

Research Article

About Advances in Tensor Data Denoising Methods

Julien Marot, Caroline Fossati, and Salah Bourennane

Institut Fresnel CNRS UMR 6133, Ecole Centrale Marseille, Université Paul Cézanne, D.U. de Saint Jérôme, 13397 Marseille Cedex 20, France

Correspondence should be addressed to Salah Bourennane, salah.bourennane@fresnel.fr

Received 15 December 2007; Revised 15 June 2008; Accepted 31 July 2008

Recommended by Lisimachos P. Kondi

Tensor methods are of great interest since the development of multicomponent sensors. The acquired multicomponent data are represented by tensors, that is, multiway arrays. This paper presents advances on filtering methods to improve tensor data denoising. Channel-by-channel and multiway methods are presented. The first multiway method is based on the lower-rank (K_1, \dots, K_N) truncation of the HOSVD. The second one consists of an extension of Wiener filtering to data tensors. When multiway tensor filtering is performed, the processed tensor is flattened along each mode successively, and singular value decomposition of the flattened matrix is performed. Data projection on the singular vectors associated with dominant singular values results in noise reduction. We propose a synthesis of crucial issues which were recently solved, that is, the estimation of the number of dominant singular vectors, the optimal choice of flattening directions, and the reduction of the computational load of multiway tensor filtering methods. The presented methods are compared through an application to a color image and a seismic signal, multiway Wiener filtering providing the best denoising results. We apply multiway Wiener filtering and its fast version to a hyperspectral image. The fast multiway filtering method is 29 times faster and yields very close denoising results.

Copyright © 2008 Julien Marot et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Tensor data modelling and tensor analysis have been improved and used in several application fields. These application fields are quantum physics, economy, psychology, data analysis, chemometrics [1]. Specific applications are the characterization of DS-CDMA systems [2], and the classification of facial expressions. For this application, a multilinear independent component analysis [3] was created. Another specific application is in particular the processing and visualization of medical images obtained through magnetic resonance imaging [4].

Tensor data generalize the classical vector and matrix data to entities with more than two dimensions [1, 5, 6]. In signal processing, there was a recent development of multicomponent sensors, especially in imagery (color or multispectral images, video, etc.) and seismic fields (an antenna of sensors selects and records signals of a given polarization). The digital data obtained from these sensors are fundamentally multiway arrays, which are called, in the signal processing community and in this paper in particular, higher-order tensor objects, or tensors. Each multiway array entry corresponds to any quantity. The

elements of a multiway array are accessed via several indexes. Each index is associated with a dimension of the tensor generally called “ n th-mode” [5, 7–10]. Measured data are not fully reliable since any real sensor will provide noisy and possibly incomplete and degraded data. Therefore, all problems dealt with in conventional signal processing such as filtering, restoration from noisy data must also be addressed when dealing with tensor signals [6, 11].

In order to keep the data tensor as a whole entity, new signal processing methods have been proposed [12–15]. Hence, instead of adapting the data tensor to the classical matrix-based algebraic techniques [16, 17] (by rearrangement or splitting), these new methods propose to adapt their processing to the tensor structure of the multicomponent data. Multilinear algebra is adapted to multicomponent data. In particular, it involves two tensor decomposition models. They generalize that the matrix SVD has been initially developed in order to achieve a multimode principal component analysis and recently used in tensor signal processing. They rely on two models: PARAFAC and TUCKER3 models.

(1) The PARAFAC model and the CANDECOMP model developed in [18, 19], respectively. In [20], the link was

set between CANDECOMP and PARAFAC models. The CANDECOMP/PARAFAC model, referred to as the CP model [21], has recently been applied to food industry [22], array processing [23], and telecommunications [2]. PARAFAC decomposition of a tensor containing data received on an array of sensors yields strong identifiability results. Identifiability results depend firstly on a relationship between the rank, in the sense of PARAFAC decomposition, of the data tensor, secondly on the Kruskal rank of matrices which characterize the propagation and source amplitude.

In particular, nonnegative tensor factorization [24] is used in multiway blind source separation, multidimensional data analysis, and sparse signal/image representations. Fixed point optimization algorithm proposed in [25] and more specifically fixed-point alternating least squares [25] can be used to achieve such a decomposition.

(2) The TUCKER3 model [10, 26] adopted in higher-order SVD (HOSVD) [7, 27] and in LRTA- (K_1, \dots, K_N) (lower-rank (K_1, \dots, K_N) tensor approximation) [8, 28, 29]. We denote by HOSVD- (K_1, \dots, K_N) the truncation of HOSVD, performed with ranks (K_1, \dots, K_N) , in modes $1, \dots, N$, respectively. This model has recently been used as multimode PCA in seismics for wave separation based on a subspace method, in image processing for face recognition and expression analysis [30, 31]. Indeed tensor representation improves automatic face recognition in an adapted independent component analysis framework. “Multilinear independent component analysis” [30] distinguishes between different factors, or modes, inherent to image formation. In particular, this was used for classification of facial expressions. The TUCKER3 model is also used for noise filtering of color images [14].

Each decomposition method corresponds to one definition of the tensor rank. PARAFAC decomposes a tensor into a summation of rank one tensors. The HOSVD- (K_1, \dots, K_N) and the LRTA- (K_1, \dots, K_N) rely on the n th-mode rank definition, that is, the matrix rank of the tensor n th-mode flattening matrix [7, 8]. Both methods perform data projection onto a lower-rank subspace. In this paper, we focus on data denoising [6, 11] by HOSVD- (K_1, \dots, K_N) , lower-rank (K_1, \dots, K_N) approximation, and multiway Wiener filtering [6]. Lower-rank (K_1, \dots, K_N) approximation and multiway Wiener filtering were further improved in the past two years. Some crucial issues were recently solved to improve tensor data denoising. Statistical criteria were adapted to estimate the values of signal subspace ranks [32]. A particular choice of flattening directions improves the results in terms of signal to noise ratio [33, 34]. Multiway filtering algorithms rely on alternating least squares (ALS) loops, which include several costly SVD. We propose to replace SVD by the faster fixed point algorithm proposed in [35]. This paper is a synthesis of the advances that solve these issues. The motivation is that by collecting papers from a range of application areas (including hyperspectral imaging and seismics), the field of tensor signal denoising can be more clearly presented to the interested scientific community, and the field itself may be cross-fertilized with concepts coming from statistics or array processing.

Section 2 presents the tensor model and its main properties. Section 3 states the tensor filtering issue. Section 4 presents classical channel-by-channel filtering methods. Section 5 reminds the principles of two multiway tensor filtering methods, namely lower-rank tensor approximation (LRTA) and multiway Wiener filtering (MWF), developed over the past few years. Section 6 presents all recently proposed improvements for multiway tensor filtering methods which permit an adequate choice of several parameters for multiway filtering methods. The parameter choice is performed as follows: the signal subspace ranks are estimated by a statistical criteria, nonorthogonal tensor flattening for the improvement of tensor data denoising when main directions are present, and fast versions of LRTA and MWF obtained by adapting fixed point and inverse power algorithms for the estimation of leading eigenvectors and smallest eigenvalue. Section 7 exemplifies the presented algorithms by an application to color image and seismic signal denoising; we study the computational load of LRTA and MWF and their fast version by an application to hyperspectral images.

