

FAST SUBSPACE-BASED TENSOR DATA FILTERING

Julien Marot, Caroline Fossati and Salah Bourennane

Ecole Centrale Marseille-Institut Fresnel CNRS UMR 6133 13397 Marseille Cedex 20

ABSTRACT

Subspace-based methods rely on dominant element selection from second order statistics. They have been extended to tensor processing, in particular to tensor data filtering. For this, the processed tensor is flattened along each mode successively, and singular value decomposition of the flattened matrix is classically performed. Data projection on the dominant singular vectors results in noise reduction. The numerical cost of SVD is elevated. Now, tensor processing methods include an ALS (Alternating Least Squares) loop, which implies that a large number of SVDs are performed. Fixed point algorithm estimates an *a priori* fixed number of singular vectors from a matrix. In this paper, we generalize fixed point algorithm as a higher-order fixed point algorithm to the estimation of only the required dominant singular vectors in a tensor processing framework. We compare the proposed method in terms of denoising quality and speed through an application to color image and hyperspectral image denoising.

Index Terms— Tensor filtering, subspace-based method, fixed point algorithm.

1. INTRODUCTION

Subspace-based methods consider significant and remaining parts of the data. They are based on data most significant feature selection. Starting from signal realizations, subspace-based methods rely on second order statistics. In particular, the eigenstructure of the covariance matrix of signal realizations provides eigenvectors which span the measurement space. Within the measurement space, dominant eigenvectors span the so-called "signal subspace" and the remaining eigenvectors span the so-called "noise subspace". Subspace-based methods are applied to source characterization in array processing [1], image denoising. Subspace-based methods were adapted to multidimensional -also called tensor- data [2, 3, 4]. The tensor data extend the classical vector data [2, 5]. A tensor is a multiway array, each array entry corresponding to a physical quantity. Tensor models were adopted in chemometrics [5], for DS-CDMA system characterization [6], multilinear independent component analysis [7]. In particular, subspace-based methods are employed for data denoising [2]. They rely, for each mode, on the flattening matrix singular value decomposition and on data projection upon dominant

singular vectors. Section 2 states the problem. The proposed method is described in Section 3. In Section 4, we evaluate the performances of the proposed method.

2. PROBLEM STATEMENT

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} of order N from $\mathbb{R}^{I_1 \times \dots \times I_N}$ such that: $\mathcal{R} = \mathcal{X} + \mathcal{N}$. 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{R} . By definition, $E^{(n)}$ is generated by the column vectors of the n th-mode flattening matrix. The n th-mode flattening matrix \mathbf{R}_n of tensor $\mathcal{R} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ is defined as a matrix from $\mathbb{R}^{I_n \times M_n}$, where: $M_n = I_{n+1} I_{n+2} \dots I_N I_1 I_2 \dots I_{n-1}$. The goal of various studies is to estimate the expected signal \mathcal{X} thanks to a multidimensional filtering of the data [2, 8]:

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

For all $n = 1$ to N , $\mathbf{P}^{(n)}$ is the n th-mode filter applied to the n th-mode of the data tensor \mathcal{R} . In this paper, we assume that noise \mathcal{N} is independent from 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 signal subspace [9, 10, 11].
- $E_2^{(n)}$ 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 noise subspace [9, 10, 11].

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} . For this, the classical method consists in performing the truncation of SVD of the flattening matrix of \mathcal{R} in each mode. For each n th-mode, the columns of $\mathbf{P}^{(n)}$ are the projectors on the subspace spanned by the dominant singular vectors of the flattening matrix. Filtering is then called truncation of the

HOSVD. SVD numerical cost is elevated. Moreover, multidimensional signal processing methods include an iterative ALS loop which implies multiple SVD processings. We seek for a faster method, which avoids singular value decomposition. In [4], higher order power method and higher order orthogonal iterative algorithms are proposed to compute the signal subspace vectors. However, the former method is limited to a rank one signal subspace whereas, in general, signal subspace dimension is larger than one. The latter method proposes a simultaneous estimation of leading eigenvectors. For a fast estimation of possibly multiple dominant singular vectors in each mode, we propose the fixed point method [12].

3. PROPOSED ALGORITHM FOR FAST MULTIWAY SUBSPACE-BASED FILTERING METHOD

We present in the general case the fast fixed-point algorithm for computing leading eigenvectors, and show how, in particular, this algorithm can be inserted in an ALS loop to compute signal subspace projectors for each n th-mode.

