Coupled matrix/tensor decomposition for multi-block datasets

Xiulin Wang

School of Biomedical Engineering

Dalian University of Technology

1. Background

Data characteristics: multi-block datasets

1. Multi-set/multi-modal

    Multiple datasets collected from the same type of data

    Multiple datasets collected from different types of data

2. Multi-way
3. Multi-coupling

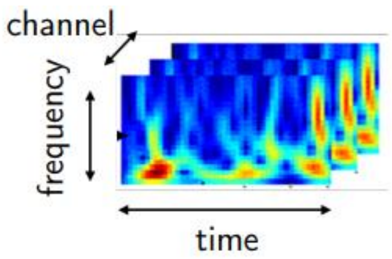


Fig. 1 EEG tensor data[1]

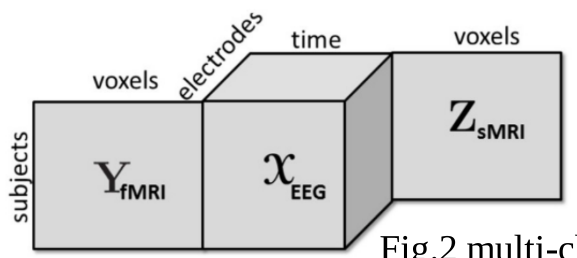


Fig.2 multi-channel EEG signals is coupled with fMRI and sMRI data in the form of matrices in the subject mode (Acar et al., 2019)

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

SM: spatial map

TC: temporal courese

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.

Spatial consistency

Intra-set independence

Inter-set dependence

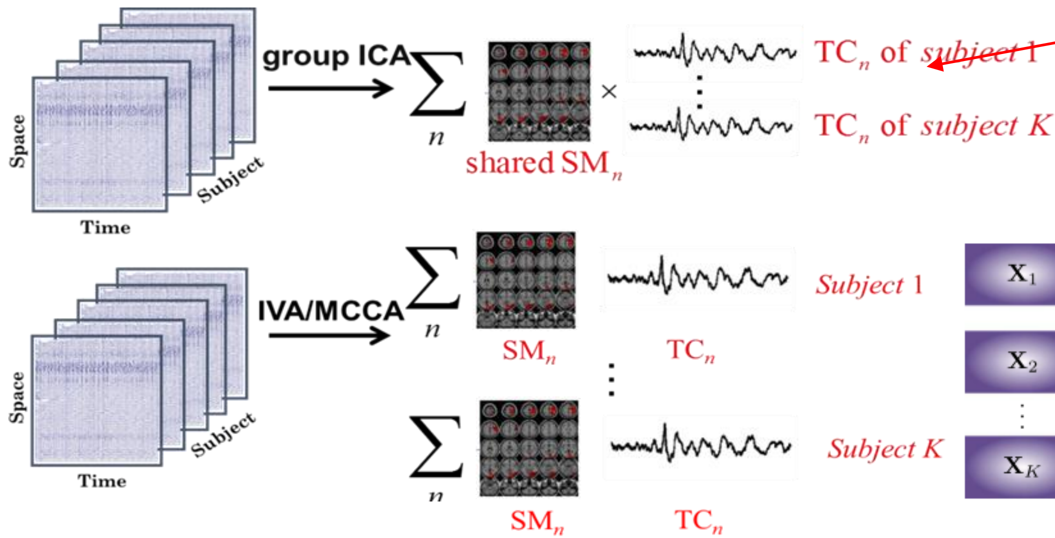


Fig. 3 (a-b) Joint analysis of multi-subject fMRI data

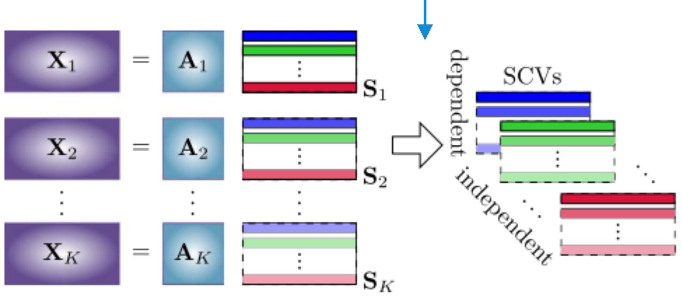
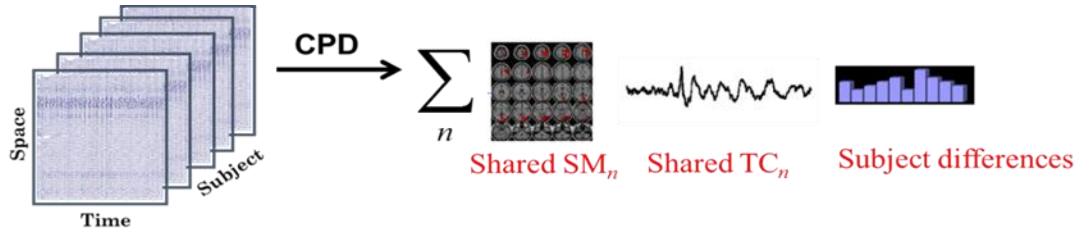


Fig. 4 Diagram of the IVA model (Zhou et al., 2016)

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. Individual tensor methods can not utilize the prior coupling information shared by datasets.

e.g.



Spatial consistency

Temporal consistency

Fig. 3 (c) 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. Individual tensor methods can not utilize the prior coupling information shared by datasets.

e.g.

