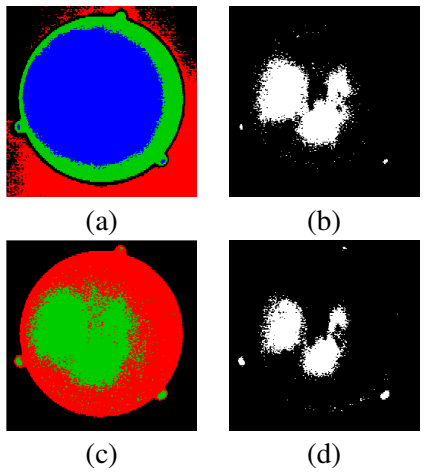
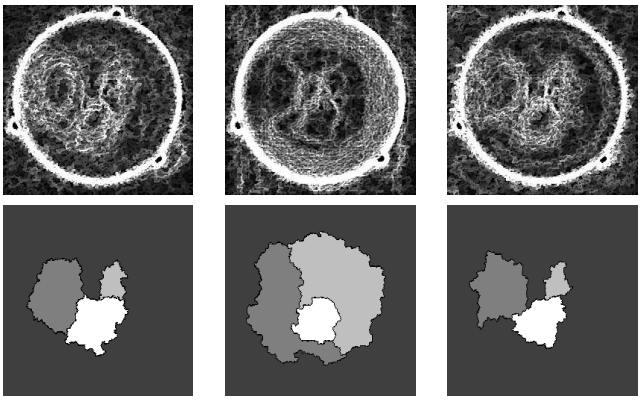


Fig. 8. *(a) Clara segmentation in 4 clusters on parameter b, blue cluster is selected, (b) Clara on parameters a and m: parameters markers, (c) Clara segmentation in 3 clusters on first axis $c^{\mathbf{p}}_{\beta_1}$, green cluster is selected, (d) Clara on axes 2 and 3 ($c^{\mathbf{p}}_{\beta_2}$, $c^{\mathbf{p}}_{\beta_3}$): parameters factors markers.*

To improve the quality of gradients, each image channel is again leveled using as image markers the corresponding channels, filtered by a gaussian kernel. Then, several gradients are tested. The morphological gradient on each channel g, the supremum and the sum of morphological gradients, $\nabla_\vee$ and ∇_+, are evaluated on the image of parameters $\mathbf{p}(x)$ and on the image of PCA factors parameters $\mathbf{c}^{\mathbf{p}}_{\beta}(x)$. Besides the Euclidean distance in PCA factorial space is equivalent to the Mahalanobis distance in parameters space. Therefore, the Mahalanobis distance based gradient $\nabla_M \mathbf{p}(x)$ is performed on the parameters $\mathbf{p}(x)$ to ensure the equal statistical value for the different parameters. For the PCA factors parameters $\mathbf{c}^{\mathbf{p}}_{\beta}(x)$, the most indicated is the Euclidean distance based gradient.

Watershed on parameters

The morphological gradient is computed on each channel of image $\mathbf{p}(x)$. Markers (fig. 8 (b)) are filtered by opening with an hexagonal structuring element of size 5. With the gradient and the markers, a watershed segmentation is obtained for each parameter.

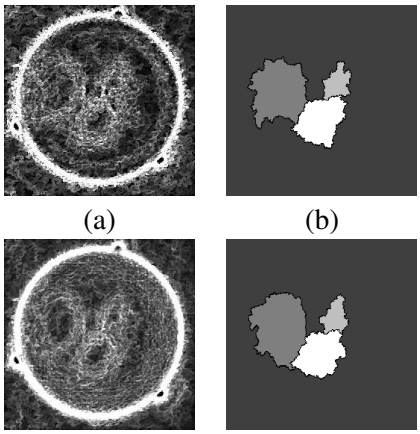


a *b* *m*

Fig. 9. *First line: morphological gradient, second line: watershed on parameters: slope a, intercept b and rise m. For visualization, gradients are scaled by a factor.*

Comparing to the reference, segmentations with morphological gradients on parameters slope a and rise m are good. However, the segmentation on intercept b is not satisfactory, due to the leaks on the gradient. Although these segmentations are pertinent, they are only marginal segmentations, i.e. only one parameter is taken into account in one segmentation. Therefore, we have tested the use of a vectorial gradient.

A metric-based gradient with the Mahalanobis distance $\nabla_M \mathbf{p}(x)$, and a gradient supremum of channels gradient $\nabla_\vee \mathbf{p}(x)$ are tested on the parameters image (fig. 10). In this case the gradient sum $\nabla_+ \mathbf{p}(x)$ is approximatively equivalent to the supremum gradient because it is not possible to define specific weights for the parameters. For each case, the same markers are used. Moreover, to compute the Mahalanobis distance, we assume that parameters are uncorrelated.