3.1. Fast fixed-point algorithm for computing leading eigenvectors

One way to compute the K orthonormal basis vectors is to use Gram-Schmidt method.

1. Choose K , the number of principal axes or eigenvectors required to estimate. Consider covariance matrix \mathbf{C} and set $p \leftarrow 1$.
2. Initialize eigenvector \mathbf{u}_p of size $d \times 1$, e. g. randomly;
3. Update \mathbf{u}_p as $\mathbf{u}_p \leftarrow \mathbf{C}\mathbf{u}_p$;
4. Do the Gram-Schmidt orthogonalization process $\mathbf{u}_p \leftarrow \mathbf{u}_p - \sum_{j=1}^{p-1} (\mathbf{u}_p^T \mathbf{u}_j) \mathbf{u}_j$;
5. Normalize \mathbf{u}_p by dividing it by its norm: $\mathbf{u}_p \leftarrow \frac{\mathbf{u}_p}{\|\mathbf{u}_p\|}$.
6. If \mathbf{u}_p has not converged, go back to step 3.
7. Increment counter $p \leftarrow p + 1$ and go to step 2 until p equals K .

The eigenvector with dominant eigenvalue will be measured first. Similarly, all the remaining $K - 1$ basis vectors (orthonormal to the previously measured basis vectors) will be measured one by one in a reducing order of dominance. The previously measured $(p - 1)^{\text{th}}$ basis vectors will be utilized 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 $\mathbf{u}_p^{+T} \mathbf{u}_p = 1$. It is usually economical to use a finite tolerance error to satisfy the convergence criterion $|\mathbf{u}_p^{+T} \mathbf{u}_p - 1| < \eta$ where η is a prior fixed threshold. Let $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K]$ 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 dominant eigenvectors. This subspace is also called "signal subspace".

3.2. Higher-order fixed point algorithm for the estimation of projectors onto signal subspaces

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. In a tensor case, the lower rank- (K_1, \dots, K_N) approximation of \mathcal{R} is represented by tensor $\mathcal{B}^{K_1, \dots, K_N}$ which minimizes the quadratic tensor Frobenius norm $\|\mathcal{R} - \mathcal{B}\|^2$ 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. We propose to replace HOSVD in tensor lower rank approximation by a higher-order fixed point algorithm (HOFP). We obtain a fast rank- (K_1, \dots, K_N) approximation:

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 HOFP:

- (a) n th-mode flatten \mathcal{R} into matrix \mathbf{R}_n ;
- (b) Compute matrix $\mathbf{U}_0^{(n)}$ formed by the K_n eigenvectors associated with the K_n largest singular values of \mathbf{R}_n . For this, use fixed point algorithm that selects dominant singular vectors (see 3.1). $\mathbf{U}_0^{(n)}$ is the initial matrix of the n th-mode signal subspace orthogonal basis vectors;
- (c) 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;
- (d) Compute the HOSVD- (K_1, \dots, K_N) of tensor \mathcal{R} given by $\mathcal{B}_0 = \mathcal{R} \times_1 \mathbf{P}_0^{(1)} \times_2 \dots \times_N \mathbf{P}_0^{(N)}$;

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

- (a) For $n = 1$ to N :
 - i. Form $\mathcal{B}^{(n),k}$: $\mathcal{B}^{(n),k} = \mathcal{R} \times_1 \mathbf{P}_{k+1}^{(1)} \times_2 \dots \times_{n-1} \mathbf{P}_{k+1}^{(n-1)} \times_{n+1} \mathbf{P}_k^{(n+1)} \times_{n+2} \dots \times_N \mathbf{P}_k^{(N)}$;
 - 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; To compute

all vectors of $\mathbf{U}_k^{(n)}$, use fixed point algorithm (see 3.1).

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

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

(c) Increment k .

4. **Output:** the estimated signal tensor is obtained through $\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{P}_{k_{stop}}^{(1)} \times_2 \cdots \times_N \mathbf{P}_{k_{stop}}^{(N)}$. $\hat{\mathcal{X}}$ is the 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.

By using fixed-point algorithm in place of the singular value decomposition in step 3(a)iv to compute the projector for each mode, a faster algorithm is expected.

