Coupled matrix/tensor decomposition for joint analysis of multi-block datasets

Xiulin Wang

Faculty of Information Technology

University of Jyväskylä

1. Background

Multi-block datasets

Data characteristics:

1. Multi-set/multi-modal (chen et al., 2016)

   Multiple datasets collected from the same type of data

   Multiple datasets collected from different types of data

2. Multi-way

3. Multi-coupling

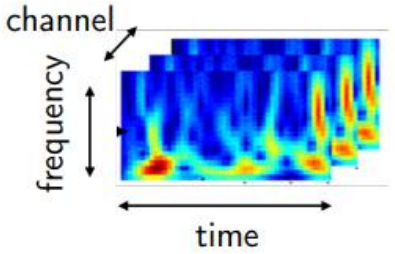


Fig. 1 EEG tensor data[1]

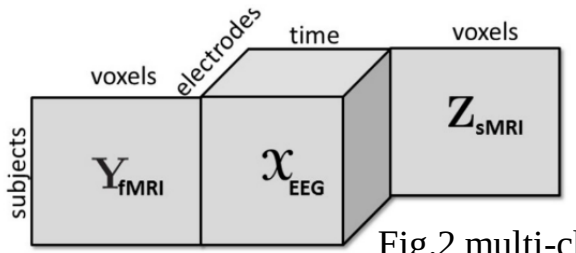


Fig.2 multi-channel EEG signals is coupled with fMRI and sMRI data in the form of matrices in the subject mode (Acar et al., 2019)

[1]https://qibinzhao.github.io/resource/slides/

Existing methods

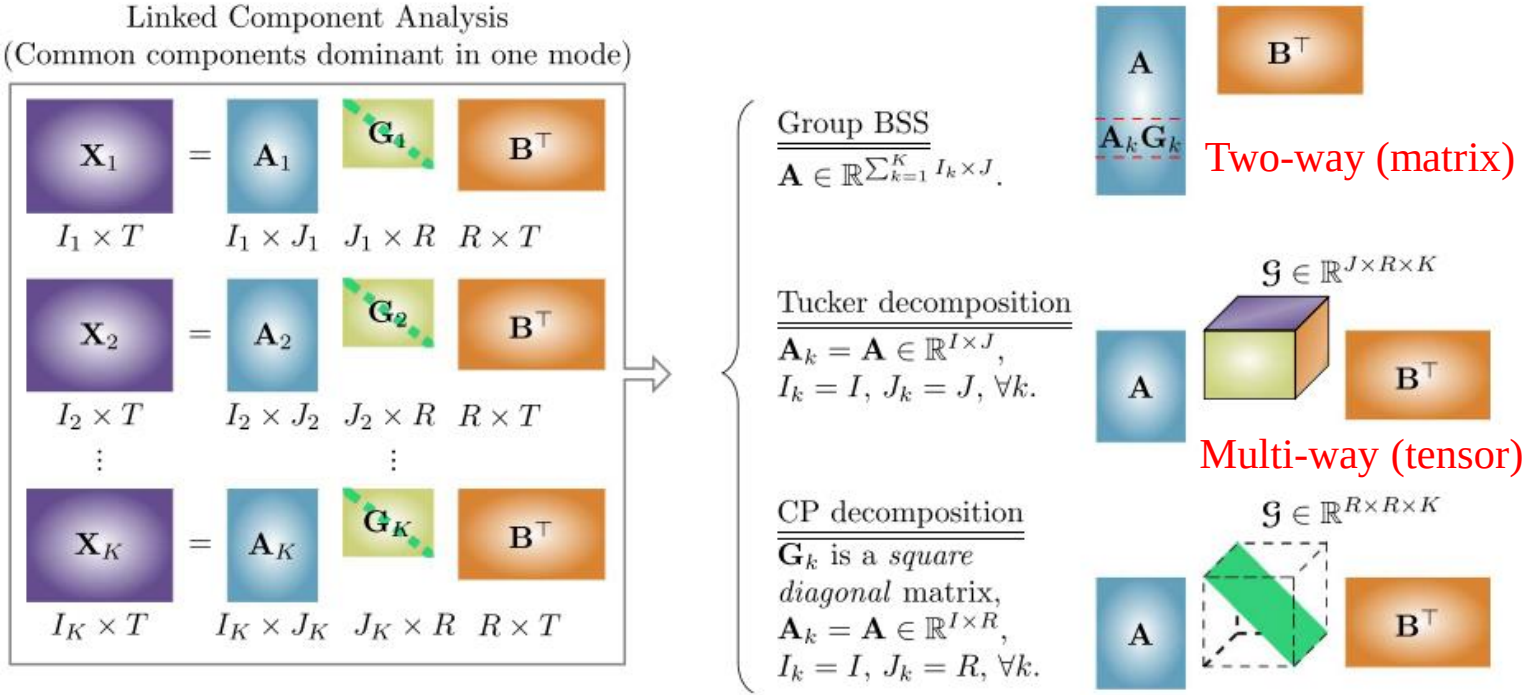


Fig. 3 Decomposition models interpreted as linked component analysis of multi-block datasets (Zhou et al. 2016).

Tips: group BSS includes joint ICA, group ICA, IVA and their variants.

Existing problems

- **Two-way matrix methods**: **concatenate** or **stack** the extra modes besides two modes to generate a two-way matrix, such unfolding inevitably loses some potentially existing interactions between/among the folded modes (multi-way structure). Generally, they can not guarantee unique solutions.

e.g.