2. DATA TENSOR PROPERTIES

We define a tensor of order N as a multidimensional array whose entries are accessed via N indexes. A tensor is denoted by $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_N}$, where each element is denoted by $a_{i_1 \dots i_N}$, and \mathbb{C} is the complex manifold. An order N tensor has size I_n in mode n , where n refers to the n th index. In signal processing, tensors are built on vector spaces associated with quantities such as length, width, height, time, color channel, and so forth. Each mode of the tensor is associated with one quantity. For example, seismic signals can be modelled by complex valued third-order tensors. Tensor elements can be complex values, to take into account the phase shifts between sensors [6]. The three modes are associated, respectively, with sensor, time, and polarization. In image processing, multicomponent images can be modelled as third-order tensors: two dimensions for rows and columns, and one dimension for the spectral channel. In the same way, a sequence of color images can be modelled by a fourth-order tensor by adding to the previous model one mode associated with the time sampling. 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. The n th-mode flattening matrix \mathbf{A}_n of tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ is defined as a matrix from $\mathbb{R}^{I_n \times M_n}$, where $M_n = I_{n+1} I_{n+2} \dots I_N I_1 I_2 \dots I_{n-1}$. 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 [7, 21, 36]. When dealing with a 1st-mode flattening of dimensionality $I_1 \times (I_2 I_3)$, we formally assume that the index i_2 values vary more slowly than index i_3 values. 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} . In the following, we use the operator “ \times_n ” as the “ n th-mode product” that generalizes the matrix product to tensors. Given $\mathcal{A} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ and a matrix

$\mathbf{U} \in \mathbb{R}^{I_n \times I_n}$, 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 $\mathbb{R}^{I_1 \times \dots \times I_{n-1} \times I_n \times I_{n+1} \times \dots \times I_N}$, whose entries are

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

Next section presents the principles of subspace-based tensor filtering methods.

3. TENSOR FILTERING PROBLEM FORMULATION

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

$$\mathcal{R} = \mathcal{X} + \mathcal{N}. \quad (2)$$

\mathcal{R} , \mathcal{X} , and \mathcal{N} are tensors of order N from $\mathbb{R}^{I_1 \times \dots \times I_N}$. 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 [6, 11, 13, 14]:

$$\widehat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \mathbf{H}^{(2)} \times_3 \dots \times_N \mathbf{H}^{(N)}, \quad (3)$$

Equation (3) performs n th-mode filtering of data tensor \mathcal{R} by n th-mode filter $\mathbf{H}^{(n)}$.

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 ($K_n < I_n$, $\forall n = 1$ 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 $E_2^{(n)}$, defined as

- (i) $E_1^{(n)}$ is the subspace of dimension K_n , spanned by the K_n singular vectors and associated with the K_n largest singular values of matrix \mathbf{X}_n ; $E_1^{(n)}$ is called the signal subspace [37–40];
- (ii) $E_2^{(n)}$ is the subspace of dimension $I_n - K_n$, spanned by the $I_n - K_n$ singular vectors and associated with the $I_n - K_n$ smallest singular values of matrix \mathbf{X}_n ; $E_2^{(n)}$ is called the noise subspace [37–40].

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 tensor channel-by-channel filtering methods based on n th-mode signal subspaces. We present further a method to estimate the dimensions K_1, K_2, \dots, K_N .

4. CHANNEL-BY-CHANNEL FILTERING

The classical algebraic methods operate on two-dimensional data matrices and are based on the singular value decomposition (SVD) [37, 41, 42], and on Eckart-Young theorem

concerning the best lower-rank approximation of a matrix [16] in the least-squares sense. Channel-by-channel filtering consists first of 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 [43], 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. Let us consider the slice matrix $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$ of data tensor \mathcal{R} . Projectors \mathbf{P} on the left signal subspace and \mathbf{Q} on the right signal subspace are built from, respectively, the left and the right singular vectors associated with the K largest singular values of $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$. The parameter K simultaneously defines the dimensions of the left and right signal subspaces. Applying the projectors \mathbf{P} and \mathbf{Q} on the slice matrix $\mathcal{R}(:, :, i_3, \dots, i_j, \dots, i_N)$ amounts to compute its best lower-rank K matrix approximation [16] in the least-squares sense. The filtering of each slice matrix of data tensor \mathcal{R} separately is called in the following “channel-by-channel” SVD-based filtering of \mathcal{R} . It is detailed in [5].

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.

Other channel-by-channel processings are the following: consecutive Wiener filtering of each channel (*2D-Wiener*), PCA followed by *2D-Wiener* (*PCA-2D Wiener*), or soft wavelet threshold (SWT). PCA aims at decorrelating the data (*PCA-2D SWT*) [44–46].

Channel-by-channel filtering methods exhibit a major drawback; they do not take into account the relationships between the components of the processed tensor. Next section presents multiway filtering methods that process jointly all data ways.

5. REVIEW OF MULTIWAY FILTERING METHODS

Multiway filtering methods process jointly all slice matrices of a tensor, which improves the denoising results compared to channel-by-channel processings [6, 11, 13, 14, 32].

5.1. Lower-rank tensor approximation

The LRTA- (K_1, \dots, K_N) of \mathcal{R} minimizes the tensor Frobenius norm (square root of the summation of squared modulus of all terms) $\|\mathcal{R} - \mathcal{B}\|$ subject to the condition that $\mathcal{B} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ is a rank- (K_1, \dots, K_N) tensor. The description of TUCKALS3 algorithm, used in lower-rank (K_1, \dots, K_N) approximation is provided in Algorithm 1.

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 $\mathbf{P}_l^{(m)}$, with $m \neq n$, $l = k$ if $m > n$ and $l = k + 1$ if $m < n$. TUCKALS3 algorithm has recently been used to process

- (1) **Input:** data tensor \mathcal{R} and dimensions K_1, \dots, K_N 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- (K_1, \dots, K_N) :
- n th-mode flatten \mathcal{R} into matrix \mathbf{R}_n ,
 - compute the SVD of \mathbf{R}_n ,
 - 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,
 - form the initial orthogonal projector $\mathbf{P}_0^{(n)} = \mathbf{U}_0^{(n)} \mathbf{U}_0^{(n)T}$ on the n th-mode signal subspace,
 - compute the truncation of HOSVD, with signal subspace ranks (K_1, \dots, K_N) , of tensor \mathcal{R} given by
 $\mathcal{B}_0 = \mathcal{R} \times_1 \mathbf{P}_0^{(1)} \times_2 \cdots \times_N \mathbf{P}_0^{(N)}$.
- (3) **ALS loop**
Repeat until convergence, that is, for example, while $\|\mathcal{B}_{k+1} - \mathcal{B}_k\|^2 > \varepsilon$, $\varepsilon > 0$, being a prior fixed threshold,
- for $n = 1$ to N ,
 - form $\mathcal{B}^{(n),k}$:
 $\mathcal{B}^{(n),k} = \mathcal{R} \times_1 \mathbf{P}_{k+1}^{(1)} \times_2 \cdots \times_{n-1} \mathbf{P}_{k+1}^{(n-1)} \times_{n+1} \mathbf{P}_k^{(n+1)} \times_{n+2} \cdots \times_N \mathbf{P}_k^{(N)}$,
 - n th-mode flatten tensor $\mathcal{B}^{(n),k}$ into matrix $\mathbf{B}_n^{(n),k}$,
 - compute matrix $\mathbf{C}^{(n),k} = \mathbf{B}_n^{(n),k} \mathbf{R}_n^T$,
 - 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,
 - compute $\mathbf{P}_{k+1}^{(n)} = \mathbf{U}_{k+1}^{(n)} \mathbf{U}_{k+1}^{(n)T}$,
 - compute $\mathcal{B}_{k+1} = \mathcal{R} \times_1 \mathbf{P}_{k+1}^{(1)} \times_2 \cdots \times_N \mathbf{P}_{k+1}^{(N)}$,
 - increment k .
- (4) **Output**
The estimated signal tensor is obtained through $\widehat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{P}_{k_{\text{stop}}}^{(1)} \times_2 \cdots \times_N \mathbf{P}_{k_{\text{stop}}}^{(N)}$. $\widehat{\mathcal{X}}$ is the lower-rank (K_1, \dots, K_N) approximation of \mathcal{R} , where k_{stop} is the index of the last iteration after the convergence of TUCKALS3 algorithm.

ALGORITHM 1: Lower-rank (K_1, \dots, K_N) approximation—TUCKALS3 algorithm.

a multimode PCA in order to perform white noise removal in color images, and denoising of multicomponent seismic waves [11, 14].

