Removing Noise from Event-Related Potentials using a Probabilistic Generative Model with Grouped Covariance Matrices

Hayato Maki[1], Tomoki Toda[2,3], Sakriani Sakti[2], Graham Neubig[2], and Satoshi Nakamura[2]

Abstract— **Analysis of electroencephalograms (EEG) usually suffers from a variety of noises. In this paper, we propose a new method for background noise removal from single-trial event-related potentials (ERPs) recorded with a multi-channel EEG. An observed signal is separated into multiple signals with a multi-channel Wiener filter, whose coefficients are estimated based on a probabilistic generative model in the time-frequency domain. The main contribution is a method to estimate covariance matrices for each frequency bins of short-time Fourier transform (STFT) representing different spatial spread of a multi-channel EEG signal according to frequencies. An experiment using a pseudo-ERP data set demonstrates the effectiveness of our proposed method.**

I. INTRODUCTION

Event-related potentials (ERPs) are a brain response to a specific stimulus recorded by electroencephalograms (EEGs). ERPs are usually averaged in time domain to weaken task-unrelated background activities. However, a number of variabilities exist across trials exist in latency, amplitude, and scalp distribution [1]. This across-trial averaging procedure conceals all the information concerning across-trial variability of brain response. Moreover, to perform reliable analysis, at least 20 trials for each experimental condition are generally averaged, resulting in long experimental time and subject fatigue. Therefore, methods that allow for single-trial analysis without averaging are an important research topic with the potential of making it possible to study event-related brain dynamics more precisely [2]. Throughout this paper, we refer to the ERPs as the *signal* to preserve and task-unrelated neural activities as well as non-neural artifacts as *noise* to be removed.

Within the paradigm of EEG signal processing, there is a large amount of research regarding external artifact removal, especially eye blink removal [3]. However, removal of background EEGs is less studied, although there is some research proposing methods based on independent component analysis (ICA) [4], [5].

In this paper, we propose a new approach to remove background activities from single-trial ERPs based on an approach that has been proposed originally in the field of sound source separation [6] and further applied to EEG signals [7]. In this method, model parameters are estimated by the maximum likelihood criterion without prior knowledge of target signal. We have extended this approach, exploiting prior knowledge of the target signal to further improve the effectiveness of model parameter estimation[8]. This approach has a virtue that, in contrast to ICA, it doesn't assume the number of sources because it does not estimate sources, but contributions, of each event signal to each EEG electrode. While ICA aims to separate an observed signal into its sources, our method separates it into signals elicited by each event such as ERP or an eye blinks.

We denote the group of source indexes that contributes to the k-th event by E_k and denote the EEG signal that is elicited by E_k and captured at the scalp by $\mathbf{c}_k(t)$ and call it k-th event signal. In the time-frequency domain, short-time Fourier transform (STFT) coefficients of the event signal $\mathbf{c}_k(n,f)$ are locally modeled by a multivariate complex Gaussian distribution whose parameters are a function of (n,f) where n is the index of the time frame and f is the index of the frequency bin. A multi-channel Wiener filter is constructed using estimated model parameters to separate an observed EEG signal into some event signal.

Model parameters includes full rank time-invariant covariance matrices called *spatial correlation matrices* encoding the spatial spread of the event signals. In the previous works [7], [8], spatial correlation matrices were assumed independent from frequency. In this paper, we propose a generative model with frequency dependent spatial correlation matrices because EEG signals usually have different spatial spread according to frequencies.

II. EXISTING TECHNIQUES FOR SIGNAL ENHANCEMENT

1) Observation Model: The STFT coefficients of the k-th event signal in time-frequency slot (n,f) are expressed as $\mathbf{c}_k(n,f)$:

$$\mathbf{c}_k(n,f) = \sum_{l \in E_k} \mathbf{h}_l s_l(n,f) = [c_{k,1}(n,f), \cdots, c_{k,L}(n,f)]^\top \quad (1)$$

where $\mathbf{h}_l$ is the transfer function vector and $s_l(n,f)$ is the l-th source activity in the slot (n,f) . The observed multi-channel EEG signal $\mathbf{x}(n,f)$ is expressed as

$$\mathbf{x}(n,f) = [x_1(n,f), \cdots, x_L(n,f)]^\top = \sum_{k=1}^{K} \mathbf{c}_k(n,f), \quad (2)$$

where K is a hyper parameter for the total number of events.

