AUTOMATIC RANK ESTIMATION OF PARAFAC DECOMPOSITION AND APPLICATION TO MULTISPECTRAL IMAGE WAVELET DENOISING

Abir Zidi[1], *Julien Marot*[2], *Salah Bourennane*[2], *Klaus Spinnler*[1]

[1] Fraunhofer-Entwicklungszentrum Rontgentechnik EZRT, Fraunhofer-Institut IIS
Flugplatzstr. 75, 90768 Furth, Germany
[2] Groupe GSM, Institut Fresnel, École Centrale Marseille, Aix Marseille Université
D. U. St Jérôme Av. Normandie-Niemen, 13397, Marseille, France

ABSTRACT

There are two main contributions in this paper. Firstly, we estimate the rank for the truncation of the Parafac decomposition in an optimal sense. For this, we propose a least squares criterion and justify the choice of the fast Nelder-Mead method to minimize this criterion. Secondly, we combine the truncation of the Parafac decomposition with multidimensional wavelet packet transform. A single rank value is estimated for each decomposition level, which simplifies the implementation. We exemplify the proposed method with an application to multispectral image denoising: we study the performance of the proposed method based on Parafac decomposition, compared to ForWaRD.

Index Terms— Multispectral imaging; Image denoising; Optimization.

1. INTRODUCTION

In the frame of multidimensional data denoising, tensor methods were proposed to take into account the relationships between dimensions, as opposed to slice-by-slice methods [1]. For instance, Multiway Wiener Filtering (MWF) has been derived from the minimization of a mean square error criterion between the expected noise-free tensor and the denoised tensor [1]. Since this seminal work, some extensions have been proposed, for the minimum rank approximation of matrices [2] or tensors, based either on Tucker decomposition [3] or on Parafac decomposition [4]. It has been shown that the truncation of the Parafac decomposition permits to minimize the square error (SE) between the 'raw-data' tensor and the estimated tensor [4]. Another field of applied mathematics, Wavelets [5], has provided efficient algorithms for image denoising. For instance, ForWaRD method [6] performs image deconvolution and denoising in the Fourier domain, and refined image denoising in the wavelet domain.

Relation with previous work in the field Owing to the ability of the wavelet transform to separate high frequency components from low frequency ones, wavelets are privileged tools for image denoising [7]. That is why, recently, the wavelets and tensor framework got recently 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 [3]. This method, called MWF-MWPT (multiway Wiener filtering-multidimensional wavelet packet transform), preserves well the spatial details, but exhibits a drawback: the elevated number of parameters to estimate, that is, for each wavelet decomposition level, and for each tensor mode, the adequate signal subspace rank value. The truncation of the Parafac decomposition exhibits good performances in reducing low amplitude, possibly signal dependent noise [4], and there is only one parameter to estimate: a single rank for the truncation. However, it has never been introduced in a wavelet framework. For rank estimation, DIFFIT [8], Convex Hull [9], and Minimum Description Length [10] have much merit, especially for SNR (signal to noise ratio) values higher than 10 dB, but results seem to be restricted to artificially generated data.

Goal and contributions In this paper, our objectives are as follows: reduce the low-amplitude noise in multispectral images, with a low number of parameters and their automatic tuning. For this, we focus on the truncation of the Parafac decomposition. The originality of our work relies on the automatic estimation of the rank for the truncation of the Parafac decomposition, and on the introduction of the Parafac decomposition in the wavelet framework.

Outline Section 2 sets the problem of the Parafac rank estimation. In Section 3 we propose a criterion to estimate the rank in an optimal sense, and the Nelder-Mead method [11] to optimize it; in Section 4, the truncation of Parafac decomposition is set in a wavelet framework. Section 5 presents a flowchart of the proposed method; in Section 6, multispectral image denoising results are obtained with the proposed method and comparative ones [6, 12]. Conclusions are given in Section 7.

Notation $\mathcal{X}$ denotes a tensor, that is, a multidimensional array, $\mathbf{X}$ denotes a 2-D matrix, $\mathbf{x}$ denotes a 1-D vector, and x a scalar.

