Morphological scale-space for hyperspectral images and dimensionality exploration using tensor modeling

Santiago Velasco-Forero, Jesus Angulo

HAL Id: hal-00458687

https://minesparis-psl.hal.science/hal-00458687

Submitted on 22 Feb 2010

MORPHOLOGICAL SCALE-SPACE FOR HYPERSPECTRAL IMAGES AND DIMENSIONALITY EXPLORATION USING TENSOR MODELING

Santiago Velasco-Forero, Jesús Angulo

Mines ParisTech, Center de Morphologie Mathématique, Fontainebleau, France

ABSTRACT

This paper proposes a framework to integrate spatial information into unsupervised feature extraction for hyperspectral images. In this approach a nonlinear scale-space representation using morphological levelings is formulated. In order to apply feature extraction, Tensor Principal Components are computed involving spatial and spectral information. The proposed method has shown significant gain over the conventional schemes used with real hyperspectral images. In addition, the proposed framework opens a wide field for future developments in which spatial information can be easily integrated into the feature extraction stage. Examples using real hyperspectral images with high spatial resolution showed excellent performance even with a low number of training samples.

Index Terms— Unsupervised feature extraction, mathematical morphology, tensor analysis, dimensional reduction, classification, principal component, hyperspectral imagery.

1. INTRODUCTION

Hyperspectral images (HSI) are acquired by a sensor that captures radiant flux over hundreds of contiguous bands, producing a high dimensional signature for each pixel in the image. It is accepted that HSI include many highly correlated bands resulting in spectral redundancy. This increases computational complexity and degrades classification accuracy [1]. The difficulties in HSI using conventional supervised classification methods originally developed for low-dimensional multispectral data are illustrated in [2]. Additionally, identification for relatively small objects that require high spatial resolution is often lost when using these methods with HSI. Due to its simplicity, principal component analysis (PCA) using singular value decomposition (SVD) over the zero mean covariance matrix is the most popular approach to dimensional reduction in HSI. Unfortunately, PCA requires that images are first vectorized which retains spectral information but neglects the spatial information presented in the original array. Recently, several authors have pointed out that the natural representation for hyperspectral images is a three dimensional array instead of a matrix of vectorized images

Our approach presents a nonlinear scale-space representation (NSSR) based on morphological levelings, which include openings and closings by reconstruction. NSSR creates a fourth dimension in HIS which is analyzed as a forth tensor dimension. Instead of focusing exclusively on PCA in morphological operators, we focus on the analysis of spatial and spectral patterns: Simultaneously using high-order or multilinear singular value decomposition (HO-SVD) as is proposed for tensors in [7]. Classification is performed using Support Vector Machines (SVM) since they have shown good results with limited training sets [8]and the "kernel trick" allows us to define natural distances after tensor dimensionality reduction.

The paper is organized as follows: Section 2 introduces tensor notation and basic concepts being used later on. Section 3 introduces nonlinear scale-space representations with morphological levelings. Section 4 presents our approach using the previous framework. Section 5 shows the effectiveness of the modified approach via practical examples with comparison to standard approach. And section 6 concludes the paper.

2. TENSOR MODELING

2.1. Notation

A HSI has a natural representation as a tensor, denoted by $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$. Each index is called *mode*: the first two are spatial and the third is spectral. Thus, $\mathcal{X}$ is a sample of I_3 images of size $I_1 \times I_2$. Our approach is apply multilinear algebra to the whole tensor structure instead of rearranging the data tensor to adapt it to classical matrix-based algebraic technique.

The following are the notation and the basic definitions we use for multilinear algebra. Scalars are denoted by lower case letters $(a, b, \ldots)$, vectors by bold lower case letters $(\mathbf{a}, \mathbf{b}, \ldots)$, matrices by bold upper-case letters $(\mathbf{X}, \mathbf{Y}, \ldots)$, and higher-order tensors by calligraphic upper-case letters $(\mathcal{X}, \mathcal{Y}, \ldots)$. The order of tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \ldots \times I_N}$ is N. We utilize subscripts to illustrate the tensor order, for example $\mathcal{X}_{ijkl}$ is a tensor of order 4.

