



Comparative study on morphological principal component analysis of hyperspectral images

Gianni Franchi, Jesus Angulo

► To cite this version:

Gianni Franchi, Jesus Angulo. Comparative study on morphological principal component analysis of hyperspectral images. Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 6th Workshop on, Jun 2014, Lausanne, Switzerland. hal-01256947

HAL Id: hal-01256947

<https://hal.science/hal-01256947>

Submitted on 15 Jan 2016

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.

COMPARATIVE STUDY ON MORPHOLOGICAL PRINCIPAL COMPONENT ANALYSIS OF HYPERSPECTRAL IMAGES

Gianni Franchi Jesús Angulo

CMM-Centre for Mathematical Morphology, MINES ParisTech; France

ABSTRACT

This paper deals with a problem of reducing the dimension of hyperspectral images using the principal component analysis. Since hyperspectral images are always reduced before any process, we choose to do this reduction by adding spatial information that can be useful then for classification process; to do it we choose to project our data in new spaces thanks mathematical morphology.

Index Terms— Statistical Learning, reduction of dimension, Hyperspectral images, Morphology.

1. INTRODUCTION

Conventionally, hyperspectral images which allow us to reconstruct the spectral profiles of objects imaged by the acquisition of several tens or several hundred of narrow spectral bands are often reduced in dimension before any treatment. Many hyperspectral reduction methods are linear and do not care of the multiple sources of nonlinearity presented in [1]. However lately, non linear reduction techniques have been developed, and some of them have been used in hyperspectral images [2]. However, most of these techniques present disadvantages [3] in comparison to the principal component analysis that is why we choose to work with the PCA and to add spatial information, which are nonlinear. Since mathematical morphology is a nonlinear image processing methodology based on the application of complete lattice theory to spatial structures, we choose to explore this kind of processing like [4, 5], however in our case we use mathematical morphology to improve the reduction of the dimension.

2. NOTATIONS AND NOTIONS OF MATHEMATICAL MORPHOLOGY

2.1. Morphological Decomposition

In this section we are going to introduce the notation that would be used on the rest of this paper. Let E be a subset of the discrete space \mathbb{Z}^2 , which represents the support space of a 2D image and $F \subseteq \mathbb{R}^D$ be a set of pixels values in dimension D . Hence, in our case the value of a pixel $x \in E$ is represented by a vector $v \in F$ of dimension D . Additionally,

we will write higher order tensors by a calligraphic upper-case letters ($\mathcal{I}, \mathcal{S}, \dots$). Moreover if $\mathcal{I} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, for all $i \in [1, n_3]$ $\mathcal{I}_{::,i}$ represents a matrix of size $n_1 \times n_2$ where the third component is equal to i . We can associate a tensor to the hyperspectral image, we called this tensor $\mathcal{F} \in \mathbb{R}^{n_1 \times n_2 \times D}$.

Let f be a grey scale image which can be represented by a function. Area openings (resp. area closings) are morphological filters that remove from an image the bright (resp. dark) connected components having a surface area smaller than the parameter $s_l \in \mathbb{N}$ [6]: $\gamma_{s_l}^a(f) = \bigvee_i \{\gamma_{B_i}(f) | B_i \text{ is connected and } \text{card}(B_i) \leq s_l\}$ and $\varphi_{s_l}^a(f) = \bigwedge_i \{\varphi_{B_i}(f) | B_i \text{ is connected and } \text{card}(B_i) = s_l\}$, where $\gamma_B(f)$ and $\varphi_B(f)$ represent respectively the morphological flat opening and closing according to structuring element B .