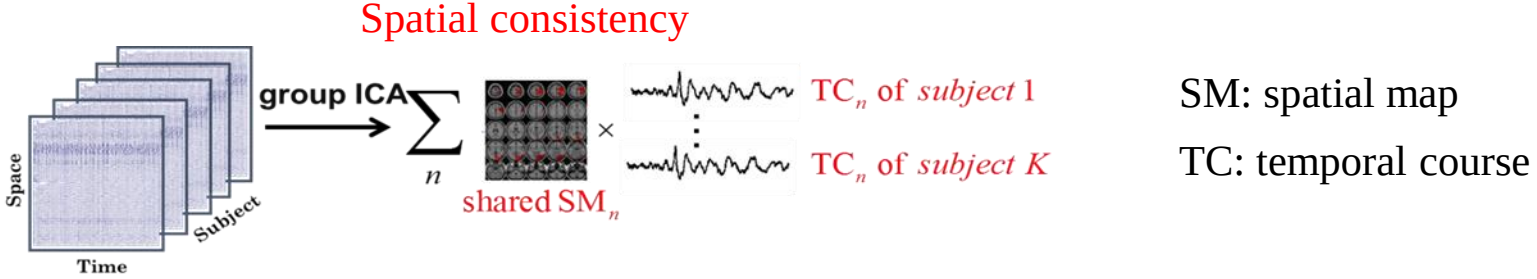


Fig. 4 Joint analysis of multi-subject fMRI data

Existing problems

- **Multi-way tensor methods**: stack the matrices or tensors to generate the higher-order tensor and assume the complete consistency among corresponding modes of multi-block datasets.

e.

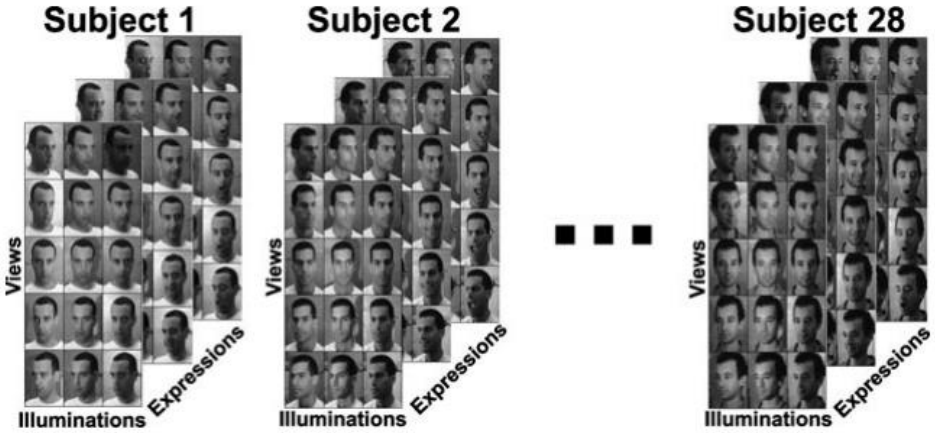


Fig. 5 Illustration of the Weizmann face database used in the analysis of TensorFaces (Mørup et al., 2011)

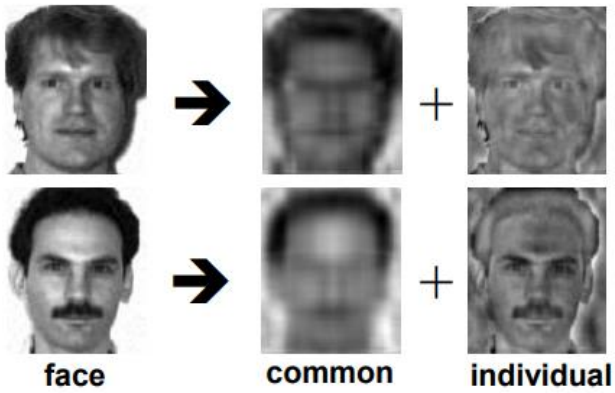


Fig. 6 Illustration of face images obtained by coupled tensor decomposition (Yokota et al., 2012)

2. New methods

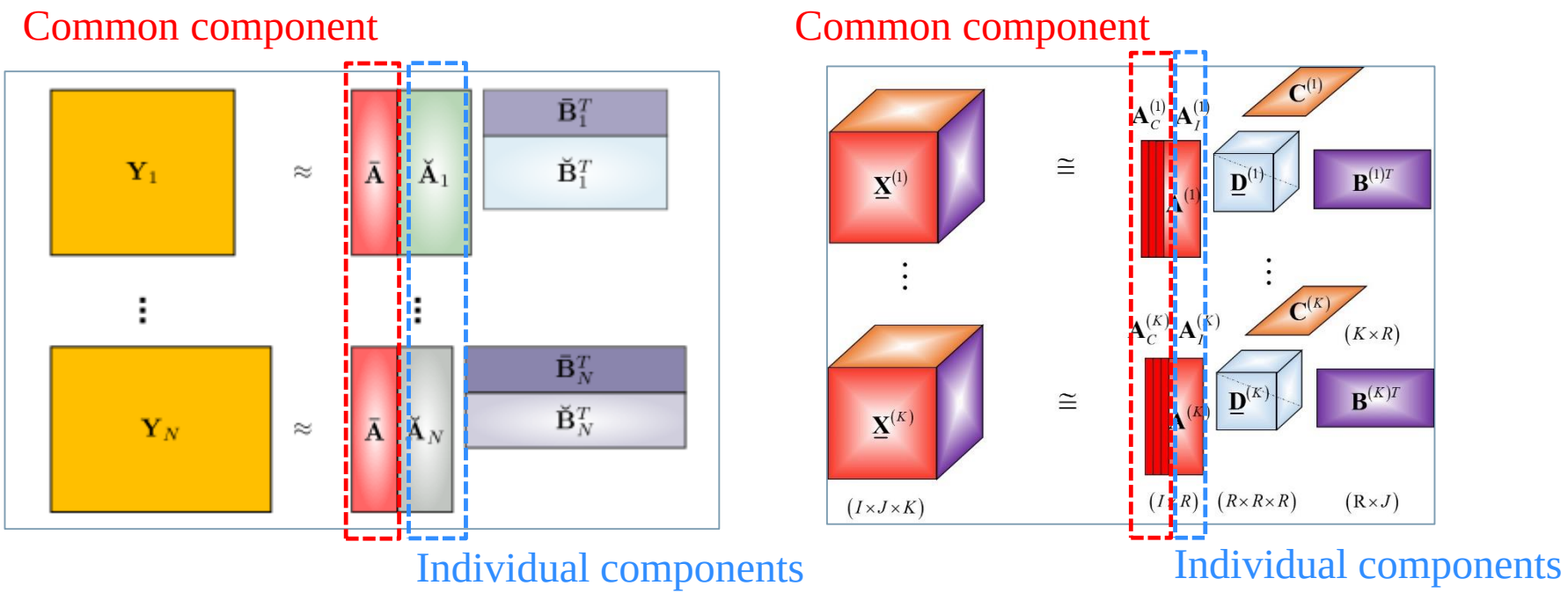


Fig. 7 Conceptual illustration of coupled matrix/tensor decomposition model (Zhou et al., 2016&Cichocki et al.,2013)

