Multiway Filtering Based on Multilinear Algebra Tools

Salah Bourennane and Caroline Fossati

Ecole Centrale Marseille,
Institut Fresnel-UMR CNRS 61333 13397 Marseille cedex 20 France
Salah.bourennane@fresnel.fr, Caroline.fossati@fresnel.fr

Abstract

This paper presents some recent filtering methods based on the lower-rank tensor approximation approach for denoising tensor signals. In this approach, multicomponent data are represented by tensors, that is, multiway arrays, and the presented tensor filtering methods rely on multilinear algebra. First, the classical channel-by-channel SVD-based filtering method is overviewed. Then, an extension of the classical matrix filtering method is presented. It is based on the lower rank- $\left(K_1,\ldots,K_n\right)$ *truncation of the HOSVD which performs a multimode Principal Component Analysis (PCA) and is implicitly developed for an additive white Gaussian noise. Two tensor filtering methods recently developed by the authors are also overviewed. The performances and comparative results between all these tensor filtering methods are presented for the cases of noise reduction in color images.*

Keywords: *Tensor, Multiliear Algebra, Filtering, Image.*

1. Introduction

Tensor data modelling and tensor analysis have been improved and used in several application fields such as quantum physics, economy, chemometrics, psychology, data analysis, etc. Nevertheless, only recent studies focus their interest on tensor methods in signal processing applications. Tensor formulation in signal processing has received great attention since the recent development of multicomponent sensors, especially in imagery (color or multispectral images, video, etc.) and seismic fields (antenna of sensors recording waves with polarization properties). Indeed, the digital data obtained from these sensors are fundamentally higher order tensor objects, that is, multiway arrays whose elements are accessed via more than two indexes. Each index is associated with a dimension of the tensor generally called "n^{th}-mode" [7, 8, 15, 16]. For the last decades, the classical algebraic processing methods have been specifically developed for vector and matrix representations. They are usually based on the covariance matrix, the cross-spectral matrix, or more recently, on the higher order statistics. Their overall aim is classically to determine a subspace associated with the signal or the parameters to estimate. They mainly rely on 2 Matrix algebra-based DR methods three algebraic tools such as:

(1) The Singular Value Decomposition (SVD) [12], which is used in Principal Component Analysis (PCA).

(2) Penrose-Moore matrix inversion [12].

(3) The matrix lower rank approximation, which, according to Eckart-Young theorem [9], can be achieved thanks to a simple SVD truncation.

These methods have proven to be very efficient in several applications. When dealing with multicomponent data represented as tensors, the classical processing techniques consist in rearranging or splitting the data set into matrices or vectors in order for the previously quoted classical algebraic processing methods to be applicable. The original data structure is then built anew, after processing. In order to keep the data tensor as a whole entity, new signal processing methods have been proposed [19, 20]. Hence, instead of adapting the data tensor to the classical matrix-based algebraic techniques (by rearrangement or splitting), these new methods propose to adapt their processing to the tensor structure of the multicomponent data. This new approach implicitly implies the use of multilinear algebra and mathematical tools that extend the SVD to tensors. Two main tensor decomposition methods that generalize the matrix SVD have been initially developed in order to achieve a multimode Principal Component Analysis and recently used in tensor signal processing. They rely on two models, which are the TUCKER3 model and the PARAFAC model. These two decomposition methods differ in the tensor rank definition on which they are based. The HOSVD-$(K_1,...,K_n)$ and the rank-$(K_1,...,K_n)$ approximation rely on the n^{th}-mode rank definition, that is, the rank of the tensor n^{th}-mode flattening matrix [7, 8]. The rank-$(K_1,...,K_n)$ approximation [8] relies on an optimization algorithm which is initialized by the HOSVD-$(K_1,...,K_n)$ [7]. The rank-$(K_1,...,K_n)$ approximation improves the approximation obtained with the HOSVD-$(K_1,...,K_n)$. The goal of this paper is to present an overview of the principal results concerning this new approach of data tensor filtering. More details on the algorithms presented in this survey can be found in [19-22]. These algorithms are analogous to multilinear ICA, but were developed independently for image filtering. The presented algorithms are based on a signal subspace approach, so they are efficient when the noise components are uncorrelated, the signal and the additive noise are uncorrelated, and when some rows or columns of the image are redundant. In this case it is possible to distinguish between a signal subspace and a noise subspace, as for the traditional SVD-based filtering and Wiener filtering algorithms. Wiener filtering requires prior knowledge on the expected noise-free signal or image. However, multiway filtering methods provide the following advantage over traditional filtering methods: by apprehending a multiway data set as a whole entity, they take into account the dependence between modes thanks to ALS algorithms. The goal of the paper is also to present some simulations and comparative results concerning color images and multicomponent seismic signal filtering.

The paper is organized as follows: Section 2 presents the tensor data and a short overview of its main properties. Section 3 introduces the tensor formulation of the classical noise-removal problem as well as some new tensor filtering notations. Firstly, we explain how the channel-by-channel SVD-based method processes successively each component of the data tensor. Secondly, we consider two methods that take into account the relationships between each component of the considered tensor. These two methods are based on the n^{th}-mode signal subspace. The first method for signal tensor estimation is based on multimode PCA achieved by rank-$(K_1,...,K_n)$ approximation. The second method is a new tensor version of Wiener filtering. Section 4 presents some comparative results where the overviewed multiway filtering methods are applied to noise reduction in color images. Section 5 concludes the paper. The following notations are used in the rest of the paper: scalars are denoted by italic lowercase roman, like a; vectors by boldface lowercase roman, like $\mathbf{a}$; matrices by boldface uppercase roman, like $\mathbf{A}$; tensors by uppercase calligraphic, like $\mathcal{A}$. We distinguish a random vector, like $\mathbf{a}$, from one of its realizations, by using a supplementary index, like $\mathbf{a}_i$.