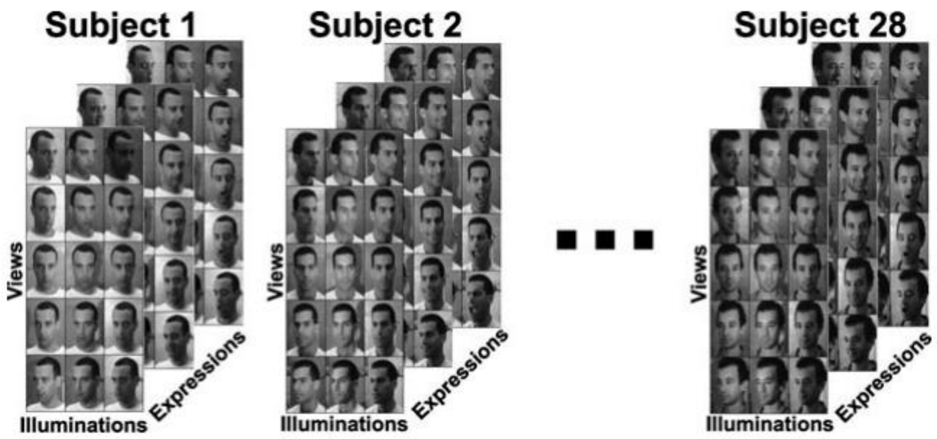


Fig. 5 Illustration of the Weizmann face database used in the analysis of TensorFaces (Mørup et al., 2011)

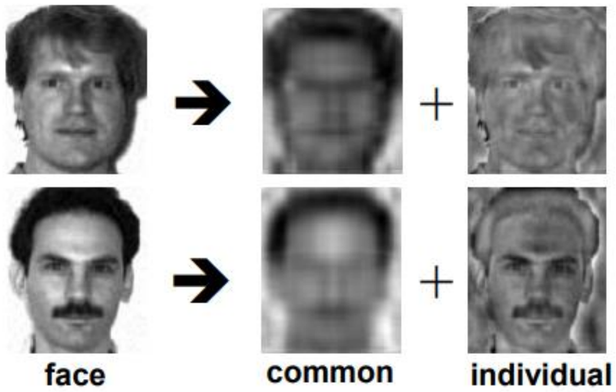


Fig. 6 Illustration of face images obtained by coupled tensor decomposition (Yokota et al., 2012)

Solutions: Coupled matrix/tensor methods

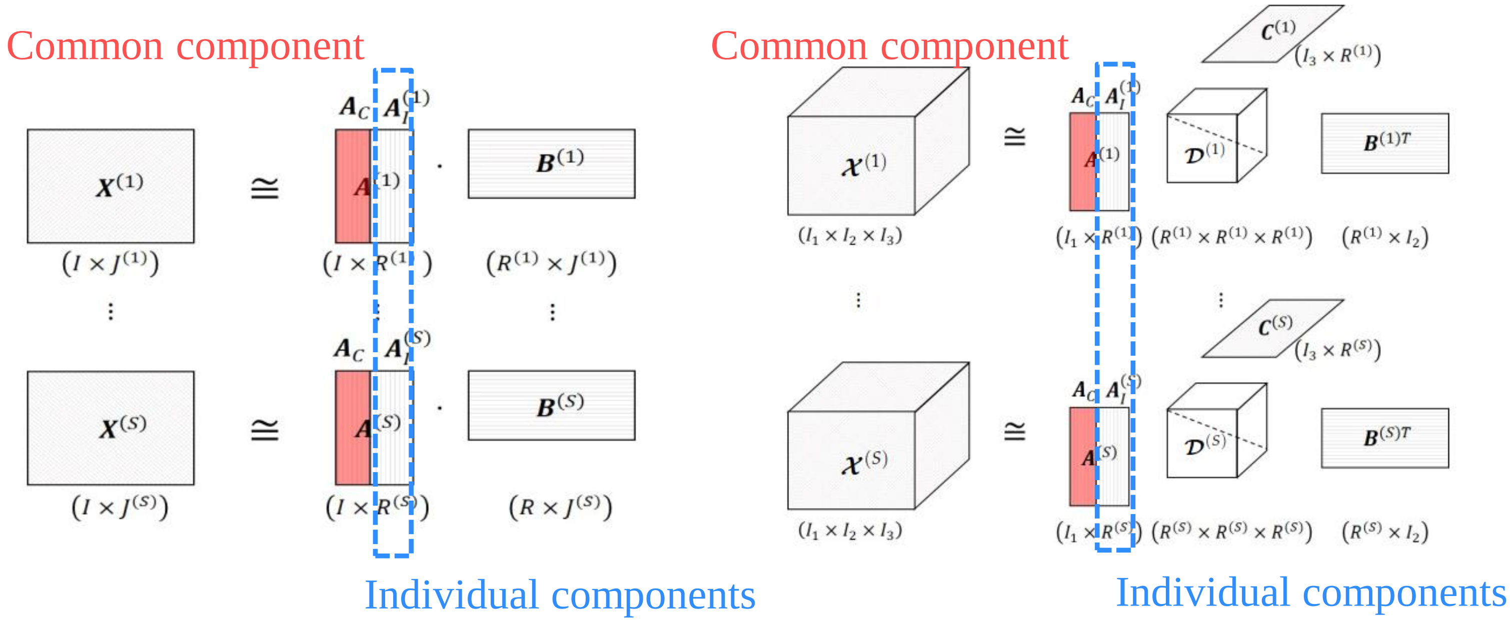


Fig. 7 Conceptual illustration of mode-1coupled matrix/tensor decomposition model

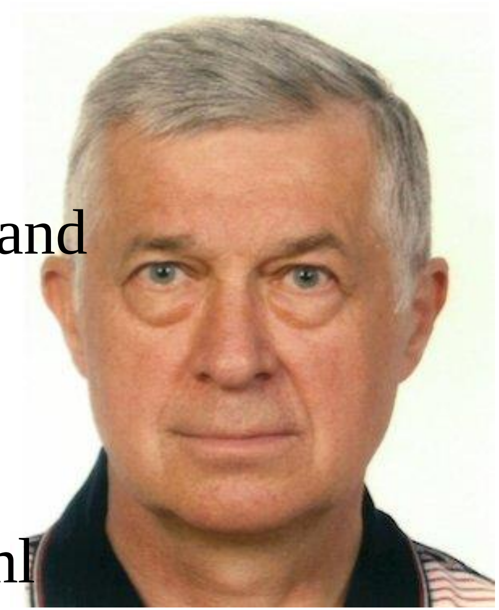
- Common component shared by all blocks corresponds to the same or maximally correlated components.
- Individual components correspond to different individual characteristics

[3] Zhou G. et al. Group Component Analysis for Multiblock Data: Common and Individual Feature Extraction
[4] Cichocki A. et al. Tensor Decompositions: A New Concept in Brain Data Analysis?