Let us consider $\{\gamma_{s_l}^a\}$, $l = 1 \dots S$ and $\{\varphi_{s_l}^a\}$, $l = 1 \dots S$, two families indexed one of openings and one of closing. Typically, the index l is associated to the size of the structuring element or in our case to the surface area. The notion of morphological decomposition is related to granulocytic axiomatic [7]. Namely, based on [8] we have :

$$f = \frac{1}{2} \left(\sum_{l=1}^S (\gamma_{s_{l-1}}^a(f) - \gamma_{s_l}^a(f)) - \sum_{l=1}^S (\varphi_{s_l}^a(f) - \varphi_{s_{l-1}}^a(f)) + \gamma_S(f) + \varphi_S(f) \right)$$

Therefore we have a decomposition of the initial image into S scales, together with the last opening and closing. We remark that residue $(\gamma_{s_{l-1}}^a(f) - \gamma_{s_l}^a(f))$ represents bright details between levels s_l and s_{l-1} . Identically $(\varphi_{s_l}^a(f) - \varphi_{s_{l-1}}^a(f))$ stands for dark details between levels s_l and s_{l-1} . We can now calculate a decomposition of each band of the hyperspectral cube. However there are some issues to be taken into account : by decomposing an image into scales we have now to deal with an object of bigger dimensionality, and also the decomposition may not be optimal: it depends on the discretization of the S scales, the size of the pixel scales, etc... Then in next section we are going to introduce the pattern spectrum and how it can be used to deal with the problem of discretization.

2.2. Pattern Spectrum (PS)

The pattern spectrum provides the probability density function (pdf) of the granulometry. The area normalized pattern spectrum of f is at size s_l : $PS^a(f, l) = \frac{\text{Mes}(\gamma_{s_l}^a(f)) - \text{Mes}(\gamma_{s_{l+1}}^a(f))}{\text{Mes}(f)}$,

$PS^a(f, -l) = \frac{\text{Mes}(\varphi_{s_l+1}^a(f)) - \text{Mes}(\varphi_{s_l}^a(f))}{\text{Mes}(f)}$, where "Mes" represents the integral of the image. Based on the analogy between the pdf and the PS, we can calculate the cumulative distribution function of the spectrum for the opening part and also one for the closing part. Afterwards we sample these cumulative distribution functions where the number of sample is fixed and is equal to S , under the constrain that the sampled cumulative distribution function must be as much as possible similar to the original function. Since in probability, the

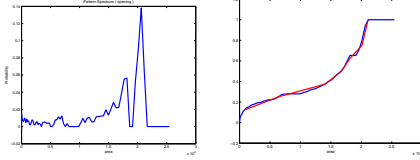


Fig. 1. On the left the Pattern Spectrum with opening of a grey scale image, on the right in blue its corresponding cumulative distribution function of the spectrum, and in red its approximation with $S = 8$.

cumulative distribution function is a way to characterised a distribution such as the pdf then thanks to this property, we can be sure that our discretization represents the original image.

2.3. Distance function

Let us consider a binary image b_f , and X a close set of this image, the distance function corresponding to this close set X gives to each point $x \in X$ a real number that depends on the position of x with respect to X such as : $d(x, X) = \min\{\rho(x, y) : y \in X^c\}$, where $\rho(x, y)$ is the Euclidean distance between x and y , and where X^c is the complemented of set X . Now let us consider a gray-scale image f , and $\{X_s(f)\}_s$ its upper level sets. Then, according to [9], the gray-scale distance transform of f is : $d(x, f) = \frac{1}{255} \times \sum_{s=a}^b d(x, X_s(f))$, where $a = \min\{f(x), x \in E\}$, and $b = \max\{f(x), x \in E\}$, where E is the index of the position of the pixel of f .

3. MORPHOLOGICAL PRINCIPAL COMPONENT ANALYSIS (MPCA)

3.1. Classical PCA

Principal Component Analysis (PCA) starts with a set vector $v_i \in \mathbb{R}^D$, $1 \leq i \leq n$ where n represents the number of vectors, in our case it corresponds to the number of image pixels, i.e., $n = n_1 n_2$. The goal of the PCA is to reduce the dimension of this vector space, namely $F = \{v_i\}_{i=1}^n \rightarrow F' = \{v'_i\}_{i=1}^n$, with $v'_i \in \mathbb{R}^d$, where $d \ll D$. In our case, $F \in M_{n,D}(\mathbb{R})$ represents the hyperspectral image \mathcal{F} , where

each column $F_k \in \mathbb{R}^n$, $1 \leq k \leq D$ corresponds to a vectorized spectral band. PCA maximizes the objective function : $\mathbf{L}(w_j, \lambda) = w_j^T V w_j - \lambda(w_j^T w_j - 1)$, where $\lambda \in \mathbb{R}$ and $V = n^{-1}(F^T F)$, $V \in M_{D,D}(\mathbb{R})$, is the covariance of F . This reduction can be done by projecting the data on d eigenvectors of V corresponding to the d higher eigenvalues of V .

