［研究助成（A）］

ブラインド信号源分離に適用する非線形混合のマルチ部分空間表現に関する研究

A Study on Multi-Subspace Representation of Nonlinear Mixture with Application in Blind Source Separation

2201034

研究代表者　慶應義塾大学大学院 理工学部・理工学研究科　特任助教　Wang Lu

［研 究 の 目 的］

Development and wide deployment of various sensor technologies have led to the generation of large amounts of multi-block data (i. e., multi-set, multi-relational, or multi-view), which need to be jointly analyzed to extract meaningful information. The technology called tensor refers to a multi-way array of numbers, which encapsulates high-dimensional big data information in a compact format. Benefiting from the power of multi-linear algebra of their mathematical backbone, data analysis techniques using tensor decomposition shows great flexibility in the choice of constraints that reflect data diversity. Thus, multi-block tensor decomposition and multi-way analysis allow us to discover meaningful hidden spatial and temporal structures of complex data and perform generalizations by capturing multi-linear and multi-aspect relationships. As a higher-order generalization of matrices and vectors, theoretical tools and numerical algorithms for tensor manipulation have been developed across many fields including tensor flow for neural networks, and tensor decomposition for big data.

Multi-block data are being explored in many applications, due to inevitably imperfection of potential interactions. For instance, to predict hand movement through brain responses recorded by multiple electrodes, it is usually necessary to consider the object pattern in addition to the two patterns of time and space. This means the brain recording can be naturally fit into a multi-way array with multiple modes. However, the mostly applied computing tools for brain research are oriented for one-way or two-way data. Consequently, the facilitation and availability of the multi-way signal processing are attributed to the ability of the model interactions between complex latent components. Moreover, most methods are suffered from the fact that the observed recordings usually have a low signal-to-noise ratio (SNR), caused by the inevitable contamination from the Electrooculography (EOG) signal, body movement, and artificial noise during the propagation process. It is a typical problem to as blind source separation (BSS). Therefore, it is crucial to extract accurate target signals from the mixture recordings, which reflects the accuracy of prediction when capturing multiple interactions and couplings.

［研究の内容，成果］

In this topic, we propose to apply the tensor theory to develop a novel structure for the joint BSS technique that allows us to discover the multi-set components and meaningful spatio-temporal structures of data. In our model, the tensor computing is regularized by a low-rank prior to exploring a stack of cross covariance matrices along with temporal information, which in terms of limited storage, providing the efficiently encapsulate and compress large-scale data into a compact format. Thus, the objective function is formulated by taking the coupling information into account, which explores not only the statistical independence within multiple datasets but also dependence among the different datasets. Thus, the objective function is formulated by taking the coupling information into account, which explores not only the statistical independence within multiple datasets but also dependence among the different datasets. The approach is beneficial to decompose the tensor format into multiple low-rank tensors of matrices, thus providing an efficient way for analyzing big and complex data. In addition, a practical problem for the mixing matrices is its uniqueness and identifiability. We discuss how our results provide the uniqueness decomposition under some conditions used for estimating the mixing matrices.

In this research, a joint BSS model is considered on an ensemble of K datasets. The N observations in each dataset are generated by the linear mixture of N source signals. The mixing process is modeled by

$$\mathbf{x}^{[k]}=\mathbf{A}^{[k]}\mathbf{s}^{[k]}+\mathbf{e}^{[k]},\quad k=1,2,\cdots,K,$$

where $\mathbf{x}^{[k]}=[x_1^{[k]},\cdots,x_N^{[k]}]^{\mathrm{T}}\in\mathbf{R}^N$ and $\mathbf{s}^{[k]}=[s_1^{[k]},\cdots,s_N^{[k]}]^{\mathrm{T}}\in\mathbf{R}^N$ represent the N -dimensional observations and source signals in the k th dataset. The mixing matrix $\mathbf{A}^{[k]}\in\mathbf{R}^{N\times N}$ represents channel effect between signals and sensors which is unknown. $\mathbf{e}^{[k]}\in\mathbf{R}^N$ denotes the possible additive noise, which is generally assumed to be zero mean, temporally white and uncorrelated with the source signals.

The purpose of BSS is to find a set of invertible matrices$\mathbf{A}^{-[1]}, \mathbf{A}^{-[2]}, \cdots, \mathbf{A}^{-[K]}$ such that to estimate the source signals from the observations in multiple datasets K, where $\mathbf{A}^{-[1]}$ denotes $(\mathbf{A}^{[k]})^{-1}$ for convenience.