978-1-4673-9961-6/16/$31.00 ©2016 IEEE

2. PROBLEM SETTING

2.1. Truncation of the parafac decomposition: overview

We model a noisy multispectral image as a tensor resulting from a multidimensional signal $\mathcal{X}$ impaired by an additive white noise $\mathcal{N}$ [1, 13]. The multispectral image can be expressed as :

$$\mathcal{R} = \mathcal{X} + \mathcal{N}. \quad (1)$$

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. We aim at estimating the desired tensor signal $\mathcal{X}$ from data tensor $\mathcal{R}$. The PARAFAC model factorizes a tensor into a sum of rank-1 tensors [14]: a tensor $\mathcal{R} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ can be expressed as

$$\mathcal{R} = \sum_{k=1}^{T} \mathcal{R}_k = \sum_{k=1}^{T} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} \quad (2)$$

$\mathcal{R}_k \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is rank-1 tensor; $\mathbf{a}_k^{(1)}, \mathbf{a}_k^{(2)}, \mathbf{a}_k^{(3)} \in \mathbb{R}^{I_n}$ are normalized vectors of the *n*-mode space of $\mathcal{R}$ normalized by $\mathbf{a}_k^{(n)} = \mathbf{a}_k^{(n)} / \|\mathbf{a}_k^{(n)}\|$, $n = 1, 2, 3$; and $\lambda_k = \|\mathbf{a}_k^{(1)}\| \cdot \|\mathbf{a}_k^{(2)}\| \cdot \|\mathbf{a}_k^{(3)}\|$, $k = 1, 2, \cdots, T$.

It has been proved in [15] that truncating the PARAFAC decomposition to the K terms weighted by the largest values of λ_k 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 values λ_k are correctly ordered, the truncation of the Parafac decomposition consists in selecting the K first terms in Eq. (3), so that we obtain the estimate:

$$\hat{\mathcal{X}} = \sum_{k=1}^{K} \mathcal{R}_k = \sum_{k=1}^{K} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} \quad (3)$$

where K is the rank, $\hat{\mathcal{X}}$ is the truncation of the Parafac decomposition of $\mathcal{R}$, also called rank-K approximation of $\mathcal{R}$. As shown in in [3], $\hat{\mathcal{X}}$ is a close estimate of the noise-free tensor $\mathcal{X}$, especially when the noise magnitude is low.

2.2. The rank estimation issue

In [4], an iterative 'brute-force' method is proposed to obtain a guess on a convenient Parafac truncation rank. the covariance matrix of the residual noise is computed, and the non-diagonal coefficients are summated. As it should tend to 0 when only decorrelated noise is present, a threshold is set empirically on this summation. Unfortunately, the optimality of this method as a function of any criterion is not proved. For these reasons, we propose a criterion to minimize, and an adequate optimization algorithm, to estimate the rank of the truncation which is optimal in the sense of this criterion.

3. NELDER-MEAD OPTIMIZATION METHOD FOR THE ESTIMATION OF THE PARAFAC TRUNCATION RANK

For the first time, we propose a method to estimate an optimal rank for the truncation of the Parafac decomposition. For this, firstly, we propose a squared error criterion; secondly, we justify the use of Nelder-Mead optimization method to minimize this criterion.

3.1. Proposed criterion

The proposed criterion is as follows:

$$J(K) = ||\mathcal{X}_1 - \hat{\mathcal{X}}||^2, \quad (4)$$

where $||.||$ represents the Frobenius norm, $\mathcal{X}_1$ is a first, gross estimate of the expected tensor $\mathcal{X}$ or the noisy tensor $\mathcal{R}$ itself, and $\hat{\mathcal{X}}$ its final estimate. It is worth noticing that we remain close to the framework of the Parafac decomposition, while choosing the SE as a criterion. Now we wish to solve:

$$\hat{K} = \underset{K}{argmin}(J(K)) \quad (5)$$

It is rather complex to minimize the criterion J as it is a nonlinear function of the parameter K. We wish to estimate a single value -the rank for the truncation of the Parafac decomposition-, but the optimization method used for this purpose must be global. Moreover, the multispectral image considered in this paper exhibit a large number of voxels. Therefore, the proposed optimization method must be fast.

3.2. Why the Nelder-Mead method ?

Some methods from the literature may be available to perform the minimization in Eq. (5), but they exhibit some drawbacks: the DIRECT method [16], for instance, requires the objective function to be Lipschitzian. Particle swarm optimization (PSO) [12] provides the global minimum of a scalar function of several variables but is not assumed, to the best of our knowledge, to be particularly fast.

The Nelder-Mead Simplex method [11] is a global optimization method, which is meant to minimize a scalar-valued nonlinear function of several real variables, without any derivative information. It is known to yield a rapid decrease in cost function values [11]. It has been shown that, in dimension two, the Nelder-Mead method may converge to a non-critical point of the minimized function [17]. However, as specified in [11], the global convergence of the Nelder-Mead method is ensured in a one-dimensional problem, which is the case in the current work. For these reasons, we select the Nelder-Mead Simplex Method.

