RECENT ADVANCES ON TENSOR MODELS AND THEIR RELEVANCE FOR MULTIDIMENSIONAL DATA PROCESSING

Julien Marot, Salah Bourennane

Groupe GSM, Institut Fresnel, École Centrale Marseille, Aix Marseille Université
Domaine Universitaire de Saint Jérôme Av. Escadrille Normandie, 13397, Marseille, France

ABSTRACT

This paper reviews the last advances which concerned tensor methods based on three main decompositions: Tucker, Parafac, and Paratuck. We show how they improved the processing of multidimensional data such as hyperspectral images and multiple input multiple output signals. First, we show how multiway Wiener filtering, based on Tucker decomposition, was set in a wavelet framework. Secondly, we remind how signal dependent noise is handled while applying the truncation of the Parafac decomposition. Thirdly, we review the sequential Parafac Paratuck decomposition and exemplify its interest for a fast characterization of channel and symbols in a MIMO framework.

***Index Terms*—** Tensor decomposition; Tucker; Parafac; Paratuck

1. INTRODUCTION

In the frame of multidimensional data processing, tensor methods were proposed to take into account the relationships between modes of the data, as opposed to slice-by-slice methods [1]. These tensor methods rely on tensor decompositions which extend the singular value decomposition which is valid for two-dimensional data: these are the Tucker, and the Parafac-CANDECOMP [2, 3, 4] decompositions. Recently, an 'ad hoc' decomposition, called Paratuck, has been proposed to characterize MIMO (multiple input, multiple output) signals [5].

Goal and contributions This paper proposes an overview of three main tensor decompositions (Tucker, Parafac, and Paratuck) and emphasizes the related last advances and their interest for multidimensional data processing, namely hyperspectral images (HSI) and MIMO signals.

Relation with previous work in the field An overview of lower-rank tensor approximations, which is focused on the Tucker decomposition, was proposed in [6]. In [6], tensor filtering is presented as a method to retrieve an expected multidimensional signal from its noisy version. A more general review of tensor decompositions and their applications was proposed in [7]. We propose a brief overview of some work performed since then.

Outline Since the publication of [6] and [7], some extensions have been proposed, for the minimum rank approximation of matrices [8] or tensors, based either on Tucker decomposition [9, 10] or on Parafac decomposition [11]. Also, a fast method combining Parafac and Paratuck decomposition has been proposed to estimate blindly the channel and symbols of a multidimensional MIMO signal [12]. We present these last advances in Sections 2, 3, and 4 respectively. The conclusions are given in Section 5. A common problematic to the three main parts of the paper is the following: How to retrieve the underlying useful-signal part from multidimensional data ?

Notation $\mathcal{X}$ denotes a multidimensional array also called 'tensor', **X** a 2-D matrix, **x** a 1-D vector, and x a scalar.

We model a noisy HSI as a tensor resulting from a multidimensional signal $\mathcal{X}$ impaired by an additive white noise $\mathcal{N}$ [6]. The HSI tensor can be expressed as: $\mathcal{R} = \mathcal{X} + \mathcal{N}$. Tensors $\mathcal{R}$, $\mathcal{X}$, and $\mathcal{N}$ are of size $I_1 \times I_2 \times I_3$. For each spectral band indexed by $i = 1, \ldots, I_3$, the noise $\mathbf{N}(:,:,i)$ is assumed stationary zero-mean, possibly signal dependent (SD). In sections 2 and 3 we aim to get an estimate $\hat{\mathcal{X}}$ of the desired tensor signal $\mathcal{X}$ from the raw tensor $\mathcal{R}$.

2. TUCKER DECOMPOSITION IN WAVELET FRAMEWORK FOR TARGET DETECTION

The Tucker decomposition has proved to be effective while performing dimensionality reduction, that is, data compression along each mode of the data, through the higher-order singular value decomposition (HOSVD) where the singular vectors associated with the dominant singular values are selected from the decomposition along each mode of the multidimensional data [13]. Each scalar component of tensor $\mathcal{X}$ is expressed as follows:

$$x_{i_1,i_2,i_3} = \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \sum_{k_3=1}^{K_3} \left(a_{i_1,k_1} b_{i_2,k_2} c_{i_3,k_3}\right) g_{k_1,k_2,k_3}, \quad (1)$$

where the components a, b, and c stand for the factor terms, the components g are the scalar components of the core tensor [6], and K_1, K_2, K_3 are the ranks for each mode. The Tucker decomposition leads to an optimal retrieval of the useful part of the data in terms of mean square error, thus performing data denoising [6]. The specific problematic in this section is how to denoise an HSI while preserving small features to eventually detect targets even when they are small. Owing to the ability of the wavelet transform to separate high frequency components from low frequency ones, wavelets are privileged tools for image denoising [14]. That is why, recently, the wavelets and tensor framework got closer to each other: the multiway Wiener filtering, based on the Tucker decomposition, has been introduced in a wavelet framework to preserve small targets while denoising hyperspectral images [9]. This method, called MWF-MWPT (multiway Wiener filtering-multidimensional wavelet packet transform), preserves well the spatial details.