- Common component shared by all blocks corresponds to the same or maximally correlated components.
- Individual components correspond to different individual characteristics

Coupled matrix/tensor decomposition

- It can be considered as a generalization of matrix/tensor decomposition to multi-block datasets, provides a natural framework for the simultaneous analysis of multi-block tensors with coupled information
- It considers incomplete consistency among data and enables the simultanous decomposition of common and individual components.
- It can potentially reveal underlying structures and inner-relationships among data with keeping the original data structure.
- It can take full advantage of prior information to achieve the higher accuracy and uniqueness of solutions, while circumventing the independence constraint.

3. Applications

- 1. Video sequence modeling
- 2. Image denoising and reconstruction
- 3. Joint analysis of multi-subject fMRI-like data
- 4. Joint analysis of multi-subject ongoing EEG data

3.1 Video sequence modeling

Objective: to separate a moving foreground (a rotating radar antenna) from a nearly static background (a radar tower) in a color video sequence.

(Zdunek et al., 2019)

Video sequence was selected from dreamstime repository: https://pl.dreamstime.com/.

3.1 Video sequence modeling

Method: Coupled tensor decomposition (unconstrained), optimized by **Fast HALS**

Preprocessing:

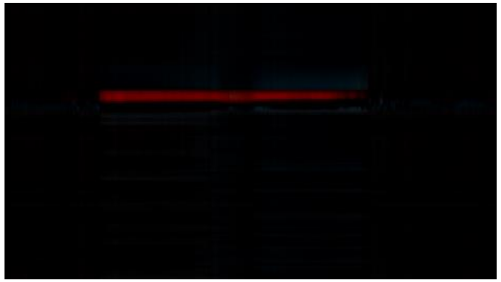
1. convert to RGB format (resolution: 163*290).
2. choose 30 frames that present one full rotation (No. 1, 18 and 24 are given).
3. generate the coupled tensors (30): 163*290*3
4. number of components: 50
5. number of common components: 45

Fig. 8 Selected frames from the video sequence of a rotating radar antenna.

3.1 Video sequence modeling

Results: successfully extract common and individual features simultaneously

Common part



Individual part: frame-1

Individual part: frame-18

Individual part: frame-24

3.2 Face denoising and reconstruction

- **Yale face database**: 165 gray-scale images from 15 individuals, and each individual has 11 images of different face expressions
- **Salt & pepper noise**
- **Decomposition model**:

  LTD: linked (coupled) tensor decomposition (common + individual features)

  ITD: individual tensor decomposition (individual features)

  STD: simultaneous tensor decomposition (common features)
- **Decomposition algorithm:**

  Hierarchical alternative least squares (HALS)

3.2 Face denoising and reconstruction

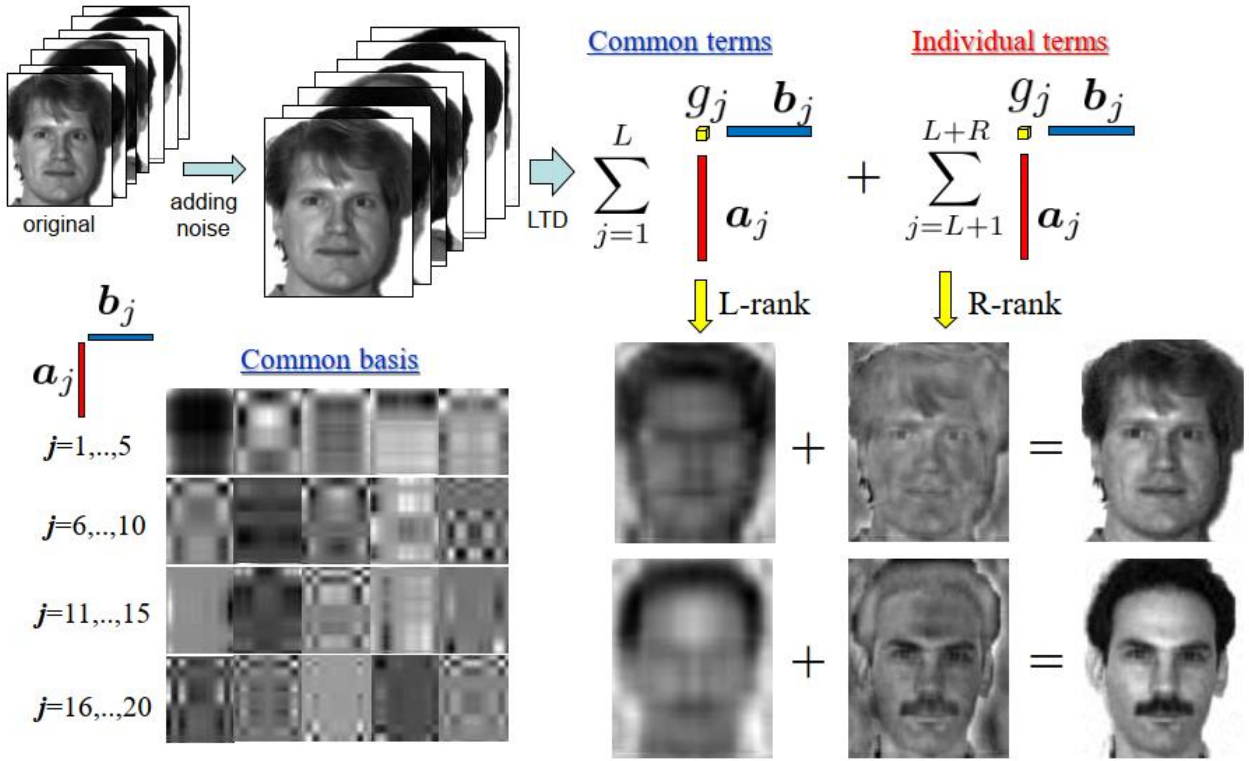
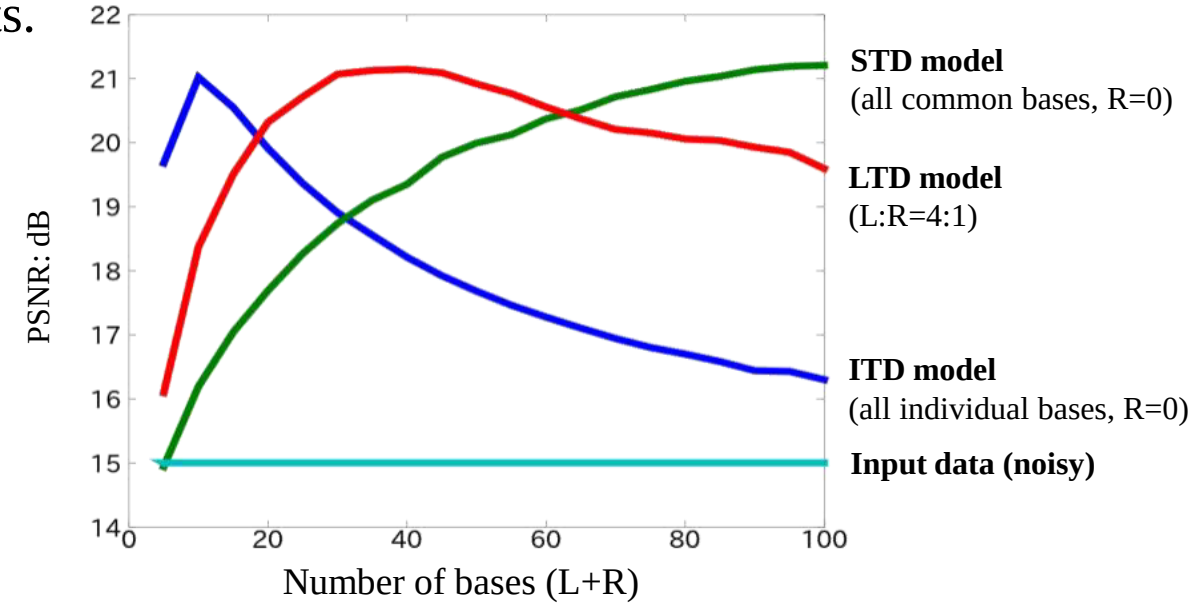