Nelder-Mead algorithm falls into five steps: ordering, reflection, expansion, contraction, and shrinkage. At each iteration, a simplex is modified, following these five steps. Four scalar parameters must be specified to run the Nelder-Mead method: coefficients of reflection (ρ), expansion (χ), contraction (γ), and shrinkage (σ). They are such that $\rho > 0$, $\chi > 1$, $\chi > \rho$, $0 < \gamma < 1$, $0 < \sigma < 1$. The proposed improved truncation of the Parafac decomposition, with an automatic estimation of the rank is now introduced in a wavelet framework.

4. PARAFAC TRUNCATION OF THE MULTIDIMENSIONAL WAVELET PACKET TRANSFORM COEFFICIENTS

After the MWPT, abundant and rare features can be separated into different components, therefore the 'useful part' of each component can be estimated more accurately than using the entire dataset [3]. Furthermore, a better estimation of the 'useful part' can improve the performance of the Parafac rank-K truncation of each component. We propose a combination, named Parafac-MWPT. This subsection proves the ability of Parafac-MWPT to minimize the SE between $\mathcal{R}$ and $\hat{\mathcal{X}}$. By performing multidimensional wavelet transform to tensor $\mathcal{R}$, $\mathcal{X}$ and $\mathcal{N}$, we obtain:

$$\begin{aligned} &\mathcal{R} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \\ &= (\mathcal{X} + \mathcal{N}) \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \\ &= \mathcal{X} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \\ &+ \mathcal{N} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \end{aligned} \tag{6}$$

where $\times_n$ denotes the *n*-mode product for each mode *n*: wavelet filtering is performed successively after flattening tensor $\mathcal{R}$ along each mode. The coefficient tensor of each part:

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

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

$$\mathcal{C}_{\mathbf{l}}^{\mathcal{N}} = \mathcal{N} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \tag{9}$$

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

$$\hat{\mathcal{C}}_{\mathbf{l}}^{\mathcal{X}} = \hat{\mathcal{X}} \times_1 \mathbf{W}_1 \times_2 \mathbf{W}_2 \times_3 \mathbf{W}_3 \tag{10}$$

With the extraction process proposed in [3], 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:

$$\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}} = \mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}} + \mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{N}} \tag{11}$$

From Parseval's theorem, the following expression can be obtained:

$$\|\mathcal{R} - \hat{\mathcal{X}}\|^2 = \|\mathcal{C}_{\mathbf{l}}^{\mathcal{R}} - \hat{\mathcal{C}}_{\mathbf{l}}^{\mathcal{X}}\|^2 = \sum_{\mathbf{m}} \|\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}} - \hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}}\|^2 \tag{12}$$

which means that minimizing the SE between $\mathcal{R}$ and its estimate $\hat{\mathcal{X}}$ is equivalent to minimizing the SE between $\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}}$ and $\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}}$ for each $\mathbf{m}$. If we apply Nelder-Mead optimization method while performing the truncation of the Parafac decomposition (see section 3) to estimate the rank K, we get the optimal rank for the Parafac truncation of $\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}}$:

$$\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{X}} = \sum_{k=1}^{K} \mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R},k} \tag{13}$$

After estimating $\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}}$ for each $\mathbf{m}$, we obtain $\hat{\mathcal{C}}_{\mathbf{l}}^{\mathcal{R}}$ by concatenating $\hat{\mathcal{C}}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}}$. Furthermore, the estimate $\hat{\mathcal{X}}$ can be obtained by inverse MWPT:

$$\hat{\mathcal{X}} = \hat{\mathcal{C}}_{\mathbf{l}}^{\mathcal{X}} \times_1 \mathbf{W}_1^T \times_2 \mathbf{W}_2^T \times_3 \mathbf{W}_3^T \tag{14}$$

The process composed of wavelet transform, Parafac truncation, and inverse wavelet transform yields an estimate $\hat{\mathcal{X}}$ of the expected tensor. In Eq. (4), we notice that the first reference tensor $\mathcal{X}_1$ which is used in the criterion $J(K)$ is the noisy tensor $\mathcal{R}$ itself. The wavelet coefficient tensor $\mathcal{C}_{\mathbf{l},\mathbf{m}}^{\mathcal{R}}$ is supposed to be noisy, which is not convenient for a reference. Therefore, we introduce the denoising process in a loop: Let $\hat{\mathcal{X}}_r$ be the estimated tensor at iteration 'r'. At iteration 'r+1', the new reference becomes tensor $\hat{\mathcal{X}}_r$. The loop is stopped at a given iteration 'r' when $||\hat{\mathcal{X}}_r - \hat{\mathcal{X}}_{r-1}|| < \delta$ where δ is an *a priori* set very low threshold.