5.2. 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 multimode PCA achieved by lower-rank (K_1, \dots, K_N) approximation. In spite of the good results provided 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 (3), which optimize an estimation criterion. The most classical method is to minimize the mean square error between the expected signal tensor \mathcal{X} and the estimated signal tensor $\widehat{\mathcal{X}}$ given in (3):

$$e(\mathbf{H}^{(1)}, \dots, \mathbf{H}^{(N)}) = E[\|\mathcal{X} - \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \cdots \times_N \mathbf{H}^{(N)}\|^2]. \quad (4)$$

Due to the criterion which is minimized, filters $\mathbf{H}^{(n)}$, $n = 1$ to N , can be called “ n th-mode Wiener filters” [6].

According to the calculations presented in [6], the minimization of (4) 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 [6]:

$$\mathbf{H}^{(n)} = \boldsymbol{\gamma}_{\mathbf{X}\mathbf{R}}^{(n)} \boldsymbol{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)-1}. \quad (5)$$

The expressions of $\boldsymbol{\gamma}_{\mathbf{X}\mathbf{R}}^{(n)}$ and $\boldsymbol{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)}$ can be found in [6]. $\boldsymbol{\gamma}_{\mathbf{X}\mathbf{R}}^{(n)}$ depends on data tensor \mathcal{R} and on signal tensor \mathcal{X} . $\boldsymbol{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)}$ only depends on data tensor \mathcal{R} .

In order to obtain $\mathbf{H}^{(n)}$ through (5), we suppose that the filters $\{\mathbf{H}^{(m)}, m = 1 \text{ to } N, m \neq n\}$ are known. Data tensor \mathcal{R} is available, but signal tensor \mathcal{X} is unknown. So, only the term $\boldsymbol{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)}$ can be derived, and not the term $\boldsymbol{\gamma}_{\mathbf{X}\mathbf{R}}^{(n)}$. Hence, some more assumptions on \mathcal{X} have to be made in order to overcome the indetermination over $\boldsymbol{\gamma}_{\mathbf{X}\mathbf{R}}^{(n)}$ [6, 13]. 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, [6, 13] 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 $E_1^{(n)}$:

$$\mathbf{X}_n = \mathbf{V}_s^{(n)} \mathbf{O}^{(n)}, \quad (6)$$

with $\mathbf{X}_n \in \mathbb{R}^{I_n \times M_n}$, and $\mathbf{V}_s^{(n)} \in \mathbb{R}^{I_n \times K_n}$ being the matrix containing the K_n orthonormal basis vectors of n th-mode signal subspace $E_1^{(n)}$. Matrix $\mathbf{O}^{(n)} \in \mathbb{R}^{K_n \times M_n}$ 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 (2) is white, Gaussian, and independent from signal \mathcal{X} , and introducing the signal model equation (6) in (5) leads to a computable expression of n th-mode Wiener filter $\mathbf{H}^{(n)}$ (see [6]):

$$\mathbf{H}^{(n)} = \mathbf{V}_s^{(n)} \boldsymbol{\gamma}_{\text{OO}}^{(n)} \boldsymbol{\Lambda}_{\Gamma_s}^{(n)-1} \mathbf{V}_s^{(n)T}. \quad (7)$$

We define matrix $\mathbf{T}^{(n)}$ as

$$\mathbf{T}^{(n)} = \mathbf{H}^{(1)} \otimes \dots \otimes \mathbf{H}^{(n-1)} \otimes \mathbf{H}^{(n+1)} \otimes \dots \otimes \mathbf{H}^{(N)}, \quad (8)$$

where \otimes stands for Kronecker product, and matrix $\mathbf{Q}^{(n)}$ as

$$\mathbf{Q}^{(n)} = \mathbf{T}^{(n)T} \mathbf{T}^{(n)}. \quad (9)$$

In (7), $\boldsymbol{\gamma}_{\text{OO}}^{(n)} \boldsymbol{\Lambda}_{\Gamma_s}^{(n)-1}$ is a diagonal weight matrix given by

$$\boldsymbol{\gamma}_{\text{OO}}^{(n)} \boldsymbol{\Lambda}_{\Gamma_s}^{(n)-1} = \text{diag} \left[\frac{\beta_1}{\lambda_1^\Gamma}, \dots, \frac{\beta_{K_n}}{\lambda_{K_n}^\Gamma} \right], \quad (10)$$

where $\lambda_1^\Gamma, \dots, \lambda_{K_n}^\Gamma$ are the K_n largest eigenvalues of $\mathbf{Q}^{(n)}$ -weighted covariance matrix $\boldsymbol{\Gamma}_{\text{RR}}^{(n)} = E[\mathbf{R}_n \mathbf{Q}^{(n)} \mathbf{R}_n^T]$. Parameters $\beta_1, \dots, \beta_{K_n}$ depend on $\lambda_1^\gamma, \dots, \lambda_{K_n}^\gamma$ which are the K_n largest eigenvalues of $\mathbf{T}^{(n)}$ -weighted covariance matrix

$\boldsymbol{\gamma}_{\text{RR}}^{(n)} = E[\mathbf{R}_n \mathbf{T}^{(n)} \mathbf{R}_n^T]$, according to the following relation:

$$\beta_{k_n} = \lambda_{k_n}^\gamma - \sigma_\Gamma^{(n)2}, \quad \forall k_n = 1, \dots, K_n. \quad (11)$$

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 eigenvalue of noise $\mathbf{T}^{(n)}$ -weighted covariance matrix $\boldsymbol{\gamma}_{\text{NN}}^{(n)} = E[\mathbf{N}_n \mathbf{T}^{(n)} \mathbf{N}_n^T]$. Thanks to the additive noise and the signal independence assumptions, the $I_n - K_n$ smallest eigenvalues of $\boldsymbol{\gamma}_{\text{RR}}^{(n)}$ are equal to $\sigma_\Gamma^{(n)2}$, and thus, can be estimated by the following relation:

$$\hat{\sigma}_\Gamma^{(n)2} = \frac{1}{I_n - K_n} \sum_{k_n=K_n+1}^{I_n} \lambda_{k_n}^\gamma. \quad (12)$$

In order to determine the n th-mode Wiener filters $\mathbf{H}^{(n)}$ that minimizes the mean square error (see (4)), the alternating least squares (ALSs) algorithm has been proposed in [6, 13]. It can be summarized in Algorithm 2.

Both lower-rank tensor approximation and multiway tensor filtering methods are based on singular value decomposition. We propose to adapt faster methods to estimate only the needed leading eigenvectors and dominant eigenvalues.

6. CHOICE OF PARAMETERS FOR MULTIWAY FILTERING METHODS

6.1. n th-mode signal subspace rank estimation by statistical criteria

The subspace-based tensor methods project the data onto a lower-dimensional subspace of each n th-mode. For the

LRTA- (K_1, K_2, \dots, K_N) , the (K_1, K_2, \dots, K_N) -parameter is the number of eigenvalues of the flattened \mathbf{R}_n (for $n = 1$ to N) which permits an optimal approximation of \mathcal{R} in the least squares sense. For the multiway Wiener filter, it is the number of eigenvalues which permits an optimal restoration of \mathcal{X} in the least mean squares sense. In a noisy environment, it is equivalent to the useful n th-mode signal subspace dimension. Moreover, because the eigenvalue distribution of the n th-mode flattened matrix \mathbf{R}_n depends on the noise power of \mathcal{N} , the K_n -value decreases when noise power increases.

Finding the correct K_n -values which yield an *optimum* restoration appears, for two reasons, as a good strategy to improve the denoising results [32]. Actually, for all n th-modes, if K_n is too small, some information is lost after restoration, and if K_n is too large, some noise may be included in the restored information. Because the number of feasible (K_1, K_2, \dots, K_N) combinations is equal to $I_1 \cdot I_2 \cdot \dots \cdot I_N$ which may be large, an estimation method is chosen rather than empirical method. We review a method, for the K_n -value estimation for each n th-mode, which adapts the well-know minimum description length (MDL) detection criterion [47]. The optimal signal subspace dimension is obtained by minimizing MDL criterion. The useful signal subspace dimension is equal to the lower n th-mode rank of the n th-mode flattened matrix \mathbf{R}_n .

