

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Wang, Xiulin; Zhang, Chi; Ristaniemi, Tapani; Cong, Fengyu

Title: Generalization of Linked Canonical Polyadic Tensor Decomposition for Group Analysis

Year: 2019

Version: Accepted version (Final draft)

Copyright: © Springer Nature Switzerland AG 2019.

Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Wang, X., Zhang, C., Ristaniemi, T., & Cong, F. (2019). Generalization of Linked Canonical Polyadic Tensor Decomposition for Group Analysis. In H. Lu, H. Tang, & Z. Wang (Eds.), *ISNN 2019 : Advances in Neural Networks : 16th International Symposium on Neural Networks, Proceedings, Part II* (pp. 180-189). Springer International Publishing. *Lecture Notes in Computer Science*, 11555. https://doi.org/10.1007/978-3-030-22808-8_19

Generalization of Linked Canonical Polyadic Tensor Decomposition for Group Analysis

Xiulin Wang^{1,2}, Chi Zhang¹, Tapani Ristaniemi², and Fengyu Cong^{1,2} *

¹School of Biomedical Engineering, Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China
xiulin.wang@foxmail.com, chizhang@dlut.edu.cn, cong@dlut.edu.cn

²Faculty of Information Technology,
University of Jyväskylä, Jyväskylä 40100, Finland
tapani.e.ristaniemi@jyu.fi

Abstract. Real-world data are often linked with each other since they share some common characteristics. The mutual linking can be seen as a core driving force of group analysis. This study proposes a generalized linked canonical polyadic tensor decomposition (GLCPTD) model that is well suited to exploiting the linking nature in multi-block tensor analysis. To address GLCPTD model, an efficient algorithm based on hierarchical alternating least squares (HALS) method is proposed, termed as GLCPTD-HALS algorithm. The proposed algorithm enables the simultaneous extraction of common components, individual components and core tensors from tensor blocks. Simulation experiments of synthetic EEG data analysis and image reconstruction and denoising were conducted to demonstrate the superior performance of the proposed generalized model and its realization.

Key words: Linked tensor decomposition · Hierarchical alternating least squares · Canonical polyadic · Simultaneous extraction

1 Introduction

Linked tensor decomposition (LTD) is an emerging technique for group analysis in recent years, specially designed for simultaneous analysis of multi-block tensor data. It has been successfully applied in the fields of neuroscience [1], multi-dimensional harmonic retrieval [2], array signal processing [3] and metabolic physiology [4].

Linked tensor decomposition can be seen as an extension of tensor decomposition applied to single-block tensor [5–7] in multi-block data analysis, e.g., analysis of electrophysiological (EEG) data collected from different subjects under a certain stimulus, which can be naturally linked together for sharing the similar brain activities [1]. LTD method can take full advantage of such linking/coupling information among data blocks to improve the decomposition identifiability [3]. In addition, LTD method has its advantage in imposing constraints

* Corresponding author

on particular modes or components compared to its matrix counterpart [9, 10]. Any combination of constraints including independence, sparsity, smoothness and non-negativity can be added more easily and flexibly [11]. Moreover, imposing specific constraints on different modes or components would contribute to obtaining more reasonable decomposition solutions with convincing interpretations [6, 8, 11]. For example, the constraint of non-negativity is applied in the processing of ERP data with time-frequency representation [6]. Furthermore, tensor decomposition is superior to two-way matrix factorization such as solution uniqueness and component identification in some cases [12]. To unfold some of the modes in matrix factorization will inevitably lose the potential interactions under the multiway structure [13]. Therefore, it is reliable to take the high-order characteristics of tensors into consideration in data analysis.

With the LTD model, simultaneous extraction of common components, individual components and core tensors can be obtained. The notion ‘linked’ is based on the assumption that different data blocks share the same or highly correlated components while retaining individual information [14]. In group data analysis, e.g. face images collected from different subjects with the same expression [14], or EEG data collected from different participants under the same stimulus [8], all subjects may share the similar or even identical information, which can be regarded as linking factors among tensors. However, individual characteristics will exist in particular subjects at the same time, which may lead to inconsistent number of components for tensors. Obviously, this inconsistency does not match the linked canonical polyadic tensor decomposition (LCPTD) model in [14]. Therefore, this study aims to develop a more generalized and flexible model with inconsistent component number for linked tensor decomposition. To obtain the solution of the new model, we propose a generalized linked canonical polyadic tensor decomposition algorithm based on HALS strategy [7], which is termed as GLCPTD-HALS algorithm. The experiment results show that the generalized model is more practical in multi-block data analysis, and its realization can achieve better performance.