Why coupled?

- Coupled matrix/tensor decompostion 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 factor matrices, individual factor matrices and core tensors.
- 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.

2. Top research Labs (1/3)



- **SKOLTECH-Laboratory:** Tensor networks and deep learning for applications in data mining
- **Leader:** Prof. **Andrzej Cichocki**
- Google scholar **H-index: 97**
- **Website:** http://www.deeptensor.ml/index.html
- **Research Interests**

1) multi-way analysis, tensor decompositions and factorizations, group and multi-block analysis in applications to processing and mining of biomedical and geophysical massive data, big data analytics in biomedical engineering, computational neuroscience;
2) tensor Networks and their applications in big data analytics;
3) early detection of Alzheimer's disease;
4) blind source separation (BSS), especially ICA, SCA, NMF, multiway BSS, Linked;
5) multi-block, multilinear ICA, nonnegative tensor factorizations. 6) Intelligent signal processing and massive data analysis and their applications;
7) learning theories and optimization techniques;
8) inverse problems and their biomedical applications;
9) brain computer interface (BCI), Brain Robot Interface and noninvasive recording and visualization of brain signals (EEG/ MEG, fMRI);
10) neural computation and nonlinear adaptive systems;
11) optimization and operations research and their biomedical applications;
12) neuroinformatics and bioinformatics.

2. Top research Labs (2/3)

- **Tensor Laboratory in ESAT-STADIUS Center**
- **Leader:** Prof. **Lieven De Lathauwer**
- Google scholar **H-index: 56**
- **Website**

  https://www.esat.kuleuven.be/stadius/person.php?id=22
- **Research Interests**

  linear and multilinear algebra, higher-order tensor, matrix/tensor decomposition, blind source separation, etc.
- **Software Toolbox**

  Tensorlab: MATLAB toolbox for tensor computations and decompositions

  https://www.tensorlab.net/

2. Top research Labs (3/3)

- **Tensor group in Sandia National Laboratories**
- **Leader:** Dr. **Tamara G. Kolda**
- Google scholar **H-index: 47**
- **Website:** http://www.kolda.net/
- **Research Interests**

  multilinear algebra and tensor decompositions,

  graph models and algorithms, data mining, optimization, etc.

  "Tensor decompositions and applications", SIAM review, 2009
- **Software Toolbox**

  Tensor Toolbox: Higher-order operations of multidimensional arrays

  http://www.tensortoolbox.org/

2. Top research Labs (4/4)

- **Simula Research Laboratory (Machine Intelligence Department)**
- **Leader:** Dr. **Evrim Acar**
- Google scholar **H-index:** 26
- **Website:** http://www.models.life.ku.dk/sites/default/files/EAA.html
- **Research Interests**

  Data mining and mathematical modeling: tensor decompositions, data fusion using coupled factorizations of higher-order tensors and matrices, applications in metabolomics, neuroscience, etc.
- **Software Toolbox**

  CMTF Toolbox: Coupled Matrix and Tensor Factorization

  http://www.models.life.ku.dk/joda/CMTF_Toolbox

Our researches (1/2)

- Constrained joint blind source separation by **group nonnegative matrix factorization** (GNMF) with sparse regularization

- part-based representation of nonnegative data
- constrained joint analysis of data from multiple sources allows us to explore potential connections (coupling information) and extract meaningful hidden components.
- common patterns can be aligned naturally

3. Our researches (1/2)

- Constrained joint blind source separation by group nonnegative matrix factorization (GNMF) with sparse regularization (GNMF-SR)

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

$$\boldsymbol{X}^{(s)} = \boldsymbol{A}^{(s)}\boldsymbol{B}^{(s)}, s = 1,2, \dots, S$$

$$\boldsymbol{A}^{(s)} = \begin{bmatrix} \boxed{\boldsymbol{A}_C^{(s)}} & \boxed{\boldsymbol{A}_I^{(s)}} \end{bmatrix}$$

$\boldsymbol{A}^{(s)}$ - variable matrix

$\boldsymbol{B}^{(s)}$ - coefficient matrix

$$\boldsymbol{A}_C^{(1)} = \boldsymbol{A}_C^{(2)} = \cdots \boldsymbol{A}_C^{(S)} = A_C$$

3. Our researches (1/2)

- Constrained joint blind source separation by group nonnegative matrix factorization (GNMF) with sparse regularization (GNMF-SR)

  - GNMF-SR model

$$\underset{\boldsymbol{A}^{(s)},\boldsymbol{B}^{(s)}}{\text{minimize}}\ \frac{1}{2}\sum_{s=1}^{S}\left\|\boldsymbol{X}^{(s)}-\boldsymbol{A}^{(s)}\boldsymbol{B}^{(s)}\right\|_F^2+\sum_{s=1}^{S}\beta^{(s)}\sum_{r=1}^{R(s)}\left\|\boldsymbol{a}_r^{(s)}\right\|_1$$

$$\text{subject to}\ \ \boldsymbol{A}^{(s)}>0,\boldsymbol{B}^{(s)}>0,\boldsymbol{A}_C^{(1)}=\cdots\boldsymbol{A}_C^{(S)}=\boldsymbol{A}_C$$