2. Tensor representation and properties

We define a tensor of order N as a multidimensional array whose entries are accessed via N indexes. A tensor is denoted by $\mathcal{A} \in \mathbf{R}^{I_1 \times \ldots \times I_N}$, where each element is denoted by $a_{i_1 \ldots i_N}$, and $\mathbf{R}$ is the real manifold. Each dimension of a tensor is called n^{th}-mode, where n refers to the n^{th} index. Fig. 1 shows how a color image can be represented by a third order tensor $\mathcal{A} \in \mathbf{R}^{I_1 \times I_2 \times I_3}$, where I_1 is the number of rows, I_2 is the number of columns, and I_3 is the number of color channels. In the case of a color image, we have $I_3 = 3$. Let us define $E^{(n)}$ as the n^{th}-mode vector space of dimension I_n associated with the n^{th}-mode of tensor $\mathcal{A}$. By definition, $E^{(n)}$ is generated by the column vectors of the n^{th}-mode flattening matrix.

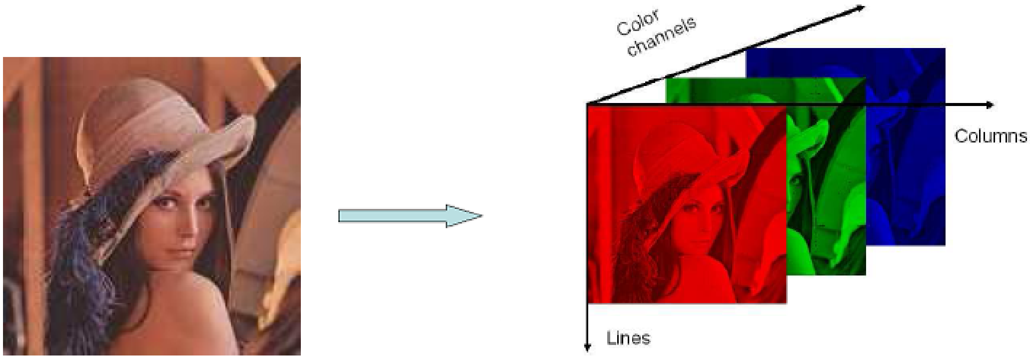

Figure 1. “Lena” standard color image and its tensor representation

The n^{th}-mode flattening matrix $\mathbf{A}_n$ of tensor $\mathcal{A} \in \mathbf{R}^{I_1 \times \ldots \times I_N}$ is defined as a matrix from $\mathbf{R}^{I_1 \times M_n}$, where:

$$M_n = I_{n+1} I_{n+2} \ldots I_N I_1 I_2 \ldots I_{n-1}, \quad (1)$$

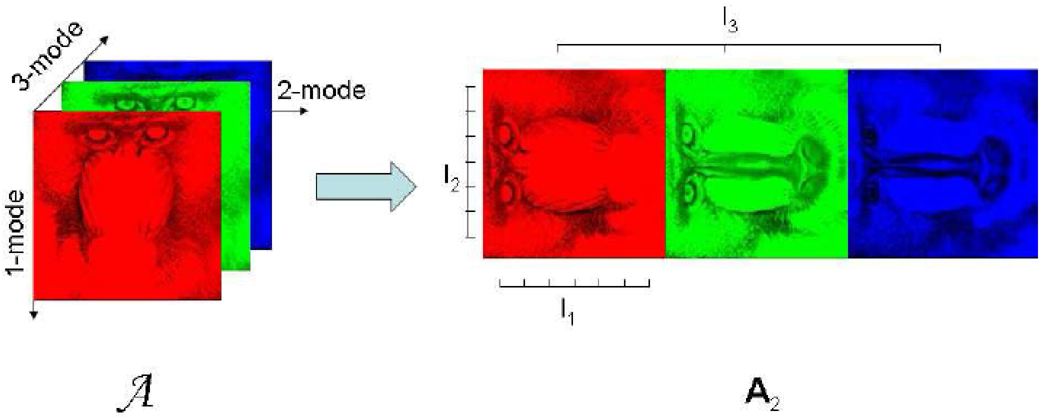

Figure 2. 2^{nd}-mode flattening of tensor A: A2

For example, when we consider a third-order tensor, the definition of the matrix flattening involves the dimensions I_1, I_2, I_3 in a backward cyclic way [4, 7,14]. When dealing with a 1^{st}-mode flattening of dimensionality $I_1 \times (I_2 I_3)$, we formally assume that the index i_2 varies more slowly than i_3. For all n = 1 to 3, $\mathbf{A}_n$ columns are the I_n-dimensional vectors obtained from $\mathcal{A}$ by varying the index i_n from 1 to I_n and keeping the other indexes fixed. These vectors are called the n^{th}-mode vectors of tensor $\mathcal{A}$. An illustration of the 2^{nd}-mode flattening of a color image is presented in Fig. 2.

In the following, we use the operator "$\times_n$" as the "n^{th}-mode product", that generalizes the matrix product to tensors. Given $\mathcal{A} \in \mathbf{R}^{I_1 \times \ldots \times I_N}$ and a matrix $\mathbf{U} \in \mathbf{R}^{J_n \times I_n}$, the n^{th}-mode product between tensor $\mathcal{A}$ and matrix $\mathbf{U}$ leads to the tensor $\mathcal{B} = \mathcal{A} \times_n \mathbf{U}$, which is a tensor of, $\mathbf{R}^{I_1 \times \ldots I_{n-1} \times J_n \times I_{n+1} \times \ldots \times I_N}$,whose entries are given by:

$$b_{i_1 \ldots i_{n-1} j_n i_{n+1} \ldots i_N} = \sum_{i_n=1}^{I_n} a_{i_1 \ldots i_{n-1} i_n i_{n+1} \ldots i_N} u_{j_n i_n} \ . \quad (2)$$

Next section presents the recent filtering methods for tensor data.