Considering the separation criterion of the joint BSS that are similar to several existingmethod, two assumptions are satisfied regarding the sources in this paper, which is illustrate in Fig. 1.

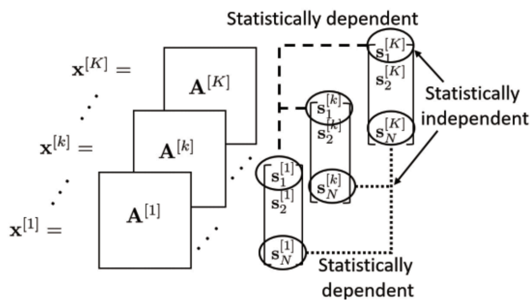
(1) The sources are statistical uncorrelated within each dataset k formed as

$$\mathbf{E}[s_i^{[k]}(t)s_j^{[k]}(t+\tau)^H]=0,\ 1\leq i\neq j\leq N,$$

where τ is the time delay and $(\cdot)^H$ denotes the complex conjugate transpose.

(2) Only the corresponding sources are dependence from two different datasets, formed as

$$\begin{aligned}\mathbf{D}^{[k]}&=\mathbb{E}[\mathbf{s}^{[k_1]}(t)(\mathbf{s}^{[k_2]}(t+\tau))^H]\\&=diag[\rho_1(\tau),\cdots,\rho_N(\rho)],\end{aligned}$$

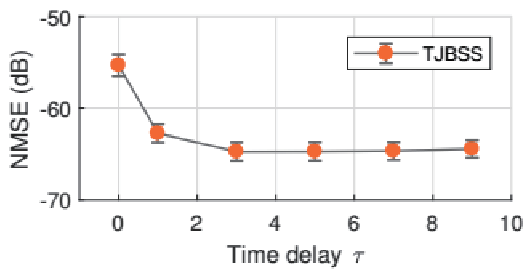


The assumptions regarding the sources : the sources are statistical independence within each dataset, but that are dependence among the different datasets

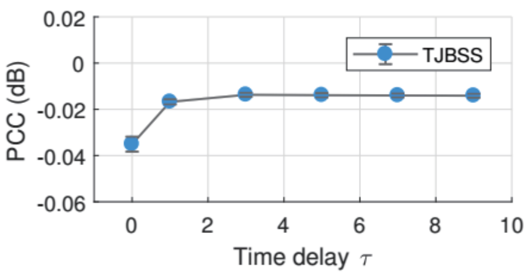
Fig. 1　Illustration of a joint BSS model where $1 \leq k_1 \neq k_2 \leq K$ and the diagonal matrix $\rho_n(\tau)$ define a covariance between $s_n^{[k_1]}(t)$ and $s_n^{[k_2]}(t+\tau)$. Here, we take into account for the case of $K=3$ datasets as an example. Using the proposed TJBSS algorithm, the underdetermined problem also can be effective to deal with when the number of dataset more than 3.

Separation of Synthetically Generated Data

We first cross-validated the time delay $¥tau$ on the synthetic data used for learning parameter of our model. It is clear that the more variable as estimation but the longer to access the model. i. e., the tensor structure is stacked by the covariance matrices with a time delay τ. The greater value of τ are used to constrain the separation model, the better performance can be achieved. However, we can see from Fig. 2, the performance of the proposed model becomes saturated as the parameter τ achieves 3 approximately. To prevent the potential model over-saturation, τ set to 3 suggested by the trends of convergence in both NMSE and PCC measurements.



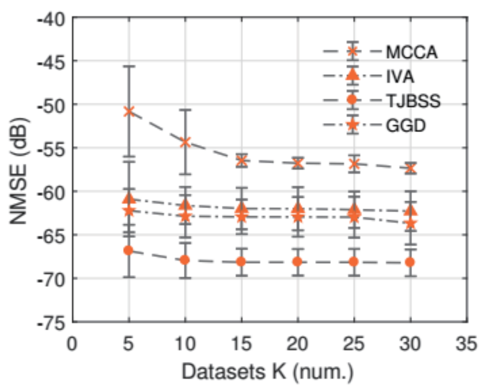
(a) NMSE of the estimated matrices and mixing matrices



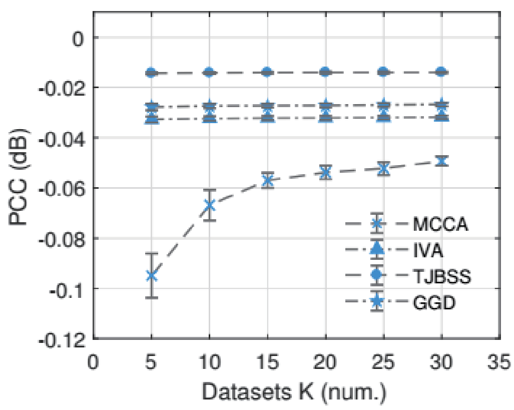
(b) PCC of the recovered sources and underlying sources