2) Generative Model: To model the generation of observed signals probabilistically, we make the following two assumptions:

Assumption1

The amplitude of the l-th source that contributes to the k-th event in each slot (n,f) follows a complex normal distribution with mean 0. The variance is given by the product of the degree of the event activity $v_k(n,f)$ and the source activity λ_l:

$$\begin{aligned} p(s_l(n,f)) &= \mathcal{N}_c(s_l(n,f); 0, v_k(n,f)\lambda_l) \\ &= \frac{1}{\pi v_k(n,f)\lambda_l} \exp\left(\frac{|s_l(n,f)|^2}{v_k(n,f)\lambda_l}\right), \end{aligned} \quad (3)$$

where $l \in E_k$ and the group E_k contains source indexes associated with the k-th event signal.

Assumption2

Sources are not correlated with each other within each

*Part of this work was supported by JSPS KAKENHI Grant Number 26540117.

*The Ethical Review Board of Nara Institute of Science and Technology (NAIST) approved all experimental procedures involving human subjects.

[1]Hayato Maki is with Graduate School of Information Science, NAIST, Japan maki.hayato.lt3@is.naist.jp

[2]Tomoki Toda, Sakriani Sakti, Graham Neubig and Satoshi Nakamura are Faculty of Graduate School of Information Science, NAIST, Japan

[3]Tomoki Toda is also Faculty of the Information Technology Center, Nagoya University, Japan

time-frequency slot (n,f):

$$E[s_{l_1}(n,f)s^*_{l_2}(n,f)] = 0 \text{ for } l_1 \neq l_2, \tag{4}$$

where $l_1, l_2 \in E_k$ and $\{\cdot\}^*$ indicates the complex conjugate of $s_{l_2}(n,f)$.

From the assumptions above, the probability density function that each event signal follows is the multivariate complex normal distribution with zero mean,

$$\begin{aligned} p(\mathbf{c}_k(n,f)) &= \mathcal{N}_c\big(\mathbf{c}_k(n,f); \mathbf{0}, \mathbf{R}_{c_k}(n,f)\big) \\ &= \frac{\exp\big(-\mathbf{c}_k(n,f)^H \mathbf{R}_{c_k}^{-1}(n,f)\mathbf{c}_k(n,f)\big)}{\pi^I \det(\mathbf{R}_{c_k}(n,f))}. \end{aligned} \tag{5}$$

For simplicity, we restrict the covariance matrix $\mathbf{R}_{c_k}(n,f)$ to the product of the time-invariant covariance matrix $\mathbf{R}_k$ that encodes spatial spread of the k-th event signal, and the time-frequency variant scalar $v_k(n,f)$ that encodes time-frequency power of the signal:

$$\begin{aligned} \mathbf{R}_{c_k}(n,f) &= E\left[\mathbf{c}_k(n,f)\mathbf{c}_k(n,f)^H\right] \\ &= E\left[\sum_{l_1 \in E_k} \mathbf{h}_{l_1} s_{l_1}(n,f) \left(\sum_{l_2 \in E_k} \mathbf{h}_{l_2} s_{l_2}(n,f)\right)^H\right] \\ &= \sum_{l \in E_k} E[|s_l(n,f)|^2] \mathbf{h}_l \mathbf{h}_l^\top \\ &= \sum_{l \in E_k} v_k(n,f)\lambda_l \mathbf{h}_l \mathbf{h}_l^\top = v_k(n,f)\mathbf{R}_k, \end{aligned} \tag{6}$$

where $\{\cdot\}^H$ is the complex conjugate transpose. We call $\mathbf{R}_k$ the *spatial correlation matrix* of the k-th event signal.