Consequently, for each mode, the MDL criterion can be expressed as

$$\text{MDL}(k) = -\log \left(\frac{\prod_{i=k+1}^{I_n} \lambda_i^{1/(I_n-k)}}{(1/(I_n-k))^{\sum_{i=k+1}^{I_n} \lambda_i}} \right)^{(I_n-k)M_n} + \frac{1}{2} k (2I_n - k) \log M_n. \quad (13)$$

When we consider lower-rank tensor approximation, $(\lambda_i)_{1 \leq i \leq I_n}$ are either the I_n singular values of \mathbf{R}_n (see step 2c of Algorithm 1), or the I_n eigenvalues of $\mathbf{C}^{(n),k}$ (see step (3)(a)iv). When we consider multiway Wiener filtering, $(\lambda_i)_{1 \leq i \leq I_n}$ are the I_n eigenvalues of either matrix $\boldsymbol{\gamma}_{\text{RR}}^{(n)}$ or matrix $\boldsymbol{\Gamma}_{\text{RR}}^{(n)}$ (see steps 2(a)iiB and 2(a)iiE).

The n th-mode rank K_n is the value of k ($k \in [1, \dots, I_n - 1]$) which minimizes MDL criterion.

The estimation of the signal subspace dimension of each mode is performed at each ALS iteration.

6.2. Flattening directions for SNR improvement

To improve denoising quality, flattening is performed along main directions in the image, which are estimated by SLIDE algorithm [48].

6.2.1. Rank reduction and flattening directions

Let us consider a matrix \mathbf{A} of size $I_1 \times I_1$ which could represent an image containing a straight line. The rank of this matrix is closely linked to the orientation of the line: an image with a horizontal or a vertical line has rank 1, else it is more than one. The limit case is when the straight line is along

- (1) **Initialization** $k = 0$: $\mathcal{R}^0 = \mathcal{R} \otimes \mathbf{H}_0^{(n)} = \mathbf{I}_n$, identity matrix, for all $n = 1$ to N .
- (2) **ALS loop**:
 repeat until convergence, that is, $\|\mathcal{R}^{k+1} - \mathcal{R}^k\|^2 < \varepsilon$, with $\varepsilon > 0$ a prior fixed threshold,
 (a) for $n = 1$ to N ,
 (i) form $\mathcal{R}^{(n),k}$:

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

 (ii) determine $\mathbf{H}_{k+1}^{(n)} = \arg \min_{\mathbf{Z}^{(n)}} \|\mathcal{X} - \mathcal{R}^{(n),k} \times_n \mathbf{Z}^{(n)}\|^2$ subject to $\mathbf{Z}^{(n)} \in \mathbb{R}^{I_n \times I_n}$ thanks to the following procedure:
 (A) n th-mode flatten $\mathcal{R}^{(n),k}$ into $\mathbf{R}_n^{(n),k} = \mathbf{R}_n(\mathbf{H}_{k+1}^{(1)} \otimes \cdots \otimes \mathbf{H}_{k+1}^{(n-1)} \otimes \mathbf{H}_k^{(n+1)} \otimes \cdots \otimes \mathbf{H}_k^{(N)})^T$, and \mathcal{R} into \mathbf{R}_n ,
 (B) compute $\boldsymbol{\gamma}_{\mathbf{R}\mathbf{R}}^{(n)} = E[\mathbf{R}_n \mathbf{R}_n^{(n),k T}]$,
 (C) determine $\lambda_1^{\gamma}, \dots, \lambda_{K_n}^{\gamma}$, the K_n largest eigenvalues of $\boldsymbol{\gamma}_{\mathbf{R}\mathbf{R}}^{(n)}$,
 (D) for $k_n = 1$ to I_n , estimate $\sigma_{\Gamma}^{(n)2}$ thanks to (12) and for $k_n = 1$ to K_n , estimate β_{k_n} thanks to (11),
 (E) compute $\boldsymbol{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)} = E[\mathbf{R}_n^{(n),k} \mathbf{R}_n^{(n),k T}]$,
 (F) determine $\lambda_1^{\Gamma}, \dots, \lambda_{K_n}^{\Gamma}$, the K_n largest eigenvalues of $\boldsymbol{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)}$,
 (G) determine $\mathbf{V}_s^{(n)}$, the matrix of the K_n eigenvectors associated with the K_n largest eigenvalues of $\boldsymbol{\Gamma}_{\mathbf{R}\mathbf{R}}^{(n)}$,
 (H) compute the weight matrix $\boldsymbol{\gamma}_{\mathbf{O}\mathbf{O}}^{(n)} \boldsymbol{\Lambda}_{\Gamma_s}^{(n)-1}$ given in (10),
 (I) compute $\mathbf{H}_{k+1}^{(n)}$, the n th-mode Wiener filter at the $(k+1)$ th iteration, using (7),
 (b) form $\mathcal{R}^{k+1} = \mathcal{R} \times_1 \mathbf{H}_{k+1}^{(1)} \times_2 \cdots \times_N \mathbf{H}_{k+1}^{(N)}$,
 (c) increment k .
- (3) **output**: $\widehat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}_{k_{\text{stop}}}^{(1)} \times_2 \cdots \times_N \mathbf{H}_{k_{\text{stop}}}^{(N)}$, with k_{stop} being the last iteration after convergence of the algorithm.

ALGORITHM 2

a diagonal, in this case, the rank of the matrix is I_1 . This is also true for tensors. If a color image has been corrupted by a white noise, a lower-rank approximation performed with the rank of the n th-mode signal subspace leads to the reconstruction of initial signal. In the case of a straight line along a diagonal of the image, the signal subspace is equal to the minimum dimension of the image. In this case, no truncation can be done without losing information and the image cannot be restored this way. If the line is either horizontal or vertical, the truncation to rank- $(K_1 = 1, K_2 = 1, K_3 = 3)$ leads to a good restoration [34].

6.2.2. Estimation of main directions

To retrieve main directions, a classical method is the Hough transform [49]. In [48, 50], an analogy between straight line detection and sensor array processing has been drawn. This method can be used to provide main directions of an image. The whole algorithm is called subspace-based Line DEtection (SLIDE). The number of main directions is given by MDL criterion [47]. The main idea of SLIDE is to generate virtual signals out of the image to set the analogy between localization of sources in array processing and recognition of straight lines in image processing. Principles of SLIDE are detailed in [48]. In the case of a noisy image containing d straight lines, the signal measured at the l th row of the image is [48]

$$z_l = \sum_{k=1}^d e^{j\mu(l-1)\tan\theta_k} \cdot e^{-j\mu x_{0k}} + n_l, \quad l = 1, \dots, N, \quad (14)$$

where μ is a propagation parameter [48], n_l is the noise resulting from outlier pixels at the l th row. Starting from this

signal, the SLIDE method [48, 50] estimates the orientation θ_k of the d straight lines. Defining

$$a_l(\theta_k) = e^{j\mu(l-1)\tan\theta_k}, \quad s_k = e^{-j\mu x_{0k}}, \quad (15)$$

we obtain

$$z_l = \sum_{k=1}^d a_l(\theta_k) s_k + n_l, \quad \forall l = 1, \dots, N. \quad (16)$$

Thus, the $N \times 1$ vector \mathbf{z} is defined by

$$\mathbf{z} = \mathbf{A}\mathbf{s} + \mathbf{n}, \quad (17)$$

where \mathbf{z} and \mathbf{n} are $N \times 1$ vectors corresponding, respectively, to received signal and noise, \mathbf{A} is a $N \times d$ matrix and \mathbf{s} is the $d \times 1$ source signal vector. This relation is the classical equation of an array processing problem.

SLIDE algorithm uses TLS-ESPRIT algorithm, which splits the array into two subarrays [48]. SLIDE algorithm [48, 50] provides the estimation of the angles θ_k :

$$\theta_k = \tan^{-1} \left[\frac{1}{\mu\Delta} \text{Im} \left(\ln \frac{\lambda_k}{|\lambda_k|} \right) \right], \quad k = 1, \dots, d, \quad (18)$$

where Δ is the displacement between the two subarrays, $\{\lambda_k, k = 1, \dots, M\}$ are the eigenvalues of a diagonal unitary matrix that relates the measurements from the first subarray to the measurements resulting from the second subarray, and ‘‘Im’’ stands for ‘‘imaginary part.’’ Details of this algorithm can be found in [48].

