FAST TENSOR SIGNAL FILTERING USING FIXED POINT ALGORITHM

J. Marot and S. Bourennane

Institut Fresnel / UMR-CNRS, D. U. de Saint-Jérôme, 13397 Marseille cedex 20 France

ABSTRACT

Subspace-based methods rely on the selection of leading eigenvectors, associated with dominant eigenvalues. They have been extended to tensor data processing, such as denoising. Usually EVD (Eigenvalue Decomposition) is performed and data projection on leading eigenvectors results in noise reduction. Tensor processing methods, in particular multiway Wiener filtering algorithm, include an ALS (Alternating Least Squares) loop, which involves several EVDs. Fixed point algorithm is a faster method than EVD to estimate a fixed number of eigenvectors. In this paper, we adapt fixed point algorithm to the estimation of only the required leading eigenvectors in a tensor processing framework. We adapt inverse power method to estimate the required noise variance. We provide a comparative study in terms of speed through an application to hyperspectral image denoising.

***Index Terms*—** Algebra, Algorithms, Multidimensional signal processing, Wiener filtering, Image restoration.

1. INTRODUCTION

Subspace-based methods consider significant and remaining parts of the data. In particular, the eigenstructure of the covariance matrix of signal realizations provides eigenvectors which span the measurement space. Within the measurement space, leading eigenvectors span the so-called "signal subspace" and the remaining eigenvectors span the so-called "noise subspace". Subspace-based methods were adapted to multidimensional -also called tensor- data [1, 2, 3]. Tensor data generalize the classical vector and matrix data to entities with more than two dimensions [2, 4]. In this paper, the term tensor denotes a multiway (or multidimensional) array, each array entry corresponding to any quantity. Tensor models were adopted in chemometrics [4], for DS-CDMA system characterization [5], for facial expression classification by multilinear independent component analysis [6]. In particular, subspace-based methods are employed for data denoising [2]. They include ALS (Alternating Least Squares) loops whose computational load is dominated by EVDs' one. Section 2 reminds the principles of multiway Wiener filtering (MWF). Section 3 states the issue solved in this paper, that is, how to reduce the computational load of the ALS loop in multiway Wiener filtering. Section 4 shows how to adapt fixed point algorithm and inverse power method to obtain a fast version of multiway Wiener filtering. Section 5 evaluates the performances of the proposed method, especially in terms of computational load, and with an application to hyperspectral image (HSI) denoising.

2. OVERVIEW OF MULTIWAY WIENER FILTERING

The measurement of a multiway signal $\mathcal{X}$ by multicomponent sensors with additive noise $\mathcal{N}$ results in a data tensor $\mathcal{R}$ of order N from $\mathbb{C}^{I_1\times\cdots\times I_N}$ such that : $\mathcal{R} = \mathcal{X} + \mathcal{N}$. Let us define $E^{(n)}$ as the nth-mode vector space of size I_n, associated with the nth-mode of tensor $\mathcal{R}$. By definition, $E^{(n)}$ is generated by the column vectors of the nth-mode flattening matrix. The nth-mode flattening matrix $\mathbf{R}_n$ of tensor $\mathcal{R} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ is defined as a matrix from $\mathbb{C}^{I_n\times M_n}$, where : $M_n = I_{n+1}I_{n+2}\cdots I_N I_1 I_2\cdots I_{n-1}$. The goal of various studies is to estimate the expected signal $\mathcal{X}$ thanks to a multidimensional filtering of the data [2] :

$$\widehat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \cdots \times_N \mathbf{H}^{(N)}, \tag{1}$$

For all $n = 1$ to N, $\mathbf{H}^{(n)}$ is the nth-mode filter applied to the nth-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 nth-mode rank K_n is smaller than the nth-mode size 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 nth-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 [7, 8, 9].

• $E_2^{(n)}$ 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 [7, 8, 9].