2.1. Overview of the multiway Wiener filtering

The optimization criterion classically used to derive the Wiener filter is the minimization of the mean square error e between the desired and the estimated data. It was extended to $third$-order tensors:

$$e(\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \mathbf{H}^{(3)}) = \mathrm{E}\left[\left\|\mathcal{X} - \hat{\mathcal{X}}\right\|^2\right] \tag{2}$$

where $\mathbf{H}^{(1)}$, $\mathbf{H}^{(2)}$, $\mathbf{H}^{(3)}$ are the first, second, and third-mode filters. Through the minimization of Eq. (2), the estimated data $\hat{\mathcal{X}}$ can be expressed as follows [6]:

$$\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}^{(1)} \times_2 \mathbf{H}^{(2)} \times_3 \mathbf{H}^{(3)}, \tag{3}$$

The n-mode filters $\mathbf{H}^{(n)}$ are obtained using an Alternating Least Squares (ALS) algorithm which solves the linear optimization problem of Eq. (2). Their computation involves the Tucker decomposition of tensor $\mathcal{R}$; the expression of each n^{th}-mode filter can be found in [15]. It depends on the singular vectors associated with the dominant singular values and on the signal to noise ratio. A last advance in this field consisted in introducing the MWF in a wavelet framework.

2.2. Multiway Wiener filter in multidimensional wavelet packet domain

Multidimensional wavelet packet transform exhibits the ability to separate the high frequency components from the low frequency components when applied to multidimensional data. As the abundant signals are concentrated in the low-frequency components, and the rare and small signals are concentrated in the high-frequency components [9], we expect from performing MWF in the wavelet packet domain that we will better preserve the rare features in the data, such as targets when hyperspectral images are considered. This subsection proves the ability of this method, named MWPT-MWF (multidimensional wavelet packet transform - multiway Wiener filter), to minimize the MSE between $\mathcal{X}$ and $\hat{\mathcal{X}}$ (see Eq. (2)).

By performing MWPT to tensor $\mathcal{R}$, $\mathcal{X}$ and $\mathcal{N}$, we obtain:

$$\begin{aligned} &\mathcal{R} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \\ &= \mathcal{X} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \\ &+ \mathcal{N} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \end{aligned} \tag{4}$$

The coefficient tensor of each part is denoted by:

$$\mathcal{C}_{\mathbf{l}}^{\mathcal{R}} = \mathcal{R} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \tag{5}$$

$$\mathcal{C}_{\mathbf{l}}^{\mathcal{X}} = \mathcal{X} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \tag{6}$$

$$\mathcal{C}_{\mathbf{l}}^{\mathcal{N}} = \mathcal{N} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \tag{7}$$

and the coefficient tensor of the estimate $\hat{\mathcal{X}}$ is denoted by:

$$\hat{\mathcal{C}}_{\mathbf{l}}^{\mathcal{X}} = \hat{\mathcal{X}} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \tag{8}$$

With the extraction process proposed in [9], we obtain the components of each frequency $\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}}$, $\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}}$ and $\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{N}}$ from $\mathcal{C}_{\mathbf{l}}^{\mathcal{R}}$, $\mathcal{C}_{\mathbf{l}}^{\mathcal{X}}$ and $\mathcal{C}_{\mathbf{l}}^{\mathcal{N}}$ respectively. We obtain:

$$\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}} = \mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}} + \mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{N}} \tag{9}$$