Fig. 2 The NMSE and PCC performances on the different τ

Then, we show the average performances in the different dataset K, where in Fig. 3 (a) measures the NMSE between the estimated matrices and the mixing matrices, and (b) evaluates the PPC between the estimated sources and the underlying sources. As we can seen, MCCA algorithm presents the poorest performance even K is increased. Although MCCA algorithm performs on the same criterion, i. e., second-order statistics that shows poorer performance mainly due to the error accumulation from deflation-based separation approach. In addition, the performances are not primarily dependent on number of datasets when N=10. For IVA, TJBSS, and GGD algorithms, the NMSE is slightly improved in the point of K=10. MCCA algorithm achieves reliable separation performance when the pa-



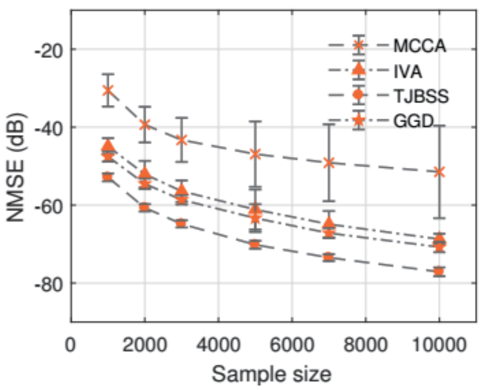
(a) NMSE of the estimated matrices and mixing matrices



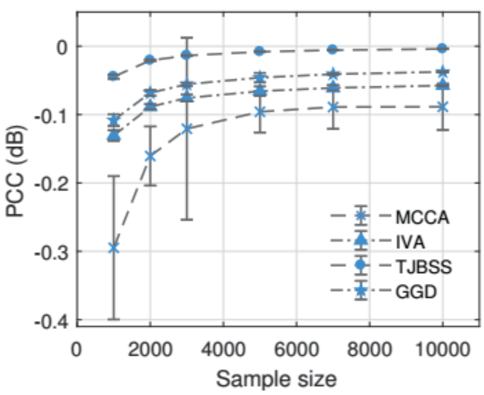
(b) PCC of the recovered sources and underlying sources

Fig. 3 The NMSE and PCC performance on the different K rameter K is greater than 15. A large enough parameter K can provide reliable separation even if the number of sources is large, e. x., we use N=10 here. The proposed TJBSS consistently presents more accuracy on NMSE and PCC, because the approach exploits not only the statistical independence within multiple datasets, but also dependence among the different datasets.

In general, the performances do not depend on the sample size, because the target dataset is given for consistency. However, the enough sample size are need to ensure that the covariance matrices used for the proposed algorithms are positive definite. It is sufficient in theoretical, however in some scenarios, a small number of samples can achieve the good performance. The averaged performances in the different sample sizes showed in Fig. 4, where (a) measures the NMSE between the estimated matrices and the mixing matrices, and (b) evaluates the PPC between the estimated sources and the underlying sources. It can be seen that all of the algorithms show similar and reasonable performance, in which the larger number of samples leads to the performance increase. The poor performance of MCCA is due to using gradient search algorithm may converge locally, as shown in the simulation results here and the speech separation results. GGD as extended approach from MCCA seems reliably solve this problem that achieves the better performance.



(a) NMSE of the estimated matrices and mixing



(b) PCC of the recovered sources and underlying sources

Fig. 4 The NMSE and PCC performance on the different sample size T

Separation of Real-world Biological Data

In this section, we demonstrate an application of the proposed algorithm, which the separation of electroencephalogram (EEG) and electrooculogram (EOG) signals from the real-world electrode recordings. The recordings are collected from six subjects. Fig. 5 shows the time scheme of the experimental paradigm in one trial. The subjects sit the chair in front of a computer screen. For each trial, the experiment begins at t=0 s when a fixation cross appears on the black screen. The prompt starts at the 2 seconds, i. e., t=2 s, the cue in the form of an arrow points appear on the screen during 1.25 s, either to the scenarios of left, right, up or down. The prompt allows subjects to perform the

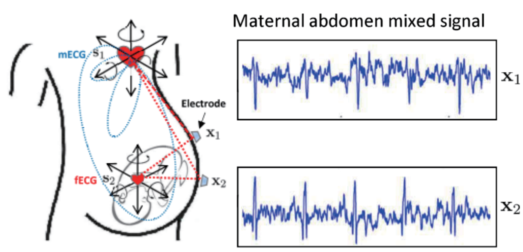


Fig. 5 The Mixture Models of the FECG and MECG Measurements

corresponding motor imagery task until t＝6 s. The next trial repeats after a 2 seconds break with a black screen. 360 trials for each motor imagery task are recorded from each subject. We exploit all the 22 channels available.