2.2. Tensor Decomposition

A matrix $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2}$ is a two-mode mathematical object that has two associated vector spaces, a row space and a column space. SVD orthogonalizes these two spaces and decomposes the matrix as $\mathbf{X} = \mathbf{U}_1 \boldsymbol{\Sigma} \mathbf{U}_2^T$, where $\mathbf{U}_1$ and $\mathbf{U}_2^T$ represent the orthogonal column space and $\boldsymbol{\Sigma}$ is a diagonal singular value matrix. In terms of the mode-n products, this SVD decomposition can be rewritten as $\mathbf{X} = \boldsymbol{\Sigma} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2$, where $\times_n$ is the mode-n product [7] [9]. Extension to an N-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times \ldots \times I_N}$ were presented by [7] orthogonalizing N spaces and expresses the tensor as the mode-n product of N-orthogonal spaces

$$\mathcal{X} = \mathcal{C} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \ldots \times_N \mathbf{U}_N \qquad (1)$$

Tensor $\mathcal{C}$, known as the core tensor, is analogous to the diagonal singular value matrix in conventional matrix SVD. It is important to realize, however, that the core tensor does not have a diagonal structure; rather, $\mathcal{C}$ is in general a full tensor [7]. The core tensor governs the interaction between the mode matrices $\mathbf{U}_n$, for $n = 1, \ldots, N$. Mode matrix $\mathbf{U}_n$ contains the orthonormal vectors spanning the column space of the matrix $\mathbf{X}_n$ that results from the mode-n flattening of $\mathcal{X}$ [7]. The HOSVD is performed using Alternative Least Square algorithm used to jointly find n-mode matrices $\mathbf{U}_n$. An excellent compendium about tensor decomposition is presented in [9].

2.3. Tensor Principal Component Analysis

In high-dimensional problem as HSI, it is of great interest to reduce the spectral dimension in order to address problems as "Curse of Dimensionality" in distance-based or nonparametric analysis and "Hughes phenomenon" in linear classifiers. Commonly a pre-processing step consists in performing a PCA to reduce the dimensional space. We present a tensor version for PCA that minimize the reconstruction error in the Frobenius norm. Thus, the best lower rank tensor approximation of $\mathcal{X}$ [7], denoted by $\tilde{\mathcal{X}}$ is:

$$\tilde{\mathcal{X}} = \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2 \times_3 \ldots \times_N \mathbf{P}_N \qquad (2)$$

where $\mathbf{P}_n = \mathbf{U}_n \mathbf{U}_n^T$, and $\mathbf{U}_n$ is found using expression 1. Our motivation is to reduce the high dimension correspondent to some mode-n product, i.e. yield a projection $\mathbb{R}^{I_1 \times I_2 \ldots \times I_N} \rightarrow \mathbb{R}^{I_1 \times I_2 \ldots \times k}$. Thus, we present the *Tensor PCA* as follows:

$$\widehat{X}_{I_1, I_2, \ldots, k} = \mathcal{X} \times_1 \mathbf{P}_1 \times_2 \mathbf{P}_2 \times_3 \ldots \times_N \widehat{\mathbf{U}}_N^T \qquad (3)$$

where $\widehat{\mathbf{U}}_N^T$ contains the k eigenvectors associated with the k largest eigenvalues holding of the unfolding matrix $\mathbf{X}_N$ [7]. Tensor PCA was previously presented for HSIs in [3] including transformation for sphericity.