(a) (b)

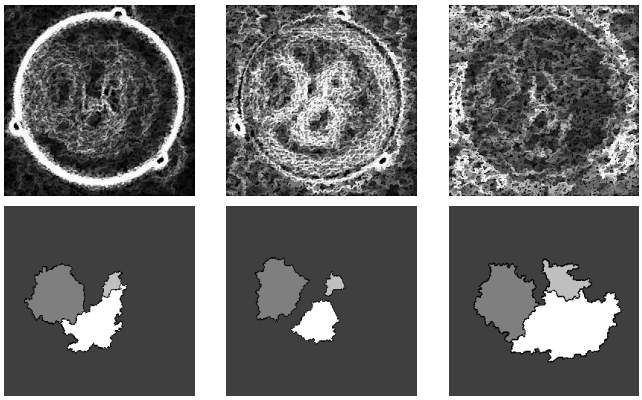
(c) (d)

Fig. 10. *Segmentations on parameters: (a) Mahalanobis distance based gradient, (b) associated watershed, (c) gradients supremum, (d) associated watershed. For visualization, gradients are scaled by a factor.*

Both segmentations from a vectorial gradient are good, compared to the reference. Consequently vectorial gradients are good in both cases.

Watershed on PCA axes of parameters

As for parameters, a morphological gradient is computed on each channel of $\mathbf{c}^{\mathbf{p}}_{\beta}(x)$. The markers (fig. 8 (d)) are again regularized with an opening followed by the corresponding watershed of each gradient.



axis 1 axis 2 axis 3

Fig. 11. *Morphological gradient and watershed on factors of PCA parameters: axis 1* $c^{\mathbf{p}}_{\beta_1}$*, axis 2* $c^{\mathbf{p}}_{\beta_2}$ *and axis 3* $c^{\mathbf{p}}_{\beta_3}$*. For visualization, gradients are scaled by a factor.*

The resulting segmentations on the parameters factors on axes 1, $c^{\mathbf{p}}_{\beta_1}$, and 2, $c^{\mathbf{p}}_{\beta_2}$, are good, as compared to the reference. In fact their gradients have distinct contours without leaks. However, it is not the case for morphological gradient on the parameters factors on axes 3, $c^{\mathbf{p}}_{\beta_3}$. Moreover, as for parameters, vectorial gradients are also tested.

First of all, an Euclidean gradient $\nabla_E \mathbf{c}^{\mathbf{p}}_{\beta}(x)$ is tested. Then the supremum of channels gradient $\nabla_{\vee} \mathbf{c}^{\mathbf{p}}_{\beta}(x)$ and the weighted sum gradient $\nabla_{+} \mathbf{c}^{\mathbf{p}}_{\beta}(x)$ are used (fig. 12). The weights for $\nabla_{+} \mathbf{c}^{\mathbf{p}}_{\beta}(x)$ are equal to the inertia contributions of axes: 0.97, 0.021 and 0.0066. For each case the same markers are used.

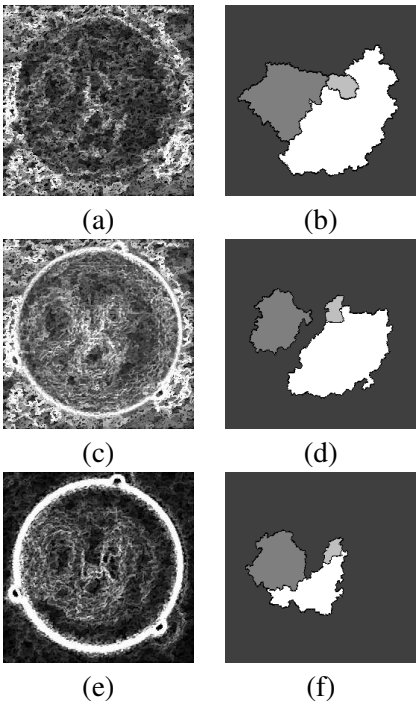


Fig. 12. *Segmentations on PCA parameters factors on the 3 axes: (a) Euclidean gradient, (b) associated watershed, (c) gradient supremum, (d) associated watershed, (e) weighted sum gradient, (f) associated watershed. For visualization, gradients are scaled by a factor.*

Both segmentations, with the gradient supremum and the Euclidean gradient, are not good, compared to the reference. This is due to the leaks on the gradients for the glue occlusions. In fact the morphological gradient on the factors on the third axis of PCA parameters is not relevant (fig. 11). However, the weighted sum gradient is much better because the weight the third axis is very low compared to the other ones. In this case the weighted sum gradient is more adapted.