Data description

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

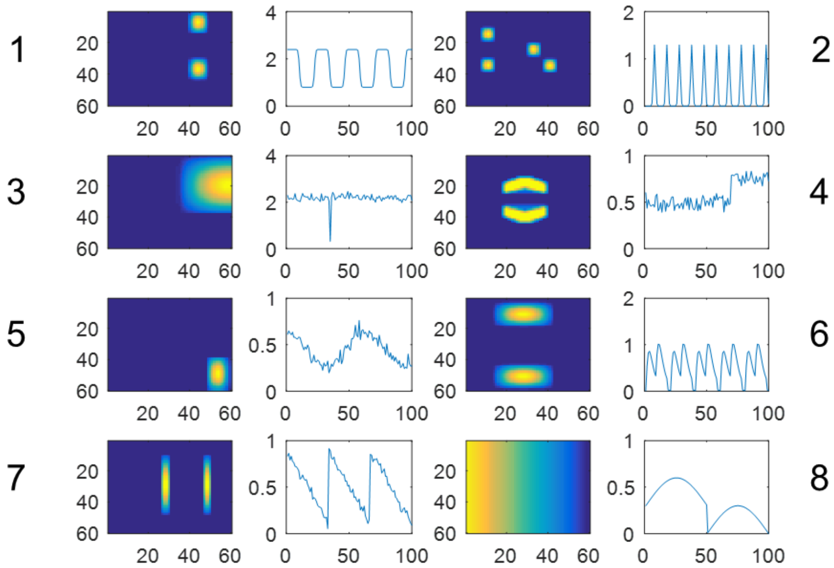


Fig. 8 The amplitude of the simulated fMRI spatial maps (SM: 3600×8) and their corresponding time courses (TC: 100×8).

Data construction

- 6 groups: $\text{Mixed Signal}\{k\}=\text{SM}\{k\}\times\text{TC}\{k\}, k=1,2,\ldots,6$
- Index: [1 2 5 6 7], [1 2 4], [1 2 4 5], [1 2 8], [1 2 3 5], [1 2 3 4]

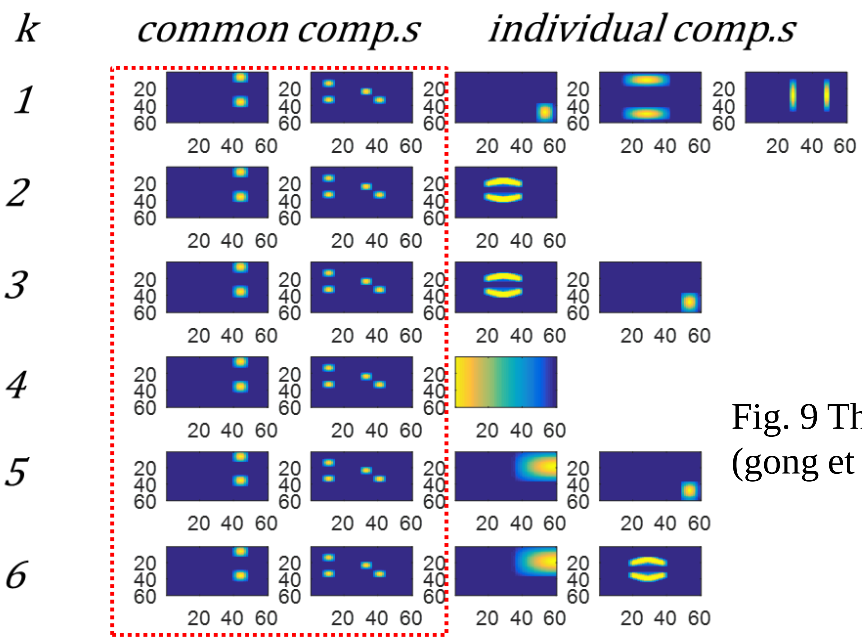
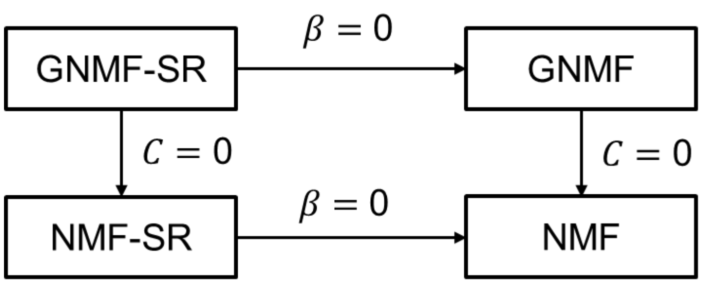


Fig. 9 The amplitude of SMs of the constructed matrices (gong et al., 2018)

Experiment settings

- Real domain (the amplitude of SMs and TCs)
- R = {5, 3, 4, 2, 4, 4} , C = 2 or 0 (common or not)
- SNR=20dB, 30 independent runs
- Sparsity β: 25 values ranging from 0 to 5
- 4 models are considered



