Cortical Patterns for Prediction of Subjective Preference Induced by Chords*

Hajime Yano[1], Tetsuya Takiguchi[2] and Seiji Nakagawa[3]

Abstract— **To extract an effective feature in prediction of subjective impressions from single-trial neurophysiological recordings, the spatial filter that extracts brain activities related to impressions were constructed using the common spatial pattern (CSP). We focus on subjective preference induced by chords composed of 3 notes with different frequency ratio. Magnetic cortical activities while hearing chords and comparative judgment on pair of them were measured. The predictive model that predicts the scale value of preference was trained using the CSP-based feature for each participant. The result of the evaluation experiment shows that the CSP-based feature improved the mean prediction accuracy in all participants, compared with the other features without spatially filtering. Furthermore, the capability of construction of a spatial filter that extracts cortical activities varying with degree of preference using the comparative judgments was indicated.**

I. Introduction

Sensory assessment of sound is essential to improve the comfortableness of sound environment and the quality of the product. While the psychological methods were conventionally used for assessment of sound, the neurophysiological methods are promising because they are potentially less influenced by the cognitive bias, and they can assess even the subconscious impressions. To create a practical neurophysiological index, the relations between sensory scales indicating subjective impressions and brain cortical activities have been investigated in the past [1], [2], [3]. Although these studies indicated the possibility of assessment of auditory impressions, the neurophysiological indices were not sufficiently robust for practical use: the found correlations were small, the found relations were complex, or they were not consistent in any kinds of sounds. Most of the relations between brain activities and subjective impressions are still unclear.

On the other hand, creating a neurophysiological index for assessment of impressions is regarded as finding a mapping from the space of brain activities to sensory scales. Using the machine learning techniques, complex relations between brain activities and subjective impressions can be learned. Additionally, the prediction of the scale value from a single-trial recording is desirable for a practical index.

*This work was supported in part by JSPS KAKENHI (Grant No. JP18K19820).

[1]Hajime Yano is with Graduate School of System Informatics, Kobe University, Kobe, Hyogo, Japan (e-mail: h-yano@stu.kobe-u.ac.jp), and with National Institute of Advanced Industrial Science and Technology (AIST), Ikeda, Osaka, Japan.

[2]Tetsuya Takiguchi is with Kobe University, Kobe, Hyogo, Japan (email: takigu@kobe-u.ac.jp).

[3]Seiji Nakagawa is with Center for Frontier Medical Engineering, Chiba University, Chiba, Japan (e-mail: s.nakagawa@chiba-u.ac.jp).

The common spatial pattern (CSP) is well-known as a successful method to extract neural activities related to the motor imagery from the single-trial electroencephalography (EEG) in application of brain-computer interface [4]. The spatial filter obtained by the CSP can enhance/attenuate the activities under a condition while attenuating/enhancing the activities under the other condition. The signal-to-noise ratio of the single-trial recordings under assessment of subjective impressions can be improved using the CSP.

In this paper, we focus on subjective preference induced by chords composed of 3 notes with different frequency ratio. It is considered that consonance and dissonance of the chords affect subjective preference of them. First, magnetoencephalography (MEG) while hearing chords and the comparative judgments on pair of them were measured. Owing to MEG that has high accuracy of source localization, accurate spatial patterns can be estimated. Paired-comparison enables a participant to judge the preference relation of stimuli more easily and accurately. Second, the CSP-based spatial filter that extracts brain activities varying with degree of subjective preference was constructed using the preference relation of stimuli. Third, the model that predicts the scale value of preference from brain activities was trained for each participant and its performance was evaluated.

II. MEG Recordings

A. Stimuli

Seven kinds of chords composed of 3 notes, i.e. triad, were generated as stimuli, using the methods in previous studies on interaction between degree of consonance and brain activities [5], [6]. These chords had different frequency ratios: 2:3:4, 4:5:6, 6:7:8, 8:9:10, 10:11:12, 12:13:14, and 14:15:16. Each note consists of 6 harmonics with equal amplitude. The lowest frequency of each chord was fixed at 220 Hz. The duration of stimuli was 200 ms including 20 ms rise and fall time. The sound pressure level of the stimuli was 73.0 ± 0.5 dB SPL.