Fig. 9 Flow diagram of face denoising and reconstruction via linked tensor decomposition (Yokota et al., 2012)

3.2 Face denoising and reconstruction

- **Results**: measuring **PSNR** of reconstructed images for various number of components.



ITD ： good area is narrow, it is difficult to select the optimal number of bases ☹

STD ： we can see large number is good ☺, but computational cost is very high ☹

LTD ： good area is wide and in small number of bases, computational efficient ☺

(Yokota et al., 2012)

3.3 Joint analysis of multi-subject fMRI-like data

- nonnegative matrix factorization (NMF)

  --- single data

  --- part-based representation of nonnegative data

- Group nonnegative matrix factorization (GNMF)

  --- multi-set data: multi-subject/multi-modal biomedical data

  --- Constrained joint analysis of data from multiple sources allows us to explore potential connections (coupling information) and extract meaningful hidden components (sparse+nonnegative).

  --- Common patterns can be aligned naturally

Data description

- 'complex_fMRI_data.mat’
- url: http://mlsp.umbc.edu/simulated_complex_fmri_data.html

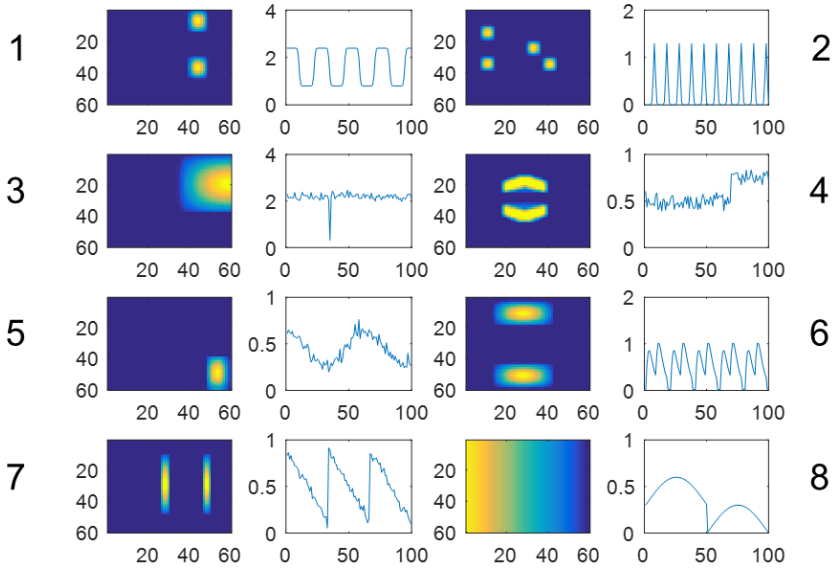


Fig. 10 The amplitude of the simulated fMRI spatial maps (SM: 3600×8) and their corresponding time courses (TC: 100×8).

Data reconstruction

$\boldsymbol{SM}$ - spatial map

$\boldsymbol{TC}$ - time course

- 6 groups: $\boldsymbol{X}\{k\}=\boldsymbol{SM}\{k\} \times \boldsymbol{TC}\{k\}, k = 1,2, \ldots, 6$ (Gong et al., 2018)
- Index: [1 2 5 6 7], [1 2 4], [1 2 4 5], [1 2 8], [1 2 3 5], [1 2 3 4]