Hence, one way to estimate signal tensor $\mathcal{X}$ from noisy data tensor $\mathcal{R}$ is to estimate $E_1^{(n)}$ in every nth-mode of $\mathcal{R}$. The nth-mode filters $\mathbf{H}^{(n)}$, $n = 1$ to N, of Eq. (1) optimize an estimation criterion. The most classical method consists in minimizing the mean squared error criterion, given in Eq. (1), between the expected signal tensor $\mathcal{X}$ and the estimated signal tensor $\widehat{\mathcal{X}}$:

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

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

The minimization of Eq. (2) with respect to filter $\mathbf{H}^{(n)}$, for fixed $\mathbf{H}^{(m)}$, $m \neq n$, leads to the following expression of nth-mode Wiener filter :

$$\mathbf{H}^{(n)} = \gamma_{\mathbf{XR}}^{(n)}{\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}}^{-1}, \quad (3)$$

where $\gamma_{\mathbf{XR}}^{(n)} = \mathrm{E}\left[\mathbf{X}_n\mathbf{T}^{(n)}\mathbf{R}_n^T\right]$. The expression of $\mathbf{T}^{(n)}$ is given in [2], and : $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)} = \mathrm{E}\left[\mathbf{R}_n\mathbf{Q}^{(n)}\mathbf{R}_n^T\right]$, with : $\mathbf{Q}^{(n)} = {\mathbf{T}^{(n)}}^T\mathbf{T}^{(n)}$. To obtain $\mathbf{H}^{(n)}$ through Eq. (3), 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 $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$ can be derived, and not the term $\gamma_{\mathbf{XR}}^{(n)}$. Hence, some more assumptions on $\mathcal{X}$ have to be made in order to overcome the indetermination over $\gamma_{\mathbf{XR}}^{(n)}$ [2]. 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, [2] proposed to consider that the nth-mode flattening matrix $\mathbf{X}_n$ can be expressed as a weighted combination of K_n vectors from the nth-mode signal subspace $E_1^{(n)}$:

$$\mathbf{X}_n = \mathbf{V}_s^{(n)}\mathbf{O}^{(n)}, \quad (4)$$

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 nth-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 nth-mode flattening matrix $\mathbf{X}_n$ is orthogonal to nth-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 $\mathcal{N}$ is white and independent from signal $\mathcal{X}$, and introducing the signal model of Eq. (4) in Eq. (3) leads to a computable expression of nth-mode Wiener filter $\mathbf{H}^{(n)}$:

$$\mathbf{H}^{(n)} = \mathbf{V}_s^{(n)}\gamma_{\mathbf{OO}}^{(n)}{\mathbf{\Lambda}_{\mathbf{\Gamma s}}^{(n)}}^{-1}{\mathbf{V}_s^{(n)}}^T, \quad (5)$$

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

$$\gamma_{\mathbf{OO}}^{(n)}{\mathbf{\Lambda}_{\mathbf{\Gamma s}}^{(n)}}^{-1} = \mathrm{diag}\left[\frac{\beta_1}{\lambda_1^{\Gamma}},\cdots,\frac{\beta_{K_n}}{\lambda_{K_n}^{\Gamma}}\right], \quad (6)$$

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

$$\beta_{k_n} = \lambda_{k_n}^{\gamma} - {\sigma_{\Gamma}^{(n)}}^2, \ \forall\, k_n = 1,\ldots,K_n \quad (7)$$

Superscript γ refers to the $\mathbf{T}^{(n)}$-weighted covariance, and subscript $\mathbf{\Gamma}$ to the $\mathbf{Q}^{(n)}$-weighted covariance. ${\sigma_{\Gamma}^{(n)}}^2$ is the degenerated eigenvalue of noise $\mathbf{T}^{(n)}$-weighted covariance matrix $\gamma_{\mathbf{NN}}^{(n)} = \mathrm{E}\left[\mathbf{N}_n\mathbf{T}^{(n)}\mathbf{N}_n^T\right]$. Thanks to the additive noise and 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 :

$$\widehat{\sigma}_{\Gamma}^{(n)^2} = \frac{1}{I_n - K_n}\sum_{k_n=K_n+1}^{I_n}\lambda_{k_n}^{\gamma}. \quad (8)$$

To determine the nth-mode Wiener filters $\mathbf{H}^{(n)}$ that minimize the mean squared error in Eq. (2), the ALS algorithm has been proposed in [2].

3. PROBLEM STATEMENT

Among other computations, the ALS algorithm used in MWF involves the computation of several EVDs, in the following steps, for each mode and each iteration index :

- Compute $\gamma_{\mathbf{RR}}^{(n)} = \mathrm{E}[\mathbf{R}_n{\mathbf{R}_n^{(n),k}}^T]$, determine $\lambda_1^{\gamma},\ldots,\lambda_{K_n}^{\gamma}$, the K_n largest eigenvalues of $\gamma_{\mathbf{RR}}^{(n)}$;
- Compute $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)} = \mathrm{E}[\mathbf{R}_n^{(n),k}{\mathbf{R}_n^{(n),k}}^T]$, determine $\lambda_1^{\Gamma},\ldots,\lambda_{K_n}^{\Gamma}$, the K_n largest eigenvalues of $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$;
- For $k_n = 1$ to I_n, estimate ${\sigma_{\Gamma}^{(n)}}^2$ thanks to (8) and for $k_n = 1$ to K_n, estimate β_{k_n} thanks to (7).