This paper is organized as follows. Section 2 introduces LCPTD model and its generalization. In section 3, GLCPTD-HALS algorithm is proposed. In section 4, simulation experiments are conducted to verify the performance of proposed algorithm. The last section summarizes this paper.

2 Problem formulation

In this section, we mainly introduce the linked canonical polyadic tensor decomposition (LCPTD) model [14] and its generalization. CP model [15] is also called parallel factor analysis (PARAFAC) [16] and canonical composition (CANDECOMP) [17]. CP decomposition (CPD) can decompose a tensor into a minimal number of rank-1 tensors, and the minimum number R is termed as the rank of a tensor. It can achieve good unique identification under some mild conditions without any special constraints. Please refer to [18] for a detailed description of standard notations and basic tensor operations.

2.1 Review of LCPTD Model

To deal with multi-block tensors with coupling information, researchers in [14] proposed a model of simultaneous decomposition, namely LCPTD model, which is defined as follows:

$$\begin{aligned} \underline{\mathbf{X}}^{(s)} &\approx \hat{\underline{\mathbf{X}}}^{(s)} = \sum_{r=1}^R \lambda_r^{(s)} \mathbf{u}_r^{(1,s)} \circ \mathbf{u}_r^{(2,s)} \circ \dots \circ \mathbf{u}_r^{(N,s)} \\ &= \left[\underline{\mathbf{G}}^{(s)}; \mathbf{U}^{(1,s)}, \mathbf{U}^{(2,s)}, \dots, \mathbf{U}^{(N,s)} \right], \end{aligned} \quad (1)$$

where $\underline{\mathbf{X}}^{(s)} \in \mathfrak{R}^{I_1 \times I_2 \times \dots \times I_N}$ and $\hat{\underline{\mathbf{X}}}^{(s)} \in \mathfrak{R}^{I_1 \times I_2 \times \dots \times I_N}$ denote the original and estimated tensors, respectively. $\mathbf{U}^{(n,s)} = [\mathbf{u}_1^{(n,s)}, \mathbf{u}_2^{(n,s)}, \dots, \mathbf{u}_R^{(n,s)}] \in \mathfrak{R}^{I_n \times R}$ denotes the n -mode factor matrix of s th tensor. S, R, N are denoted as the number, rank and order of tensors, respectively. $\underline{\mathbf{G}}^{(s)} \in \mathfrak{R}^{R \times R \times \dots \times R}$ denotes the s th core tensor with non-zero entries only on the super-diagonal. $\lambda_r^{(s)}$ is the (r, r, \dots, r) th element of $\underline{\mathbf{G}}^{(s)}$. The LCPTD model assumes that each factor matrix $\mathbf{U}^{(n,s)} = [\mathbf{U}_C^{(n)} \mathbf{U}_I^{(n,s)}] \in \mathfrak{R}^{I_n \times R}$ consists of two parts: $\mathbf{U}_C^{(n)} \in \mathfrak{R}^{I_n \times L_n}$, $0 \leq L_n \leq R$ and $\mathbf{U}_I^{(n,s)} \in \mathfrak{R}^{I_n \times (R-L_n)}$. The former shared by all tensor blocks represents the coupling (same or highly correlated) information, whereas the latter corresponds to the individual characteristics of each tensor block.