Minimizing the MSE between $\mathcal{X}$ and its estimate $\hat{\mathcal{X}}$ is equivalent to minimize the MSE between $\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}}$ and $\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}}$ for each $\mathbf{m}$ [9]. The estimate of each wavelet coefficient $\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}}$ is computed by multiway Wiener filtering of $\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}}$:

$$\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}} = \mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}} \times_1 \mathbf{H}_{1,\mathbf{m}} \times_2 \mathbf{H}_{2,\mathbf{m}} \times_3 \mathbf{H}_{3,\mathbf{m}} \tag{10}$$

where $\mathbf{H}_{1,\mathbf{m}}, \mathbf{H}_{2,\mathbf{m}}, \mathbf{H}_{3,\mathbf{m}}$ are the n-mode filters of the multiway Wiener [9], differing for each $\mathbf{m}$. After estimating $\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}}$ for each $\mathbf{m}$, we obtain $\hat{\mathcal{C}}_{\mathbf{l}}^{\mathcal{X}}$ by concatenating all coefficients $\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}}$. Furthermore, the estimate $\hat{\mathcal{X}}$ can be obtained by inverse MWPT:

$$\hat{\mathcal{X}} = \hat{\mathcal{C}}_{\mathbf{l}}^{\mathcal{X}} \times_1 \mathbf{W}_1^T \times_2 \mathbf{W}_2^T \times_3 \mathbf{W}_3^T \tag{11}$$

Figure 1 presents a target detection result obtained after denoising by MWF, and by the novel-most MWPT-MWF. This result (extracted from [10]) shows that MWPT-MWF, by processing separately the rare and the abundant data, preserves better the small targets.

In the next section, we show how the truncation of the Parafac decomposition permits to remove signal dependent noise from tensor data.

3. TRUNCATION OF PARAFAC DECOMPOSITION FOR SD NOISE REDUCTION IN HSI

Following the results obtained by Kruskal [4], a three-dimensional array can be decomposed into a sum of triads, and the so-called 'Parafac rank' is the minimum number of triads which permits to reconstruct this three-dimensional

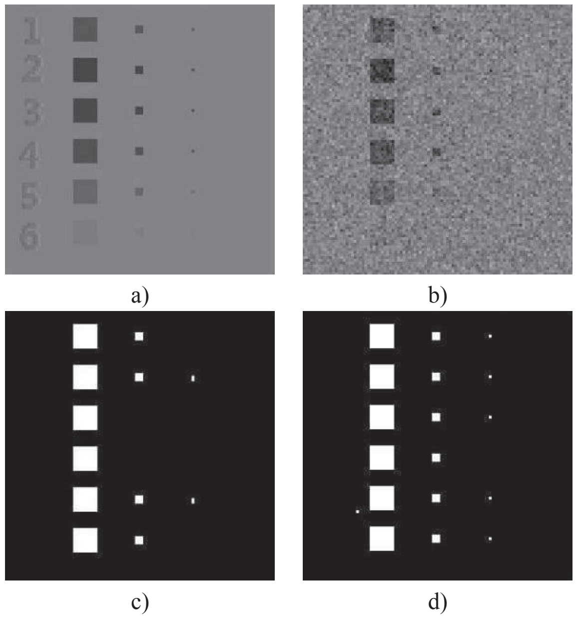

Fig. 1. a) Pure image; b) Noised image; c) Detection after MWF; d) Detection after MWPT-MWF. (extracted from [16])

array. Following the Parafac model, the scalar components of a tensor $\mathcal{X}$ with Parafac rank T can be expressed as follows:

$$x_{i_1,i_2,i_3} = \sum_{k=1}^{T} (a_{i_1,k} b_{i_2,k} c_{i_3,k}) \lambda_k, \tag{12}$$

where the terms a, b and c stand for the factor terms, and coefficients λ weight the contribution of each term of the decomposition. We notice that Parafac decomposition is a special case of Tucker decomposition (see Eq. (1)) where the core tensor is diagonal. In the field of HSI processing, we consider the truncation of the Parafac decomposition, that is, we select adequately part of the triads of the decomposition, to remove the undesired components such as noise. The specific problematic in this section is how to remove signal dependent noise from hyperspectral images, while minimizing a square error between estimated and raw tensor.

3.1. Truncation of the parafac decomposition: overview

In Eq. (12), it is assumed that the rank of tensor $\mathcal{X}$ is T. Selecting K factors out of $T > K$ factors required for an exact tensor reconstruction is called 'truncation' of the Parafac decomposition.