5. FLOWCHART OF THE ALGORITHM

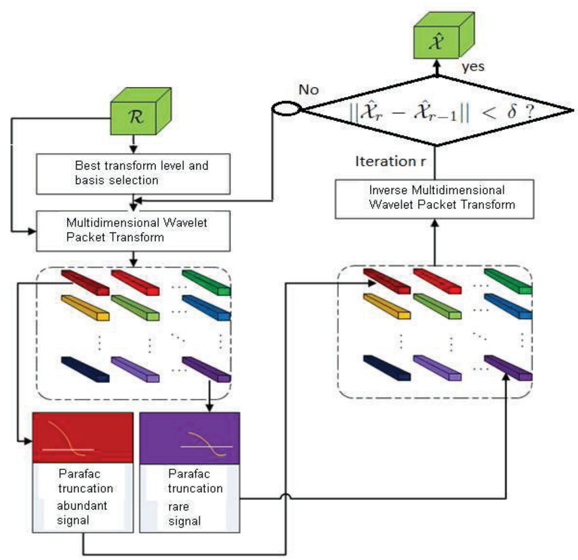

Fig. 1. Truncation of the Parafac decomposition in wavelet packet transform domain: iterative process

6. RESULTS

In this section, we apply the proposed method based on the truncation of the Parafac decomposition with automatic es-

Method \ Criterion	SNR	PSNR	SSIM
Noised image	10.0	23.7	0.68
ForWaRD	$2.4\ 10^{-2}$	13.6	0.75
Parafac-MWPT:			
• Nelder-Mead	12.8	26.4	0.74
• PSO	13.1	26.7	0.73

Table 1. Numerical criteria obtained with the proposed method based on Parafac decomposition, with a rank estimation by Nelder-Mead or PSO optimization algorithms; and ForWaRD algorithm.

timation of the rank by Nelder-Mead and multidimensional wavelet packet to multispectral image denoising. We compare the results obtained with the comparative ForWaRD method [6] on a real-world multispectral image. Programmes were written in $\mathrm{Matlab}^{®}$, and executed on a PC computer running Windows, with a 3GHz double core and 3GB RAM. The multispectral image is artificially impaired with white, identically distributed random noise. This image is of size $128\times128\times16$: it includes 128 columns, 128 rows and 16 spectral bands. Before impairing the image, we artificially fix the Parafac decomposition rank in the original noise-free image by applying the truncation of the Parafac decomposition with a known rank, $K = 40$. In the proposed method, multidimensional wavelet packet is implemented with Daubechies wavelets and two decomposition. This has been empirically shown to yield the best results for this type of data [3]. This holds also for the comparative ForWaRD algorithm [6], applied successively slice-by-slice to the image. Nelder-Mead and PSO require respectively 200 and 2000 iterations for convergence; the parameters in Nelder-Mead are chosen as $\rho = 1$, $\chi = 2$, $\gamma = 1/2$, $\sigma = 1/2$ [11]. We initialize the rank value K to $\min(I_1,\ I_2,\ I_3)$. The numerical results are computed from images truncated to the size $120\times120\times16$ to avoid the border issues. The parameter δ is set to 10^{-6}. Table 1 presents the results obtained in terms of SNR (signal to noise ratio), PSNR (peak signal to noise ratio), and slice-by-slice SSIM (structural similarity) [18] with the proposed and ForWaRD methods. The rank of the Parafac decomposition is estimated either with Nelder-Mead or with PSO. The number of particles in PSO are such that the results in terms of SNR are approximately the same than when Nelder-Mead is used. Examples of estimated rank values are as follows: for the first decomposition level PSO provides $K = 41$, and Nelder-Mead provides $K = 39$. Table 1 shows that the rank value provided by PSO and Nealder-Mead yield close SNR and PSNR values, which are higher than the comparative ForWaRD method. However, one iteration of PSO lasts $8\ 10^{-2}$ sec. to get this estimation, while Nelder-Mead lasts $4\ 10^{-2}$ sec. Figure 2 displays visual results. When the truncation of the Parafac decomposition is used, we display the Nelder-Mead result: the visual aspect of the PSO result is very similar.