2.2 Generalization of LCPTD Model

Even though multiple data blocks are collected under the same condition, individual characteristics will exist in the particular blocks due to the individual differences. These characteristics may lead to inconsistent number of components for tensors. Obviously, this inconsistency does not match the LCPTD model. Therefore, we extend the LCPTD model to the generalized case with different component number $R^{(s)}$, termed as GLCPTD, which is defined as:

$$\begin{aligned} \underline{\mathbf{X}}^{(s)} &\approx \hat{\underline{\mathbf{X}}}^{(s)} = \sum_{r=1}^{R^{(s)}} \lambda_r^{(s)} \mathbf{u}_r^{(1,s)} \circ \mathbf{u}_r^{(2,s)} \circ \dots \circ \mathbf{u}_r^{(N,s)} \\ &= \left[\underline{\mathbf{G}}^{(s)}; \mathbf{U}^{(1,s)}, \mathbf{U}^{(2,s)}, \dots, \mathbf{U}^{(N,s)} \right]. \end{aligned} \quad (2)$$

The generalized LCPTD model still assumes that each factor matrix $\mathbf{U}^{(n,s)} = [\mathbf{U}_C^{(n)} \mathbf{U}_I^{(n,s)}] \in \mathfrak{R}^{I_n \times R^{(s)}}$ consists of two parts: $\mathbf{U}_C^{(n)} \in \mathfrak{R}^{I_n \times L_n}$, $0 \leq L_n \leq \min(R^{(s)})$ and $\mathbf{U}_I^{(n,s)} \in \mathfrak{R}^{I_n \times (R^{(s)} - L_n)}$, representing the same meanings with LCPTD model. $\underline{\mathbf{G}}^{(s)} \in \mathfrak{R}^{R^{(s)} \times R^{(s)} \times \dots \times R^{(s)}}$ denotes the s th core tensor.

Fig. 1 illustrates the conceptual model of dual-linked tensor decomposition based on CP model (all tensors are linked together by the common parts $\mathbf{U}_C^{(1)}$ and $\mathbf{U}_C^{(2)}$).

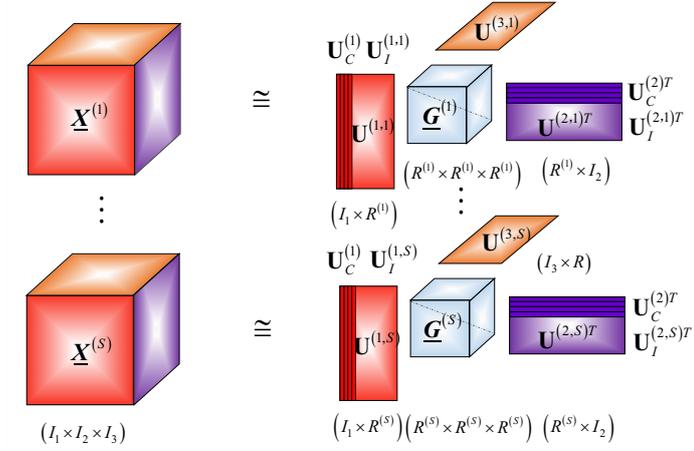


Fig. 1. Conceptual illustration of GLCPTD model with dual-linked parts [11]

3 Realization of GLCPTD Model

In this section, we aim to provide a solution of how to solve the above-mentioned GLCPTD model through HALS strategy [7]. The optimization criterion of squared Euclidean distance minimization is utilized to minimize the error between the original and estimated tensors. Therefore, the cost function can be expressed as:

$$\min \sum_{s=1}^S \left\| \underline{\mathbf{X}}^{(s)} - \sum_{r=1}^{R^{(s)}} \lambda_r^{(s)} \mathbf{u}_r^{(1,s)} \circ \mathbf{u}_r^{(2,s)} \circ \dots \circ \mathbf{u}_r^{(N,s)} \right\|_F^2 \quad (3)$$

$$s.t. \mathbf{u}_r^{(n,1)} = \dots = \mathbf{u}_r^{(n,S)}, \quad r \leq L_n,$$

$$\|\mathbf{u}_r^{(n,s)}\| = 1, n = 1 \dots N, r = 1 \dots R^{(s)}, s = 1 \dots S.$$