It has been proved in [17] that truncating the PARAFAC decomposition of a tensor $\mathcal{R}$ with rank T to the K terms weighted by the K largest values among $\{\lambda_1, \lambda_2, \ldots, \lambda_T\}$ yields the minimum square error SE= $\|\mathcal{R} - \hat{\mathcal{X}}\|^2$ between raw tensor $\mathcal{R}$ and estimate $\hat{\mathcal{X}}$. Assuming without loss of generality that the λ values are correctly ordered, the truncation of the Parafac decomposition consists in selecting the K first terms in Eq. (13). In a tensor form, we get the following estimated tensor:

$$\hat{\mathcal{X}} = \sum_{k=1}^{K} \lambda_k \mathbf{a}_k \circ \mathbf{b}_k \circ \mathbf{c}_k \tag{13}$$

where K is the rank, $\hat{\mathcal{X}}$ is the truncation of the Parafac decomposition of $\mathcal{R}$, also called rank-K approximation of $\mathcal{R}$, and vectors $\mathbf{a}_k = [a_{1,k}, \ldots, a_{I_1,k}]^T$, $\mathbf{b}_k = [b_{1,k}, \ldots, b_{I_2,k}]^T$, $\mathbf{c}_k = [c_{1,k}, \ldots, c_{I_3,k}]^T$ contain the factor terms for the k^{th} index of the decomposition. As shown in [9], $\hat{\mathcal{X}}$ is a close estimate of the noise-free tensor $\mathcal{X}$, especially when the noise magnitude is low.

3.2. Removal of signal dependent noise in HSI

Before applying an algorithm for signal dependent noise removal, we assume that signal independent noise has been removed, with any method, for instance multiway Wiener filtering (see Section 2). The signal dependent noise reduction algorithm proposed in [11] consists in applying the truncation of the Parafac decomposition of the tensor data with the appropriate rank. The novel-most aspect of this work is the algorithm which permits to estimate the best possible rank for the truncation. This algorithm is as follows:
Within $k_{SD} = \{1, \cdots, \min\{I_1 I_2, I_1 I_3, I_2 I_3\}\}$:

1. At each k_{SD}, apply Parafac truncation to $\mathcal{X} + \mathcal{N}_{SD}(\mathcal{X})$ to get $\hat{\mathcal{X}}$
2. Estimate the remaining noise
$$\hat{\mathcal{N}}_{SD} = \mathcal{X} + \mathcal{N}_{SD}(\mathcal{X}) - \hat{\mathcal{X}}$$
3. Get the 'whitened' covariance matrix:
$$\hat{\mathbf{C}}_{\hat{\mathcal{N}}_{SD}} = \mathbf{C}_{\hat{\mathcal{N}}_{SD}} \, \mu^{-1}$$
μ is the useful data mean value
4. Check if the noise covariance matrix is diagonal. For this, perform the following tests:
   - $\frac{1}{I_n} \sum_{i_n=1}^{I_n} (\hat{c}_{i_n,i_n} - \frac{1}{I_n} \sum_{i_n=1}^{I_n} \hat{c}_{i_n,i_n})^2 < \delta_1$
   - $|||\hat{\mathbf{C}}_{\hat{\mathcal{N}}_{SD}}||^2 - \sum_{i_n=1}^{I_n} \hat{c}^2_{i_n,i_n}| < \delta_2$,

   where δ_1 and δ_1 are fixed thresholds. If the two tests above hold, we can infer that only signal part remains. In this case, choose K_{SD} as k_{SD}, the optimal rank for the Parafac decomposition and use K_{SD} to estimate $\hat{\mathcal{X}}$. Otherwise $k_{SD} \leftarrow k_{SD} + 1$ and go to beginning.

From the results presented in Fig. 2, the truncation of the Parafac decomposition removes low power noise in HSI images, even when the noise is signal dependent.

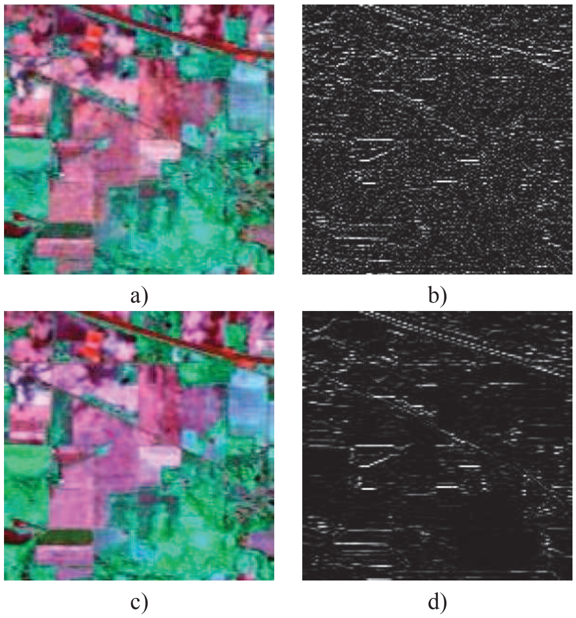

Fig. 2. a) AVIRIS HSI; b) SD noise in AVIRIS HSI; c) Denoised by Parafac truncation; d) Noise in denoised image.