To simplify the generative model, we assume that the events in each time frequency slot are sparse. In other words, we assume that we observe only one event signal in each slot. In order to express this sparseness mathematically, we introduce the latent variables $z_k(n,f)$, which take a 1-of-K representation, in which one particular element $z_k(n,f)$ is equal to 1 and all other elements are equal to 0. If we observe the l-th event signal in the slot (n,f), we can express the observed signal assuming $z_l(n,f) = 1$ as follows:

$$\mathbf{x}(n,f) = \sum_{k=1}^{K} z_k(n,f)\mathbf{c}_k(n,f) = \mathbf{c}_l(n,f), \tag{7}$$

From all assumptions above, the likelihood of the observed signals $\mathbf{x}$ in the time-frequency slot (n,f) is expressed by a Gaussian mixture model as follows:

$$p(\mathbf{x}(n,f)|\theta) = \sum_{k=1}^{K} \alpha_k \mathcal{N}_c\left(\mathbf{x}(n,f); \mathbf{0}, v_k(n,f)\mathbf{R}_k\right) \tag{8}$$

where α_k is the prior probability that the k-th event signal activates. The model parameter set θ consists of α_k, $v_k(n,f)$, and $\mathbf{R}_k$ of each mixture component.

The optimal model parameter set θ is estimated by maximizing the observation likelihood described in eq. (8). This maximization process can be effectively solved with the EM algorithm.

In the E-step, the posterior probability $m_k(n,f)$ is calculated at each time-frequency slot:

$$m_k(n,f) = \frac{\alpha_k \mathcal{N}_c\left(\mathbf{x}(n,f); \mathbf{0}, v_k(n,f)\mathbf{R}_k\right)}{\sum_{k'=1}^{K} \alpha_{k'} \mathcal{N}_c\left(\mathbf{x}(n,f); \mathbf{0}, v_{k'}(n,f)\mathbf{R}_{k'}\right)}. \tag{9}$$

In the M-step, $\hat{\alpha}_k$ and $\hat{v}_k(n,f)$ are updated as follows:

$$\hat{\alpha}_k = \frac{\sum_{n,f} m_k(n,f)}{\sum_{n,f,k'} m_{k'}(n,f)}, \tag{10}$$

$$\hat{v}_k(n,f) = \frac{1}{L}\mathbf{x}(n,f)^H \mathbf{R}_k^{-1} \mathbf{x}(n,f), \tag{11}$$

$$\hat{\mathbf{R}}_k = \frac{\sum_{n,f} \frac{m_k(n,f)}{\hat{v}_k(n,f)} \mathbf{x}(n,f)\mathbf{x}(n,f)^H}{\sum_{n,f} m_k(n,f)}, \tag{12}$$

where $\hat{v}_k(n,f)$ and $\hat{\mathbf{R}}_k(f)$ are iteratively updated because they depend on each other.

Finally, the target event signal is extracted from the observed EEG signals using a multi-channel Wiener filter as follows:

$$\hat{\mathbf{c}}_k(n,f) = m_k(n,f)\hat{v}_k(n,f)\hat{\mathbf{R}}_k \mathbf{R}_{\mathbf{x}}^{-1}(n,f)\mathbf{x}(n,f). \tag{13}$$

Utilizing Spatial Correlation Prior and MAP Estimaition

We often know which event signal we would like to enhance, such as ERP, and we can record EEG signals related to the target event beforehand and use them as prior knowledge for enhancement. In such a case, we need to enhance the target event signal from the observed EEG signal, rather than blindly separating it into multiple EEG event signals described in the previous section. Hence, we have proposed utilizing prior information of event signals to estimate spatial correlation matrices [8] utilizing Wishart distributions, whose probability density function is given by

$$\mathcal{W}_p(\mathbf{R}|\mathbf{W}, q) = B(\mathbf{W}, p, q)|\mathbf{R}|^{\frac{q-p-1}{2}} \exp\left(-\frac{\mathrm{Tr}(\mathbf{W}^{-1}\mathbf{R})}{2}\right) \tag{14}$$

where $\mathbf{W}$ is a $p \times p$ positive definite matrix, q is called the *number of degrees of freedom* and is restricted to $q \geq p$, and $B(\mathbf{W}, p, q)$ is a normalizing factor. The Wishart distribution is the conjugate prior distribution of a p-dimensional multivariate Gaussian random variable's precision matrix with a known mean vector.