The above minimized optimization problem can be transformed into $\max(R^{(s)})$ sub-problems via HALS strategy, which can be solved sequentially and iteratively as follows:

$$D_F^{(r)}(\lambda_r^{(s)}, \mathbf{u}_r^{(n,s)}) = \sum_{s=1, r \leq R^{(s)}}^S \left\| \underline{\mathbf{Y}}_r^{(s)} - \lambda_r^{(s)} \mathbf{u}_r^{(1,s)} \circ \mathbf{u}_r^{(2,s)} \circ \dots \circ \mathbf{u}_r^{(N,s)} \right\|_F^2, \quad (4)$$

where $\underline{\mathbf{Y}}_r^{(s)} \doteq \underline{\mathbf{X}}^{(s)} - \sum_{k \neq r}^{R^{(s)}} \lambda_k^{(s)} \mathbf{u}_k^{(1,s)} \circ \mathbf{u}_k^{(2,s)} \circ \dots \circ \mathbf{u}_k^{(N,s)}$. For the solution of $\mathbf{u}_r^{(n,s)}$, we only set the derivative of $D_F^{(r)}(\lambda_r^{(s)}, \mathbf{u}_r^{(n,s)})$ with respect to $\mathbf{u}_r^{(n,s)}$ to zero. The learning rule of $\mathbf{u}_r^{(n,s)}$ can be formulated as:

$$\mathbf{u}_r^{(n,s)} = \begin{cases} \sum_s \left(\underline{\mathbf{Y}}_{r,(n)}^{(s)} \lambda_r^{(s)} \{\mathbf{u}_r^{(s)}\}^{\odot -n} \right) / \sum_s \left(\lambda_r^{(s)T} \lambda_r^{(s)} \right), & r \leq L_n, \\ \underline{\mathbf{Y}}_{r,(n)}^{(s)} \{\mathbf{u}_r^{(s)}\}^{\odot -n} / \lambda_r^{(s)T}, & r > L_n, \end{cases} \quad (5)$$

where $\underline{\mathbf{Y}}_{r,(n)}^{(s)}$ is the mode- n matricization of $\mathbf{Y}_r^{(s)}$. $\{\mathbf{u}_r^{(s)}\}^{\odot-n} = \mathbf{u}_r^{(N,s)} \odot \dots \odot \mathbf{u}_r^{(n+1,s)} \odot \mathbf{u}_r^{(n-1,s)} \odot \dots \odot \mathbf{u}_r^{(1,s)}$ and ‘ \odot ’ denotes the Khatri-Rao product. If $r \leq L_n$, $\mathbf{u}_r^{(n,s)}$ will be calculated by combining all tensor information and assigned to each s . Otherwise, it needs to be calculated separately. $\mathbf{u}_r^{(n,s)}$ needs to be normalized to unit variance by $\mathbf{u}_r^{(n,s)} \leftarrow \mathbf{u}_r^{(n,s)} / \|\mathbf{u}_r^{(n,s)}\|_2$ in each iteration. After N iterations of $\mathbf{u}_r^{(n,s)}$, the (r, r, \dots, r) th element $\lambda_r^{(s)}$ of core tensors is updated as follows:

$$\lambda_r^{(s)} \leftarrow \underline{\mathbf{Y}}_r^{(s)} \times_1 \mathbf{u}_r^{(1,s)} \times_2 \mathbf{u}_r^{(2,s)} \dots \times_N \mathbf{u}_r^{(N,s)}. \quad (6)$$

These $\max(R^{(s)})$ stages are alternatively updated one after another until convergence. In order to impose non-negativity, a simple ‘‘half-rectifying’’ non-linear projection is applied as $\mathbf{u}_r^{(n,s)} \leftarrow \|\mathbf{u}_r^{(n,s)}\|_+$ or $\lambda_r^{(s)} \leftarrow \|\lambda_r^{(s)}\|_+$ after (5) and (6). We summarize the extended GLCPTD-HALS algorithm in Algorithm 1.