3. NONLINEAR SCALE-SPACE REPRESENTATION WITH MORPHOLOGICAL LEVELINGS (NSS-ML)

NSS-ML generates a multiscale representation where transformation from finer scale to a coarse scale are defined using morphological levelings [10]. The application of levelings for Multiscale imaging structure decomposition in remote sensing has been previously studied by [4]. In [4], the levelings are achieved using reconstruction of separated openings/closings. Then the residues of both operators (i.e., top-hat transformations) are thresholded to segment the image into convex/concave classes for the various scales. As we shows below, levelings can be directly computed by standard iterative geodesic operators. In addition, based on a pyramid of markers, the levelings define a nonlinear scale space which decompose simultaneously the bright/dark image structure in various scales.

3.1. Morphological Levelings

Given a grey scale image $\mathbf{X} = f(x, y)$ and a marker image $\mathbf{M} = m(x, y)$, where the marker is typically a rough simplification of reference image which determines the structures to be reconstructed and those to be simplified during the geodesic erosion/dilation procedure. The leveling $\lambda(f, m)$ can be obtained by the following iterative algorithm:

$$\lambda(f, m) = \lambda^i(f, m) = [f \wedge \delta_f^i(m)] \vee \varepsilon_f^i(m)$$

such that $\lambda^i(f, m) = \lambda^{i+1}(f, m)$ (convergence until idempotency).
The geodesic dilation of size i is given by $\delta_f^i(m) = \delta_f^1 \delta_f^{i-1}(m)$ where unitary conditional dilation is $\delta_f^1(m) = \delta_B(m) \wedge f$ with $\delta_B(m)$ an unitary dilation. The geodesic erosion of size i, $\varepsilon_f^i(m)$ can be obtained by the property of duality by complement $\varepsilon_f^i(m) = [\delta_{f^c}^i(m^c)]^c$, where f^c is the negative of image f.

(a) $\mathbf{X}^{S_0}$ (b) $\Psi_1(\mathbf{X}^{S_0})$ (c) $\mathbf{X}^{S_1}$ (d) $\mathbf{R}^{S_1}$
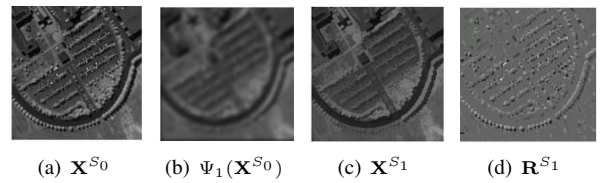
Fig. 1. Morphological decomposition steps using $\boldsymbol{\sigma} = (\sigma_1) = 3$

3.2. Morphological Multiscale Decomposition

The idea is to start from a family of transformations indexed by a scale parameter, i.e.,$\{\Psi_n(\mathbf{X})\}_{n=0}^N$, such that $\Psi_0(\mathbf{X}) = \mathbf{X}$. In this study, we typically use a pyramid of Gaussian filters, $\Psi_n(\mathbf{X}) = G_{\sigma_n} * \mathbf{X}$ such that $\sigma_n > \sigma_{n-1}$ for all n.

Now, we can define the following scheme for image structure decomposition:

.	Image ≡ Structure	Residual ≡ Texture
Scale 0	$\mathbf{X}^{S_0} = \mathbf{X}$	$\mathbf{R}^{S_0} = 0$
Scale 1	$\mathbf{X}^{S_1} = \lambda(\mathbf{X}^{S_0}; \Psi_1(\mathbf{X}^{S_0}))$	$\mathbf{R}^{S_1} = \mathbf{X}^{S_0} - \mathbf{X}^{S_1}$
Scale 2	$\mathbf{X}^{S_2} = \lambda(\mathbf{X}^{S_1}; \Psi_2(\mathbf{X}^{S_1}))$	$\mathbf{R}^{S_2} = \mathbf{X}^{S_1} - \mathbf{X}^{S_2}$
$\vdots$	$\vdots$	$\vdots$
Scale N	$\mathbf{X}^{S_N} = \lambda(\mathbf{X}^{S_N-1}; \Psi_N(\mathbf{X}^{S_N-1}))$	$\mathbf{R}^{S_N} = \mathbf{X}^{S_N-1} - \mathbf{X}^{S_N}$