The orientation values obtained enable us to flatten the data tensor along the main directions in the tensor. This first improvement reduces the blur effect induced by Wiener filtering in the result image.

6.3. Fast multiway filtering methods

We present in the general case the fast fixed-point algorithm proposed in [35] for computing K leading eigenvectors of any matrix \mathbf{C} , and show how, in particular, this algorithm can be inserted in an ALS loop to compute signal subspace projectors for each mode. We present the inverse power method which estimates the leading eigenvalues and shows how it can be inserted in multiway filtering algorithm to compute the weight matrix for each mode.

6.3.1. Fast singular vector estimation

One way to compute the K orthonormal basis vectors of any matrix \mathbf{C} is to use the fixed-point algorithm proposed in [35].

Choose K , the number of required leading eigenvectors to be estimated. Consider matrix \mathbf{C} and set iteration index $p \leftarrow 1$. Set a threshold η . For $p = 1$ to K .

- (1) Initialize eigenvector \mathbf{u}_p , whose length is the number of lines of \mathbf{C} (e.g., randomly). Set counter $it \leftarrow 1$ and $\mathbf{u}_p^{it} \leftarrow \mathbf{u}_p$. Set \mathbf{u}_p^0 as a random vector.
- (2) While $\|\mathbf{u}_p^{itT} \mathbf{u}_p^{it-1} - 1\| < \eta$,
 - (a) update \mathbf{u}_p^{it} as $\mathbf{u}_p^{it} \leftarrow \mathbf{C} \mathbf{u}_p^{it}$,
 - (b) do the Gram-Schmidt orthogonalization process $\mathbf{u}_p^{it} \leftarrow \mathbf{u}_p^{it} - \sum_{j=1}^{p-1} (\mathbf{u}_p^{itT} \mathbf{u}_j^{it}) \mathbf{u}_j^{it}$,
 - (c) normalize \mathbf{u}_p^{it} by dividing it by its norm: $\mathbf{u}_p^{it} \leftarrow \mathbf{u}_p^{it} / \|\mathbf{u}_p^{it}\|$,
 - (d) increment counter $it \leftarrow it + 1$.
- (3) Increment counter $p \leftarrow p + 1$ and go to step (1) until p equals K .

The eigenvector with dominant eigenvalue will be estimated first. Similarly, all the remaining $K - 1$ basis vectors (orthonormal to the previously estimated basis vectors) will be estimated one by one in a reducing order of dominance. The previously estimated $(p - 1)$ th basis vectors will be used to find the p th basis vector. The algorithm for p th basis vector will converge when the new value \mathbf{u}_p^+ and old value \mathbf{u}_p are such that $\mathbf{u}_p^{+T} \mathbf{u}_p$ is close to 1. The smaller η , the more accurate the estimation. Let $\mathbf{U} = [\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_K]$ be the matrix whose columns are the K orthonormal basis vectors. Then, $\mathbf{U} \mathbf{U}^T$ is the projector onto the subspace spanned by the K eigenvectors associated with dominant eigenvalues.

So fixed-point algorithm can be used in LRTA- (K_1, K_2, \dots, K_N) to retrieve the basis vectors $\mathbf{U}_0^{(n)}$ in steps (2)b, (2)c, and the basis vectors $\mathbf{U}_k^{(n)}$ in step 3(a)iv. Thus, the initialization step is faster since it does not need the I_n basis vectors but only the K_n first ones and it does not need in step (2)b the SVD of the data tensor n th-mode flattening matrix \mathbf{R}_n . In multiway Wiener filtering algorithm, fixed-point algorithm can replace every SVD to compute the K_n largest eigenvectors of matrix $\mathbf{V}_s^{(n)}$ in step 2(a)iiG.

6.3.2. Fast singular value estimation

Fixed-point algorithm is sufficient to replace SVD in lower-rank tensor approximation, but we notice that, when multiway Wiener filtering is performed, the eigenvalues of $\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)}$ are required in step 2(a)iiC, and the eigenvalues of $\boldsymbol{\Gamma}_{\mathbf{RR}}^{(n)}$ are required in step 2(a)iiF. Indeed, multiway Wiener filtering involves weight matrices which depend on eigenvalues of signal and data covariance flattening matrices $\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)}$ and $\boldsymbol{\Gamma}_{\mathbf{RR}}^{(n)}$ (see (10)). This can be achieved in steps 2(a)iiC and 2(a)iiF of multiway Wiener filtering algorithm by the following calculation involving the previously computed leading eigenvectors: $\mathbf{V}_{sy}^{(n)T} \boldsymbol{\gamma}_{\mathbf{RR}}^{(n)} \mathbf{V}_{sy}^{(n)} = \text{diag}\{\lambda_1^y, \dots, \lambda_{K_n}^y\}$, and $\mathbf{V}_{s\Gamma}^{(n)T} \boldsymbol{\Gamma}_{\mathbf{RR}}^{(n)} \mathbf{V}_{s\Gamma}^{(n)} = \text{diag}\{\lambda_1^\Gamma, \dots, \lambda_{K_n}^\Gamma\}$, respectively.

Matrix $\mathbf{V}_{sy}^{(n)}$ (resp., $\mathbf{V}_{s\Gamma}^{(n)}$) contains the K_n leading eigenvectors of $\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)}$ (resp., $\boldsymbol{\Gamma}_{\mathbf{RR}}^{(n)}$) associated with the K_n largest eigenvalues. These eigenvectors are obtained by fixed point algorithm.

Here is the detail of the way we obtained the K_n eigenvalues of matrices $\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)}$ and $\boldsymbol{\Gamma}_{\mathbf{RR}}^{(n)}$.

We give some details concerning matrix $\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)}$:

$$\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)} = \mathbf{V}_{sy}^{(n)} \boldsymbol{\Lambda}_{sy}^{(n)} \mathbf{V}_{sy}^{(n)T} + \mathbf{V}_{ny}^{(n)} \boldsymbol{\Lambda}_{ny}^{(n)} \mathbf{V}_{ny}^{(n)T}. \quad (19)$$

When we multiply $\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)}$ left by $\mathbf{V}_{sy}^{(n)T}$ and right by $\mathbf{V}_{sy}^{(n)}$, we obtain

$$\mathbf{V}_{sy}^{(n)T} \boldsymbol{\gamma}_{\mathbf{RR}}^{(n)} \mathbf{V}_{sy}^{(n)} = \boldsymbol{\Lambda}_{sy}^{(n)} + \mathbf{0} = \boldsymbol{\Lambda}_{sy}^{(n)} = \text{diag}\{\lambda_1^y, \dots, \lambda_{K_n}^y\}. \quad (20)$$

Similarly are obtained the dominant eigenvalues of matrix $\boldsymbol{\Gamma}_{\mathbf{RR}}^{(n)}$.

Thus, β_{k_n} can be computed following (11). But multiway Wiener filtering also requires the $I_n - K_n$ smallest eigenvalues of $\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)}$, equal to $\sigma_\Gamma^{(n)}$ (see step 2(a)iiD of Wiener algorithm and (12)). Thus, we adapt the inverse power method to retrieve $\boldsymbol{\gamma}_{\mathbf{RR}}^{(n)}$ smallest eigenvalue.

- (1) Initialize randomly \mathbf{x}_0 of size $K_n \times 1$.
- (2) While $\|\mathbf{x} - \mathbf{x}_0\| / \|\mathbf{x}\| \leq \varepsilon$ do
 - (a) $\mathbf{x} \leftarrow \boldsymbol{\gamma}_{\mathbf{RR}}^{(n)-1} \cdot \mathbf{x}_0$,
 - (b) $\lambda \leftarrow \|\mathbf{x}\|$,
 - (c) $\mathbf{x} \leftarrow \mathbf{x} / \lambda$,
 - (d) $\mathbf{x}_0 \leftarrow \mathbf{x}$,
- (3) $\sigma_\Gamma^{(n)} = 1 / \lambda$.