Consequently, the Euclidean gradient and the supremum of morphological gradients are tested on the two first PCA parameters axes (fig. 13). In this case, the segmentations are much better. Therefore, the choice of axes for which the gradient is better must be emphasized. In fact, adding another noisy axis creates leaks on the gradient.

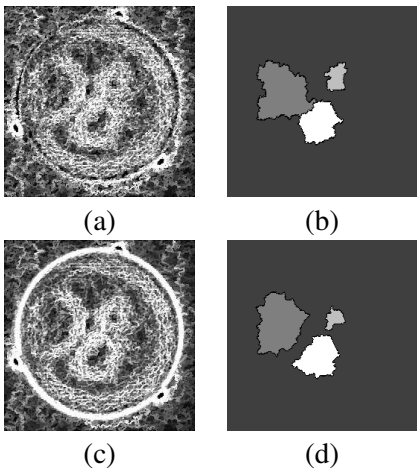


Fig. 13. *Segmentations on PCA parameters factors on axes 1 and 2: (a) Euclidean gradient, (b) associated watershed, (c) gradient supremum, (d) associated watershed. For visualization, gradients are scaled.*

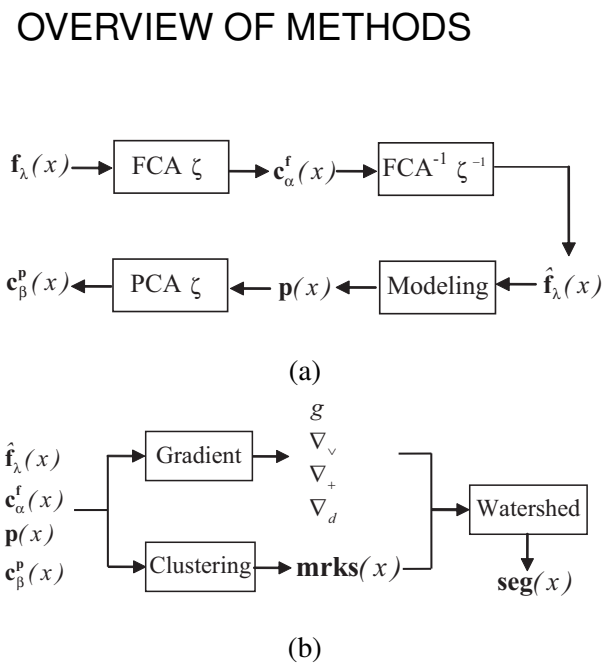


Fig. 14. *(a) Flowchart of different transformations: data denoising and reduction, (b) diagram for Watershed segmentation.*

In this paper we have considered four multidimensional spaces for image segmentation (fig. 14):

- **space 1:** the image space $\widehat{\mathbf{f}}_\lambda(x)$ after image filtering by FCA;
- **space 2:** the factorial space of image $\mathbf{c}^{\mathbf{f}}_\alpha(x)$ after image filtering and data reduction using FCA;
- **space 3:** the parameters space $\mathbf{p}(x)$ after image filtering by FCA and data reduction by model fitting;
- **space 4:** the factors parameters space $\mathbf{c}^{\mathbf{p}}_\beta(x)$ after image filtering by FCA, data reduction by model fitting and parameters orthogonalisation by PCA.

These spaces can be grouped in two different approaches. The first one is data reduction by FCA. Space 2 belongs to this approach. The second approach reduces data by model fitting. Spaces 3 and 4 belongs to this approach. Space 1 provides a direct approach to be compared to the others.

On each space, the same method of segmentation can be applied. First of all, a filtering is done on image $\mathbf{f}_\lambda(x)$ using FCA. Then the components of the spaces are generated: $\widehat{\mathbf{f}}_\lambda(x)$, $\mathbf{c}^{\mathbf{f}}_\alpha(x)$, $\mathbf{p}(x)$, $\mathbf{c}^{\mathbf{p}}_\beta(x)$. These components generate new hyperspectral images corresponding to the components space. In each space, the same method is applied on hyperspectral images. The segmentation combines a spectral and a spatial part. The spectral part consists of a classification in the considered space to obtain the markers. The spatial part consists in computing a gradient on hyperspectral images (fig. 14). Then, with the markers of the spectral part and the gradient of the spatial part, a watershed segmentation is performed on the considered space of hyperspectral images.

CONCLUSION AND PERSPECTIVES

This paper has presented a watershed-based segmentation for hyperspectral images. This approach combines a spectral part (the markers) and a spatial part (the gradient).