EEG and EOG signals extraction are a typical BSS problem, because of the collecting EEG signals are inevitably contaminated by some propagation inside the body, such as EOG signals and artifact. In practice, the distribution of channels is different due to the location of the different electrodes. Although channels are comparable dependent, it is difficult to identify which channel will more contribute to estimation in the different subjects. In this case, the joint BSS is considered an attractive approach.

[今後の研究の方向，課題]

Blind separation of source signals have received wide attention in various fields such as speech enhancement, image recognition, wireless communication, and thus have been thoroughly studied in the signal processing community. Among, it is a non-trivial task to obtain an accurate and reliable fetal electrocardiogram (FECG) in a non-invasive fashion. Problems develop due to the facts, that the electrocardiogram (ECG) also contains a maternal electrocardiogram (MECG), which can be from one-half to one-thousandth the magnitude of the MECG. Moreover, the FECG will occasionally overlap the MECG and make it normally impossible to detect. Along with the MECG, extensive electrocardiography (EMG) noise also interferes with the FECG and it can completely mask the FECG.

[成果の発表，論文等]

[1] L. Wang, T. Ohtsuki, K. Ohwada, N. Honma, H. Hayashi, “Joint Multiple Subspace-Based BSS Method for Fetal Heart Rate Extraction from Non-Invasive Recordings,” 42st International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2020), Montreal, Canada, Jul. 2020.

[2] L. Wang, T. Ohtsuki, K. Owada, N. Honma, and H. Hayashi, “Fetal Heart Rate Monitoring from Non-invasive ElectrodeRecordings,” 電子情報通信学会ソサイエティ大会， オンライン， 2020 年 9 月 16 日.

[3] A. A. Qureshi, L. Wang, T. Ohtsuki, K. Owada, N. Honma, and H. Hayashi, “Reconstructing Maternal ECG from Abdominal ECG Using Autoencoder,” 電子情報通信学会ソサイエティ大会，オンライン，2020 年 9 月 16 日.

[4] A. A. Qureshi, L. Wang, and T. Ohtsuki, K. Owada, N. Honma, and H. Hayashi, “Reconstructing Maternal ECG from Abdominal ECG Using Autoencoder,” 2020 International Conference on Emerging Technologies for Communications (ICETC 2020), Tokyo, Japan, Dec. 2020.

[5] 新飯田夏帆，大槻知明，王路，大和田一成，本間直樹，林隼人，“心拍波形の一次導関数を用いた胎児 R ピーク検出と RR 間隔の分布に応じた複数の重み付けによる胎児心拍数推定，” 電子情報通信学会総合大会，オンライン，2021 年 3 月 12 日.

[6] A. A. Qureshi, L. Wang, and T. Ohtsuki, K. Owada, N. Honma, and H. Hayashi, “Improved Fetal Heart Rate Detection Via Maternal ECG Removal Using Autoencoder in Frequency Domain,” 電子情報通信学会総合大会， オンライン，2021 年 3 月 12 日.

[7] L. Wang, T. Ohtsuki, K. Ohwada, N. Honma, H. Hayashi, “A Multi-Layer Hybrid Network with Its Application in Fetal Heart Rate Monitoring,” IEEE Signal Processing Letters. (Under Review)

[8] L. Wang, T. Ohtsuki, K. Ohwada, N. Honma, H. Hayashi, “Improved Fetal Heart Rate Detection by Removing Maternal Cycle Using Adaptive Template,” IEEE Global Communications Conference. (Under Review)

[9] A. A. Qureshi, L. Wang, and T. Ohtsuki, K. Owada, N. Honma, and H. Hayashi, “An Autoencoder-Based Heart Rate Detector or Non-invasive Recordings,” International Conference of the IEEE Engineering in Medicine and Biology Society. (Under Review)

[10] N. Niida, L. Wang, and T. Ohtsuki, K. Owada, N. Honma, and H. Hayashi, “Fetal Heart Rate Detection UsingFirst Derivative of ECG Waveformand Multiple Weighting Functions,” International Conference of the IEEEEngineering in Medicine and Biology Society.（Under Review）