From Bayes' theorem, the posterior probability density function of the time-frequency invariant spatial covariance matrices $\mathbf{R} = \{\mathbf{R}_1, \cdots, \mathbf{R}_K\}$ is proportional to the product of the likelihood of the observed data and their prior as follows:

$$p(\mathbf{R}|\mathbf{x}, \theta_{\backslash \mathbf{R}}, \Psi_k, q_k) \propto \prod_{n,f} p\left(\mathbf{x}(n,f)|\theta\right) \prod_{k=1}^{K} \mathcal{W}_L\left(\mathbf{R}_k^{-1}|\Psi_k^{-1}, q_k\right) \tag{15}$$

where $\theta_{\backslash \mathbf{R}}$ is the model parameter set except for $\mathbf{R}$.

The optimal model parameter set θ is estimated by maximizing the posterior probability described in eq. (15) instead of observation likelihood described in eq. (8). This maximization process can also be effectively solved with the EM algorithm.

In the E-step, the posterior probability $m_k(n,f)$ is calculated at each time-frequency slot as shown in eq. (10). In the M-step, $\hat{\alpha}_k$ and $\hat{v}_k(n,f)$ are updated as shown in eq. (10) and eq. (11). The spatial correlation matrices are updated as follows:

$$\hat{\mathbf{R}}_k = \frac{\left(\Psi_k + 2\sum_{n,f} \frac{m_k(n,f)}{\hat{v}_k(n,f)} \mathbf{x}(n,f)\mathbf{x}(n,f)^H\right)}{(q_k - L - 1) + 2\sum_{n,f} m_k(n,f)}. \tag{16}$$

III. Proposed Method

A. Frequency Dependent Spatial Correlation Matrix

Spatial covariance matrices, which represent the spatial spread of multi-channel EEG signals in the generative model of the existing work, are independent from frequency. However, EEG signals usually have different spatial spread over the scalp according to the frequency, as shown in Fig. 2, so a probabilistic model considering

the frequency difference has the potential to be more effective. In this paper, we propose a generative model with frequency dependent spatial correlation matrices $\mathbf{R}_k(f)$ instead of frequency independent matrix $\mathbf{R}_k$ of the previous works. In this manner, the estimation of spatial covariance matrices in the EM algorithm is done as follows:

$$\hat{\mathbf{R}}_k(f) = \frac{\left(\Psi_k(f) + 2\sum_n \frac{m_k(n,f)}{\hat{v}_k(n,f)} \mathbf{x}(n,f)\mathbf{x}(n,f)^H\right)}{(q_k(f) - L - 1) + 2\sum_n m_k(n,f)}. \quad (17)$$

Note that the range of summation is over only time frame n.

B. Grouped Spatial Correlation Matrix

In the frequency dependent covariance model described in the previous section, the number of samples $\mathbf{x}(n,f)$ to estimate a covariance matrix is reduced to $1/F$ compared to the conventional method described in Section II, as the samples that belong to only one frequency bin are involved in estimating the corresponding covariance matrix as shown in eq. (17). This has an adverse effect on robust covariance matrix estimation although it can effectively represent a frequency dependent spatial spread of a multi-channel EEG signal, while all samples $\mathbf{x}(n,f)$ in the time-frequency grid are involved in the conventional method. Therefore, there is a tradeoff between frequency specific covariance matrix modeling and robust covariance matrix estimation. To solve this problem, we group covariance matrices from multiple frequency bins together to share their statistics. The difference of the normalized $\ell 1$ norm of covariance matrices for the adjacent frequencies of a multi-channel EEG signal are shown in Fig. 2 to show how much covariance matrices of frequencies differ. We can see that covariance matrices corresponding to higher frequencies, except for the Nyquist frequency have a relatively small difference between adjacent matrices. Based on this, we estimate one covariance matrix corresponding to higher frequencies while individually estimating covariance matrices corresponding to lower frequencies. In this manner, the frequency bins are grouped as lower and higher groups as:

$$\begin{aligned}\{f_1, \cdots, f_F\} &= \{\{f_1, \cdots, f_g\}, \{f_{g+1}, \cdots, f_F\}\} \\ &= \{g_{\text{low}}, g_{\text{high}}\}\end{aligned} \quad (18)$$

where f_g is an index of a frequency bin. The estimation of spatial covariance matrices in the EM algorithm is done as follows:

$$\hat{\mathbf{R}}_k(f) = \frac{\left(\Psi_k(f) + 2\sum_n \sum_{f \in g_{\text{high}}} \frac{m_k(n,f)}{\hat{v}_k(n,f)} \mathbf{x}(n,f)\mathbf{x}(n,f)^H\right)}{(q_k(f) - L - 1) + 2\sum_n \sum_{f \in g_{\text{high}}} m_k(n,f)} \quad (19)$$

if $f \in g_{\text{high}}$, otherwise $\mathbf{R}_k(f)$ is updated by eq. (17).

IV. Experimental Evaluation

A. Experimental Paradigm

EEG data were collected from a healthy male volunteer aged 24 years without any neurological disorders. All EEG signals were recorded from 20 scalp electrodes at locations based on a modified International 10-20 system [9], digitally sampled at 250 Hz.

We conducted three sessions. In the first and second sessions, we conducted auditory oddball paradigm experiments [10] using 1kHz and 2kHz sound. A random sequence of auditory stimuli including 2kHz and 1kHz sine waves was presented to the subject. 1kHz sounds were presented 200 times as non-target stimuli and 2kHz sine wave were presented 50 times as target stimuli for each session. The subjects were told to count the number of target stimuli without any body movement. It is widely known that the ERP component of P300 will be elicited when a subject hears target stimuli. In the third session, he was told to relax without any body movement to record resting state EEG for two minutes.

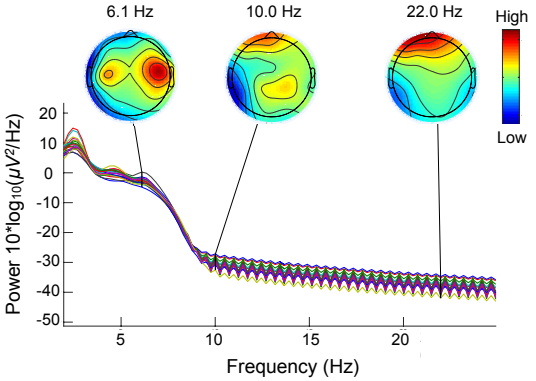

Fig. 1. Scalp topographies within the different frequency bins. Each line stands for an EEG channel.

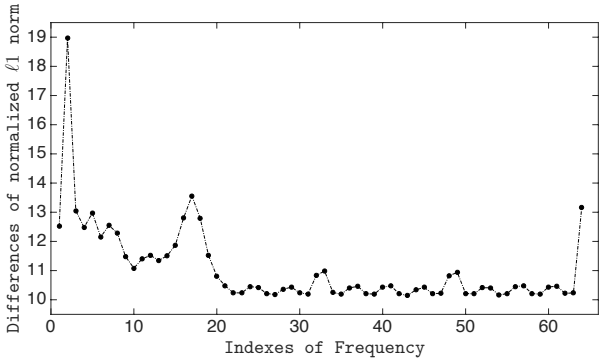

Fig. 2. The differences of normalized ℓ 1 norm of adjacent covariance matrices. The last point stands for the Nyquist frequency.