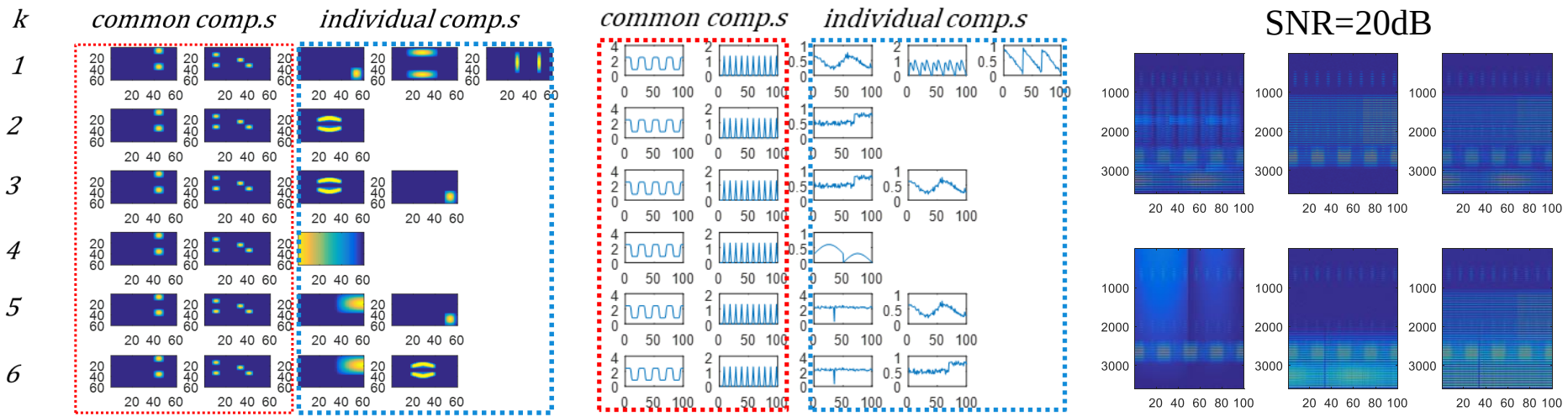


Fig. 11 The amplitude of SMs, TCs and mixtures

Parameter Settings

- R = {5, 3, 4, 2, 4, 4} , C = 2 or 0 (common or not)
- 30 independent runs, 20 dB
- Sparsity β: 25 values ranging from 0 to 5
- Method: ADMM
- 4 models are considered

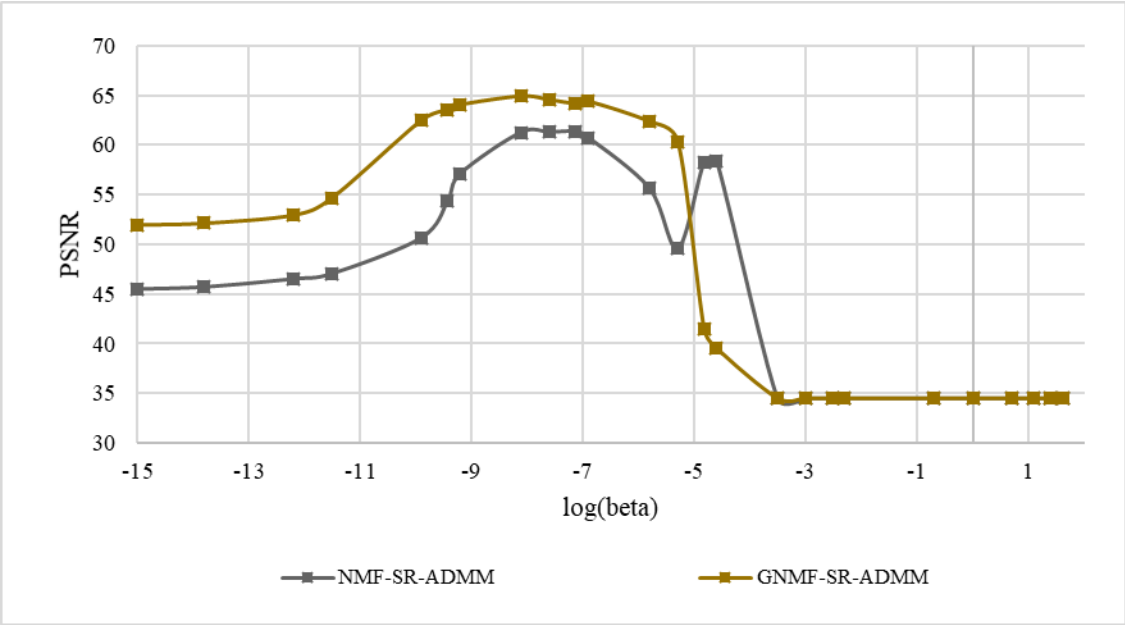
NMF: nonnegative matrix factorization (C=0, β=0)

GNMF: group nonnegative matrix factorization (C=2, β=0)

NMF-SR: NMF with sparse regularization (SR) (C=0, β>0)

GNMF-SR: GNMF with sparse regularization (SR) (C=2, β>0)

Results-1

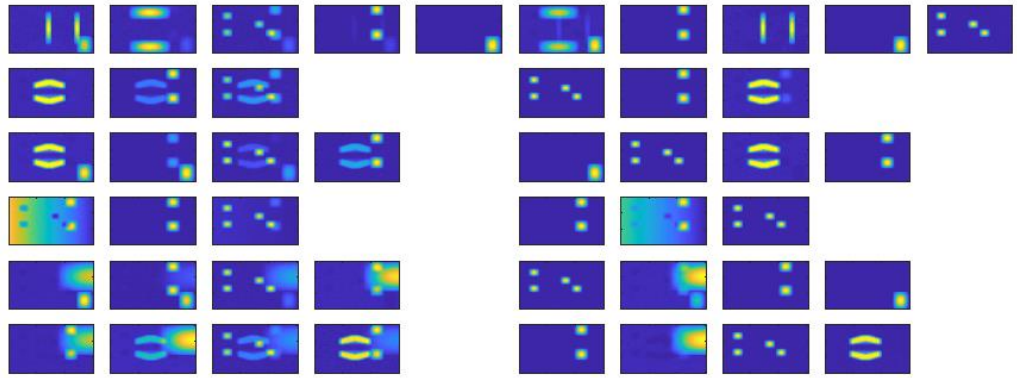


(1) A moderate sparse penalty will improve the performance of the algorithm, and then increasing it may have a negative impact;