4. FAST PARAFAC-PARATUCK DECOMPOSITION FOR MIMO SIGNAL CHARACTERIZATION

A remarkable interest of the Parafac decomposition relies in its uniqueness properties [4, 18]: the uniqueness of three-way array Parafac decompositions guarantees the identifiability of signal and propagation parameters in MIMO systems. However, contrary to the Tucker decomposition, the Parafac decomposition allows no crossed interaction between factor terms: in Eq. (12), for any triplet $\{i1, i2, i3\}$, the triad $a_{i1,k}b_{i2,k'}c_{i3,k''}$ contributes to the decomposition only if $k = k' = k''$. That is one of the reasons why, in the frame of MIMO signal characterization, the Paratuck decomposition was introduced [5]. We consider a MIMO wireless communication system with K receive antennas. The transmited signal contains P data-blocks with data streams composed of N information symbols. Data are transmitted with M 'transmit' antennas. The signal component received on the k^{th} receive antenna, for the n^{th} symbol, and the p^{th} data-block is denoted by $x_{k,n,p}$, and all components are set in the multidimensional array $\mathcal{X} \in \mathbb{C}^{K \times N \times P}$. We denote by $s_{n,r}$ the n^{th} symbol of the r^{th} data stream, by $h_{k,m}$ the fading coefficients between transmit and receive antennas, and by $w_{m,r}$ the precoding coefficients between transmit antennas and data streams. Moreover, $\phi_{p,m}$ are antenna allocation terms from transmit antenna to data-block, and $\psi_{p,r}$ allocation terms for the allocation of any stream to any data-block. The signal components are expressed as:

$$x_{k,n,p} = \sum_{m=1}^{M} \sum_{r=1}^{R} (h_{k,m} s_{n,r}) \phi_{p,m} \psi_{p,r} w_{m,r}, \qquad (14)$$

This expression was named Paratuck decomposition of $\mathcal{X}$ [5]. As the Tucker and the Parafac decomposition, it contains factor terms, namely $h_{k,m}$ and $s_{n,r}$, and it allows an interaction between all these factor terms, with a pattern controled by coefficients $\phi_{p,m}$ and $\psi_{p,r}$ (either 0 or 1). The precoder term $w_{m,r}$ accounts for the contribution of the factor pair $\{h_{k,m}, s_{n,r}\}$. Contrary to the more general Tucker decomposition, there are only two (though a *third* order tensor is decomposed) sets of factor terms for which all interactions are possible. A specific problematic in this section is how to reach a fast, blind estimation of the channel and the transmitted symbols, while minimizing the bit error rate. In [5], the so-called Paratuck-2 decomposition is introduced, allowing for the first time both temporal allocation and temporal multiplexing. Uniqueness and identifiability properties of Paratuck-2 are studied. In [19], the so-called tensor space time coding is introduced. A Paratuck-(2,4) model for fourth-order tensors is proposed: an extra time diversity induces a BER improvement. Fig. 3a) presents the results obtained with Paratuck-2 decomposition in terms of bit error rate (BER) with respect to signal to noise ratio (SNR). This exemplifies the superiority of the Paratuck-2 model for high SNR values. Last advances consist in combining Parafac and Paratuck decompositions [12]. The obtained algorithm, called sequential Parafac Paratuck-2 (SPP), does not need training sequences.

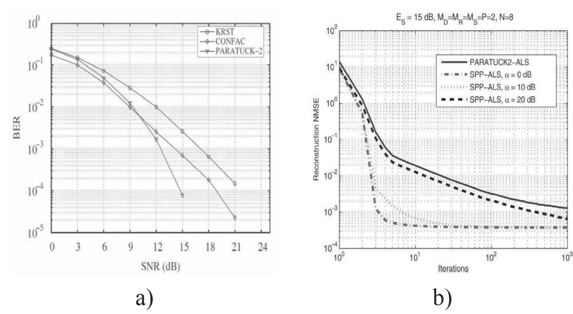

Fig. 3. a) BER vs SNR of Paratuck-2, KRST (Khatri-Rao space time), and CONFAC (constrained factor) MIMO transceivers (extracted from [5]); b) NMSE vs iterations for SPP, and Paratuck-2 (extracted from [12])

Fig. 3b) presents the results obtained with the SPP method in terms of NMSE (normalized mean square error), with respect to iterations. A faster convergence is reached, owing to a reduced computational complexity.