Algorithm 1: GLCPTD-HALS algorithm

Input: $\underline{\mathbf{X}}^{(s)}$, L_n and $R^{(s)}$, $n = 1, \dots, N$, $s = 1, \dots, S$
Initialization: $\underline{\mathbf{G}}^{(s)}$, $\mathbf{U}^{(n,s)}$, $\mathbf{u}_r^{(n,s)} \leftarrow \mathbf{u}_r^{(n,s)} / \|\mathbf{u}_r^{(n,s)}\|_2$
 $\underline{\mathbf{E}}^{(s)} = \underline{\mathbf{X}}^{(s)} - \sum_r^{R^{(s)}} \lambda_r^{(s)} \mathbf{u}_r^{(1,s)} \circ \mathbf{u}_r^{(2,s)} \circ \dots \circ \mathbf{u}_r^{(N,s)}$
while not convergence do
 for $r = 1, 2, \dots, \max(R^{(s)})$ **do**
 $\underline{\mathbf{Y}}^{(s)} = \underline{\mathbf{E}}^{(s)} + \lambda_r^{(s)} \mathbf{u}_r^{(1,s)} \circ \mathbf{u}_r^{(2,s)} \circ \dots \circ \mathbf{u}_r^{(N,s)}$, $r \leq R^{(s)}$, $s = 1, 2, \dots, S$
 for $n = 1, 2, \dots, N$ **do**
 | update $\mathbf{u}_r^{(n,s)}$, $r \leq R^{(s)}$, $s = 1, 2, \dots, S$ via equation (5)
 end
 update $\lambda_r^{(s)}$, $s = 1, 2, \dots, S$ via equation (6)
 $\underline{\mathbf{E}}^{(s)} = \underline{\mathbf{Y}}^{(s)} - \lambda_r^{(s)} \mathbf{u}_r^{(1,s)} \circ \mathbf{u}_r^{(2,s)} \circ \dots \circ \mathbf{u}_r^{(N,s)}$, $r \leq R^{(s)}$, $s = 1, 2, \dots, S$
 end
end
Output: $\underline{\mathbf{G}}^{(s)}$, $\mathbf{U}^{(n,s)}$, $n = 1, \dots, N$, $s = 1, \dots, S$

4 Simulation Results

4.1 Synthetic EEG Data Analysis

In this part, we synthetically generate three types of factor matrices based on brain activities, respectively presenting topography, waveform and power spectrum, as shown in the Fig. 2 (a)-(c). Through the back projection of factor matrices, four tensor blocks representing four subjects are constructed with the SNR of 10dB, as shown in Fig. 2 (d). SNR refers to the signal-to-noise ratio, which is defined as $\text{SNR} = 10\log_{10}(\sigma_s/\sigma_n)$. σ_s and σ_n denote the levels of signal and noise, respectively. To prove the usefulness of GLCPTD model, we set the

number of components for four tensors as $\{4, 4, 3, 3\}$. Furthermore, factor matrices of topography and power spectrum consist of two common bases and one or two individual bases ($L_1 = L_2 = 2$), while the components of waveform are completely individual ($L_3 = 0$). The common bases represent that the occipital region in the mid-line and left-hemisphere of four subjects are activated with the alpha oscillations (8~13Hz).

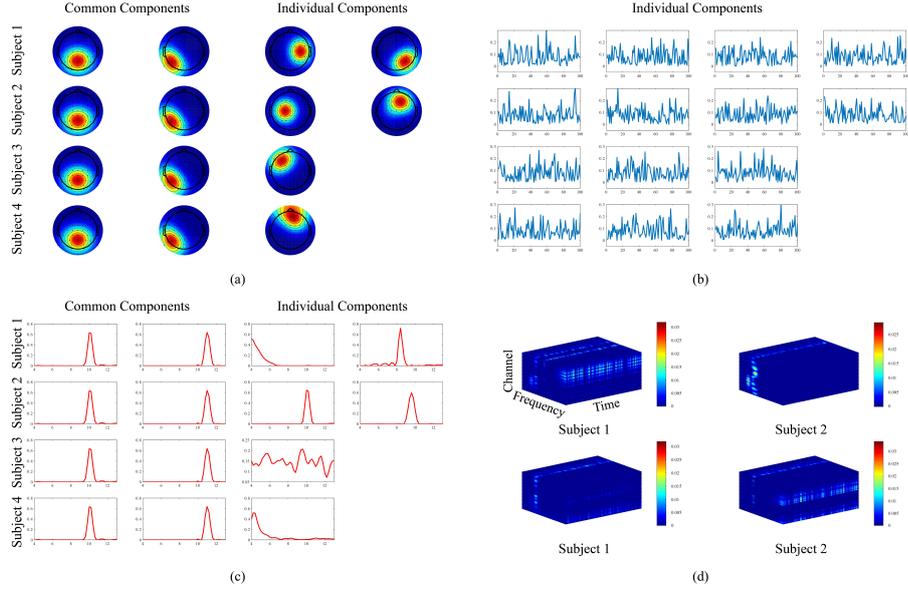


Fig. 2. Illustration of factor matrices of (a) topography, (b) waveform and (c) power spectrum and (d) tensors (frequency \times time \times channel) for four subjects. Factor matrices of topography and power spectrum for each subject consist of two common components and one/two individual components, while temporal components are individual for each subject.