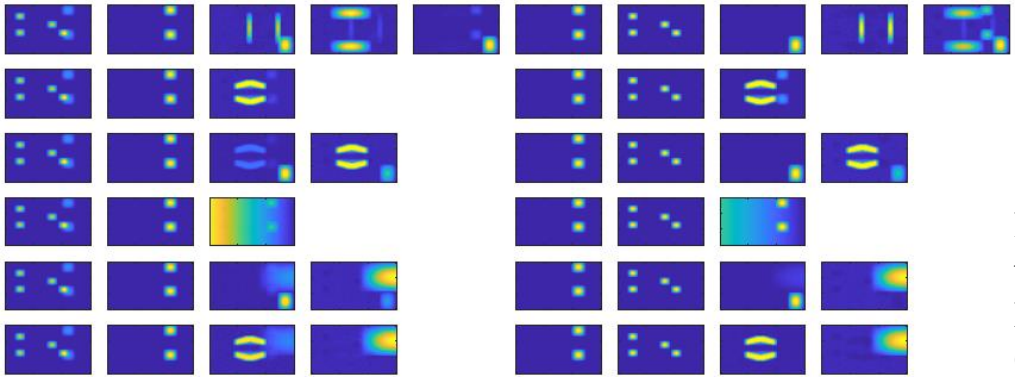
(2) The performance of the GNMF-based methods is superior to that of NMF-based ones;

Fig. 12 Mean PSNR of SM estimates for 6 groups under NMF-SR (C= 0) and GNMF-SR (C= 2) models with the βs of 25 values varying from 0 to 5, SNR=20dB.

Results-2



(a) NMF-ADMM, PSNR=45.52dB (b) NMF-SR-ADMM, PSNR=61.55dB



(c) GNMF-ADMM, PSNR=52.07dB (d) GNMF-SR-ADMM, PSNR=64.31dB

(1) By imposing the sparse regularization, small outliers or shadows can be eliminated.
(2) GNMF-based algorithm can extract both the common and individual patterns for all the datasets, and also successfully correct the disorder scenario of common patterns

Fig. 13 SM images of constructed data and that of estimated ones via NMF-ADMM, NMF-SR-ADMM, GNMF-ADMM and GNMF-SR-ADMM under SNR=20dB & $\beta = 3e-4$

3.4 Joint analysis of multi-subject ongoing EEG data

Data description (cong et al., 2013):

- 14 subjects, 20~46 years old, right-handed and healthy adults.
- 8.5-minute tango music of Piazzolla
- 64 electrodes
- Sampling rate: 2048 Hz, down-sampled to 256 Hz in the preprocess.
- Short-time Fourier transform (STFT)
- Tensor: 64 spatial channels × 46 frequency bins (4-13Hz) × 510 temporal samples × 14 subjects

Existing problems

- Problems of processing ongoing EEG data:
  1. High dimensionality (space, time and frequency)
  2. Nonnegative nature (time-frequency reprensentation)
  3. Heavy time consumption load (due to 1 and 2)
  4. Incomplete couplings (double-coupled)

- Solutions: In order to discover the reliable links between brain responses and musical stimulus, we proposed a comprehensive framework based on Fast double-coupled nonnegative CP decomposition (FDC-NCPD) for ongoing EEG data processing and analysis (Wang et al., 2020).

Why double-coupled?

In ongoing EEG data, we found correlations of temporal components between subjects are almost non-existent. (ps. Correlations of spectral components are not pronounced due to the sparse nature)

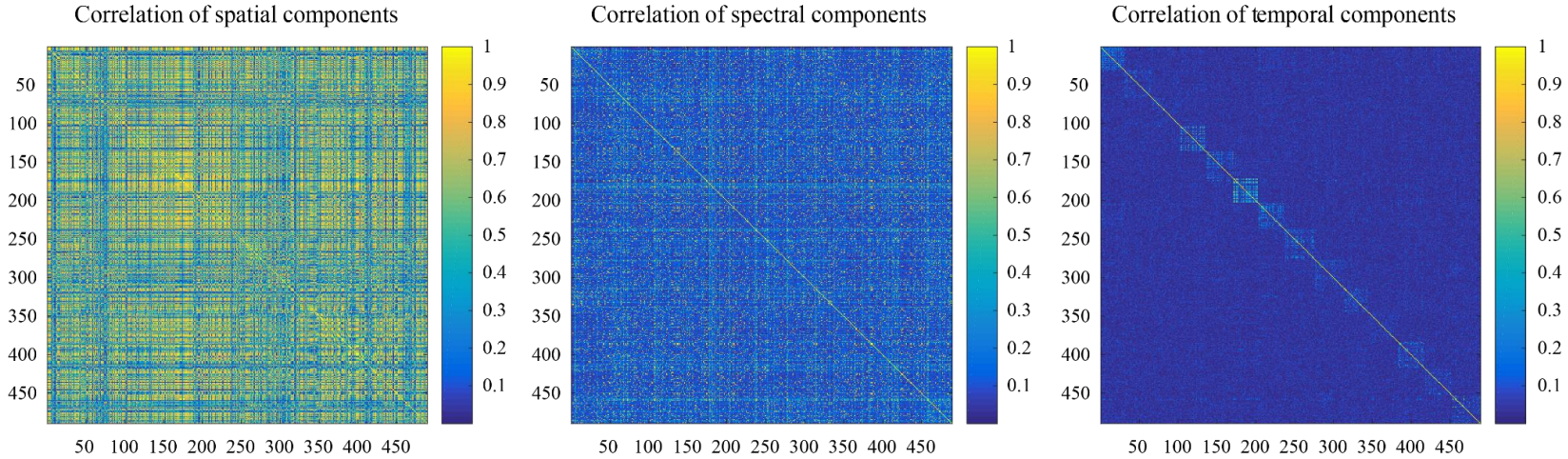
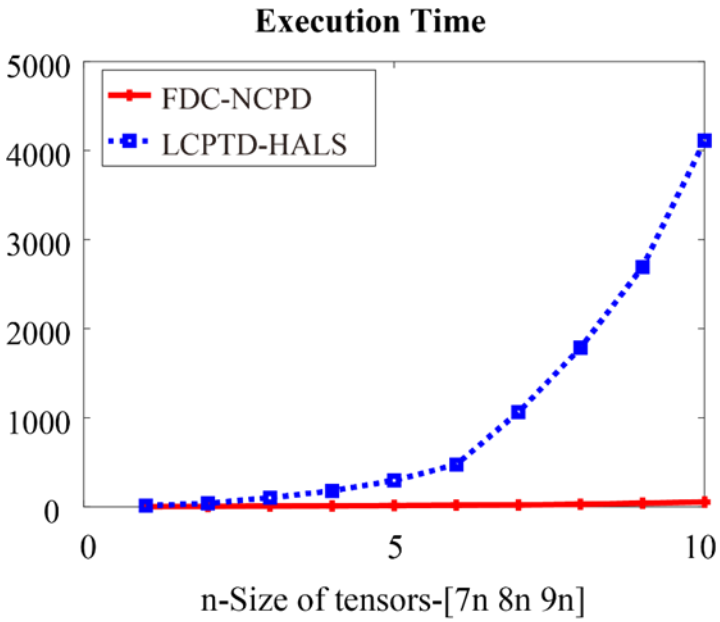
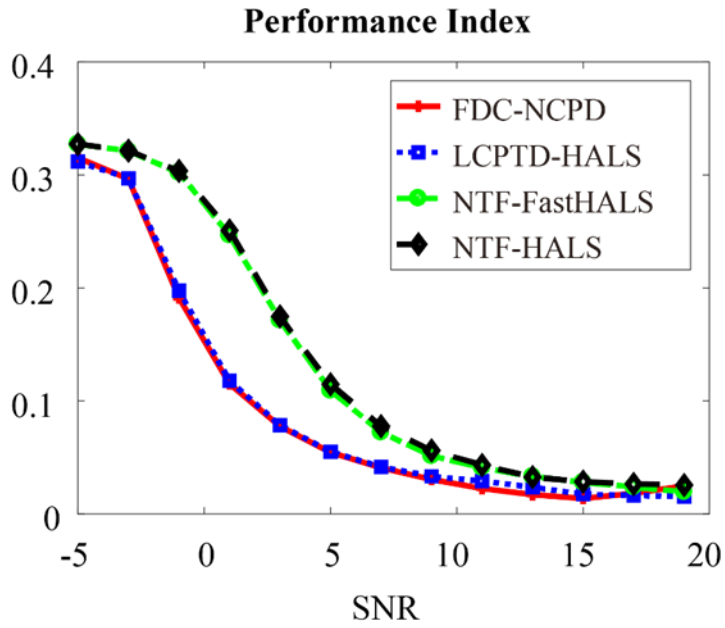