The numerical cost of the ALS loop is dominated by the one of the EVDs which are performed therein. We seek for a faster method, which avoids eigenvalue decomposition. In [3], higher order power method and higher order orthogonal iterative algorithms are proposed instead of EVD to compute the signal subspace vectors. For a fast estimation of possibly multiple leading eigenvectors in each mode, various methods were proposed [10, 11, 12]. For instance QR factorization [10] is a fast method to estimate all eigenelements of any matrix, but as we seek for only a few of them. So we propose the fixed point algorithm which is presented in [13]. MWF requires not only leading eigenvectors but also dominant eigenvalues and the noise variance. We propose to use the inverse power method to obtain the noise power. With these two methods, we show that the computational load is reduced while keeping the same denoising performances.

4. PROPOSED FAST ALGORITHM

One way to compute the K orthonormal basis eigenvectors of any matrix $\mathbf{C}$ is to use the fixed-point algorithm proposed in [13].

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 $\left|\left|\mathbf{u}_p^{it^T}\mathbf{u}_p^{it-1} - 1\right|\right| < \eta$:
   (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 $\mathbf{u}^{it}{}_p \leftarrow \mathbf{u}^{it}{}_p - \sum_{j=1}^{j=p-1}(\mathbf{u}^{it}{}_p^T\mathbf{u}^{it}{}_j)\mathbf{u}^{it}{}_j$;
   (c) Normalize $\mathbf{u}_p^{it}$ by dividing it by its norm : $\mathbf{u}_p^{it} \leftarrow \frac{\mathbf{u}_p^{it}}{||\mathbf{u}_p^{it}||}$.
   (d) Increment counter $it \leftarrow it + 1$
3. Increment counter $p \leftarrow p + 1$ and go to step 1 until p equals K.

However, the fixed-point algorithm itself is not enough to replace EVD in the MWF ALS loop. MWF requires not only leading eigenvectors for each mode but also the associated dominant eigenvalues, and the weight matrix requires eigenvalues of signal and data covariance flattening matrices $\gamma_{\mathbf{RR}}{}^{(n)}$ and $\mathbf{\Gamma_{RR}}{}^{(n)}$, see Eq. (6). This can be achieved in MWF algorithm by the following computation which yields the expected eigenvalues :
$\mathbf{V}_s^{(n)^T}\gamma_{\mathbf{RR}}^{(\mathbf{n})}\mathbf{V}_s^{(n)} = diag\{[\lambda_1^{\gamma}, \ldots, \lambda_{K_n}^{\gamma}]\}$.
In the same way, we have :
$\mathbf{V}_s^{(n)^T}\mathbf{\Gamma}_{\mathbf{RR}}^{(\mathbf{n})}\mathbf{V}_s^{(n)} = diag\{[\lambda_1^{\Gamma}, \ldots, \lambda_{K_n}^{\Gamma}]\}$.
Thus, β_{k_n} can be computed following Eq. (7). But it also requires the $I_n - K_n$ smallest eigenvalues of $\gamma_{\mathbf{RR}}{}^{(n)}$, equal to $\sigma_{\Gamma}^{(n)}$, see Eq. (8). Thus, we adapt the inverse power method to retrieve the smallest eigenvalue of $\gamma_{\mathbf{RR}}{}^{(n)}$ in the following algorithm :

1. Initialize randomly $\mathbf{x}_0$ of size $K_n \times 1$.
2. While $\frac{\|\mathbf{x}-\mathbf{x_0}\|}{\|\mathbf{x}\|} \leq \epsilon$ do :
   (a) $\mathbf{x} = \gamma_{\mathbf{RR}}{}^{(n)^{-1}} \cdot \mathbf{x}_0$
   (b) $\lambda = \|\mathbf{x}\|$
   (c) $\mathbf{x} = \frac{\mathbf{x}}{\lambda}$
   (d) $\mathbf{x_0} \leftarrow \mathbf{x}$
3. $\sigma_{\Gamma}^{(n)} = \frac{1}{\lambda}$