B. MEG Measurements

Five males (mean age $\pm$SD: 30.5$\pm$9.05) participated in the MEG measurements. They had normal hearing and no history of neurological diseases. Informed consent was obtained from each participant after explanation of the experiment. The experiment was approved by the Institutional Review Board on Ergonomic Research of AIST.

The measurements were performed in a magnetically shielded room using a 122-channel whole-head neuromagnetometer (Neuromag-122$^{\mathrm{TM}}$Neuromag Ltd.). In each trial, paired different stimuli were presented to participant's both ears sequentially with the inter-stimulus interval of 1000 ms, using an insertion-type earphone. After presentation of the latter stimulus, a participant judged which of stimuli he/she felt preferred using two buttons corresponding to judgment. All permutations of 2 different stimuli were presented randomly and repetitively until each stimulus was presented at least 100 times. Magnetic signals were passed through the analog filter with a passband from 0.03 to 100 Hz and sampled at 400 Hz.

III. Feature Extraction

A. Data Preprocessing

Artifacts and unnecessary components were removed from measured raw data. Measured signals on some bad channels of the magnetometer were ignored in subsequent analyses. Low frequency components less than 2 Hz, and power line noise of 60 Hz and its harmonic of 120 Hz were cut off using the 4th-order and 2nd-order zero-phase Butterworth filters. Subsequently, trials including large absolute amplitudes more than 1000 fT/cm were removed as artifact. The independent component analysis (ICA) was applied to large-artifact-free trials to remove remaining artifacts and magnetic response of cardiac activities and eye movements. ICA algorithm used in this phase was FastICA [7].

Seven kinds of oscillatory activities: theta (4–8 Hz), alpha (8–13 Hz), low-beta (13–20 Hz), high-beta (20–30 Hz), low-gamma (30–50 Hz), mid-gamma (50–70 Hz), and high-gamma (70–100 Hz), were obtained from denoised trials using the 8th-order zero-phase Butterworth filter.

B. Common Spatial Pattern

The CSP is a successfully used technique to design optimal filter for discrimination between EEG signals in two conditions [4], [8]. A band-passed multichannel signal of a single trial is denoted as $\boldsymbol{E} \in \mathbb{R}^{N \times T}$, where N is the number of channels and T is the number of temporal samples. The estimate of the covariance matrix under the condition $c \in \{1, 2\}$ is denoted by $C_c \in \mathbb{R}^{N \times N}$, and it was computed by following equation:

$$\boldsymbol{C}_c = \frac{1}{|\mathcal{I}_c|} \sum_{i \in \mathcal{I}_c} \frac{\boldsymbol{E}_i \boldsymbol{E}_i^\top}{\mathrm{tr}(\boldsymbol{E}_i \boldsymbol{E}_i^\top)} \tag{1}$$

where $\mathcal{I}_c$ is the set of indices labeled as the condition c, $^\top$ is the transpose operator, and $\mathrm{tr}(\cdot)$ is sum of the diagonal elements in a square matrix. The spatial filter of the CSP is obtained by solving the following generalized eigenvalue problem [4]:

$$\boldsymbol{C}_1 \boldsymbol{w} = \lambda(\boldsymbol{C}_1 + \boldsymbol{C}_2)\boldsymbol{w}. \tag{2}$$

The generalized eigenvector corresponding to the largest generalized eigenvalue maximizes the variance under the condition 1 and minimizes the variance under the condition 2, and vice versa.