Fig. 14 Inter- and intra-subject correlations of spatial, spectral and temporal components. The spatial (spectral or temporal) components decomposed from ongoing EEG data of 14 subjects by tensor decomposition individually are concatenated together, and then the correlation coefficients are calculated

Synthetic data





Execution Time (s): FDC-NCPD << LCPTD-HALS, when n becomes larger
Performance Index: Coupled methods (LCPTD-HALS and FDC-NCPD) are better than non-coupled ones

FDC-NTD-based ongoing EEG analysis

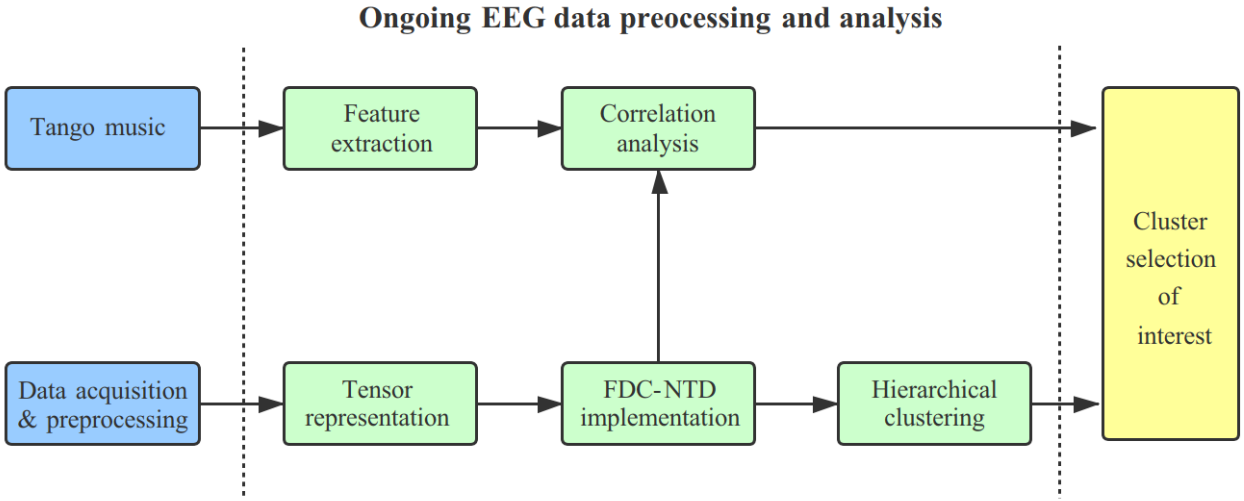


Fig. 15 Flow diagram of FDC-NTD-based ongoing EEG analysis

Results-1

- Interested components

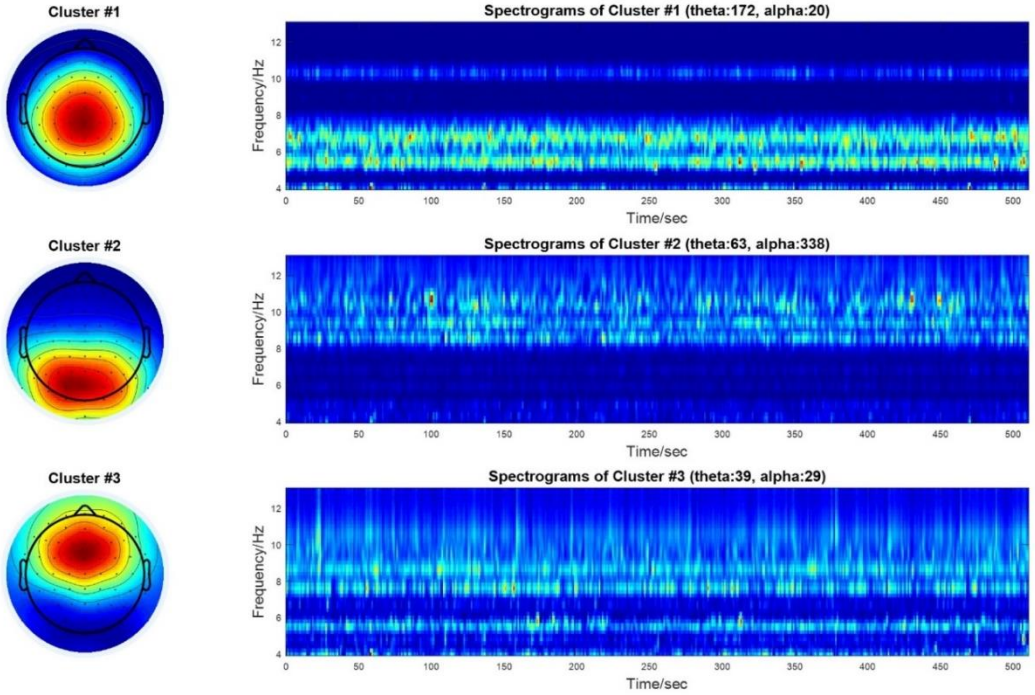
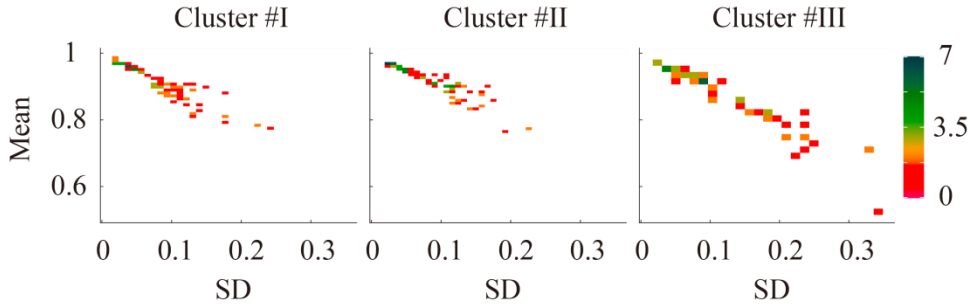


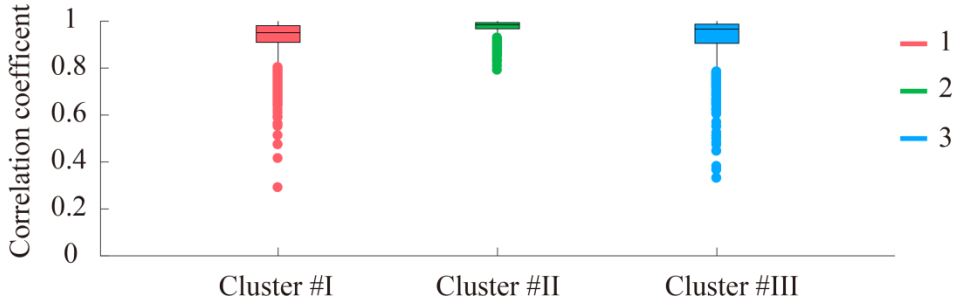
Fig. 16 Illustrations of averaged clusters of interest #I, #II and #III obtained from 100 runs. Cluster #1, #II and #III appeared 83, 100 and 96 times in 100 runs, respectively. Averaged topographies (left column) indicate the activations of centro-parietal, occipito-parietal and frontal regions of the brain elicited by musical stimulus, respectively. Overall spectrograms of clusters #I, #II and #III (right column) from 100 runs illustrate the frequency oscillations (theta, alpha, theta-alpha) over the entire period.

Results-2

- Clustering evaluation



(a) Correlations within runs



(b) Correlations between runs

Fig. 17 Correlations within and between runs for clusters #I, #II and #III in 100 runs.
(a) Distribution of means and SDs of correlation coefficients calculated by the internal spatial components in each run for clusters #I (mean-0.9236, SD-0.0881), #II (mean-0.9058, SD-0.0873) and #III (mean-0.8668, SD-0.1264).