Therefore, $\sigma_\Gamma^{(n)2}$ can be estimated in step 2(a)iiD, and the calculation of (10) can be performed in a fast way.

7. APPLICATION OF MULTIWAY FILTERING METHODS

We apply the reviewed methods to the denoising of a color image and of a hyperspectral image. In the first case, we compare multiway tensor data denoising methods with channel-by-channel SVD. In the second case, we concentrate

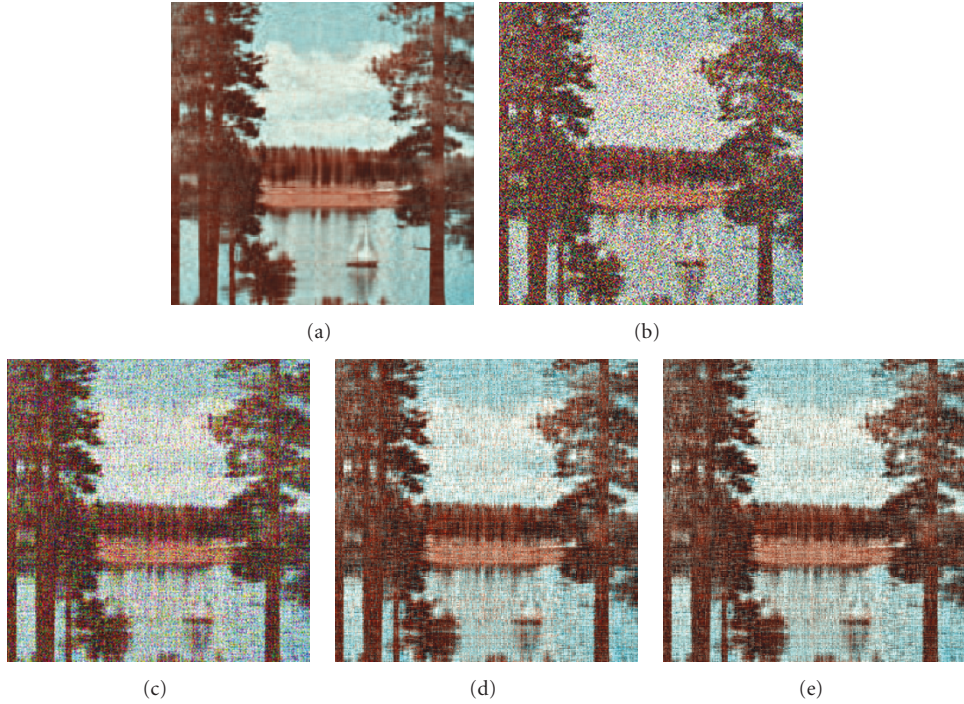


FIGURE 1: (a) Nonnoisy image. (b) Image to be processed, impaired by an additive white noise, with SNR = 8.1 dB. (c) Channel-by-channel SVD-based filtering of parameter $K = 30$. (d) Lower-rank (30, 30, 2) approximation. (e) MWF-(30, 30, 2) filtering.

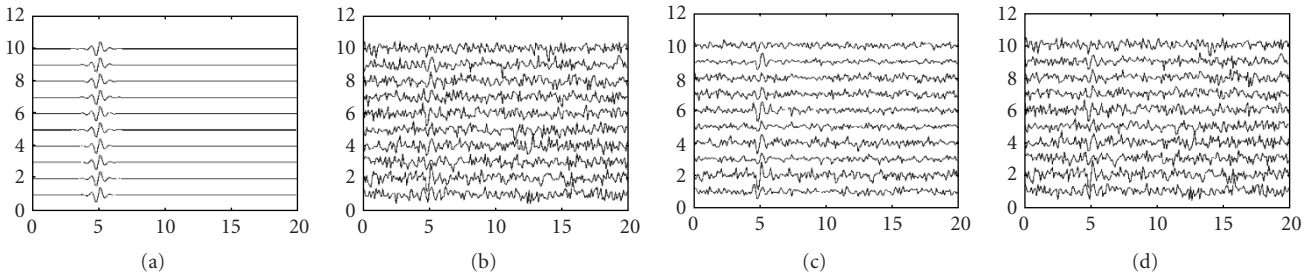


FIGURE 2: Polarization component 1 of a seismic signal: nonnoisy impaired results with LRTA-(8, 8, 3), and result with MWF-(8, 8, 3).

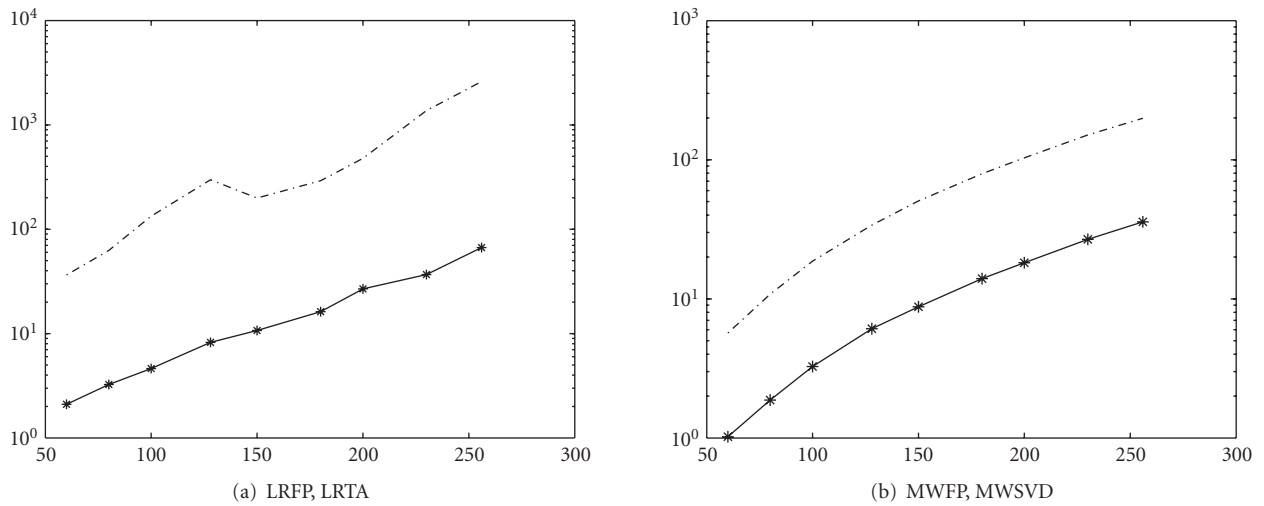


FIGURE 3: Computational times (s) as a function of the number of rows and columns: tensor filtering using (a) LRFP (-*-), LRTA (-.-); (b) MWF (-*-), MWSVD (-.-).

on the required computational times. The subspace ranks are estimated by MDL criterion unless it is specified.

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

$$\mathcal{N} = \alpha \cdot \mathcal{G}, \quad (21)$$

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

To evaluate quantitatively the results obtained by the presented methods, we define the signal to noise ratio (SNR, in dB) in the noisy data tensor by $\text{SNR} = 10 \log(\|\mathcal{X}\|^2 / \|\mathcal{N}\|^2)$, and to a posteriori verify the quality of the estimated signal tensor, we use the normalized quadratic error (NQE) criterion defined as follows: $\text{NQE}(\hat{\mathcal{X}}) = \|\hat{\mathcal{X}} - \mathcal{X}\|^2 / \|\mathcal{X}\|^2$.

7.1. Denoising of a color image impaired by additive noise

Let us consider the ‘‘sailboat’’ standard color image of Figure 1(a) represented as a third-order tensor $\mathcal{X} \in \mathbb{R}^{256 \times 256 \times 3}$. The ranks of the signal subspace for each mode are set as 30 for the 1st mode, 30 for the 2nd mode, and 2 for the 3rd mode. This is fixed thanks to the following process. For Figure 1(a), we took the standard nonnoisy ‘‘sailboat’’ image and we artificially reduced the ranks of the nonnoisy image, that is, we set the parameters (K_1, K_2, K_3) to $(30, 30, 2)$, thanks to the truncation of HOSVD. This permits to ensure that, for each mode, the rank of the signal subspace is lower than the corresponding dimension. This also permits to evaluate the performance of the filtering methods applied, independently from the accuracy of the estimation of the values of the ranks by MDL or AIC criterion.