Therefore, the morphological multiscale decomposition of a grey scale image $\mathbf{X}$ as a tensor of order 3 denoted by $\mathcal{X}^{\sigma}_{ijk}$ can be defined by:

$$M(\mathbf{X}_{ij}, \boldsymbol{\sigma}) = \mathcal{X}^{\boldsymbol{\sigma}}_{ijn} = \left[\left\{ \mathbf{R}^{S_n}_{ij} \right\}^N_{n=1}, \mathbf{X}^{S_N}_{ij} \right] = \mathcal{X}^{\boldsymbol{\sigma}} \quad (4)$$

Additionally, $\mathcal{X}^{\boldsymbol{\sigma}}_{ijn}$ has the following structure+texture property:

$$\bigoplus^{N+1}_{n=1} \mathcal{X}^{\boldsymbol{\sigma}}_{ijn} = \bigoplus^N_{n=1} \mathbf{R}^{S_n}_{ij} + \mathbf{X}^{S_N}_{ij} = \mathbf{X}_{ij} = \mathbf{X}$$

Fig. 1 illustrate morphological decomposition of the original image $\mathbf{X}^{S_0}$. In remote sensing scenario, given a HSI $\mathcal{X}$ with K bands, our approach use a marginal framework to obtain a decomposition using expression 4 for each spectral band denoted by $\mathbf{X}_k$ for $k = 1, \ldots, K$ as follows:

$$M(\mathcal{X}_{ijk}, \boldsymbol{\sigma}) = \{\mathcal{X}^{\boldsymbol{\sigma}}_{ijk}\}^K_{k=1} = \mathcal{X}^{\boldsymbol{\sigma}}_{ijkn} = \mathcal{X}^{\boldsymbol{\sigma}} \quad (5)$$

4. REDUCTION AND CLASSIFICATION

The main motivation for this paper is to incorporate spatial information in the HSI classification. Previous researches have shown that mathematical morphology operators as opening and closing by reconstruction help us to set up spatial information in analysis [6]. Our approach also utilizes mathematical morphology to yield a nonlinear decomposition presented in Section 3. For a hypersepctral image $\mathcal{X}$, the operator $M(\mathcal{X}, \boldsymbol{\sigma}) = \mathcal{X}^{\boldsymbol{\sigma}}$ produces a fourth dimension that corresponds to different scales in the image and has the property that summing by scale produces the exact original hyperspectral image $(\bigoplus \mathcal{X}^{\boldsymbol{\sigma}} = \mathcal{X})$. Using tensor reduction presented in section 2 we obtained a smaller tensor preserving information both in bands and scales $\tilde{\mathcal{X}}^{\boldsymbol{\sigma}}_{I_1, I_2, k_b, k_s}$, where k_b and k_s are the number of feature to reduce in bands and scale respectively. Directly from tensor reduction, we obtain a matrix of size $k_b \times k_s$ for each pixel. Taking advantage of flexibility in support vector machine problem formulation, we solve the classification problem using gaussian kernel in matrix form.

5. EXPERIMENTS

In this section, the proposed method is compared to classical dimensional reduction in a standard hyperspectral image. Before classification using SVM-KM [11], each feature was standardized. The kernel parameter $\tilde{\rho}$ was tuned to

range $\tilde{\rho} = \{1, 2, \ldots, 10\}$ and the regularization parameter for SVMs was varied in $C = \{10^0, \ldots, 10^5\}$.