5. CONCLUSION

This paper is an review of recent works concerning tensor decompositions and their application to the retrieval of useful features from multidimensional data. The paper dealt firstly with small target detection using the Tucker decomposition in the wavelet packet framework; and secondly with signal dependent noise reduction through an iterative algorithm for Parafac decomposition where the diagonality of the covariance matrix of the remaining noise is checked. Thirdly, we reminded some work presented recently [12], where the Parafac decomposition is combined with the so-called Paratuck decomposition, to estimate in a fast way the channel and the symbols of multidimensional MIMO data.

6. REFERENCES

[1] L. Tucker, “The extension of factor analysis to three-dimensional matrices.” Holt, Rinehart and Winston, NY, 1964, pp. 109–127.

[2] J. Carroll and J. Chang, “Analysis of individual differences in multidimensional scaling via an N-way generalization of ”Eckart-Young” decomposition,” *Psychometrika*, vol. 35, pp. 283–319, 1970.

[3] R. Harshman and M. Lundy, “Research methods for multilmode data analysis,” H. Law, J. Snyder, J. Hattie, and R. McDonald, Eds. Praeger, New York, 1970, pp. 122–215.

[4] J. Kruskal, *Rank, decomposition, and uniqueness for 3-way and N-way arrays*, elsevier, Amsterdam, The Netherlands: North Holand ed. Multiway Data Analysis, 1988.

[5] A. L. de Almeida, G. Favier, and J. C. Mota, “Space-time spreading-multiplexing for {MIMO} wireless communication systems using the paratuck-2 tensor model,” *Signal Processing*, vol. 89, no. 11, pp. 2103 – 2116, 2009.

[6] D. Muti and S. Bourennane and J. Marot, “Lower-rank tensor approximation and multiway filtering,” *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 3, pp. 1172–1204, 2008.

[7] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM Review*, vol. 51, no. 3, pp. 455–500, Sept. 2009.

[8] M. Wei and D. Shen, “Minimum rank solutions to the matrix approximation problems in the spectral norm,” *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 3, pp. 940–957, 2012.

[9] T. Lin and S. Bourennane, “Survey of hyperspectral image denoising methods based on tensor decompositions,” *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, 2013.

[10] ——, “Hyperspectral image processing by jointly filtering wavelet component tensor,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 6, pp. 3529–3541, June 2013.

[11] X. Liu, S. Bourennane, and C. Fossati, “Reduction of signal-dependent noise from hyperspectral images for target detection,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 9, pp. 5396–5411, Sept 2014.

[12] L. Ximenes, G. Favier, A. de Almeida, and Y. Silva, “Parafac-paratuck semi-blind receivers for two-hop cooperative mimo relay systems,” *Signal Processing, IEEE Transactions on*, vol. 62, no. 14, pp. 3604–3615, July 2014.

[13] L. De Lathauwer, B. De Moor, and J. Vandewalle, “A multilinear singular value decomposition,” *SIAM Journal on Matrix Analysis and Applications*, vol. 21, pp. 1253–1278, April 2000.

[14] B. Tekin, U. Kamilov, E. Bostan, and M. Unser, “Benefits of consistency in image denoising with steerable wavelets,” in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, May 2013, pp. 1355–1358.

[15] D. Muti and S. Bourennane, “Multidimensional filtering based on a tensor approach,” *Signal Proceesing Journal, Elsevier*, vol. 85, no. 12, pp. 2338–2353, Dec. 2005.

[16] T. Lin, J. Marot, and S. Bourennane, “Small target detection improvement in hyperspectral image,” in *Advanced Concepts for Intelligent Vision Systems (ACIVS’13), 2013 International Conference on*, vol. 8192, Oct. 2013, pp. 460–469.

[17] X. Liu, S. Bourennane, and C. Fossati, “Nonwhite noise reduction in hyperspectral images,” *IEEE trans. Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 368–372, 2012.

[18] N. Sidiropoulos, R. Bro, and G. Giannakis, “Parallel factor analysis in sensor array processing,” *IEEE Trans. on Signal Processing*, vol. 48, no. 8, pp. 2377–2388, August 2000.

[19] G. Favier, M. Da Costa, A. De Almeida, and J. M. Romano, “Tensor spacetime (tst) coding for mimo wireless communication systems,” *Signal Processing*, vol. 92, no. 4, pp. 1079–1092, April 2012.