We apply LCPTD-HALS [14], GLCPTD-HALS, and NTF-HALS [7] algorithms with nonnegative constraint to analyze the four tensor blocks. Solutions of topography learned by these algorithms are shown in Fig. 3 (a)-(d). We can see that, GLCPTD-HALS and NTF-HALS algorithms can successfully extract the common components as well as individual components. The difference is that the components learned by NTF-HALS algorithm are disordered. Clustering and other post-ordering methods need to be applied to obtain the common bases. Although LCPTD-HALS algorithm can extract all the common components, only 3 components are extracted from subject 1 or subject 2 shown in Fig. 3 (c) and 4 components are recovered from subject 3 or subject 4 shown in Fig. 3 (d). The former causes potential components to be omitted (subject 1) or merged (subject 2). The latter depends on the magnitude of the particular component

being redistributed in a certain way driven by algorithm (e.g. the 1st component of subject 3 in Fig. 2 (a) is recovered to the 1st and 4th components in Fig. 3 (d) corresponding to that its magnitude is divided into two parts from predefined 1), which makes group analysis more complicated especially when the number of components increases.

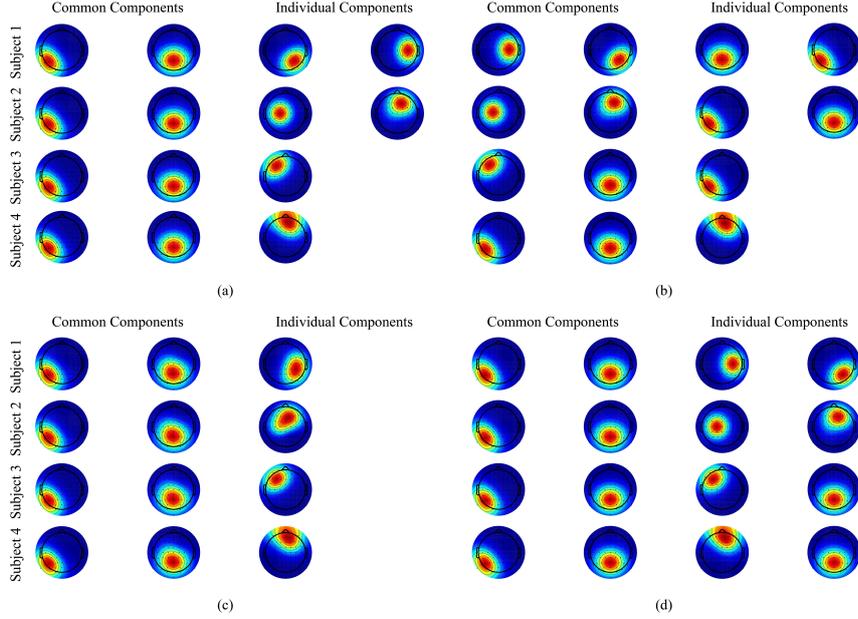


Fig. 3. Illustration of factor matrices of topography for four subjects under four conditions. (a)-(d) show the components learned by GLCPTD-HALS, NTF-HALS, LCPTD-HALS ($R = 3$) and LCPTD-HALS ($R = 4$) algorithms, respectively.

4.2 Image Reconstruction and Denoising

In this part, to examine and demonstrate the performance of the proposed algorithm, we apply the LCPTD and GLCPTD models to image reconstruction and denoising. There are 165 gray-scale images from 15 individuals in the Yale face database. Each individual has 11 images of different face expressions ('centerlight', 'glasses', 'happy', 'leftlight', 'noglasses', 'normal', 'rightlight', 'sad', 'sleepy', 'surprised', 'wink'), and the size of each image is 215×171 pixels. We construct the multi-block tensors by stacking corresponding face images under two conditions: (1) Face images from the same subject with different expressions, $I_1 = 215$, $I_2 = 171$, $I_3 = 11$, $S = 15$; (2) Face images from different subjects with the same expression, $I_1 = 215$, $I_2 = 171$, $I_3 = 15$, $S = 11$. Furthermore, 5% salt-

and-pepper noises are added to all face images. We use the peak-signal-to-noise ratio (PSNR) to measure the quality of reconstructed images.