β : sparse parameter

C : the number of coupled comp.s

Experiment results-1

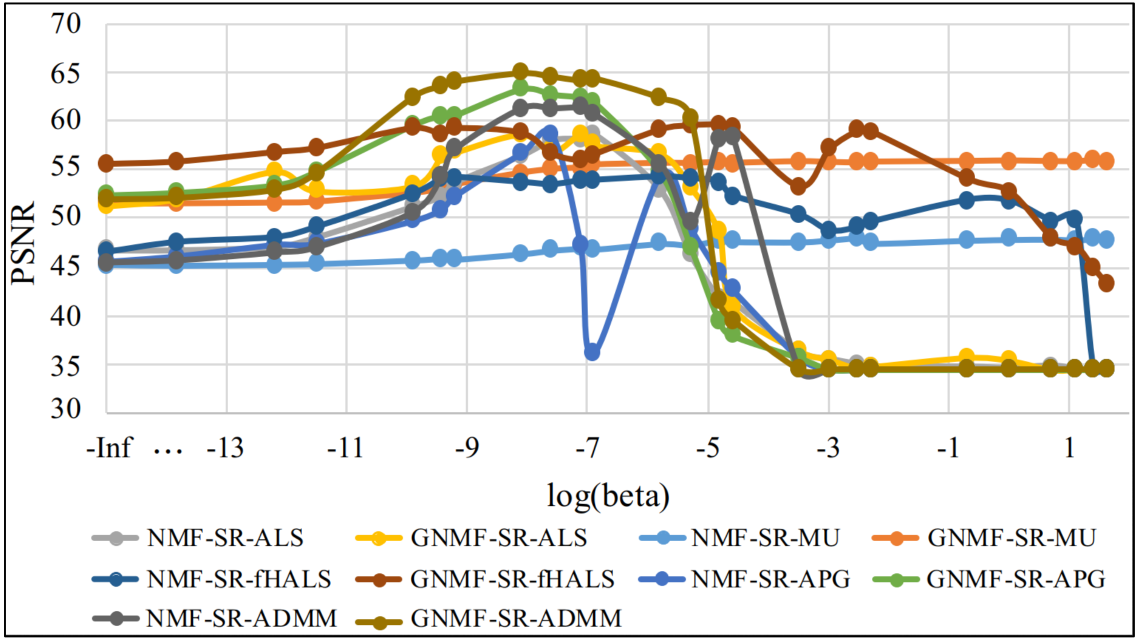
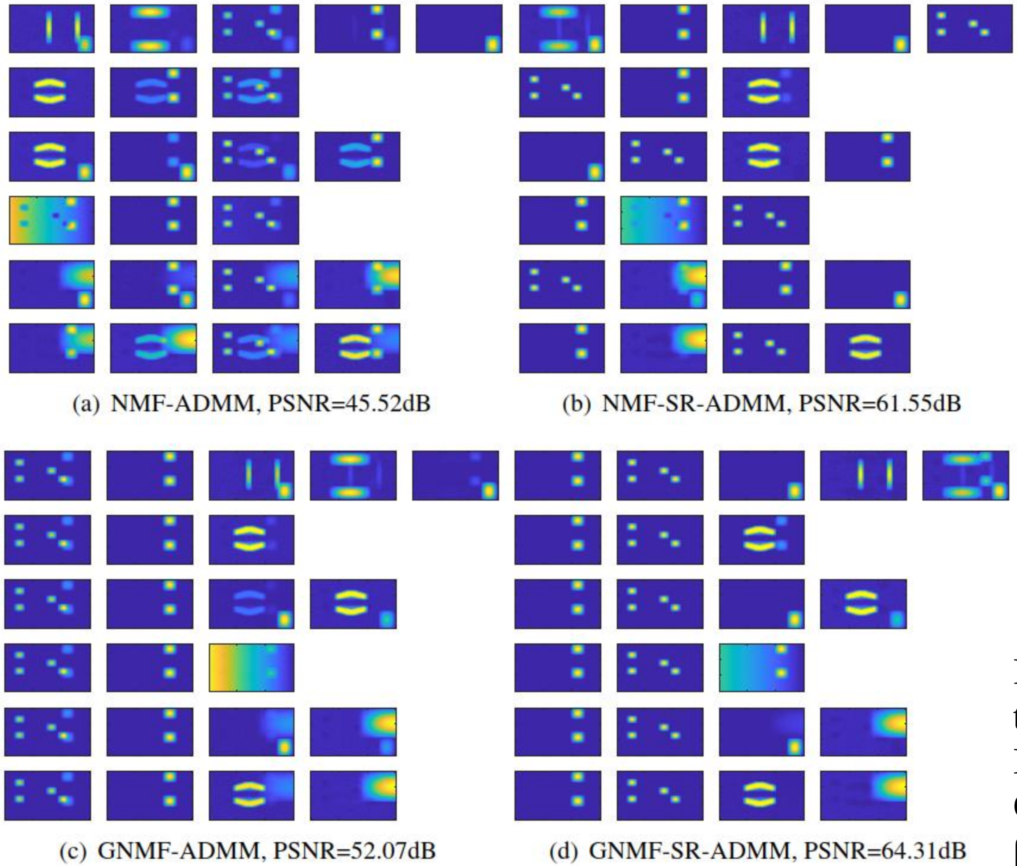


Fig. 10 Mean PSNR of SM estimates for 6 subjects under NMF-SR (L = 0) and GNMF-SR (L = 2) models with the βs of 25 values varying from 0 to 5, SNR=20dB.

(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;
(3) With sparse regularization, the performance of NMF-based and GNMF-based methods can be both significantly improved;
(4) Sparse penalty yields better performance improvements than group constraint for NMF-based methods;
(5) ADMM-based methods achieve better performance

Experiment results-2



(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. 11 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 & $\beta = 3e-4$

Our researches (2/2)

- Joint analysis of ongoing EEG data via **couped tensor decomposition**

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 × 46frequency bins (4-13Hz) × 510 temporal samples × 14 subjects

Our researches (2/2)

- Joint analysis of ongoing EEG data

  - 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 tensor decomposition for ongoing EEG data processing and analysis (Wang et al., 2019&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)