A temporal series example is used to illustrate our methodology. Comparing the results obtained for the various segmentations and the reference of Legrand *et al.* (2002), it is obvious that a data reduction approach is necessary. For the current image, the data reduction based on a model is better than the one based on factor correspondence analysis. In fact, for an hyperspectral image, it is better to use a model, when it can be fitted, because it reduces the image entropy while it keeps the inner data structure. Besides, the choice of pertinent parameters with low noise contribution is crucial for segmentation quality.

Moreover, multivariate gradients behave better than any marginal gradient on parameters. The multivariate gradients are based on an adapted distance to the considered space and on the supremum, or weighted sum, of morphological gradients on channels. The two kinds of gradients give similar results. Besides, to obtain relevant segmentations, they must be applied on a space with a low level of noise. This underlines the importance of a pertinent choice for parameters factors. About the markers, the corresponding spaces to compute them must be also with a low level of noise, to get pertinent ones.

In conclusion, a relevant multivariate segmentation requires an adapted data reduction, which gives parameters or factors with a low level of noise, is crucial ; a necessary condition to get pertinent markers and gradients.

In the future, we will develop multivariate filtering on spectral bands. More precisely, we will focus on the levelings. They are usually necessary to enhance the functions on which are computed the gradients. As for greyscale images, we will define new types of multivariate levelings. They will be adapted to peculiarities of these functions and they will also simultaneously filter all the spectral channels of hyperspectral images.

In the present example we have only used a simple approach for markers extraction, i.e. a clustering by "Clara". It is necessary to test other classification methods combining spectral and spatial information in order to improve markers detection. We are considering to do clustering with lambda flat zones and clustering after a principal coordinates analysis, using a weight/distance matrix (Benzécri, 1973; Gower, 1966).

REFERENCES

Angulo J, Serra J (2003). Color segmentation by ordered mergings. In Proc. of the IEEE International Conference on Image Processing (ICIP'2003) 2:125–128.

Benzécri JP (1973). L'Analyse des Données. L'Analyse des Correspondances. Vol. 2. Paris. Dunod.

Beucher S, Meyer F (1992). The Morphological Approach to Segmentation: The Watershed Transformation. In (E. Dougherty Ed.), Mathematical Morphology in Image Processing, Marcel Dekker, 433—481.

Evans AN, Liu XU (2006). A morphological gradient approach to color edge detection. IEEE Transactions on Image Processing, 15(6): 1454–1463.

Flouzat G, Amram O, Cherchali S (1998). Spatial and spectral segmentation of satellite remote sensing imagery using processing graphs by mathematical morphology. IEEE Geoscience and Remote Sensing Symposium Proceedings. IGARSS '98, 4:1769–1771.

Gower JC (1966). Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika, 53:325–338.

Green AA, Berman M, Switzer P, Craig MD (1988). A Transformation for Ordering Multispectral Data in terms of Image Quality with Implications for Noise Removal. IEEE Trans. Geosc. Rem. Sens. 26:65–74.

Kaufman L, Rousseeuw PJ (1990). Finding Groups in Data. An Introduction to Cluster Analysis. Ch. 2 and 3. John Wiley and Sons, 28–160.

Legrand AC, Meriaudeau F, Gorria P (2002). Active infrared non-destructive testing for glue occlusion detection within plastic lids. NDT&E International, 35:177–187.

Li P, Xiao X (2004). Evaluation of multiscale morphological segmentation of multispectral imagery for land cover classification. IEEE Geoscience and Remote Sensing Symposium. IGARSS '04. Proceedings, 4:2676–2679.

Meyer F (1992). Color image segmentation. In Proc. of 4th the International Conference on Image Processing and its Applications, Maastrich, Holland, 303–306.

Meyer F (2001). An Overview of Morphological Segmentation. International Journal of Pattern Recognition and Artificial Intelligence, 15(7): 1089–1118.

Meyer F (2004). Levelings, Image Simplification Filters for Segmentation. Journal of Mathematical Imaging and Vision, 20:59–72.

Paclík P, Duin RPW, Van Kempen GMP, Kohlus R (2003). Segmentation of multi-spectral images using the combined classifier approach. Image and Vision Computing, 21:473–482.

Scheunders P (2001). Multivalued image segmentation based on first fundamental form. IEEE Image Analysis and Processing. 11th International Conference. Proceedings, 185–190.

Serra J (1982). Image Analysis and Mathematical Morphology. Vol. 1. Academic Press. London.

Soille P (1996). Morphological partitioning of multispectral images. J. of Electronic Imaging, 5(3):252–263.

Soille P (2003). Morphological Image Analysis: Principles and Applications. Springer-Verlag, 2nd edition.