3. Tensor filtering problem formulation

The tensor data extend the classical vector data. The measurement of a multidimensional and multiway signal $\mathcal{X}$ by multicomponent sensors with additive noise $\mathcal{N}$, results in a data tensor $\mathcal{R}$ such that:

$$\mathcal{R}\mathcal{X} + \mathcal{N} \ . \quad (3)$$

$\mathcal{R}$, $\mathcal{X}$ and $\mathcal{N}$ are tensors of order N from $\mathbf{R}^{I_1 \times \ldots \times I_N}$. Tensors $\mathcal{N}$ and $\mathcal{X}$ represent noise and signal parts of the data respectively. The goal of this study is to estimate the expected signal $\mathcal{X}$ thanks to a multidimensional filtering of the data [19-22]:

$$\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \mathbf{H}^{(2)} \times_3 \ldots \times_N \mathbf{H}^{(N)}, \quad (4)$$

From a signal processing point of view, the n^{th}-mode product is a n^{th}-mode filtering of data tensor $\mathcal{R}$ by n^{th}-mode filter $\mathbf{H}^{(n)}$. Consequently, for all n = 1 to N, $\mathbf{H}^{(n)}$ is the n^{th}-mode filter applied to the n^{th}-mode of the data tensor $\mathcal{R}$.

In this paper we assume that the noise $\mathcal{N}$ is independent from the signal $\mathcal{X}$, and that the n^{th}-mode rank K_n is smaller than the n^{th}-mode dimension I_n, ($K_n < I_n$, for all n = 1 to N). Then it is possible to extend the classical subspace approach to tensors by assuming that, whatever the n^{th}-mode, the vector space $E^{(n)}$ is the direct sum of two orthogonal subspaces, namely $E_1^{(n)}$ and $E_2^{(n)}$, which are defined as follows:

- $E_1^{(n)}$ is the subspace of dimension K_n, spanned by the K_n singular vectors associated with the K_n largest singular values of matrix $\mathbf{X}_n$; $E_1^{(n)}$ is called the signal subspace [1, 17, 27, 26].

- $E_2^{(n)}$ is the subspace of dimension $I_n - K_n$, spanned by the $I_n - K_n$ singular vectors associated with the $I_n - K_n$ smallest singular values of matrix $\mathbf{X}_n$; $E_2^{(n)}$ is called the noise subspace [1, 17, 27, 26].

The dimensions $K_1, K_2, ..., K_n$ can be estimated by means of the well-known AIC (Akaike Information Criterion) or MDL (Minimum Description Length) criteria [25], which are entropy-based information criteria. Hence, one way to estimate signal tensor $\mathcal{X}$ from noisy data tensor $\mathcal{R}$ is to estimate $E_1^{(n)}$ in every n^{th}-mode of $\mathcal{R}$.

The following section presents three tensor filtering methods based on n^{th}-mode signal subspaces. The first method is an extension of classical matrix filtering algorithms. It consists of a channel-by-channel SVD-based filtering. The second filtering method is based on multimode PCA achieved by rank-($K_1, K_2, ..., K_n$) approximation.

3.1. Channel-by-channel SVD-based filtering

The classical algebraical methods operate on two-dimensional data matrices and are based on the Singular Value Decomposition (SVD) [1-3], and on Eckart-Young theorem concerning the best lower rank approximation of a matrix [9] in the least-squares sense. In the first method, a preprocessing is applied to the multidimensional and multiway data. It consists in splitting data tensor $\mathcal{R}$, representing the noisy multicomponent image into two-dimensional "slice matrices" of data, each representing a specific channel. According to the classical signal subspace methods [6], the left and right signal subspaces, corresponding to respectively the column and the row vectors of each slice matrix, are simultaneously determined by processing the SVD of the matrix associated with the data of the slice matrix.

Channel-by-channel SVD-based filtering is based on a common efficient method, but exhibits a major drawback: it does not take into account the relationships between the components of the processed tensor. Moreover, channel-by-channel SVD-based filtering is appropriate only on some conditions. For example, applying SVD-based filtering to an image is generally appropriate when the rows or columns of an image are redundant, that is, linearly dependent. In this case, the rank K of the image is equal to the number of linearly independent rows or columns. It is only in this case that it would be safe to throw out eigenvectors from $K+1$ on. It is only in this special case that the noise subspace is orthogonal to the signal subspace. Otherwise, the noise simply increases the variance of the signal subspace and underestimating the signal subspace dimension would result in throwing out both signal and noise information. Thus, one would lose spatial resolution.

The next subsection presents a multiway filtering method that processes jointly, and not successively, each component of the data tensor.

3.2 Tensor filtering based on multimode PCA

Assuming that the dimension K_n of the signal subspace is known for all n = 1 to N, one way to estimate the expected signal tensor $\mathcal{X}$ from the noisy data tensor $\mathcal{R}\mathcal{X} + \mathcal{N}$, is to orthogonally project, for every n^{th}-mode, the vectors of tensor $\mathcal{R}$ on the n^{th}-mode signal subspace $E_1^{(n)}$, for all n = 1 to N. This statement is equivalent to replace in (4) the filters $\mathbf{H}^{(n)}$ by the projectors $\mathbf{P}^{(n)}$ on the n^{th}-mode signal subspace:

$$\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{P}^{(1)} \times_2 \ldots \times_N \mathbf{P}^{(N)}. \tag{5}$$

In this last formulation, projectors $\mathbf{P}^{(n)}$ are estimated thanks to a multimode PCA applied to data tensor $\mathcal{R}$. This multimode PCA-based filtering generalizes the classical matrix filtering methods [10, 11, 13], and implicitly supposes that the additive noise is *white* and *Gaussian*.

In the vector or matrix formulation, the definition of the projector on the signal subspace is based on the eigenvectors associated with the largest eigenvalues of the covariance matrix of the set of observation vectors. Hence, the determination of the signal subspace amounts to determine the best approximation (in the least-squares sense) of the observation matrix or the covariance matrix.