3.2. MPCA and its variants

The fundamental idea of Morphological Principal Component Analysis (MPCA) consists in replacing the covariance matrix V of PCA, which represents the statistical interaction of spectral bands, by a covariance matrix V_{Morpho} computed from a morphological representation of the bands. Therefore, mathematical morphology is fully integrated in the dimensionality reduction problem by standard SVD computation to solve $V_{\text{Morpho}} w_j = \lambda_j w_j$. The corresponding principal components w_j provides the projection space for the hyperspectral image \mathcal{F} .

Scale-space Decomposition MPCA. Using the surface area-based nonlinear scale-space discussed in previous section, the grey-scale image of each spectral band $\mathcal{F}_{:, :, k}$ is decomposed into residues of area openings and area closings according to the discretization into S scales for each operator, i.e., $r_l(\mathcal{F}_{:, :, k}) = \gamma_{s_{l-1}}^a(\mathcal{F}_{:, :, k}) - \gamma_{s_l}^a(\mathcal{F}_{:, :, k})$ and $r_{-l}(\mathcal{F}_{:, :, k}) = \varphi_{s_l}^a(\mathcal{F}_{:, :, k}) - \varphi_{s_{l-1}}^a(\mathcal{F}_{:, :, k})$, $1 \leq l \leq S$. Thus we have increased the dimensionality of the initial dataset from a tensor (n_1, n_2, D) to a tensor $(n_1, n_2, D, 2S + 1)$. As discussed in [8], this tensor can be reduced using high order-SVD techniques. We propose here to simply compute a covariance matrix as the sum of the covariance matrices from the various scales. More precisely, we introduce $V_{\text{Morpho-1}} \in M_{D,D}(\mathbb{R})$ with : $V_{\text{Morpho-1}} = \sum_{l=1}^S (V(l)) + \sum_{l=1}^S (V(-l))$ where the covariance matrices at each scale l is obtained as $V(l)_{k,k'} = \text{Covar}(r_l(\mathcal{F}_{:, :, k}), r_l(\mathcal{F}_{:, :, k'}))$, $1 \leq k, k' \leq D$. We note that involves an assumption of independence of the various scales. We remark also that this technique is different of classical approaches of differential profiles as [10, ?] where all the images of the morphological decomposition are used as columns to compute a covariance matrix in $M_{2S+1, 2S+1}(\mathbb{R})$.

Pattern Spectrum MPCA. In fact, we can consider a much compact representation of the morphological information associated to area-based nonlinear scale-space of each spectral band. It simple involves to consider the area-based pattern spectrum of spectral bands as the variable to be used to find statistical redundancy on the data. In other words, the corresponding covariance matrix $V_{\text{Morpho-2}} \in M_{D,D}(\mathbb{R})$ is defined as : $V_{\text{Morpho-2}}_{k,k'} = \text{Covar}(PS^a(\mathcal{F}_{:, :, k}, l), PS^a(\mathcal{F}_{:, :, k'}, l))$, with $1 \leq k, k' \leq D$ and where $PS^a(\mathcal{F}_{:, :, k}, l)$, $-S \leq l \leq S$, is the area-based pattern spectrum obtained by area-openings and area-closings. We note that the pattern spectrum can be seen a size a pdf of image structures and consequently the

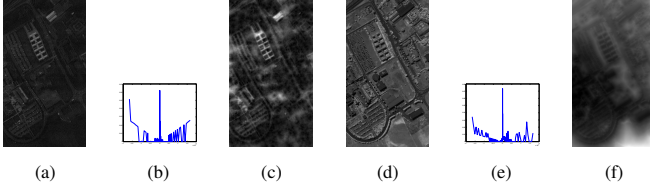


Fig. 2. (a) band 1 of Pavia, (b) PS of band 1 of Pavia (c) distance function band 1 of Pavia, idem for (d), (e) and (f) but with band 100.