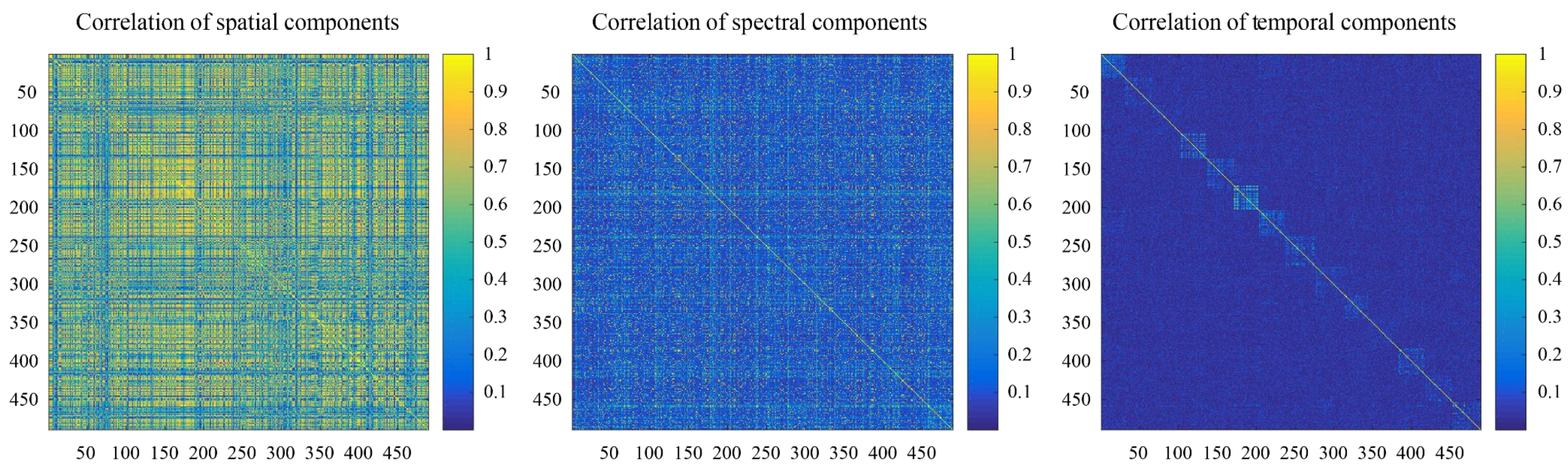
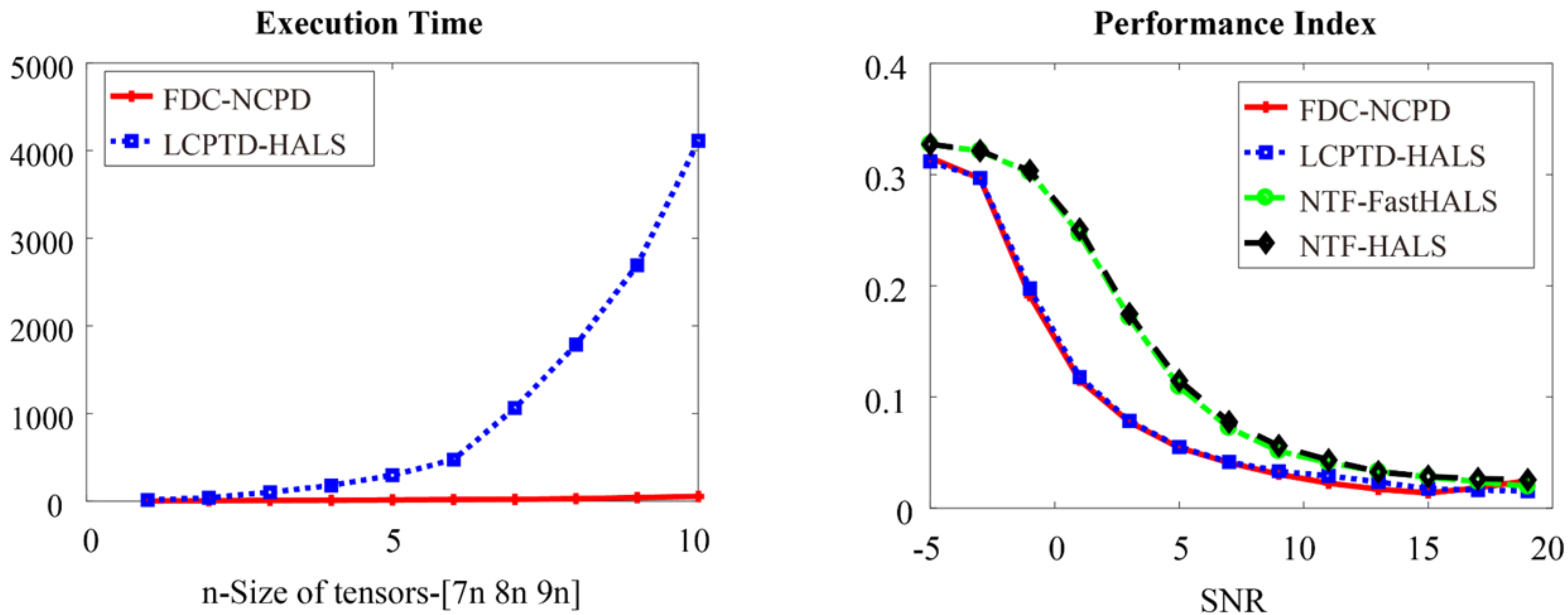


Fig. 12 Inter- and intra-subject correlations of spatial, spectral and temporal components. The spatial (spectral or temporal) components decomposed from ongoing EEG data of 14 subjects by tensor decomposition individually are concatenated together, and then the correlation coefficients are calculated

Fast double-coupled nonnegative Tensor Decomposition (FDC-NTD)



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

- data acquisition & preprocessing;
- musical feature extraction;
- tensor representation of data;
- FDC-NTD implementation (100 runs);
- correlation analysis;
- hierarchical clustering;
- cluster selection of interest