As an extension to the vector and matrix cases, in the tensor formulation, the projectors on the n^{th}-mode vector spaces are determined by computing the rank-($K_1, K_2, \ldots, K_n$) approximation of $\mathcal{R}$ in the least-squares sense. From a mathematical point of view, the rank-($K_1, K_2, \ldots, K_n$) approximation of $\mathcal{R}$ is represented by tensor $\mathcal{R}^{K_1,\ldots,K_N}$ which minimizes the quadratic tensor Frobenius norm $\left\|\mathcal{R} - \mathcal{B}\right\|^2$ subject to the condition that $\mathcal{B} \in \mathbf{R}^{I_1 \times \ldots \times I_N}$ rank-($K_1, K_2, \ldots, K_n$) tensor. The description of TUCKALS3 algorithm, used in rank-($K_1, K_2, \ldots, K_n$) approximation is provided in the following.

Rank-$\left(K_1, \ldots, K_n\right)$ approximation - TUCKALS3 algorithm:

---

1) Input: data tensor $\mathcal{R}$, and dimensions $\left(K_1, \ldots, K_n\right)$ of all n^{th}-mode signal subspaces.

2) Initialization k = 0: For n = 1 to N, calculate the projectors $\mathbf{P}_0^{(n)}$ given by HOSVD-$\left(K_1, \ldots, K_n\right)$:

(a) n^{th}-mode flatten $\mathcal{R}$ into matrix $\mathbf{R}_n$;

(b) Compute the SVD of $\mathbf{R}_n$;

(c) Compute matrix $\mathbf{U}_0^{(n)}$ formed by the K_n eigenvectors associated with the K_n largest singular values of $\mathbf{R}_n$. $\mathbf{U}_0^{(n)}$ is the initial matrix of the n^{th}-mode signal subspace orthogonal basis vectors;

(d) Form the initial orthogonal projector $\mathbf{P}_0^{(n)} = U_0^{(n)} U_0^{(n)T}$ on the n^{th}-mode signal subspace;

(e) Compute the HOSV-$\left(K_1, \ldots, K_n\right)$ of tensor $\mathcal{R}$ given by $\mathcal{B}_0 = \mathcal{R} \times_1 \mathbf{P}_0^{(1)} \times_2 \ldots \times_N \mathbf{P}_0^{(N)}$;

3) ALS loop: Repeat until convergence, that is, for example, while $\left\|\mathcal{B}_{k+1} - \mathcal{B}_k\right\|^2 < \varepsilon$, $\varepsilon > 0$ being a prior fixed threshold,

(a) For n = 1 to N:

i. Form $\mathcal{B}^{(n),k}$: $\mathcal{B}^{(n),k} = \mathcal{R} \times_1 \mathbf{P}_{k+1}^{(1)} \times_2 \ldots \times_{n-1} \mathbf{P}_{k+1}^{(n-1)} \times_{n+1} \mathbf{P}_k^{(n+1)} \times_{n+2} \ldots \times_N \mathbf{P}_k^{(N)}$

ii. n^{th}-mode flatten tensor $\mathcal{B}^{(n),k}$ into matrix $\mathbf{B}_n^{(n),k}$;

iii. Compute matrix $\mathbf{C}^{(n),k} = \mathbf{B}_n^{(n),k} \mathbf{R}_n^{\ T}$;

iv. Compute matrix $\mathbf{U}_{k+1}^{(n)}$ composed of the K_n eigenvectors associated with the K_n largest eigenvalues of $\mathbf{C}^{(n),k}.\mathbf{U}_k^{(n)}$ is the matrix of the n^{th}-mode signal subspace orthogonal basis vectors at the k^{th} iteration;

v. Compute $\mathbf{P}_{k+1}^{(n)} = \mathbf{U}_{k+1}^{(n)} \mathbf{U}_{k+1}^{(n)\,T}$;

(b) Compute $\mathcal{B}_{k+1} = \mathcal{R} \times_1 \mathbf{P}_{k+1}^{(1)} \times_2 \ldots \times_N \mathbf{P}_{k+1}^{(N)}$;

(c) Increment k.

4) Output: the estimated signal tensor is obtained through $\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{P}_{k_{stop}}^{(1)} \times_2 \ldots \times_N \mathbf{P}_{k_{stop}}^{(N)}$. $\hat{\mathcal{X}}$ is the rank-$(K_1,\ldots,K_n)$ approximation of $\mathcal{R}$, where k_{stop} is the index of the last iteration after the convergence of TUCKALS3 algorithm.

In this algorithm, the second order statistics comes from the SVD of matrix $\mathbf{R}_n$ at step 2b, which is equivalent, up to $1/M_n$ multiplicative factor, to the estimation of tensor $\mathcal{R}$ n^{th}-mode vectors [22]. The definition of M_n is given in (1). In the same way, at step 3(a)iii, matrix $\mathbf{C}^{(n),k}$ is up to $1/M_n$ multiplicative factor, the estimation of the covariance matrix between tensor $\mathcal{R}$ and tensor $\mathcal{B}^{(n),k}$ n^{th}-mode vectors. According to step 3(a)i, $\mathcal{B}^{(n),k}$ represents data tensor $\mathcal{R}$ filtered in every m^{th}-mode but the n^{th}-mode, by projection-filters $P_l^{(m)}$, with $m \neq n$, $l = k$ if $m > n$ and $l = k+1$ if $m < n$. TUCKALS3 algorithm has recently been used to process a multimode PCA in order to perform white noise removal in color images [20].

A good approximation of the rank-$(K_1,\ldots,K_n)$ approximation can simply be achieved by computing the HOSVD-$(K_1,\ldots,K_n)$ of tensor $\mathcal{R}$ [8, 18]. Indeed, the HOSVD-$(K_1,\ldots,K_n)$ of $\mathcal{R}$ consists of the initialization step of TUCKALS3 algorithm, and hence can be considered as a suboptimal solution for the rank-$(K_1,\ldots,K_n)$ approximation of tensor $\mathcal{R}$ [8].