Figure 1(b) shows the noisy image resulting from the impairment of Figure 1(a) and represented as $\mathcal{R} = \mathcal{X} + \mathcal{N}$. Third-order noise tensor \mathcal{N} is defined by (21) by choosing α such that, considering the definition above, the SNR in the noisy image of Figure 1(b) is 8.1 dB. In these simulations, the value of the parameter K of channel-by-channel SVD-based filtering, the values of the dimensions of the row, and column signal subspace are supposed to be known and fixed to 30. In the same way, parameters (K_1, K_2, K_3) of lower-rank (K_1, K_2, K_3) approximation are fixed to $(30, 30, 2)$. The channel-by-channel SVD-based filtering of noisy image \mathcal{R} (see Figure 1(b)) yields the image of Figure 1(c), and lower-rank $(30, 30, 2)$ approximation of noisy data tensor \mathcal{R} yields the image of Figure 1(d). The NQE criterion permits a quantitative comparison between channel-by-channel SVD-based filtering, LRTA- $(30, 30, 2)$, and MWF- $(30, 30, 2)$. The obtained NQE is, respectively, 0.09 with channel-by-channel SVD-based filtering, 0.025 with LRTA- $(30, 30, 2)$, and 0.01 with MWF- $(30, 30, 2)$. From the resulting image, presented on Figure 1(d), we notice that dimension reduction leads to a loss of spatial resolution. 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.

Therefore, a tradeoff should be considered between noise reduction and detail preservation. When MDL criterion

[32, 47] is applied to the left singular values of the flattening matrices computed over the successive n th-modes, the correct tradeoff is automatically reached. In the next simulation, a multicomponent seismic wave is received on a linear antenna composed of 10 sensors. The direction of propagation of the wave is assumed to be contained in a plane which is orthogonal to the antenna. The wave is composed of three components, represented as signal tensor \mathcal{X} . Each consecutive component presents a $\pi/2$ radian phase shift. Figure 2 represents nonnoisy component 1, impaired component 1 (SNR = -10 dB), the results of denoising by LRTA- $(8, 8, 3)$, and MWF- $(8, 8, 3)$ (NQE = 0.8 and 3.8, resp.).

7.2. Hyperspectral images: denoising results and compared computational loads

The proposed fast lower-rank tensor approximation, that we name lower-rank fixed point (LRFP), and the proposed fast multiway Wiener filtering, that we name multiway Wiener fixed point (MWFP), are compared with the versions of lower-rank tensor approximation and multiway Wiener filtering which use SVD, respectively, named lower-rank tensor approximation (LRTA) and multiway Wiener SVD (MWSVD).

The proposed and comparative methods can be applied to any tensor data, such as color image, multicomponent seismic signals, or hyperspectral images [6]. We exemplify the proposed method with hyperspectral image (HSI) denoising. The HSI data used in the following experiments are real-world data collected by HYDICE imaging, with a 1.5 m spatial and 10 nm spectral resolution and including 148 spectral bands (from 435 to 2326 nm). Then, HSI data can be represented as a third-order tensor, denoted by $\mathcal{R} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$. A multiway white noise \mathcal{N} is added to signal tensor \mathcal{X} . We consider HSI data with a large amount of noise, by setting SNR = 3 dB. We process images with various number of rows and columns, to study the proposed and compared algorithm speed as a function of the data size. Each band has from $I_1 = I_2 = 20$ to 256 rows and columns. Number of spectral bands I_3 is fixed to 148. Signal subspace ranks (K_1, K_2, K_3) chosen to perform lower-rank (K_1, K_2, K_3) approximation are equal to $(10, 10, 15)$. Parameter η (see Section 6.3.1) is fixed to 10^{-6} , and 5 iterations of the ALS algorithm are needed for convergence. Figure 3(a) (resp., (b)) provides the evolution of computational times for both LRFP and LRTA-based (resp., MWFP and MWSVD-based) tensor data denoising, for values of I_1 and I_2 varying between 60 and 256, in second, with a 3.0 Ghz PC running windows (same conditions are used throughout all experiments). Considering an image with 256 rows and columns, LRFP-based method leads to SNR = 17.03 dB with a computational time equal to 68 seconds and LRTA-based method leads to SNR = 17.20 dB with a computational time equal to 43 minutes, 22 seconds. Then with these image sizes, and the ratios $K_1/I_1 = K_2/I_2 = 410^{-2}$, and $K_3/I_3 = 110^{-1}$, the proposed method is 38 times faster, yielding SNR values that differ by less than 1%. MWFP-based method leads to SNR = 17.11 dB with a computational time equal to 36 seconds and

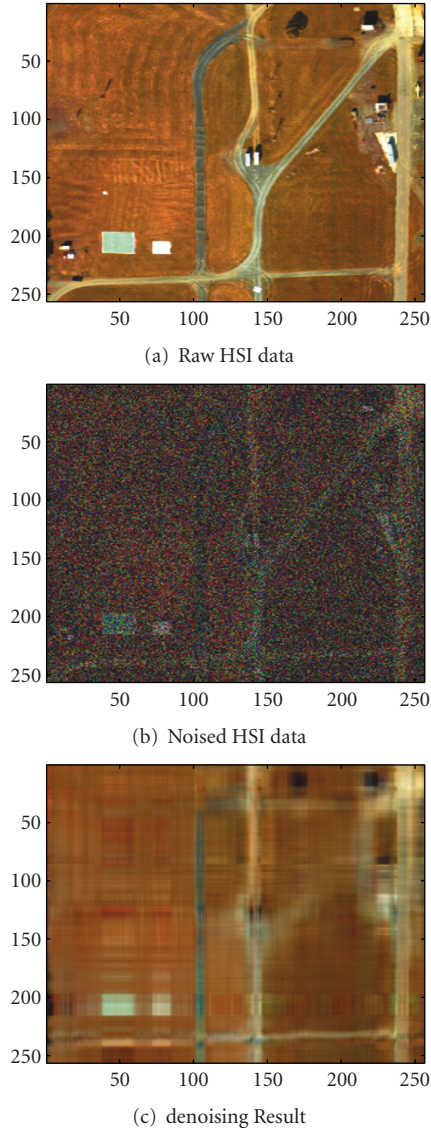


FIGURE 4: HSI image: results obtained by lower-rank tensor approximation using LRFP, LRFA, LRWF, or LRWSVD.

MWSVD-based method leads to $\text{SNR} = 17.27$ dB with a computational time equal to 17 minutes, 4 seconds. Then, the proposed method is 29 times faster, yielding SNR values that differ by less than 1%. The gain in computational times is particularly pronounced with K_1/I_1 , K_2/I_2 , and K_3/I_3 ratio values which are relatively low, which is relevant for denoising applications. Figure 4(a) is the raw image with $I_1 = I_2 = 256$; Figure 4(b) provides the noised image; Figure 4(c) is the denoising result obtained by the LRFA algorithm. Results obtained with LRFP, LRWF, or LRWSVD algorithms look very similar.

8. CONCLUSION

This paper deals with tensor data denoising methods, and last advances in this field. We review lower-rank

tensor approximation (LRFA) and multiway Wiener filtering (MWF), and remind they yield good denoising results, especially compared to channel-by-channel SVD-based processing. These methods rely on tensor flattening along each mode, and on the projection of the data upon a useful signal subspace. We propose a synthesis of the last advances in tensor signal processing methods. We show how the signal subspace ranks can be estimated by statistical criteria; we demonstrate that, by flattening tensors along main directions, output SNR is improved, and propose to use the fast SLIDE algorithm to retrieve these main directions; we adapt fixed-point algorithm and inverse power method to replace the costly SVD in lower-rank tensor approximation and multiway Wiener filtering methods, thus obtaining much faster algorithms. We exemplify the proposed improved methods on a seismic signal, color, and hyperspectral images.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers who contributed to the quality of this paper by providing helpful suggestions.