(b) Illustration of correlation coefficients calculated by the averaged spatial components between runs for clusters #I (mean-0.9760, SD-0.0255), #II (mean-0.9340, SD-0.0654) and #III (0.9282, SD-0.0928).

References

- E. Acar, C. Schenker, Y. Levin-Schwartz, V. D. Calhoun, and T. Adali, “Unraveling diagnostic biomarkers of schizophrenia through structure-revealing fusion of multi-modal neuroimaging data,” Front. Neurosci., vol. 13, no. MAY, 2019.
- G. Zhou, Q. Zhao, Y. Zhang, T. Adali, S. Xie, and A. Cichocki, “Linked Component Analysis from Matrices to High-Order Tensors: Applications to Biomedical Data,” Proceedings of the IEEE, vol. 104, no. 2. Institute of Electrical and Electronics Engineers Inc., pp. 310–331, 01-Feb-2016.
- M. Mørup, “Applications of tensor (multiway array) factorizations and decompositions in data mining,” WIREs Data Min. Knowl. Discov., vol. 1, no. 1, pp. 24–40, Jan. 2011.
- T. Yokota, A. Cichocki, and Y. Yamashita, “Linked PARAFAC/CP tensor decomposition and its fast implementation for multi-block tensor analysis,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012, vol. 7665 LNCS, no. PART 3, pp. 84–91.
- A. Cichocki, “Tensor Decompositions: A New Concept in Brain Data Analysis?,” May 2013.
- R. Zdunek, K. Fonał, and A. Wołczowski, “Linked CP tensor decomposition algorithms for shared and individual feature extraction,” Signal Process. Image Commun., vol. 73, no. April 2018, pp. 37–52, 2019.
- X. F. Gong, L. Mao, Y. L. Liu, and Q. H. Lin, “A jacobi generalized orthogonal joint diagonalization algorithm for joint blind source separation,” IEEE Access, vol. 6, no. 1, pp. 38464–38474, 2018.
- X. Wang, W. Liu, P. Toiviainen, T. Ristaniemi, and F. Cong, “Group analysis of ongoing EEG data based on fast double-coupled nonnegative tensor decomposition,” J. Neurosci. Methods, vol. 330, p. 108502, Jan. 2020.


THANKS