3.3. Multiway Wiener filtering

Let $\mathbf{R}_n$, $\mathbf{X}_n$ and $\mathbf{N}_n$ be the n^{th}-mode flattening matrices of tensors $\mathcal{R}$, $\mathcal{X}$ and $\mathcal{N}$ respectively.

In the previous subsection, the estimation of signal tensor $\mathcal{X}$ has been performed by projecting noisy data tensor $\mathcal{R}$ on each n^{th}-mode signal subspace. The n^{th}-mode projectors have been estimated thanks to the use of multimode PCA achieved by rank-$(K_1,\ldots,K_n)$ approximation. In spite of the good results given by this method, it is possible to improve the tensor filtering quality by determining n^{th}-mode filters $\mathbf{H}^{(n)}$, n = 1 to N, in (4), which optimize an estimation criterion. The most classical method is to minimize the mean squared error between the expected signal tensor $\mathcal{X}$ and the estimated signal tensor $\hat{\mathcal{X}}$ given in (4):

$$e\left(\mathbf{H}^{(1)},\ldots,\mathbf{H}^{(N)}\right) = E\left[\left\|\mathcal{X} - \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \ldots \times_N \mathbf{H}^{(N)}\right\|^2\right]. \quad (6)$$

Due to the criterion which is minimized, filters $\mathbf{H}^{(n)}$, n = 1 to N, can be called "n^{th}-mode Wiener filters" [21].

According to the calculations presented in [21], the minimization of (6) with respect to filter $\mathbf{H}^{(n)}$, for fixed $\mathbf{H}^{(m)}$, $m \neq n$, leads to the following expression of n^{th}-mode Wiener filter:

$$\mathbf{H}^{(\mathrm{n})} = \gamma_{\mathbf{XR}}^{(n)} \mathbf{\Gamma}_{RR}^{(n)-1}, \qquad (7)$$

where

$$\gamma_{\mathbf{XR}}^{(n)} = E\left[\mathbf{X}_n \mathbf{T}^{(n)} \mathbf{R}_n{}^T\right] \qquad (8)$$

is the $T^{(n)}$-weighted covariance matrix between the random column vectors of signal $\mathbf{X}_n$ and data $\mathbf{R}_n$, with:

$$\mathbf{T}^{(n)} = \mathbf{H}^{(1)} \otimes \ldots \otimes \mathbf{H}^{(n-1)} \otimes \mathbf{H}^{(n+1)} \otimes \ldots \otimes \mathbf{H}^{(N)}, \qquad (9)$$

where $\otimes$ stands for Kronecker product, and :

$$\mathbf{\Gamma}_{\mathbf{RR}}^{(n)} = E\left[\mathbf{R}_n \mathbf{Q}^{(n)} \mathbf{R}_n{}^T\right], \qquad (10)$$

is the $\mathbf{Q}^{(n)}$-weighted covariance matrix of the data $\mathbf{R}_n$, with:

$$\mathbf{Q}^{(n)} = \mathbf{T}^{(n)T} \mathbf{T}^{(n)}. \qquad (11)$$

In order to obtain $\mathbf{H}^{(n)}$ through (7), we suppose that the filters $\left\{\mathbf{H}^{(m)}, m = 1 \text{ to } N, m \neq n\right\}$ are known. Data tensor $\mathcal{R}$ is available, but signal tensor $\mathcal{X}$ is unknown. So, only the term $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$ can be derived, and not the term $\gamma_{\mathbf{XR}}^{(n)}$. Hence, more assumptions have to be made on $\mathcal{X}$ in order to overcome the indetermination over $\gamma_{\mathbf{XR}}^{(n)}$ [19, 21]. In the one-dimensional case, a classical assumption is to consider that a signal vector is a weighted combination of the signal subspace basis vectors. In extension to the tensor case, [19, 21] have proposed to consider that the n^{th}-mode flattening matrix $\mathbf{X}_n$ can be expressed as a weighted combination of K_n vectors from the n^{th}-mode signal subspace $E_1^{(n)}$:

$$\mathbf{X}_n \mathbf{V}_s^{(n)} \mathbf{O}^{(n)}, \qquad (12)$$

with $\mathbf{X}_n \in \mathbf{R}^{I_n \times M_n}$ and $\mathbf{V}_s^{(n)} \in \mathbf{R}^{I_n \times K_n}$ being the matrix containing the K_n orthonormal basis vectors of n^{th}-mode signal subspace $E_1^{(n)}$. Matrix $\mathbf{O}^{(n)} \in \mathbf{R}^{K_n \times M_n}$ is a weight matrix and contains the whole information on expected signal tensor $\mathcal{X}$. This model implies that signal n^{th}-mode flattening matrix $\mathbf{X}_n$ is orthogonal to n^{th}-mode noise flattening matrix $\mathbf{N}_n$, since signal subspace $E_1^{(n)}$ and noise subspace $E_2^{(n)}$ are supposed mutually orthogonal.

Supposing that noise $\mathcal{N}$ in (3) is white, Gaussian and independent from signal $\mathcal{X}$, and introducing the signal model (12) in (7) leads to a computable expression of n^{th}-mode Wiener filter $\mathbf{H}^{(n)}$:

$$\mathbf{H}^{(n)} = \mathbf{V}_s^{(n)} \gamma_{\mathbf{OO}}^{(n)} \Lambda_{\mathbf{\Gamma s}}^{(n)-1} \mathbf{V}_s^{(n)T}, \qquad (13)$$

where $\gamma_{\mathbf{OO}}^{(n)}\Lambda_{\mathbf{\Gamma s}}^{(n)-1}$ is a diagonal weight matrix given by:

$$\gamma_{\mathbf{OO}}^{(n)}\Lambda_{\mathbf{\Gamma s}}^{(n)-1} = diag\left[\frac{\beta_1}{\lambda_1^{\Gamma}},...,\frac{\beta_{Kn}}{\lambda_{Kn}^{\Gamma}}\right], \tag{14}$$

where , $\lambda_1^{\Gamma},...,\lambda_{Kn}^{\Gamma}$ are the K_n largest eigenvalues of $\mathbf{Q}^{(n)}$-weighted covariance matrix $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$ (see (10)). Parameters $\beta_1,...,\beta_{Kn}$ depend on , $\lambda_1^{\gamma},...,\lambda_{Kn}^{\gamma}$ which are the K_n largest eigenvalues of $\mathbf{T}^{(n)}$-weighted covariance matrix $\gamma_{\mathbf{RR}}^{(n)} = E\left[\mathbf{R}_n\mathbf{T}^{(n)}\mathbf{R}_n^{\ T}\right]$, according to the following relation:

$$\beta_{k_n} = \lambda_{k_n}^{\gamma} - \sigma_{\Gamma}^{(n)2}, \forall k_n = 1,...,K_n \tag{15}$$