4. SIMULATION RESULTS

The proposed method can be applied to any tensor data: multicomponent seismic signals, RGB color image, or hyperspectral images [2]... Color images or hyperspectral 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 multidimensional white Gaussian noise \mathcal{N} is added to signal tensor \mathcal{X} . The quality of the obtained denoising results is measured through the SNR value, with $SNR = 10 \cdot \log \frac{\|\mathcal{X}\|^2}{\|\mathcal{N}\|^2}$. We first exemplify the proposed algorithm on a low-noise 8-bit color image of size $512 \times 512 \times 3$. In this experiment we emphasize the capacity of the proposed method to preserve the expected data. We propose comparative results (in terms of SNR) between the proposed subspace-based tensor method and the wavelet-based ForWaRD algorithm [13]. A watermarked image can be considered as a raw image to which equally distributed and low-power noise composed by the watermark is added. Denoising is a watermark attack that retrieves the original non-watermarked image [14]. Fig. 1(a) and (b) present the raw and watermarked images. Watermarking is performed in the wavelet domain, by spread spectrum [15].

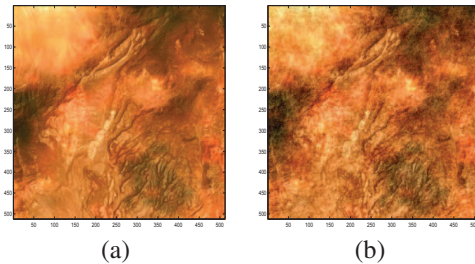


Fig. 1. (a) test image, (b) watermarked image

In the watermarked image, $SNR = 17.6$ dB. Fig. 2(a) provides the result obtained with the proposed subspace-based tensor filtering method. Fig. 2(b) provides the result

obtained by ForWaRD method when applied matrix slice by matrix slice. Signal subspace ranks (K_1, K_2, K_3) cho-

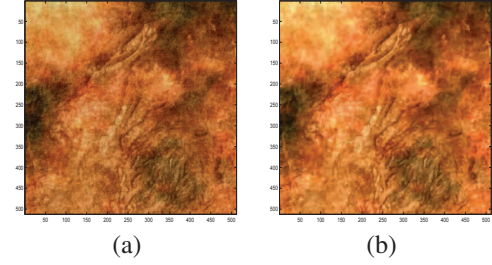


Fig. 2. denoising results: (a) proposed method, (b) ForWaRD method

sen to perform rank- (K_1, K_2, K_3) approximation are equal to $(140, 140, 3)$. 5 iterations of the ALS loop are needed for convergence. The proposed method yields 17.9 dB, and ForWaRD yields $3.08 \cdot 10^{-2}$ dB in the denoised image. The proposed method improves the SNR value although it is already elevated in the processed image, whereas ForWaRD method reduces the SNR value by smoothing details. The proposed method requires 14.4 sec. and ForWaRD 31.4 sec. ForWaRD wavelet-based method provides poor SNR results, so we next compare extensively the computational load performance of the proposed method with the performance of a subspace-based method.

We now exemplify the proposed method through hyperspectral image (HSI) denoising and compare the results obtained with another subspace-based method, namely HOSVD. 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). We consider HSI data with a large amount of noise, by setting $SNR = 3$ dB. 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 rank- (K_1, K_2, K_3) approximation are equal to $(10, 10, 15)$. Parameter η (see 3.1) is fixed to 10^{-6} , and 5 iterations of the ALS algorithm are needed for convergence. The experiments are run with a 3.0 Ghz PC running Windows. When I_1 and I_2 are equal to 128, computational loads are 8 sec. for the proposed method and 250 sec. for the comparative method. Considering an image with 256 rows and columns, HOFD-based method leads to $SNR = 17.03$ dB with a computational time equal to 68 sec. and HOSVD-based method leads to $SNR = 17.20$ dB with a computational time equal to 43 min. 22 sec. Then the proposed method is 38 times faster, yielding SNR values that differ by less than 1%. Consequently, the proposed method is particularly interesting when noise power is elevated, and data size is high. Fig. 3(a) is the raw image with $I_1 = I_2 = 256$, Fig. 3(b) is the noised image, Fig. 3(c) and (d) are the results obtained by HOFD and HOSVD algorithms.

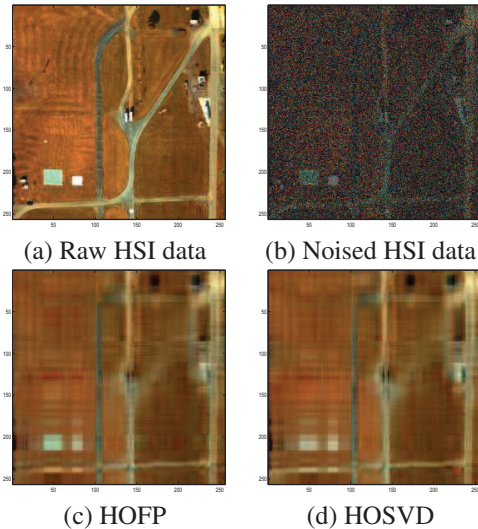


Fig. 3. HSI image : Results obtained by lower-rank tensor approximation using HOFP or HOSVD.