An image taken over the Northwest of Indiana in an Indian Pine test site in June 1992 is used in our experiment. This image contains 145x145 pixels and 200 spectral bands in the 400-2500 nm range. The whole image containing 16 classes is considered. In first scenario, we followed the procedure presented in [12]. Therefore, we present classification for 9 classes and 5 labeled samples only.
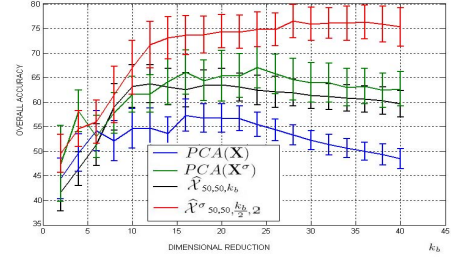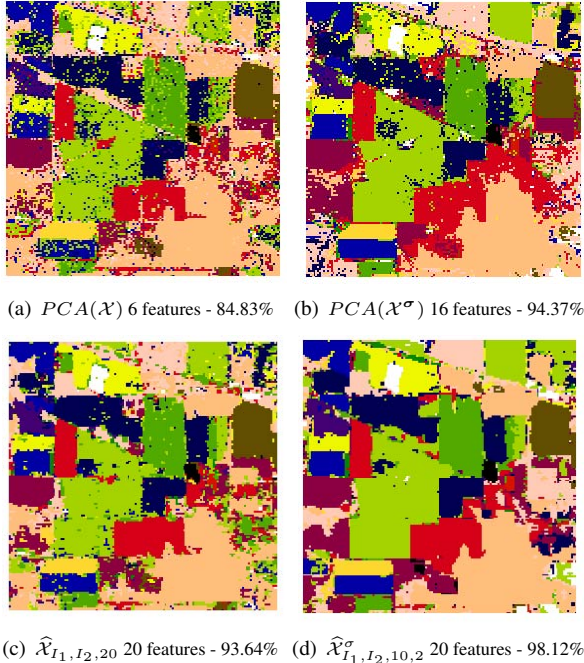

(a) *Scenario 1*: 5 training samples per class in 20 random repetitions for 9 classes.
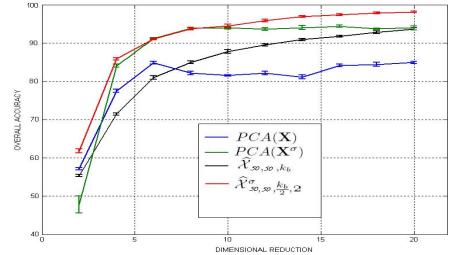
(b) *Scenario 2*: 20% training samples in 16 classes in 10 random repetitions.

Fig. 2. OA accuracy for different dimensional reduction approaches in Indian Pine.

Results for different dimensional reduction approaches are presented in Fig. 2(a). This extreme scenario shows the importance of incorporating spatial information before doing feature reduction. Our approach involving morphological information is clearly better. Additionally, our proposal involving tensor reduction has the best performance using 26 features (i.e. 13 spectral dimension (k_b) by 2 spatial dimension (k_s)) with more than 75% in overall classification. This result is better than state of the art using spatial and spectral kernel [12]. Second scenario contains 16 classes using 20% of the labeled samples for training and the rest for validation in 10 repetitions. Results are presented in Fig. 2(b). Our approach produces excellent results (98.12% and Cohen's kappa coefficient $\kappa = 0.9645$) compared to state of the art methods [13] considering 20 ($k_b = 10$ and $k_s = 2$) features selected by tensor PCA using morphological information. Classifications are presented in Fig.3 showing clear improvement when using tensor PCA with morphological information. The optimum gaussian kernel parameter was

(a) $PCA(\mathcal{X})$ 6 features - 84.83% (b) $PCA(\mathcal{X}^\sigma)$ 16 features - 94.37%

(c) $\widehat{\mathcal{X}}_{I_1,I_2,20}$ 20 features - 93.64% (d) $\widehat{\mathcal{X}}^\sigma_{I_1,I_2,10,2}$ 20 features - 98.12%

Fig. 3. Scenario 2. Classification Map by SVM using **(a)** Standard PCA **(b)** Standard PCA in Morphological Multiscale Decomposition **(c)** Tensor PCA **(d)** Tensor PCA in Morphological Multiscale Decomposition. Accuracies are presented in Figure 2.

obtained through cross-validation $\widetilde{\rho} = 4$ and regularization parameter for SVMs, $C = 10^5$. In our experiments the value $I_1 = I_2 = 50$ were fixed and $\boldsymbol{\sigma} = (2, 4, \ldots, 10)$. An optimal parameter selection analysis can be done, however that is beyond the scope of this paper.