In this study, since independent components corresponding to noise were removed, the total covariance matrix, $\boldsymbol{C}_{total} = \boldsymbol{C}_1 + \boldsymbol{C}_2$, was rank deficient. Thus, dimensionality of signal (the number of channels) was reduced using the eigenvalue decomposition, $\boldsymbol{C}_{total} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\top$, then we obtained spatial filter of CSP using the whitening and simultaneous diagonalization [8]. Let $\tilde{\boldsymbol{D}} \in \mathbb{R}^{R \times R}$ be the diagonal matrix whose diagonal elements have R $(< N)$ nonzero eigenvalues of $\boldsymbol{C}_{total}$, and let $\tilde{\boldsymbol{U}} \in \mathbb{R}^{N \times R}$ be the basis matrix composed of eigenvectors corresponding to nonzero eigenvalues arranged in $\tilde{\boldsymbol{D}}$. The transformation for dimensionally reduction and whitening is given by $\boldsymbol{P} = \tilde{\boldsymbol{U}}\tilde{\boldsymbol{D}}^{-1/2}$ where $\tilde{\boldsymbol{D}}^{-1/2}$ is the diagonal matrix such that $\tilde{\boldsymbol{D}}^{-1} = \tilde{\boldsymbol{D}}^{-1/2}\tilde{\boldsymbol{D}}^{-1/2}$. The matrix $\boldsymbol{P}$ transforms $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ into $\boldsymbol{S}_1 = \boldsymbol{P}^\top \boldsymbol{C}_1 \boldsymbol{P}$ and $\boldsymbol{S}_2 = \boldsymbol{P}^\top \boldsymbol{C}_2 \boldsymbol{P}$, respectively. $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are simultaneously diagonalized by the orthonormal matrix $\boldsymbol{V} \in \mathbb{R}^{R \times R}$:

$$\boldsymbol{V}^\top \boldsymbol{S}_1 \boldsymbol{V} = \boldsymbol{\Lambda}_1, \qquad \boldsymbol{V}^\top \boldsymbol{S}_2 \boldsymbol{V} = \boldsymbol{\Lambda}_2, \tag{3}$$

where $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ are diagonal matrices whose diagonal elements are eigenvalues of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$, respectively. The sum of the corresponding eigenvalues of $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ is always equal to 1, i.e. $\boldsymbol{\Lambda}_1 + \boldsymbol{\Lambda}_2 = \boldsymbol{I}$. Eventually, filtered signal $\boldsymbol{E}_{\mathrm{CSP}} \in \mathbb{R}^{R \times T}$ is given by

$$\boldsymbol{E}_{\mathrm{CSP}} = \boldsymbol{V}^\top \boldsymbol{P}^\top \boldsymbol{E} = \boldsymbol{W}^\top \boldsymbol{E} \tag{4}$$

where $\boldsymbol{W} = \boldsymbol{P}\boldsymbol{V}$. Each column vector of $\boldsymbol{W}$ and corresponding eigenvalue of $\boldsymbol{S}_1$ are solution of (2).

We constructed a spatial filter that discriminates between activities just after hearing more preferred stimulus (condition 1) and activities just after hearing less preferred stimulus (condition 2) using the CSP. This filter is expected to extract cortical activities increasing/decreasing with degree of subjective preference. The bandpass-filtered MEG during 200–1000 ms after stimulus onset was used as $\boldsymbol{E}$. Two generalized eigenvectors corresponding to the largest and the smallest generalized eigenvalues were chosen for spatial filtering. The square root of variance of spatially filtered signal was computed for each frequency band. The number of dimensions of the feature vector of the CSP became 14 in total.