Superscript γ refers to the $\mathbf{T}^{(n)}$-weighted covariance, and subscript Γ to the $\mathbf{Q}^{(n)}$- weighted covariance. $\sigma_{\Gamma}^{(n)2}$ is the degenerated eigen value of noise $\mathbf{T}^{(n)}$-weighted covariance matrix $\gamma_{\mathbf{NN}}^{(n)} = E\left[\mathbf{N}_n\mathbf{T}^{(n)}\mathbf{N}_n^{\ T}\right]$. Thanks to the additive noise and the signal independence assumptions, the $I_n - K_n$ smallest eigenvalues of $\gamma_{\mathbf{RR}}^{(n)}$ are equal to $\sigma_{\Gamma}^{(n)2}$, and thus, can be estimated by the following relation:

$$\hat{\sigma}_{\Gamma}^{(n)2} = \frac{1}{I_n - K_n}\sum_{k_n=K_n+1}^{I_n}\lambda_{k_n}^{\gamma} \tag{16}$$

In order to determine the nth-mode Wiener filters $\mathbf{H}^{(n)}$ that minimize the mean squared error (6), the Alternating Least Squares (ALS) algorithm has been proposed in [19, 21]. It can be summarized in the following steps:

Alternative Least Squares (ALS) Algorithm

---

1) Initialization $k = 0$: $\mathcal{R}^0 = \mathcal{R} \Leftrightarrow \mathbf{H}_0^{(n)} = \mathbf{I}_{In}$, Identity matrix $\forall n = 1,...,N$

2) ALS loop: Repeat until convergence, that is $\left\|\mathcal{R}_{k+1} - \mathcal{R}_k\right\|^2 < \varepsilon$, with $\varepsilon > 0$ prior fixed threshold,

(a) For n = 1 to N:

i. Form $\mathcal{R}^{(n),k}$: $\mathcal{R}^{(n),k} = \mathcal{R}\times_1 \mathbf{H}_{k+1}^{(1)} \times_2 ...\times_{n-1} \mathbf{H}_{k+1}^{(n-1)} \times_{n+1} \mathbf{H}_k^{(n+1)} \times_{n+2} ...\times_N \mathbf{H}_k^{(N)}$

ii. determine $\mathbf{H}_{k+1}^{(n)} = \mathbf{Z}^{(n)} \arg\min\left\|\mathcal{X} - \mathcal{R}^{(n),k} \times_n \mathbf{Z}^{(n)}\right\|^2$ subject to $\mathbf{Z}^{(n)} \in \mathbf{R}^{I_n \times I_n}$ thanks to the following procedure:

A. n^{th}-mode flattened $\mathcal{R}^{(n),k}$ into $\mathbf{R}_n^{\ (n),k} = \mathbf{R}_n(\mathbf{H}_{k+1}^{(1)} \otimes ... \otimes \mathbf{H}_{k+1}^{(n-1)} \otimes \mathbf{H}_k^{(n+1)} \otimes ... \otimes \mathbf{H}_k^{(N)})^T$, and $\mathcal{R}$ into $\mathbf{R}_n$;

B. compute $\gamma_{\mathbf{RR}}^{(n)} = E\left[\mathbf{R}_n\mathbf{R}_n^{(n),k\,T}\right]$,

C. determine $\lambda_1^{\gamma},...,\lambda_{Kn}^{\gamma}$, the K_n largest eigenvalues of $\gamma_{\mathbf{RR}}^{(n)}$;

D. for $k_n = 1$ to I_n, estimate $\sigma_{\Gamma}^{(n)2}$ thanks to (16), and for $k_n = 1$ to K_n, estimate β_{k_n} thanks to (15);

E. compute $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)} = E\left[\mathbf{R}_n^{(n),k}\mathbf{R}_n^{(n),k\,T}\right]$;

F. determine $\lambda_1^\Gamma,...,\lambda_{Kn}^\Gamma$, the K_n largest eigenvalues of $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$;

G. determine $\mathbf{V}_s^{(n)}$, the matrix of thee K_n eigenvectors associated with the K_n largest eigenvalues of $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$;

H. compute the weight matrix $\gamma_{\mathbf{OO}}^{(n)}\Lambda_{\mathbf{\Gamma s}}^{(n)-1}$ given in (14);

I. compute $\mathbf{H}_{k+1}^{(n)}$, the n^{th}-mode Wiener filter at the (k+1)$^{\text{th}}$ iteration, using (13);

(b) Form: $\mathcal{R}^{k+1} = \mathcal{R} \times_1 \mathbf{H}_{k+1}^{(1)} \times_2 ... \times_N \mathbf{H}_{k+1}^{(N)}$;

(c) Increment k.

3) **output**: $\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}_{k_{stop}}^{(1)} \times_2 ... \times_N \mathbf{H}_{k_{stop}}^{(N)}$, with k_{stop} being the last iteration after convergence of the algorithm.

4. Simulation results

In the following simulations, the channel-by-channel SVD-based filtering defined in subsection 3.1 and the rank-$(K_1,...,K_n)$ approximation based multiway and multidimensional filtering are applied to the denoising of color images and multispectral images, and to the denoising of seismic signals. Color images, multispectral images, and seismic signals can be represented by a third order tensor from $\mathbf{R}^{I_1 \times I_2 \times I_3}$, where I_1, I_2 and I_3 take different values. In all these applications, the efficiency of denoising is tested in the presence of an additive Gaussian noise.