In terms of the number of components in each tensor, we set R to 40 in the LCPTD model, which is consistent with the original parameter in [14]. Differently, in the GLCPTD model, we use the following method to calculate it: we concatenate each tensor along the first mode to generate a matrix, and perform principle component analysis (PCA) on the matrices successively; when the percentage of the total variance explained by each principle component is greater than 99.6%, the number of corresponding principle components is chosen as the number of components. In this experiment, we assume that the coupling information exists in two modes of images so that we set the number of coupled components to $L_1 = L_2, L_3 = 0$, and the values of $L_{1,2}$ are changed in $\{10, 20, 30\}$.

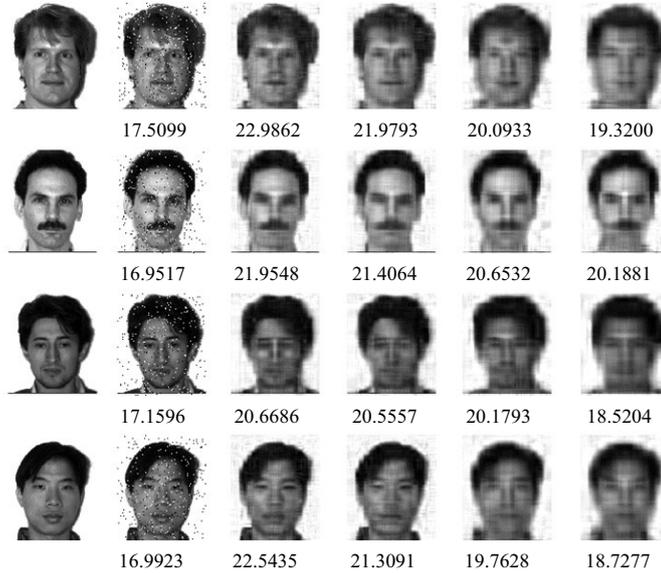


Fig. 4. Original, noisy and reconstructed face images of ‘centerlight’ from four subjects with PSNRs (dB). 1st column: original images, 2nd column: noisy images, 3rd column: GLCPTD model of condition I, 4th column: LCPTD model of condition I, 5th column: GLCPTD model of condition II, 6th column: LCPTD model of condition II.

By performing the LCPTD-HALS [14] and GLCPTD-HALS algorithms with nonnegative constraint on the above two models, we can compute the PSNRs of reconstructed images. Fig. 4 depict the original, noisy and reconstructed face images from subject 1-4 with the same expression of ‘centerlight’ ($I_{1,2} = 10$). We can see that the images reconstructed by LPCTD model/condition II are more fuzzier or distorted than those from GLCPTD model/condition I. Table 1

shows the averaged PSNRs of the reconstructed images under the two conditions. The PSNRs obtained by GLCPTD model are higher than those obtained by the LCPTD model in both conditions. It can be considered that the proposed GLCPTD model matches the real-world data more closely, which may make it more practical in real-world data analysis. The PSNRs obtained under condition I are higher than those under condition II, which means that it is more reliable to stack face images from the same subject with different expressions together. It seems that if the number of common components becomes larger, the PSNRs become smaller. The excessive number of common components may affect the fitness of the estimated tensors. However, the selection of parameter L_n is still an open issue in the current study, which will be one of our future works.