MPCA associated to it explores intrinsic dimensionality of sets of distributions instead of sets of vectors.

Distance Function MPCA. Classical PCA for hyperspectral images is based on exploring covariances between spectral intensities. The previous MPCA involves to change the covariance to a morphological scale-space representation of the images. An alternative is founded on transforming each spectral band from an intensity based map to a metric based map where at each pixel the value is associated to both the initial intensity and the spatial relationships between the image structures. This objective can be achieved using the Molchanov gray-scale distance function for each spectral band $\text{dist}(\mathcal{F}_{:,k})$. The new covariance matrix $V_{\text{Morpho-3}} \in M_{D,D}(\mathbb{R})$ is now defined as: $V_{\text{Morpho-3}}(k,k') = \text{Covar}(\text{dist}(\mathcal{F}_{:,k}), \text{dist}(\mathcal{F}_{:,k'}))$, with $1 \leq k, k' \leq D$.

Spatial/Spectral MPCA. As we have discussed, $V_{\text{Morpho-2}}$ represents a compact morphological representation of the image, however the spectral intensity information is also important for dimensionality reduction. Let us write X and Y two random variables such as X represents the spectral information of the data, and Y the morphological/spatial information. To come with a new variant of MPCA we assume the independence of the spectral and spatial information so of X and Y . This imply that $\text{var}((1 - \beta) \text{Pr}(X) + \beta \text{Pr}(Y)) = (1 - \beta)^2 \text{var} \text{Pr}(X) + \beta^2 \text{var} \text{Pr}(Y)$, with $\beta \in [0, 1]$. More concretely, that means that we can build another covariance matrix $V_{\text{Morpho-4}}$ that would represent the spectral and spatial information without increasing the dimensionality by : $V_{\text{Morpho-4}} = (1 - \beta)^2 V + \beta^2 V_{\text{Morpho-2}}$, where obviously $V_{k,k'} = \text{Covar}(\mathcal{F}_{:,k}, \mathcal{F}_{:,k'})$ and β stands for a regularization term that balances the spatial over the spectral information.

4. MPCA APPLIED TO HYPERSPECTRAL IMAGES

4.1. Criteria to evaluate PCA vs. MPCA

We can now use PCA and the four variant of MPCA to dimensionality reduction of hyperspectral images. In order to

evaluate the interest for such purpose, it is necessary to establish quantitative criteria that should be assessed. We want to show the improvement of classical PCA with respect to the following criteria.

Criterion 1. (C1) The reconstructed hyperspectral image $\tilde{\mathcal{F}}$ using the first d principal components should be regularized in order to be more spatially/spectrally homogeneous;

Criterion 2. (C2) Hyperspectral image $\tilde{\mathcal{F}}$ should preserve the edges of the main image structures;

Criterion 3. (C3) Separability of spectral classes should be improved in the dimensionality reduced space. That involves in particular a better pixel classification.

In order to assess C2, we compute for initial image \mathcal{F} the gradient of each band k , denoted $\|\nabla \mathcal{F}_k\|$, and compare it to the gradient of the reconstructed image $\tilde{\mathcal{F}}$, by measuring the Euclidean distance. In order to have an estimator of the error, we integrate for all the bands, i.e., $\text{Error}_{\text{Grad}} = \sum_{k=1}^D \sum_{i,j=1}^{n_1, n_2} \|\nabla \mathcal{F}_{i,j,k} - \nabla \tilde{\mathcal{F}}_{i,j,k}\|^2$.