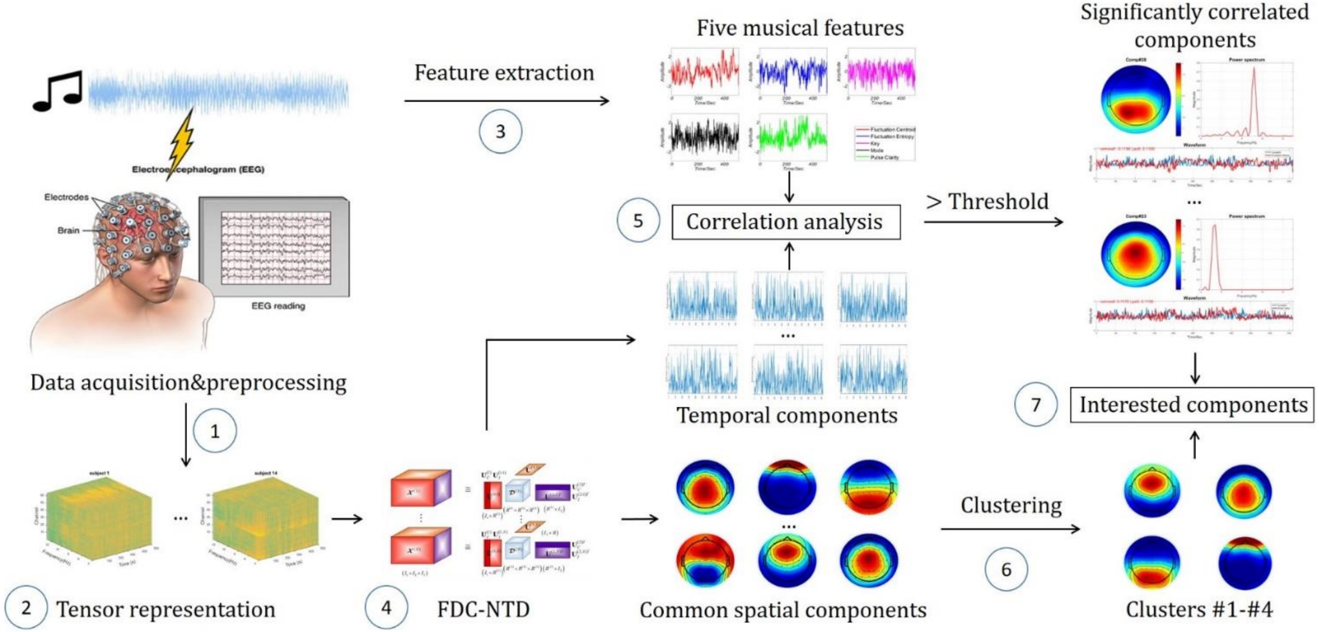


Fig. 13 Flow diagram of FDC-NTD-based ongoing EEG analysis

Experiment results-1

- An output example of correlation analysis

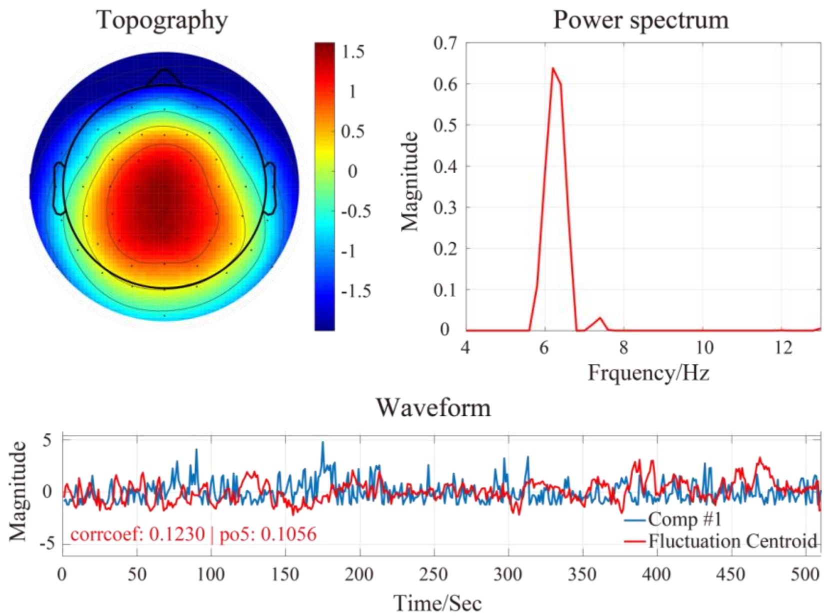


Fig. 14 The topography, power spectrum and waveform of the 1st EEG components from subject #11 of Run #7. ‘po5’ denotes the threshold of significant correlation coefficients at level $p < 0.05$ and ‘corrcoef’ denotes the correlation coefficient between temporal course of component #1 and temporal course of ‘Fluctuation Centroid’.

Experiment results-2

- An output example of hierarchical clustering

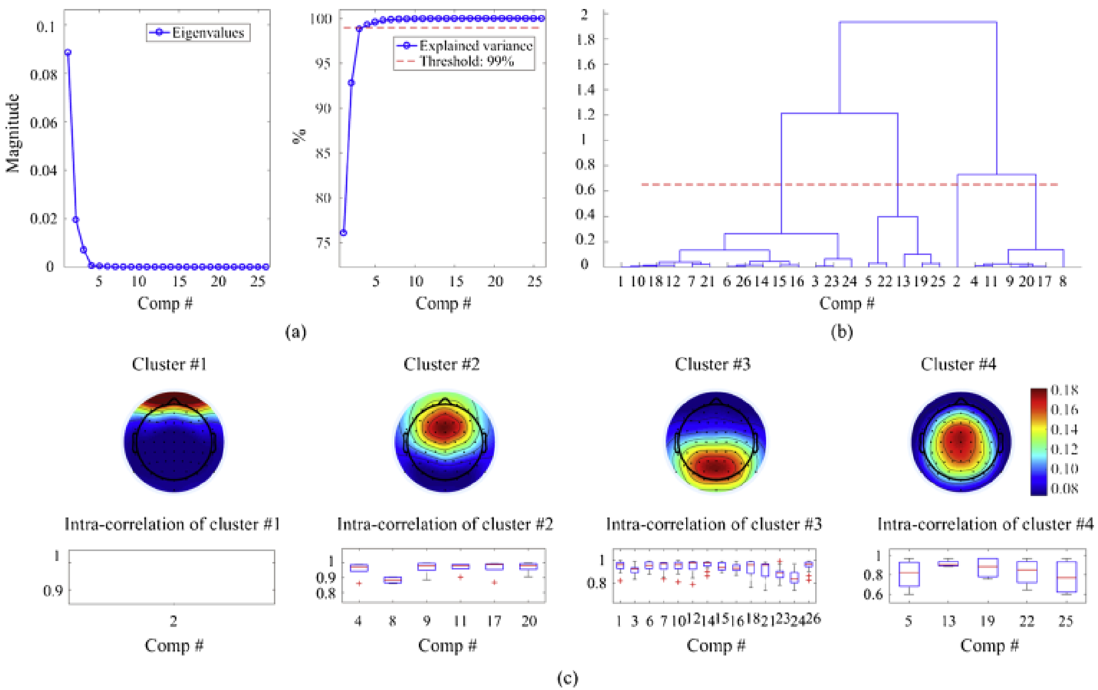


Fig. 15 Hierarchical clustering results of 26 common spatial components of Run #7.