REFERENCES

- [1] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N -way arrays," *Journal of Chemometrics*, vol. 14, no. 3, pp. 229–239, 2000.
- [2] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 810–823, 2000.
- [3] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear independent components analysis," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 547–553, San Diego, Calif, USA, June 2005.
- [4] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee, "Spatial transformations of diffusion tensor magnetic resonance images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 11, pp. 1131–1139, 2001.
- [5] D. Muti, S. Bourennane, and J. Marot, "Lower-rank tensor approximation and multiway filtering," to appear in *SIAM Journal on Matrix Analysis and Applications*.
- [6] D. Muti and S. Bourennane, "Multidimensional filtering based on a tensor approach," *Signal Processing*, vol. 85, no. 12, pp. 2338–2353, 2005.
- [7] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [8] L. De Lathauwer, B. De Moor, and J. Vandewalle, "On the best rank-1 and rank- (R_1, R_2, \dots, R_N) approximation of higher-order tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2000.
- [9] P. M. Kroonenberg, *Three-Mode Principal Component Analysis: Theory and Applications*, DSWO Press, Leiden, The Netherlands, 1983.
- [10] P. M. Kroonenberg and J. de Leeuw, "Principal component analysis of three-mode data by means of alternating least squares algorithms," *Psychometrika*, vol. 45, no. 1, pp. 69–97, 1980.

- [11] D. Muti and S. Bourennane, "Multiway filtering based on fourth-order cumulants," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 7, pp. 1147–1158, 2005.
- [12] D. Muti and S. Bourennane, "Fast optimal lower-rank tensor approximation," in *Proceedings of the 2nd IEEE International Symposium on Signal Processing and Information Technology (ISSPIT '02)*, pp. 621–625, Marrakesh, Morocco, December 2002.
- [13] D. Muti and S. Bourennane, "Multidimensional estimation based on a tensor decomposition," in *Proceedings of the IEEE Workshop on Statistical Signal Processing (SSP '03)*, pp. 98–101, St. Louis, Mo, USA, September-October 2003.
- [14] D. Muti and S. Bourennane, "Multidimensional signal processing using lower-rank tensor approximation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol. 3, pp. 457–460, Hong Kong, April 2003.
- [15] D. Muti and S. Bourennane, "Traitement du signal par décomposition tensorielle," in *Proceedings of the 19th GRETSI Symposium on Signal and Image Processing*, Paris, France, September 2003.
- [16] C. Eckart and G. Young, "The approximation of a matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936.
- [17] D. Muti, S. Bourennane, and M. Guillaume, "SVD-based image filtering improvement by means of image rotation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04)*, vol. 3, pp. 289–292, Montreal, Canada, May 2004.
- [18] R. A. Harshman and M. E. Lundy, "The PARAFAC model for three-way factor analysis and multidimensional scaling," in *Research Methods for Multimode Data Analysis*, H. G. Law, C. W. Snyder Jr., J. Hattie, and R. P. McDonald, Eds., pp. 122–215, Praeger, New York, NY, USA, 1984.
- [19] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.
- [20] J. Kruskal, "Rank, decomposition, and uniqueness for 3-way and N-way arrays," in *Multiway Data Analysis*, Elsevier/North-Holland, Amsterdam, The Netherlands, 1988.
- [21] H. A. L. Kiess, "Towards a standardized notation and terminology in multiway analysis," *Journal of Chemometrics*, vol. 14, no. 3, pp. 105–122, 2000.
- [22] R. Bro, *Multi-way analysis in the food industry*, Ph.D. thesis, Royal Veterinary and Agricultural University, Copenhagen, Denmark, 1998.
- [23] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Transactions on Signal Processing*, vol. 48, no. 8, pp. 2377–2388, 2000.
- [24] M. Welling and M. Weber, "Positive tensor factorization," *Pattern Recognition Letters*, vol. 22, no. 12, pp. 1255–1261, 2001.
- [25] A. Cichocki and R. Zdunek, "Ntflab for signal processing," Tech. Rep., Laboratory for Advanced Brain Signal Processing, BSI, RIKEN, Saitama, Japan, 2006.
- [26] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.
- [27] O. Alter and G. H. Golub, "Reconstructing the pathways of a cellular system from genome-scale signals by using matrix and tensor computations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 49, pp. 17559–17564, 2005.
- [28] L. De Lathauwer, *Signal processing based on multilinear algebra*, Ph.D. thesis, Department of Electrical Engineering, Katholieke Universiteit Leuven, Leuven, Belgium, September 1997.
- [29] A. Smilde, R. Bro, and P. Geladi, *Multi-Way Analysis: Applications in the Chemical Sciences*, John Wiley & Sons, New York, NY, USA, 2004.
- [30] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear image analysis for facial recognition," in *Proceedings of the 16th International Conference on Pattern Recognition (ICPR '02)*, vol. 2, pp. 511–514, Quebec, Canada, August 2002.
- [31] H. Wang and N. Ahuja, "Facial expression decomposition," in *Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV '03)*, vol. 2, pp. 958–965, Nice, France, October 2003.
- [32] N. Renard, S. Bourennane, and J. Blanc-Talon, "Multiway filtering applied on hyperspectral images," in *Proceedings of the 8th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS '06)*, Lecture Notes on Computer Science, pp. 127–137, Springer, Antwerp, Belgium, September 2006.
- [33] D. Letexier, S. Bourennane, and J. Blanc-Talon, "Nonorthogonal tensor matricization for hyperspectral image filtering," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 1, pp. 3–7, 2008.
- [34] D. Letexier, S. Bourennane, and J. Blanc-Talon, "Main flattening directions and Quadtree decomposition for multi-way Wiener filtering," *Signal, Image and Video Processing*, vol. 1, no. 3, pp. 253–265, 2007.
- [35] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.
- [36] B. W. Bader and T. G. Kolda, "Algorithm 862: MATLAB tensor classes for fast algorithm prototyping," *ACM Transactions on Mathematical Software*, vol. 32, no. 4, pp. 635–653, 2006.
- [37] I. Wirawan, K. Abed-Meraim, H. Maitre, and P. Duhamel, "Blind multichannel image restoration using subspace based method," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol. 5, pp. 9–12, Hong Kong, April 2003.
- [38] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, no. 3, pp. 278–305, 1991.
- [39] N. Yuen and B. Friedlander, "Asymptotic performance analysis of blind signal copy using fourth-order cumulants," *International Journal of Adaptive Control and Signal Processing*, vol. 10, no. 2-3, pp. 239–265, 1996.
- [40] N. Yuen and B. Friedlander, "DOA estimation in multipath: an approach using fourth-order cumulants," *IEEE Transactions on Signal Processing*, vol. 45, no. 5, pp. 1253–1263, 1997.
- [41] H. Andrews and C. Patterson III, "Singular value decomposition and digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 24, no. 1, pp. 26–53, 1976.
- [42] H. Andrews and C. Patterson III, "Singular value decomposition (SVD) image coding," *IEEE Transactions on Communications*, vol. 24, no. 4, pp. 425–432, 1976.
- [43] A. Bendjama, S. Bourennane, and M. Frikel, "Seismic wave separation based on higher order statistics," in *Proceedings of the 1st IEEE International Conference on Digital Signal Processing and Its Applications (DSPA '98)*, Moscow, Russia, June-July 1998.

- [44] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.
- [45] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.
- [46] J.-C. Pesquet, H. Krim, and H. Carfantan, "Time-invariant orthonormal wavelet representations," *IEEE Transactions on Signal Processing*, vol. 44, no. 8, pp. 1964–1970, 1996.
- [47] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.
- [48] H. K. Aghajan and T. Kailath, "Sensor array processing techniques for super resolution multi-line-fitting and straight edge detection," *IEEE Transactions on Image Processing*, vol. 2, no. 4, pp. 454–465, 1993.
- [49] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.
- [50] J. Sheinvald and N. Kiryati, "On the magic of SLIDE," *Machine Vision and Applications*, vol. 9, no. 5-6, pp. 251–261, 1997.