Table 1. Averaged PSNRs (dB) of reconstructed images

	Condition I			Condition II		
	$L_1, L_2 = 10$	20	30	$L_1, L_2 = 10$	20	30
LCPTD	21.3651	20.7517	19.9021	19.0809	18.8694	18.5321
GLPCTD	22.0421	21.5476	20.8134	19.9649	19.6537	19.4444

5 Conclusion

The main objective of this paper is to develop a generalized and flexible model of linked tensor decomposition which is more suitable for group analysis. We proposed the generalized LCPTD model as well as its realization, in which the common components, individual components and core tensors can be extracted simultaneously. Experiments of synthetic EEG data analysis and image reconstruction and denoising were conducted to compare the performance of proposed algorithm with LCPTD-HALS and NTF-HALS algorithms. The results illustrated the superior performance of the newly generalized model and its realization.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Grant No. 81471742), the Fundamental Research Funds for the Central Universities [DUT16JJ(G)03] in Dalian University of Technology in China, and the scholarships from China scholarship Council (No. 201706060262).

References

1. Zhou, G.X., Zhao, Q.B., Zhang, Y., et al.: Linked Component Analysis from Matrices to High-Order Tensors: Applications to Biomedical Data. Proc. IEEE. 104, 310–331 (2016)

2. Sorensen, M., De Lathauwer, L.: Multidimensional Harmonic Retrieval via Coupled Canonical Polyadic Decomposition — Part II: Algorithm and Multirate Sampling. *IEEE Trans. Signal Process.* 65, 528–539 (2017)
3. Gong, X.F., Lin, Q.H., Cong, F.Y., De Lathauwer, L.: Double coupled canonical polyadic decomposition for joint blind source separation. *IEEE Trans. on Signal Process.* 66, 3475–3490 (2016)
4. Acar, E., Bro, R., Smilde, A.K.: Data Fusion in Metabolomics Using Coupled Matrix and Tensor Factorizations. *Proc. IEEE.* 103, 1602–1620 (2015)
5. Zhou, G.X., Cichocki, A., Xie, S.L.: Fast nonnegative matrix/tensor factorization based on low-rank approximation. *IEEE Trans. Signal Process.* 60, 2928–2940 (2012)
6. Cong, F.Y., Zhou, G.X., Cichocki, A., et al.: Low-Rank Approximation Based Non-Negative Multi-Way Array Decomposition on Event-Related Potentials. *Int. J. Neural Syst.* 24, 1440005 (2014)
7. Cichocki, A., Zdunek, R., Amari, S.: Hierarchical ALS Algorithms for Nonnegative Matrix and 3D Tensor Factorization. In: 7th International Conference on Independent Component Analysis and Signal Separation, pp. 169–176. Springer, London (2007)
8. Cong, F.Y., Phan, A.H., Zhao, Q.B., et al.: Analysis of ongoing EEG elicited by natural music stimuli using nonnegative tensor factorization. In: 20th European Signal Processing Conference, pp. 494–498. Elsevier, Bucharest (2012)
9. Calhoun, V.D., Liu, J., Adali, T.: A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage.* 45, 163–172 (2009)
10. Gong, X.F., Wang, X.L., Lin, Q.H.: Generalized non-orthogonal joint diagonalization with LU decomposition and successive rotations. *IEEE Trans. on Signal Process.* 63, 1322–1334 (2015)
11. Cichocki, A.: Tensor Decompositions: A New Concept in Brain Data Analysis. *arXiv Prepr. arXiv1305.0395* (2013)
12. Mørup, M.: Applications of tensor (multiway array) factorizations and decompositions in data mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.* 1, 24–40 (2011)
13. Cong, F.Y., Lin, Q.H., Kuang, L.D., et al.: Tensor decomposition of EEG signals: A brief review. *J. Neurosci. Methods.* 248, 5969 (2015)
14. Yokota, T., Cichocki, A., Yamashita, Y.: Linked PARAFAC / CP Tensor Decomposition and Its Fast Implementation for Multi-block. In: 19th International Conference on Neural Information Processing, pp. 84–91. Springer, Doha (2012)
15. Hitchcock, F.L.: The expression of a tensor or a polyadic as a sum of products. *J. Math. Phys.* 6, 164–189 (1927)
16. Harshman, R.A.: Foundations of the PARAFAC procedure: Models and conditions for an explanatory multimodal factor analysis. *UCLA Work. Pap. Phonetics.* 16, 1–84 (1970)
17. Carroll, J.D., Chang, J.J.: Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart-Young decomposition. *Psychometrika.* 35, 283–319 (1970)
18. Kolda, T.G., Bader, B.W.: Tensor Decompositions and Applications. *SIAM Rev.* 51, 455–500 (2008)