Therefore, $\sigma_{\Gamma}^{(n)^2}$ can be estimated and Eq. (6) can be computed in a fast way.
Let's consider the computational loads of the proposed and existing algorithms : The dominant operation in the ALS loop used in MWF is the eigendecomposition of the size $I_n \times I_n$ matrices $\gamma_{\mathbf{RR}}^{(n)}$ and $\mathbf{\Gamma}_{\mathbf{RR}}^{(n)}$ (see section 3). So the computational complexity of the ALS loop is of same order of magnitude as the one of either EVD (eigenvalue decomposition) or fixed point algorithm. One well-known EVD method is the Cyclic Jacobi's method [12]. The Jacobi's method which diagonalizes a symmetric matrix requires around $O(I_n^3)$ computations [12]. Concerning the fixed point algorithm : Let It be the number of iterations.
Gram-Schmidt orthogonalization for all basis vectors $\mathbf{u}_p$ and all values $p = 1, \ldots, K$: $O(I_n K^2 It)$ operations; updating process for all $p = 1, \ldots, K$ basis vectors : $O(I_n^2 K It)$ operations. Then the estimated total number of operations is :
$O(I_n K^2 It + I_n^2 K It) \approx O(I_n^2 K + I_n K^2)$.
If I_n is large compared to K the computational complexity can be estimated to be $O(I_n^2)$. Then, replacing SVD by fixed point algorithm, the gain is of one order of magnitude.

5. SIMULATION RESULTS

The proposed method can be applied to any tensor data, such as color image, multicomponent seismic signals, or hyperspectral images [2]. We exemplify the proposed method with hyperspectral image (HSI) denoising. 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 Gaussian noise $\mathcal{N}$ is added to signal tensor $\mathcal{X}$. The SNR_{in} (in dB) in the noisy data tensor, $\mathcal{R}$, is as follows : $SNR_{in} = 10 \cdot \log\left(\frac{\|\mathcal{X}\|^2}{\|\mathcal{N}\|^2}\right)$. To quantify *a posteriori* the quality of estimation, we define the SNR_{out} (in dB) in the estimated tensor, $\widehat{\mathcal{X}}$ as : $SNR_{out} = 10 \cdot \log\left(\frac{\|\mathcal{X}\|^2}{\|\mathcal{X}-\hat{\mathcal{X}}\|^2}\right)$ where $\| \mathcal{X} - \hat{\mathcal{X}} \|^2$ is the noise power in the denoised tensor. We consider HSI data with a large amount of noise, by setting $SNR_{in} = 7.4$ 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. The number of spectral bands I_3 is fixed to 148. Signal subspace ranks (K_1, K_2, K_3) chosen to perform multiway Wiener filtering are $(10, 10, 15)$. The number of iterations for fixed point algorithm is fixed to 5, and 5 iterations of the ALS algorithm are needed for convergence in multiway Wiener filtering. We denote by MWF the multiway Wiener filtering method that uses EVD, and by MWFFP the multiway Wiener filtering method that uses fixed point algorithm and inverse power method. Fig. 1 provides the evolution of computational times for both MWFFP and MWF-based tensor denoising, for values of I_1 and I_2 varying between 60 and 256. With a 3.0 Ghz PC running Windows, considering an image with 256 rows and columns, MWFFP-based method leads to $SNR_{out} = 17.11$ dB with a computational time equal to 35 sec. and MWF-based method leads to $SNR_{out} = 17.27$ dB with a computational time equal to 17 min. 4 sec. Then the proposed method is 28 times faster, yielding SNR_{out} values that differ by less than 1%. The proposed method is 20 times faster if $K_1 = K_2 = 20$ and $K_3 = 30$. Fig. 2(a) is the raw image with $I_1 = I_2 = 256$, Fig. 2(b) provides the noised image, Fig. 2(c) and (d) are respectively the results obtained by MWFFP and MWF algorithms.