(a) Selection of the number of clusters;

(b) Dendrogram output of hierarchical clustering;

(c) Averaged topographies of clusters and correlations between components within the clusters

Experiment results-3

- Interested components

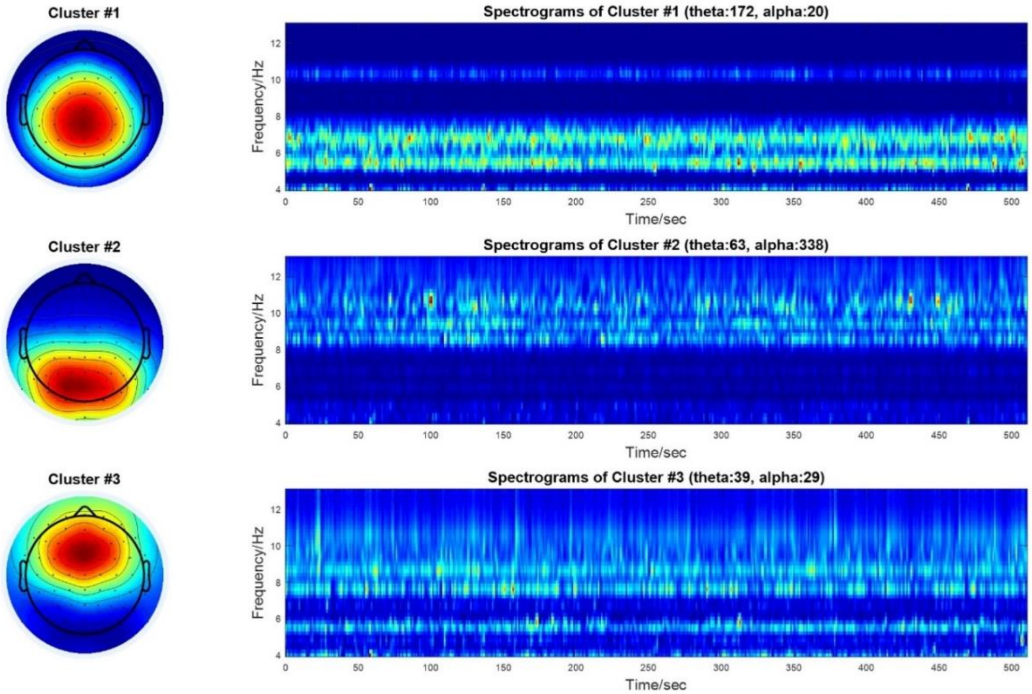
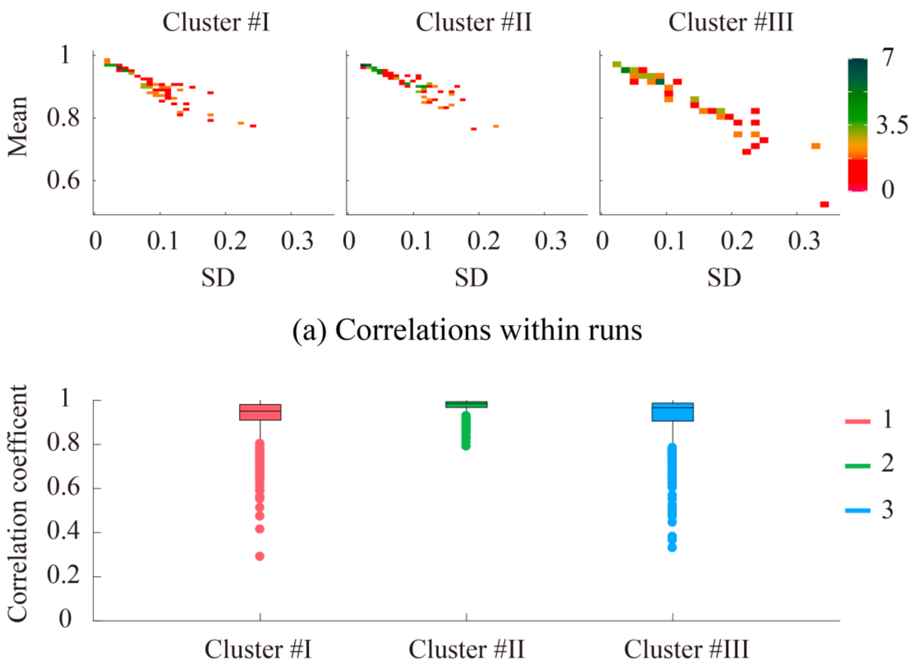


Fig. 16 Illustrations of averaged clusters of interest #I, #II and #III obtained from 100 runs. Cluster #1, #II and #III appeared 83, 100 and 96 times in 100 runs, respectively. Averaged topographies (left column) indicate the activations of centro-parietal, occipito-parietal and frontal regions of the 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.

Experiment results-4

- Clustering evaluation



(a) Correlations within runs

(b) Correlations between runs

Figure 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).

4. Publications

PI Xiulin Wang, Wenya Liu, Fengyu Cong, and Tapani Ristaniemi. Group Nonnegative Matrix Factorization with Sparse Regularization in Multi-set Data. *28th European Signal Processing Conference (EUSIPCO), Amsterdam, Netherlands*, 2020.

PII Xiulin Wang, Chi Zhang, Tapani Ristaniemi and Fengyu Cong. Generalization of Linked Canonical Polyadic Tensor Decomposition for Group Analysis. *16th International Symposium on Neural Networks (ISNN 2019), Moscow, Russia*, http://doi.org/10.1007/978-3-030-22808-8_19, 2019.

PIII Xiulin Wang, Tapani Ristaniemi and Fengyu Cong. Fast Implementation of Double-coupled Nonnegative Canonical Polyadic Decomposition. *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK*, http://doi.org/10.1109/ICASSP.2019.8682737, 2019.

PIV Xiulin Wang, Wenya Liu, Petri Toiviainen, Tapani Ristaniemi and Fengyu Cong. Group analysis of ongoing EEG data based on fast double-coupled nonnegative tensor decomposition. *Journal of neuroscience methods, 330, p.108502*, http://doi.org/10.1016/j.jneumeth.2019.108502, 2020.

PV Xiulin Wang, Tapani Ristaniemi and Fengyu Cong. Fast Learnings of Coupled Nonnegative Tensor Decomposition Using Optimal Gradient and Low-rank Approximation. *Submitted to signal processing*, 2020.


THANKS