IV. Impression Prediction

A. Predictive Model

Let f be the mapping from the feature space of brain activities into the scale of subjective preference. We call f the predictive model in this paper. The predictive model f was trained from pairs of features extracted from brain activities just after hearing pairs of stimuli and comparative judgment. We assume that a comparative judgment $y \in \{-1, 1\}$ between sequentially presented stimuli A and B is determined by a magnitude relation between $f(\boldsymbol{x}^A)$ and $f(\boldsymbol{x}^B)$, where $\boldsymbol{x}^A$ and $\boldsymbol{x}^B$ is feature vectors when A and B are presented, and that the mapping f is linear: $f(\boldsymbol{x}) = \boldsymbol{a}^\top \boldsymbol{x}$ where $\boldsymbol{a}$ denotes the linear weight. The binary variable y takes 1 if stimulus A is preferred to stimulus B, and vice versa. The predicted comparative judgment is calculated from the sign of difference between the predicted scale values as follows:

$$\hat{y} = \mathrm{sgn}(\boldsymbol{a}^\top(\boldsymbol{x}^A - \boldsymbol{x}^B)) \tag{5}$$

where $\mathrm{sgn}(\cdot)$ is the signum function. If a predicted judgment is consistent with a measured judgment, the following inequality is satisfied:

$$y\boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x}^{A}-\boldsymbol{x}^{B})\geq 1. \tag{6}$$

Training of the predictive model under inequality constraint (6) through all trials was achieved by the support vector machine (SVM)-like framework in [9]. The trained model became the maximum margin classifier that classifies the difference between paired feature vectors into two classes corresponding to comparative judgment.

B. Experiment

To verify the effectivity of the CSP-based feature, the event-related desynchronization/synchronization (ERD/ERS)-based feature and the linear discriminant analysis (LDA) feature without the spatial filtering were computed. The ERD/ERS was computed as the relative change of the mean power of the bandpass signal during 200–1000 ms after stimulus onset (denoted by P_a), from the mean power in reference interval of 500 ms before stimulus onset (denoted by P_r) [10], i.e. $\mathrm{ER}=(P_a-P_r)/P_r$. The ERD/ERS feature was obtained for each frequency band and channel. To fit its dimensionality to the CSP feature, the principal component analysis (PCA) was applied to the ERD/ERS feature for each band, and the number of dimensions reduced to 14. The LDA-based feature was extracted by the LDA instead of the PCA in extraction of the ERD/ERS-based feature. Since a pair of features was labeled by a comparative judgment, the LDA was applied to difference between paired ERD/ERS feature vectors. The dimensionality of the LDA-based feature was 7 because the number of the meaningful basis of the LDA is limited to less than the number of classes.

The predictive models were trained using these features for each participant, and their performances were evaluated with the prediction accuracy of comparative judgment. The number of paired data after artifact rejection was different for each participant, and was in the range of 312–374. All paired data were divided into the training set and the evaluation set by the procedure of the 10-fold cross-validation. The training set was divided again into the 2 sets for training of a model and tuning of the hyperparameters by the procedure of the 10-fold cross-validation.

C. Results and Discussion

Mean prediction accuracies for each feature and participant are shown in Fig. 1. Mean prediction accuracy with the CSP-feature was the highest of all features in all participants. However, there is no significant difference of the mean accuracy between the methods of feature extraction.

An example of two magnetic spatial patterns of the spatially filtered alpha activities corresponding to the largest and smallest eigenvalues are shown in Fig. 2. They were computed from the corresponding column vectors of the pseudo-inverse of the matrix $\boldsymbol{W}^{\mathsf{T}}$ in (4). Fig. 2 (a) indicates that the alpha activity in the left temporal can be related to cognitive processing on subjective preference. A distinctive

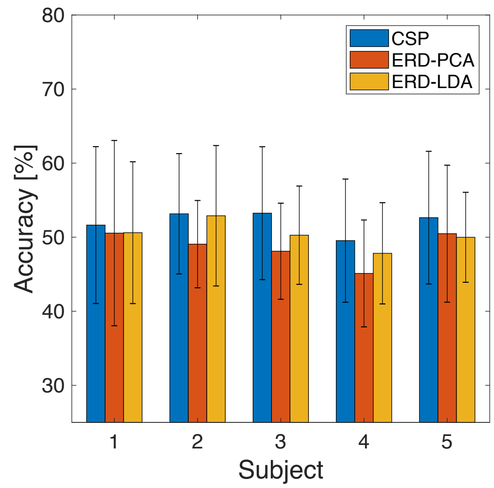


Fig. 1. Mean prediction accuracies for each feature and participant. Error bars indicate the standard deviation.