5. CONCLUSION

A novel algorithm for fast tensor processing is proposed. We adapt fixed point algorithm for the estimation of leading eigenvectors to a subspace-based denoising method. On the one hand we compare the proposed method with a wavelet-based approach for color image denoising. Results in terms of SNR are much better with the proposed subspace-based approach. We exemplify the proposed fast subspace-based tensor method on hyperspectral image denoising when few dominant singular vectors are required to perform denoising, and show that for images with 256 rows and columns, the proposed lower rank tensor approximation method using higher order fixed point (HOFP) algorithm is up to 38 times faster. Further, multicomponent seismic signals or array processing data could be considered. The proposed HOFP algorithm could be extended to multiway Wiener filtering.

6. REFERENCES

- [1] R.o. Schmidt, "Multiple emitter location and signal parameters estimation", *IEEE-ASSP*, vol. 4, no. 3, pp. 276-280, March 1983.
- [2] D. Muti and S. Bourennane, "Multidimensional Filtering Based on a Tensor Approach", *Signal Processing Journal, Elsevier*, vol. 85, no. 12, pp. 2338-2353, May 2005.
- [3] D. Muti, S. Bourennane and J. Marot, "Lower-Rank Tensor Approximation and Multiway Filtering", *SIAM Journal on Matrix Analysis and Applications (SIMAX)*, vol. 30, no. 3, pp. 1172-1204, sept. 08.
- [4] L. De Lathauwer, B. De Moor, and J. Vandewalle, "On the best rank-(1) and rank-(R_1, \dots, R_N) Approximation of Higher-Order Tensors", *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324-1342, April 2000.
- [5] N.D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays", *Journal of Chemometrics*, vol. 14, pp. 229-239, 2000.
- [6] N.D. Sidiropoulos, G. Giannakis, and R. Bro, "Blind PARAFAC Receivers for DS-CDMA Systems", *IEEE Trans. on Signal Processing*, vol. 48, no. 3, pp. 810-823, 2000.
- [7] M.A.O. Vasilescu and D. Terzopoulos, "Multilinear Independent Components Analysis", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 547-553, July 2005.
- [8] D. Muti and S. Bourennane, "Multidimensional estimation based on a tensor decomposition", *IEEE Workshop on Statistical SP*, St Louis, USA, Sept. 28 - Oct. 1, 2003.
- [9] K. Abed-Meraim, H. Maître, and P. Duhamel, "Blind multichannel image restoration using subspace based method", *IEEE ICASSP'03*, Hong-Kong (China), April 6-10, 2003.
- [10] N. Yuen and B. Friedlander, "Asymptotic performance analysis of blind signal copy using fourth order cumulant", *International Jour. Adaptative Contr. Signal Proc.*, vol. 48, pp. 239-65, Mar. 1996.
- [11] N. Yuen and B. Friedlander, "DOA estimation in multipath: an approach using fourth order cumulant", *IEEE Trans. on Signal Processing*, vol. 45, no. 5, pp. 1253-63, Mar. 1997.
- [12] A. Hyvärinen, E. Oja, "A fast fixed-point algorithm for independent component analysis", *Neural Comput.*, vol. 9, no. 7, pp. 1483-1492, 1997.
- [13] R. Neelamani, H. Choi, and R. G. Baraniuk, "ForWaRD: Fourier-Wavelet Regularized Deconvolution for ill-Conditioned Systems", *IEEE Trans. on Signal Processing*, Vol. 52, No. 2, pp. 418-433, Feb. 04.
- [14] A. Houmansadr, N. Kiyavash, N. Borisov, "Multi-flow attack resistant watermarks for network flows", *ICASSP'09*, Taiwan, Apr. 09, pp. 1497-1500.
- [15] K. Hayat, W. Puech and G. Gesquière, "An Adaptive Spread Spectrum (SS) Synchronous Data Hiding Strategy for Scalable 3D Terrain Visualization", *Electronic Imaging, Vis. and Data Analysis'09 SPIE*, San Jose, Cal., US, jan. 2009.