To assess C1, which involves image homogeneity, we propose an approach based on a partition of the image into regions. First, working on the d eigenvectors, we compute the image partition associated to the α -flat zones, the partition is denoted π_α . We remind that two neighboring pixels belong to the same α -flat zones if their distance is lower or equal to α [11]. In our case, the pixel distance corresponds to the Euclidean distance of vector values in the eigenimages and the choice of α is done in order to have a number C of α -flat zones similar for all the approaches of PCA and MPCA to be compared. By fixing the number of zones in the partition, we guarantee that the difference between a partition and another one depends exclusively on the homogeneity of the image.

Then, using the partition π_α , we compute the spectral mean value of pixels from original image \mathcal{F} in each zone, in order to produce a simplified hyperspectral image denoted $\bar{\mathcal{F}}^{\pi_\alpha}$. Finally, we assess how pixels of the original image from each α -flat zone are far from their mean by computing the following error $\text{Error}_{\text{Homg}} = \sum_{k=1}^D \sum_{i,j=1}^{n_1, n_2} |\mathcal{F}_{i,j,k} - \bar{\mathcal{F}}_{i,j,k}^{\pi_\alpha}|^2$.

Finally, C3 is related to supervised classification of hyperspectral image. We have considered SVM with a linear kernel as learning technique, where the classifier is validated using 5-fold cross validation. Sensitivity and specificity of classification are then computed.

4.2. Evaluation of Pavia hyperspectral image

The assessment of the performance of PC and MPCA was carried out using the well known University of Pavia hyperspectral image, which has dimensions $n_1 = 1610 \times n_2 = 340$ pixels, $D = 103$ spectral bands and its geometrical resolution is of 1.3 m. We have applied classical PCA and the different variants of MPCA to Pavia image. Fig. 4.2 shows the first three eigenimages, visualized as a RGB false color. We note that the PS MPCA requires $d = 5$ to represent 90% of the variance whereas the other approaches only imposes

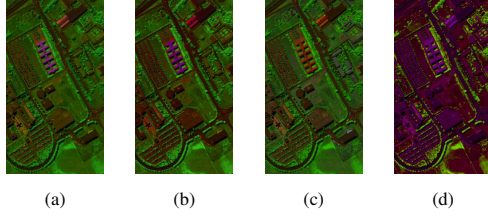


Fig. 3. RGB false color visualization of first three eigen-images from Pavia image: (a) spectral PCA, (b) scale-decomposition MPCA, (c) pattern spectrum MPCA, (d) distance function MPCA.

$d = 3$. An interesting aspect observed on the projection of the 103 spectral bands into the first two eigenvectors is how PCA and the scale-space decomposition MPCA cluster the bands linearly, since bands close in the projection are also near in the spectral domain, whereas the pattern spectrum MPCA and distance function MPCA tends to cluster spectral bands which are not necessary spectrally contiguous. Therefore the nonlinear embedding from $V_{\text{Morpho-2}}$ and $V_{\text{Morpho-3}}$ is clearly illustrated. Obviously, a similar behavior is observed for $V_{\text{Morpho-4 } \beta}$. From a quantitative viewpoint, one can see in Table 1 that globally MPCA produces a more homogenous regularization of the image than classical PCA, especially the distance function MPCA and Spatial/Spectral MPCA with an appropriate $\beta = 0.2$, which give the lowest values of $\text{Error}_{\text{Homg}}$. Finally, Table 2 summarizes the results of supervised classification of Pavia. We note the sensitivity is excellent for all the methods, however we can see how MPCA improves the results on the specificity. Once again $V_{\text{Morpho-3}}$ and $V_{\text{Morpho-4 } \beta=0.2}$ are significantly better than the others variants.

	V	$V_{\text{Morpho-1}}$	$V_{\text{Morpho-2}}$	$V_{\text{Morpho-3}}$
$\text{Error}_{\text{Homg}}$	100	100	95.9	79.3
$\text{Error}_{\text{Grad}}$	55.41	54.88	55.25	100
	$V_{\text{Morpho-4 } \beta}$ $\beta = 0.8$	$V_{\text{Morpho-4 } \beta}$ $\beta = 0.2$	$V_{\text{Morpho-4 } \beta}$ $\beta = 0.5$	
$\text{Error}_{\text{Homg}}$	93.2	83.9	88.3	
$\text{Error}_{\text{Grad}}$	38.9	98.5	78.5	

Table 1. Comparison of principal component analysis from Pavia image using results of criteria C1 and C2. The values have normalized to worst case, which gives 100.