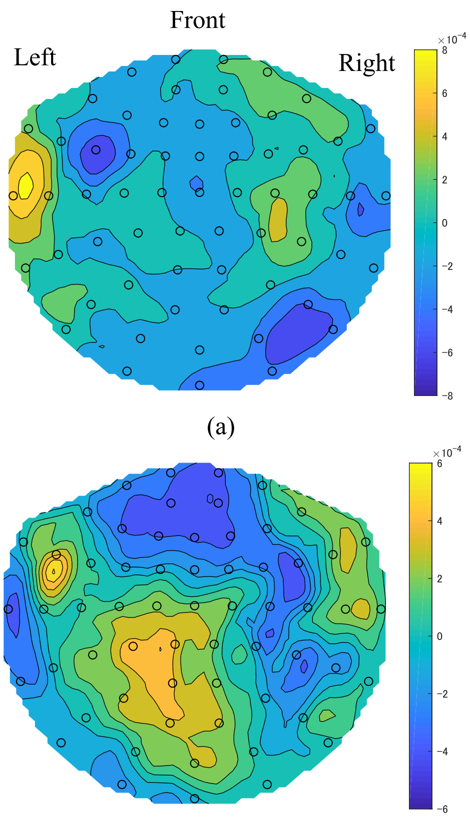
pattern in the parietal region is also shown in the in Fig. 2 (b). This region covers the motor cortex. The pattern in Fig. 2 (b) presumably did not reflect the activity related to subjective preference, but the activity in the motor cortex related to finger movement at comparative judgment which occurred in the interval of 200–1000 ms after onset of the latter stimulus.

To visualize distribution of feature vectors for each stimulus, the CSP-based features and the LDA-based features were projected onto the 2-dimensional space spanned by 2 bases corresponding to the first and second principal components of a dataset for model training. Examples of these projected features are shown in Fig. 3. The projected CSP-based features for training tended to form 2 separate clusters corresponding to relative preference of paired stimuli unlike the LDA-based feature. The CSP-based features for tuning of the hyperparameters that were not used for construction of the spatial filter tended to be projected between 2 clusters. In Fig. 3 (a), most of the CSP-based feature corresponding to the stimulus 1, which was the most preferred stimulus, belong to the left cluster.

These results indicate that the spatial filter using the CSP can extract brain activities varying with degree of subjective preference nevertheless the only binary comparative judgments were used as information on subjective preference. In light of higher mean accuracy and separability, the CSP-based feature is possibly more discriminative and robust feature for prediction of subjective preference than the LDA-based one. On the other hand, the mean accuracies were close to 50%. Such low performance was probably due to overfitting. Decrease/increase of the power of signal in either condition can be emphasized even if extracted activities are irrelevant to subjective preference.