A multidimensional and multiway white Gaussian noise $\mathcal{N}$ which is added to signal tensor $\mathcal{X}$ can be expressed as:

$$\mathcal{N}\alpha.\mathcal{G}, \tag{17}$$

where every element of $\mathcal{G} \in \mathbf{R}^{I_1 \times I_2 \times I_3}$ is an independent realization of a normalized centered Gaussian law, and α is a coefficient that permits to set the SNR in noisy data tensor $\mathcal{R}$.

4.1. Performance criterion

Following the representation of (3), the multiway noisy data tensor is expressed as $\mathcal{R}\mathcal{X} + \mathcal{N}$, where $\mathcal{X}$ is the expected signal tensor and $\mathcal{N}$ is the additive noise tensor. Let us define the Signal to Noise Ratio (SNR, in dB) in the noisy data tensor by:

$$SNR = 10\log\left(\frac{\|\mathcal{X}\|^2}{\|\mathcal{N}\|^2}\right). \tag{18}$$

4.2 Denoising of color images

Denoising of color images has already been studied in several works [5, 23, 24]. Some solutions have been brought from the field of wavelet processing, exhibiting good results in terms of output SNR. These studies only concern bidimensional data, whereas the methods that we compare are adapted to the processing of third order tensors as a whole, and in particular to three-channel images. We focus on subspace-based methods. We first consider the channel-by-channel SVD-based filtering, the rank-(K_1, K_2, K_3) approximation and multiway Wiener filtering (Wmm-(K_1, K_2, K_3)), applied to images impaired by an additive white Gaussian noise.

The channel-by-channel SVD-based filtering of noisy image $\mathcal{R}$ (see Fig. 3) yields the image of Fig. 3c, and rank-(30, 30, 2) approximation of noisy data tensor $\mathcal{R}$ yields the image of Fig. 3d. From the resulting image, presented on Fig.3, we note that dimension reduction leads to a loss of spatial resolution.

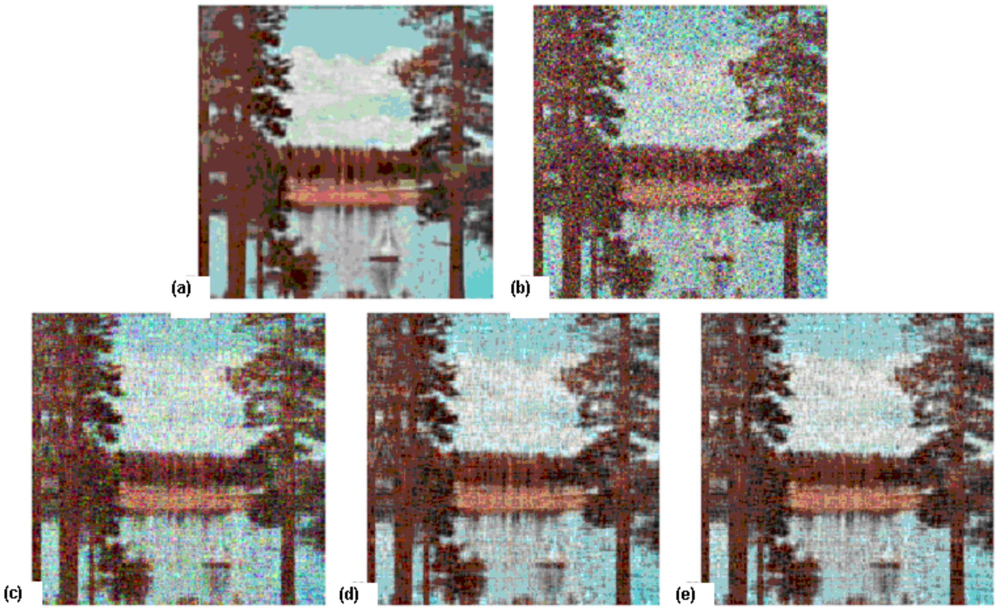

Figure 3. (a) Non-noisy image. (b) Image to be processed, impaired by an additive white Gaussian noise, with SNR=8.1dB. (c) Channel-by-channel SVD-based filtering of parameter K=30. (d) Rank-(30; 30; 2) approximation. (e) Wmm-(30; 30; 2) filtering.

However the choice of a set of values K_1, K_2, K_3 which are small enough is the condition for an efficient noise reduction effect.

Concerning the qualitative results obtained with this color image, we note that the intra-class variance of the pixel values of each component (or color mode) of the resulting image is lower for the image obtained with Wmm-(30, 30, 2) (see Fig. 3e) than for those images obtained with other methods applied in this subsection. This permits, for example, to apply after denoising a high level classification method with a higher efficiency than when classification is applied after channel-by-channel SVD-based filtering or HOSVD-(30, 30, 2).

For the $256 \times 256 \times 3$ Sailboat image of Fig. 3, the computational times needed when *Matlab*® programs are used on a 3Ghz Pentium 4 processor running Windows are as follows. HOSVD-(30, 30, 2) lasts 1.61 sec., the channel-by-channel SVD-based filtering lasts 1.94 sec., the rank-(30, 30, 2) approximation run with 25 iterations lasts 54.1 sec. and Wmm-(30, 30, 2) run with 25 iterations lasts 40.0 sec.

According to the simulations performed on a color image, it is possible to conclude that the more channels the image is composed of, the better the denoising.