From the results displayed in Fig. 2, we can assert that when the truncation of the Parafac decomposition and ForWaRD are used, the overall visual aspect of the image is well preserved, which is confirmed by the SSIM values in Table 1. However, the details are better preserved when the truncation of the Parafac decomposition is used, which is confirmed by the SNR and PSNR values in Table 1.

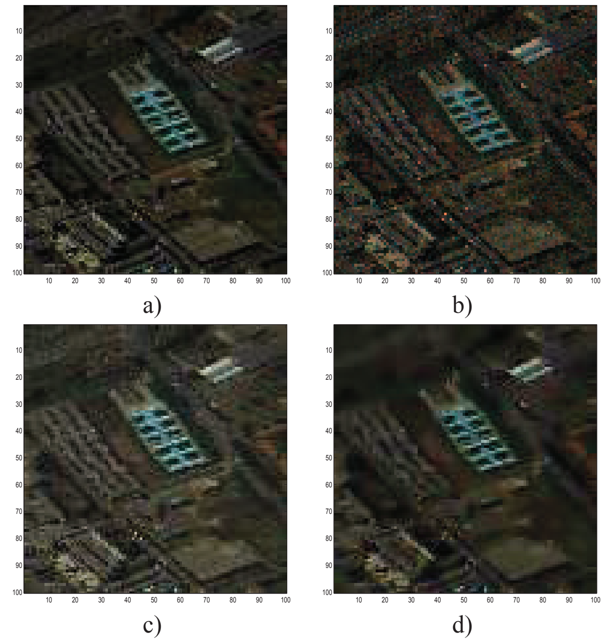

Fig. 2. a) noise-free multispectral image; b) noised multispectral image; c) proposed denoising method, using Nelder-Mead; d) ForWaRD.

7. CONCLUSION

The contribution of this paper is two-fold: firstly, we propose a least squares criterion to get an optimal evaluation of the optimal rank, with respect to this criterion, of the truncation of the Parafac decomposition; secondly, we insert the truncation of the Parafac decomposition in a wavelet framework, for the purpose of multidimensional data denoising. We illustrate the ability of the proposed method to remove low-magnitude noise in an application to multispectral image denoising. Our method based on Parafac decomposition and wavelet decomposition performs better than the wavelet-based ForWaRD method in terms of SNR and $PSNR$. It preserves the details, while preserving the visual aspect evaluated by the SSIM criterion. Also, the Nelder-Mead method is relevant for the truncation of the Parafac decomposition: in the considered application case, it is two times faster than particle swarm optimization though yielding similar results. Future works could consist in testing the behavior of the proposed method in the presence of signal-dependent noise.

8. REFERENCES

[1] 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.

[2] 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.

[3] 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.

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

[5] I. Daubechies, “Ten lectures on wavelets,” *CBMS-NSF Lecture Notes, SIAM*, no. 61, 1992.

[6] R. Neelamani, H. Choi, and R. Baraniuk, “Forward : Fourier-wavelet regularized deconvolution for i11-conditioned systems,” *IEEE Trans. on Signal Processing*, vol. 52, no. 2, pp. 418–433, Feb 2004.

[7] 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.

[8] M. E. Timmerman and H. A. L. Kiers, “Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima,” *British Journal of Mathematical and Statistical Psychology*, vol. 53, no. 1, pp. 1–16, 2000.

[9] E. Ceulemans and H. A. L. Kiers, “Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method,” *British Journal of Mathematical and Statistical Psychology*, vol. 59, no. 1, pp. 133–150, 2006.

[10] K. Liu, J. P. C. da Costa, H. C. So, L. Huang, and J. Ye, “Detection of number of components in candecomp/parafac models via minimum description length,” *Digital Signal Processing*, vol. 51, pp. 110 – 123, 2016.

[11] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, “Convergence properties of the nelder-mead simplex method in low dimensions,” *SIAM Journal of Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[12] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proc. IEEE International Conference on Neural Networks*, 1995, pp. 1942 – 1948.

[13] L. Parra, C. Spence, P. Sajda, A. Ziehe, and K. uller, “Unmixing hyperspectral data,” *Advances in Neural Information Processing Systems*, vol. 12, pp. 942–948, 2000.

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

[15] 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.

[16] D. Jones, C. Pertunen, and B. Stuckman, “Lipschitzian optimization without the lipschitz constant,” *Journal of Optimization theory and Application*, vol. 79, no. 1, pp. 157–181, October. 1993.

[17] F. Gao and L. Han, “Implementing the nelder-mead simplex algorithm withadaptive parameters,” *Computational Optimization and Applications*, vol. 51, no. 1, pp. 259–277, 2010.

[18] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600 – 612, 2004.