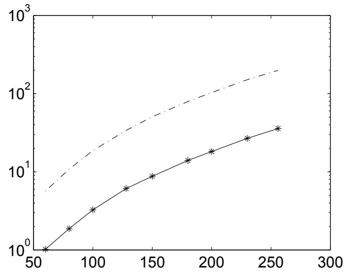
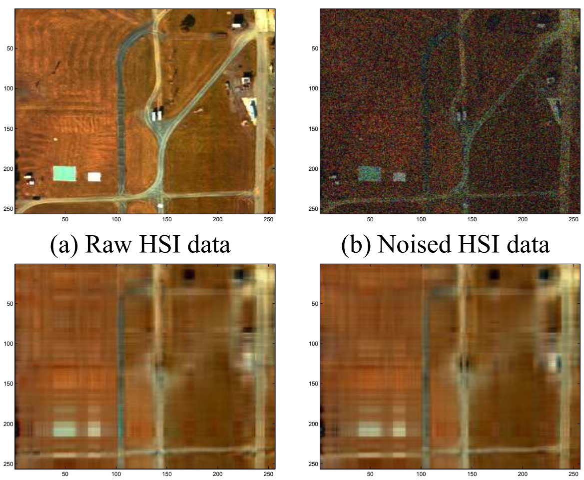


Fig. 1. *Computational times (in sec.) as a function of the number of rows and columns : tensor lower-rank approximation using MWFFP (-*-), MWF (-·-).*



(a) Raw HSI data (b) Noised HSI data

(c) MWFFP (d) MWF

Fig. 2. *HSI image : Results obtained by lower-rank tensor approximation using MWFFP or MWF.*

6. CONCLUSION

We propose an algorithm that improves multiway Wiener filtering in terms of computational load, by using fixed point algorithm and inverse power method instead of EVD. We adapt fixed point algorithm for the estimation of leading eigenvectors and inverse power method for the estimation of noise power to a subspace-based multiway denoising method. We exemplify the proposed method on hyperspectral image denoising when a low number of leading eigenvectors, compared to the image size, are required to perform denoising. We show that for images with 256 rows and columns, and compared to the considered algorithms, the proposed accelerated multiway Wiener filtering method using fixed point algorithm is up to 28 times faster than the version using EVD. The superiority of fixed point algorithm respect to EVD is especially pronounced for low values of signal subspace rank and large matrix sizes. Further, multicomponent seismic signals or array processing data could be considered.

7. REFERENCES

[1] D. Muti, S. Bourennane, and J. Marot, ”Lower-rank Tensor Approximation and Multiway Filtering”, *”SIAM Journal on Matrix Analysis and Applications”*, accepted sept. 2007.

[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] L. De Lathauwer, B. De Moor, and J. Vandewalle, ”On the best rank-(1) and rank-($R_1, \ldots, R_N$) Approximation of Higher-Order Tensors”, *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324-1342, Apr. 2000.

[4] N.D. Sidiropoulos and R. Bro, ”On the uniqueness of multilinear decomposition of N-way arrays”, *Journal of Chemometrics*, vol. 14, pp. 229-239, Mar. 2000.

[5] N.D. Sidiropoulos, G. Giannakis, and R. Bro, ”Blind PARAFAC Receivers for DS-CDMA Systems”, *IEEE Trans. on SP*, vol. 48, no. 3, pp. 810-823, 2000.

[6] M.A.O. Vasilescu and D. Terzopoulos, ”Multilinear Independent Components Analysis”, *Proc. IEEE Conf. on CVPR*, vol. 1, pp. 547-553, July 2005.

[7] K. Abed-Meraim, H. Maître, and P. Duhamel, ”Blind multichannel image restoration using subspace based method”, *IEEE-ICASSP*, Hong-Kong (China), Apr. 6-10, 2003.

[8] 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.

[9] N. Yuen and B. Friedlander, ”DOA estimation in multipath : an approach using fourth order cumulant”, *IEEE Trans. on SP*, vol. 45, no. 5, pp. 1253-63, Mar. 1997.

[10] G. H. Golub, H. A. van der Vorst, ”Eigenvalue computation in the 20th century”, *Journal of Computational and Applied Mathematics*, vol. 123, pp. 35-65, Mar. 2000.

[11] P. Comon and G.H. Golub, ”Tracking a few extreme singular values and vectors in signal processing,” *Proc. IEEE*, vol. 78, no. 8, pp. 1327-1343, 1990.

[12] G.H. Golub and C.F. Van Loan, ”Matrix Computation”, third ed., The John Hopkins University Press, 1996.

[13] A. Hyvärinen, E. Oja, ”A fast fixed-point algorithm for independent component analysis”, *Neural Comput.*, vol. 9, no. 7, pp. 1483-1492, 1997.