5. Conclusion

In this paper, an overview on new mathematical methods dedicated to multicomponent data is presented. Multicomponent data are represented as tensors, that is, multiway arrays, and the tensor filtering methods that are presented rely on multilinear algebra. First we present how to perform channel-by-channel SVD-based filtering. Then we review three methods that take into account the relationships between each component of a processed tensor. The first method consists of an extension of the classical SVD-based filtering method. In the case of an additive white Gaussian noise, the signal tensor is estimated thanks to a multimode PCA achieved by applying a lower rank-$(K_1,...,K_n)$ approximation to the noisy data tensor, or a lower rank-$(K_1,...,K_n)$ truncation of its HOSVD. This method is implicitly based on second order statistics and relies on the orthogonality between n^{th}-mode noise and signal subspaces.

Finally, the reviewed method is a multiway version of the classical Wiener filtering. In extension to the one-dimensional case, the n^{th}-mode Wiener filters are estimated by minimizing the mean squared error between the expected signal tensor and the estimated signal tensor obtained by applying the n^{th}-mode Wiener filters to the noisy data tensor thanks to the n^{th}-mode product operator. An Alternating Least Squares algorithm has been presented to determine the optimal n^{th}-mode Wiener filters. The performances of this multiway Wiener filtering and comparative results with multimode PCA have been presented in the case of additive white noise reduction in a color image.

References

[1] K. Abed-Meraim, H. Maître, and P. Duhamel. “Blind multichannel image restoration using subspace based method.” In IEEE International Conf. on Accoustics Systems and Signal Processing, Hong Kong, China, April 6-10, 2003.

[2] H.C. Andrews and C. L. Patterson. “Singular value decomposition and digital image processing.” IEEE Trans. on Acoustics, Speech, and Signal Processing, 24(1):26-53,February 1976.

[3] H.C. Andrews and C. L. Patterson. “Singular Value Decomposition (SVD) image coding.” IEEE Trans. on Communications, pages 425- 432, April 1976.

[4] B.W. Bader and T.G. Kolda. “Algorithm 862: Matlab tensor classes for fast algorithm prototyping.” ACM Transactions on Mathematical Software,, 32(4), December 2006.

[5] E. Bala and A. Ertzn. “A multivariate thresholding technique for image denoising using multiwavelets.” EURASIP Journal on ASP, pages 1205-1211, 08 2005.

[6] Benjama A. Bourennane S. and Frikel M. “Seismic wave separation based on higher order statistics.” In IEEE International Conf. on Digital Signal Processing and its Applications, Moscow, Russia, 1998.

[7] L. De Lathauwer, B. De Moor, and J. Vandewalle. “A multilinear singular value decomposition.” SIAM Journal on Matrix Analysis and Applications, 21:1253-1278, April 2000.

[8] L. De Lathauwer, B. De Moor, and J. Vandewalle. “On the best rank-(1) and rank-(r1; : : : ; rN) approximation of higher-order tensors.” SIAM Journal on Matrix Analysis and Applications, 21(4):1324-1342, April 2000.

[9] Y. Eckart and G. Young. “The approximation of a matrix by another of lower rank.” Psychometrika, 1:211-218, 1936.

[10] F.L.M. Freire and T.J. Ulrych. “Application of SVD to vertical seismic profiling.” Geophysics, 53:778-785, 1988.

[11] F. Glangeaud and J.-L. Mari. “Wave separation.” Technip IFP edition, 1994.

[12] G.H. Golub and C.F. Van Loan. “Matrix computations” 3rd ed. Baltimore, The John Hopkins University Press edition, 1996.

[13] M. Hemon and D. Mace. “The use of Karhunen-Loeve transform in seismic data prospecting.” Geophysical Prospecting, 26:600-626, 1978.

[14] H. Kiers. “Towards a standardized notation and terminology in multiway analysis.” Journal of Chemometrics, 14:105-122, 2000.

[15] P.M. Kroonenberg. “Three-mode principal component analysis.” DSWO press, Leiden, 1983.

[16] P.M. Kroonenberg and J. De Leeuw. “Principal component analysis of three-mode data by means of alternating least squares algorithms.” Psychometrika, 45(1):69-97, March 1980.

[17] J.M. Mendel. “Tutorial on higher order statistics (spectra) in signal processing and system theory: theoretical results and some applications.” In Proc. of the IEEE, volume 79, pp 278-305, March 1991.

[18] D. Muti. “Traitement du signal tensoriel. Application aux images en couleurs et aux signaux sismiques. ” Phd thesis, Institut Fresnel, Universite Paul Cezanne, Aix-Marseille III, Marseille, France, 2 december 2004.

[19] D. Muti and S. Bourennane. “Multidimensional estimation based on a tensor decomposition.” In IEEE Workshop on Statistical Signal Processing, St Louis, USA, Sept. 28 - Oct. 1, 03.

[20] D. Muti and S. Bourennane. “Multidimensional signal processing using lower rank tensor approximation.” In IEEE Int. Conf. on Accoustics, Systems and Signal Processing, Hong Kong, China, April 6-10, 2003.

[21] D. Muti and S. Bourennane. “Multidimensional filtering based on a tensor approach.” Signal Proceesing Journal, Elsevier, 85:2338-2353, May 2005.

[22] D. Muti and S. Bourennane. “Multiway filtering based on fourth order cumulants.” Applied Signal Processing, EURASIP, 7:1147-1159, May 2005.

[23] R. Neelamani, H. Choi, and R. Baraniuk. “Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems.” IEEE Transactions on SP, 52(2):418-433, Feb. 2004.

[24] L.W. She and B. Zheng. “Multiwavelets based denoising of sar images.” In Proceedings 5th international Conference on Signal Processing, pages 321-324, Beijing, China, 08 2000.

[25] M. Wax and T. Kailath. “Detection of signals information theoretic criteria.” IEEE Trans. on Acoust., Speech, Signal Processing, 33(2):387-392, April 1985.

[26] N. Yuen and B. Friedlander. “DOA estimation in multipath: an approach using fourth order cumulant.” IEEE Trans. on Signal Processing, 45(5):1253-63, 1997.

[27] N. Yuen and B. Friedlander. “Asymptotic performance analysis of blind signal copy using fourth order cumulant.” International Jour. Adaptative Contr. Signal Proc.,48:239-65, Mar. 96.