B. Creation of Validation Dataset

We created a pseudo-ERP dataset in order to evaluate noise removal methods objectively. First, we cut the multi-channel EEG data obtained from session 2 for each type of stimulus into trials of 1.2 second length, then averaged them across the trials for each type of stimulus. Next, we cut the resting state EEG obtained from session 3 into 40 signals of 1.2 second length, and then we added the averaged ERP of each corresponding channel that stands for *true* ERP to the 20 resting state trials for each type of stimuli that stand for *noise* with randomly shifted phase from -60ms to 60ms.

C. Evaluation

We compared three denoising methods, ICA [4] and two time-frequency modeling methods that have spatial correlation matrices dependent on individual frequencies and on grouped frequencies where a covariance matrix corresponding to frequencies higher than 25th frequency bins (about 48 Hz) is shared. All twenty trials for each type of stimuli are concatenated as one signal that were denoised by each compared method, and then re-epoched. An ERP signal in time domain $\mathbf{x}(t)$ is often quantified by its peak amplitude $\mathbf{x}(t_{peak})$ and latency t_{peak}, the former is its maximum value of amplitude within the given time window and the latter is the elapsed time from the onset of stimuli to reach the peak amplitude. We quantified the P300 components with the time window from 200ms to 500ms from the onset of stimuli. After quantification of the each single-trial ERP (*true* ERP + noise) and *true* ERPs or averaged

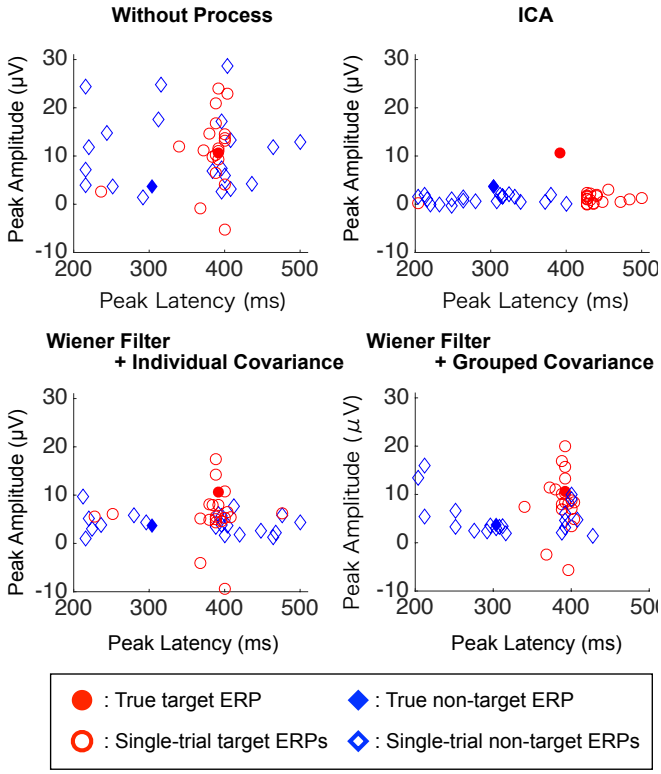

Fig. 3. Plots of peak amplitudes and latencies. Red circles stands for single-trial ERPs of target trials and blue triangles for non-target. Two filled points stands for *true* ERPs for each type of stimuli.

ERP signals of each type of stimuli obtained from session 2 and added to the resting state EEGs, we calculated amplitude (AD) and latency deviation (LD) for each single-trial from the *true* ERPs as follows:

$$\mathrm{AD}(j) = \bar{\mathbf{x}}(\bar{t}_{peak}) - \hat{\mathbf{x}}^{(j)}(t_{peak}) \tag{20}$$

$$\mathrm{LD}(j) = \bar{t}_{peak} - \hat{t}^{(j)}_{peak} \tag{21}$$

where $\bar{\mathbf{x}}$ is the 'true' ERP and $\hat{\mathbf{x}}$ is the j-th single-trial ERP. Using these deviation measures, we calculated Mahalanobis distance from each single-trial point to two *true* points to classify each single-trial point to the nearer *true* point, and then counted the number of single-trials that were correctly classified.