5. CONCLUSION

We have introduced in the paper the notion of MPCA which has been declined into four different approaches. MPCA allows to deal with a spatial/spectral representation of the image based on mathematical morphology tools for SVD-based dimensionality reduction. It is important to note that it involves only to change the covariance matrix used in the SVD to obtain the eigenvectors where the spectral bands are then

	Sensitivity $\pm \sigma(10^{-5})$	Specificity $\pm \sigma(10^{-2})$
V	0.99 ± 1.03	0.44 ± 3.44
$V_{\text{Morpho-1}}$	0.99 ± 1.03	0.45 ± 3.39
$V_{\text{Morpho-2}}$	0.99 ± 1.03	0.56 ± 3.38
$V_{\text{Morpho-3}}$	0.99 ± 0.93	0.76 ± 3.35
$V_{\text{Morpho-4 } \beta}, \beta = 0.8$	0.99 ± 0.93	0.39 ± 3.35
$V_{\text{Morpho-4 } \beta}, \beta = 0.2$	0.99 ± 0.93	0.78 ± 3.35
$V_{\text{Morpho-4 } \beta}, \beta = 0.5$	0.99 ± 0.93	0.47 ± 3.35

Table 2. Comparison of hyperspectral supervised classification on principal component space from Pavia image, linear kernel SVM and 5-fold cross validation.

projected. We have shown that the best results results are obtained when we combine spatial and spectral information ($V_{\text{Morpho-3}}$, $V_{\text{Morpho-4 } \beta}, \beta = 0.2$).

Acknowledgement

The authors would like to thank Prof. P. Gamba, University of Pavia, for providing the ROSIS data set

6. REFERENCES

- [1] Kerri Guilfoyle, Mark L. Althouse, and Chein-I Chang, "Further investigations into the use of linear and nonlinear mixing models for hyperspectral image analysis," in *AeroSense 2002*. International Society for Optics and Photonics, 2002, pp. 157–167.
- [2] C.M. Bachmann, T.L. Ainsworth, and R.A. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 3, pp. 441–454, March 2005.
- [3] Laurens JP van der Maaten, Eric O Postma, and H Jaap van den Herik, "Dimensionality reduction: A comparative review," *Journal of Machine Learning Research*, vol. 10, no. 1-41, pp. 66–71, 2009.
- [4] M. Fauvel, J.A. Benediktsson, J. Chanussot, and J.R. Sveinsson, "Spectral and spatial classification of hyperspectral data using svms and morphological profiles," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 11, pp. 3804–3814, Nov 2008.
- [5] Martino Pesaresi and J.A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 39, no. 2, pp. 309–320, Feb 2001.
- [6] L. Vincent, "Morphological area opening and closing for grayscale images," *Proc. NATO Shape in Picture Workshop, Driebergen, The Netherlands, Springer-Verlag*, pp. 197–208, September 1992.
- [7] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, Inc., Orlando, FL, USA, 1983.
- [8] S. Velasco-Forero and J. Angulo, "Classification of hyperspectral images by tensor modeling and additive morphological decomposition," *Pattern Recogn.*, vol. 46, no. 2, pp. 566–577, Feb. 2013.
- [9] Ilya S. Molchanov and Pedro Teran, "Distance transforms for real-valued functions," *Journal of Mathematical Analysis and Applications*, vol. 278, no. 2, pp. 472–484, 2003.
- [10] G. K. Ouzounis, M. Pesaresi, and P. Soille, "Differential area profiles: Decomposition properties and efficient computation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, pp. 1533–1548, Aug. 2012.
- [11] Lionel Gueguen, Santiago Velasco-Forero, and Pierre Soille, "Local mutual information for dissimilarity-based image segmentation," *Journal of Mathematical Imaging and Vision*, pp. 1–20.