6. CONCLUSION

This paper proposed a framework to integrate spatial information into unsupervised feature extraction for hyperspectral images. Nonlinear Scale-Space representations with Morphological Levelings favourably incorporate spatial information before the classification stage. Tensor structures can integrate spatial information into feature extraction to drastically reduce the dimension without diminishing classifier performance. Results in real hyperspectral images show that this tensor approach incorporates more usefully spatial information in the dimension reduction stage than its matrix version equivalent. Results show that this produces better performance than current state of the art methods used for classification with spatial and spectral information [13, 12].

7. REFERENCES

[1] D. A. Landgrebe, Hyperspectral Image data analysis as a high dimensional signal processing problem. Special issue of the IEEE Signal Processing Magazine, 19 (1), 17-28, Jan. 2002.

[2] L.O. Jiménez, E. Arzuaga, and M. Vélez, Unsupervised Linear Feature-Extraction Methods and their Effects in the Classification of High-Dim. Data, IEEE Trans. Geosci. Remote Sens., 45(2), 469-483, Feb. 2007.

[3] N.Renard and S. Bourennane, Improvement of Target Detection Methods by Multiway Filtering, IEEE Trans. Geosci. Remote Sens., 46(8), Aug. 2008.

[4] M. Pesaresi and J.A. Benediktsson, A new approach for the morphological segmentation of high-resolution satellite imagery, IEEE Trans. Geosci. Remote Sens., 39(2), 309-320, Feb. 2001.

[5] J. A. Palmason, J.A. Benediktsson, J.R. Sviensson, and J. Chanussot, Classification of hypersepctral data from urban areas using morphological preprocessing and ICA, in proc. IGARSS,176-179, Jul. 2005.

[6] M. Fauvel, J.A. Benediktsson, J. Chanussot, J. Sveinsson, Spectral and Spatial Classification of Hypersepctral Data Using SVMs and Morphological Profiles, IEEE Trans. Geosci. Remote Sens., 46(11), Nov. 2008.

[7] L. Lathauwer, B. Moor, and J.Vandewalle, A Multilinear Singular Value decomposition, SIAM J. Matrix Anal. Appl.,21(4), 1253-1278, 2000.

[8] G.M. Foody and A. Mathur, A relative evaluation of multiclass image classification by support vector machines, IEEE Trans. Geosci. Remote Sens., 42(6), 1335-1343, Jun. 2004.

[9] T. G. Kolda and B. W. Bader, Tensor Decompositions and Applications, SIAM Review, 51(3), Sept. 2009 (to appear)

[10] F. Meyer and Petros Maragos, Nonlinear Scale-Space Representation with Morphological Levelings, Journal of Visual Communication and Image Representation 11(2), 245265,2000.

[11] S. Canu, Y. Grandvalet, V. Guigue and A. Rakotomamonjy, SVM and Kernel Methods Matlab Toolbox, INSA de Rouen, France, 2005.

[12] G. Camps-Valls, T. Bandos, and D. Zhou., Semi-supervised graph-based hyperspectral image classification, IEEE Trans. Geosci. Remote Sens., 53(10), 3044 - 3054, Oct. 2007.

[13] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, J.C. Tilton and G. Triani, Recent Advances in Techniques for Hyperspectral Image Processing, Remote Sensing Environment. Accepted for publication, 2009.