D. Parameter Settings

We decomposed each raw single-trial multi-channel EEG signal into two event signals with time-frequency modeling methods. Hyper parameters of the prior distribution for the first event signal was calculated from the signal that were given by the session 1. Hyper parameters for the second event signal were calculated from the resting state EEG signal recorded in the session 3.

V. Results

The peak amplitudes and latencies of the true ERP and of the each single-trial are plotted in Fig. 3. When single-trial ERPs are without processing, the points in the figure are interspersed (upper left). By the denoising effect of proposed methods (lower), the scattered points gathered around their *true* point for each type of stimuli.

The ratio of single-trials correctly classified is shown in Fig. 4. When single-trial ERPs are without processing, only 60% of them

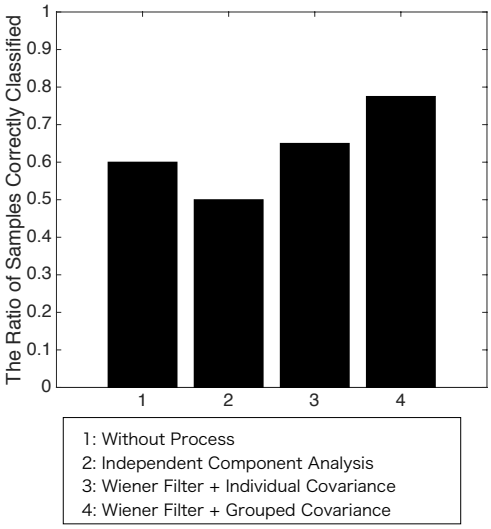

Fig. 4. The ratio of forty single-trials that are correctly classified.

are correctly classified. After denoising by the proposed method (Wiener filter + Grouped Covariance) 82.5% are correctly classified.

VI. Conclusion

We developed a new noise removal method from ERP data extending a conventional method that models generation of each event signal in the time frequency domain with frequency dependent covariance matrices with prior information of ERP. Both of the proposed methods, especially with the grouped covariance model showed effectiveness in removing background EEG signals from single-trial ERPs.

We used EEG data to obtain prior information of each event signal recorded in the same day and the same subject. Ideally, prior information collected from other days or subjects should be used for a ready-to-use system in the future work.

References

[1] M. Kutas, G. McCarthy, and E. Donchin. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science*, Vol. 197, No. 4305, pp. 792–795, 1977.

[2] T-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J Sejnowski. Analysis and visualization of single-trial event-related potentials. *Human brain mapping*, Vol. 14, No. 3, pp. 166–185, 2001.

[3] J. A. Urigüen and B. Garcia-Zapirain. EEG artifact removal - State-of-the-art and guidelines. *Journal of Neural Engineering*, 2015.

[4] D. M Groppe, S. Makeig, and M. Kutas. Independent component analysis of event-related potentials. *Cognitive science online*, Vol. 53, No. 4, pp. 601–607, 2008.

[5] S. Lemm, G. Curio, Y. Hlushchuk, and K. R. Müller. Enhancing the signal-to-noise ratio of ICA-based extracted ERPs. *IEEE Transactions on Biomedical Engineering*, Vol. 53, No. 4, pp. 601–607, 2006.

[6] N. QK Duong, E. Vincent, and R. Gribonval. Under-determined reverberant audio source separation using a full-rank spatial covariance model. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 18, No. 7, pp. 1830–1840, 2010.

[7] Y. Kurihana, S. Miyabe, T. M. Rutkowski, Y. Matsumoto, T. Yamada, and S. Makino. (in Japanese) Signal Separation of EEG using Multivariate Probabilistic Model. *Technical Report of IEICE, MBE*, Vol. 112, No. 479, pp. 161–166, 2013.

[8] H. Maki, T. Toda, S. Sakti, G. Neubig, and S. Nakamura. EEG signal enhancement using multi-channel wiener filter with a spatial correlation prior. *Proc. IEEE ICASSP*, 2015.

[9] American Clinical Neurophysiology Society. Guidelines for standard electrode position nomenclature. *Clinical Neurophysiology*, 1994.

[10] S. J Luck. *An introduction to the event-related potential technique*. MIT press, 2014.