V. Conclusion

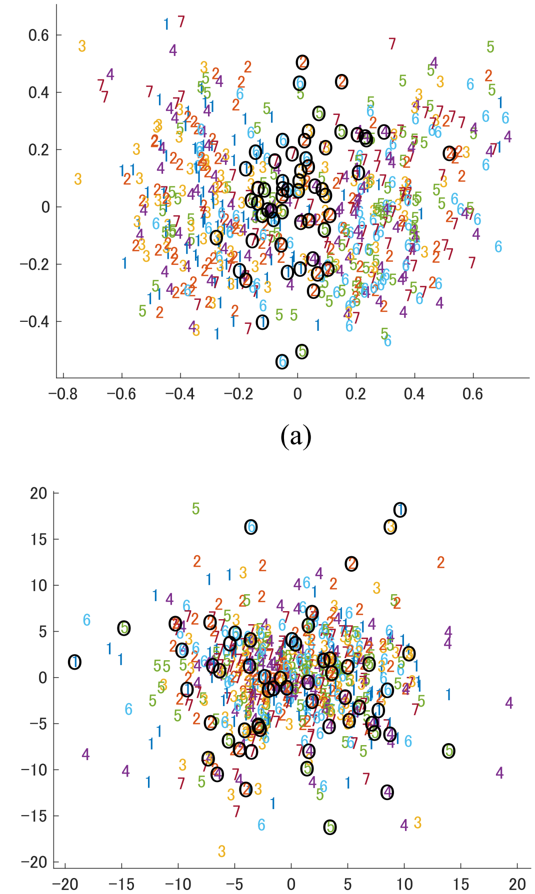
We constructed the spatial filter using the CSP to extract brain magnetic activities varying with degree of subjective



(a)

(b)

Fig. 2. An example of magnetic spatial patterns of the spatially filtered alpha activities corresponding to the largest (a) and smallest (b) eigenvalues.



(a)

(b)

Fig. 3. An example of the 2-dimensional projection of CSP-based features (a) and LDA-based features (b). Both of features were extracted from the same dataset. Each plotted number corresponds to each stimulus. Circled numbers indicate feature vectors for tuning of the hyperparameters.

preference induced by chord. In training of the filter, the only paired-comparative judgments between paired chords were used as information on subjective preference. The CSP-based feature slightly improved performance of the predictive model compared with the other features without spatial filtering.

For practical assessment of impressions using neurophysiological index, utilization of EEG and additional improvement of feature extraction are essential. Our method is applicable to EEG data recorded by same the experimental paradigm. Optimal selection of the frequency band is expected to be effective because the dominant frequency band of cortical activities related to subjective preference can be different for each subject. Regularization of the CSP is also effective to prevent the spatial filter from over-fitting [11]. The regularization that equalizes the variance of filtered activities induced by the same stimuli is considered to be suitable.

References

[1] Y. Soeta, S. Nakagawa, and M. Tonoike, "Magnetoencephalographic responses corresponding to individual subjective preference of sound fields," *Journal of Sound and Vibration*, vol. 258, pp. 419–428, 2002.

[2] Y. Soeta, S. Nakagawa, and M. Tonoike, "Magnetoencephalographic responses correspond to individual annoyance of bandpass noise," *Journal of Sound and Vibration*, vol. 277, pp. 479–489, 2004.

[3] S. Nakagawa, Y. Kanemoto, T. Hotehama, Y. Soeta, and S. Ishimitsu, "Evaluation of auditory impression of music using brain activity," *ICIC Express Letters*, vol. 7, pp. 1545–1550, 2013.

[4] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis." *IEEE Signal Process. Mag.*, vol. 25, pp. 41–56, 2008.

[5] G. M. Bidelman, and A. Krishnan, "Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem." *J. Neurosci.*, vol. 29, pp. 13165–13171, 2009.

[6] A. Otsuka, M. Yumoto, S. Kuriki, and S. Nakagawa, "Neuromagnetic auditory steady state response to chords: effect of frequency ratio." in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* (EMBC2013), pp. 4418–4421, 2013.

[7] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Netw.*, vol. 13, pp. 411–430, 2000.

[8] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehabil. Eng.*, vol.8, pp. 441–446, 2000.

[9] R. Herbrich, T. Graepel, P. B. Sdorra, and K. Obermayer, "Learning preference relations for information retrieval," in *Proc. AAAI Workshop Text Categorization and Machine Learning*, pp. 80–84, 1998.

[10] G. Pfurtscheller and F. H. Lopes, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol.*, vol. 110, pp. 1842–1857, 1999.

[11] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithm," *IEEE Trans. Biomed. Eng.*, vol. 58, pp. 355–362, 2011.