Band Selection by Criterion of Common Spatial Patterns for Motor Imagery Based Brain Machine Interfaces

Hiroshi Higashi* and Toshihisa Tanaka*
* Department of Electronic and Information Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan.
RIKEN Brain Science Institute, Saitama, Japan.
E-mail: higashi@sip.tuat.ac.jp, tanakat@cc.tuat.ac.jp Tel: +81-42-388-7123

***Abstract*—Design of spatial weights and filters that extract components related to brain activity of motor imagery is a crucial issue in brain machine interfaces. This paper proposes a novel method to design these filters. We use a similarity of the covariance matrices of narrow band observed signals over frequency bins. The similarity is defined based on the common spatial pattern method. This proposed method enables us to design the multiple bandpass filters. The experimental results of classification of EEG signals during motor imagery show that the proposed method achieves higher classification accuracy than well-known conventional methods.**

I. Introduction

Brain machine/computer interfacing (BMI/BCI) is a challenging technology of signal processing, machine learning, and neuroscience [1]. BMIs capture brain activities associated to mental tasks and external stimuli and realize non-muscular communication and control channel for conveying messages and commands to the external world [1]–[3]. Basically, noninvasively measured data such as electroencephalogram (EEG), magnetoencephalogram (MEG), and functional magnetic resonance imaging (fMRI) are widely used to observe brain activities. Among them, because of its simplicity and low cost, EEG is practical for use in engineering applications [4], [5].

Efficient decoding around motor-cortex is a crucial technique for realization of BMI associated with motor-imagery (MI-BMI) [6], [7] with the application to controlling external devices [7], prostheses [4], rehabilitation [8], and so forth. For instance it is also known that the real and imaginary movements of hands and feet evoke the change of the so-called mu rhythm in different brain regions [2], [3]. Therefore, the accurate extraction of these changes from the measured EEG signals in the presence of measurement noise and spontaneous components which are related to other brain activities enables us to classify the EEG signal associated with the different motor (imagined) actions such as movement of the right hand, left hand, or feet.

In classification of EEG signals in MI-BMI and analyzing of the brain activities during motor imagery, signal processing techniques such as bandpass filtering and spatial weighting are used [1]. For the processing, presuming the parameters such as coefficients of the filters and weights that extract the related components is a crucial issue. Moreover, the optimum parameters in classification are highly dependent on users and measurement environments [9].

In order to determine the parameters, data-driven techniques that exploit observed data are widely used [1], [2]. The observed data essentially include class labels corresponding to the tasks. The techniques should find the parameters that extract discriminative features as much as possible. For example, the well-known common spatial pattern (CSP) method finds the spatial weights by using observed signals [1], [9], [10] in such a way that the variances of the signal extracted by the linear combination of a multichannel signal and the spatial weights differ as much as possible between two classes. The standard CSP method has been extended to methods to estimate the other parameters such as the frequency bands [11]–[16], and methods to select the CSP features extracted with various parameters [17], [18].

In this paper, we propose a novel method to design the spectral filters. The idea of the method is to introduce a similarity of the covariance matrices of narrow band signals over frequency bins. The similarity is defined with the CSP method. Based on the similarity, the method determines the passbands of the filters. The advantage of the proposed method compared with the other CSP based filter design methods is that the proposed method is able to design multiple filters. Furthermore, unlike discriminative filterbank common spatial patterns (DFBCSP) [15] that can design multiple filters having one passband for each filter, the passband of each filter designed by the proposed method can comprise some passbands. Therefore, under the assumption that there are components that have some frequency bands each in EEG signals, the proposed method can design the filters extracting such components. We show experimental results of which we classified EEG signals into two motor imagery classes. Compared with some conventional methods, we suggest an effectiveness of the proposed method in MI-BMI.

This work is supported in part by KAKENHI, Grant-in-Aid for Scientific Research (B), 21360179 and 24360146, and Grant-in-Aid for JSPS Fellows, 24686.

II. Band Selection by Using Common Spatial Patterns in Frequency Bins

We introduce the proposed method in this section. First, we review the CSP method in Sec. II-A. Next, we reformulate the covariance matrices of the observed signals with discrete Fourier transform and define the CSP method with the band selected covariance matrices in Sec. II-A. In Sec. II-C, we introduce an algorithm to determine the set of the filters for the band selections.

A. Common Spatial Pattern (CSP) – Review [10]

Let $\boldsymbol{X} \in \mathbb{R}^{M\times N}$ be an observed signal, where M is the number of channels and N is the number of samples. In BMI application, we do not directly use $\boldsymbol{X}$, but use the filtered signal described as $\hat{\boldsymbol{X}} = \mathcal{H}(\boldsymbol{X})$ to find the CSP, where $\mathcal{H}$ is a bandpass filter which passes the frequency components related to brain activity of motor imagery. Denote the components of $\hat{\boldsymbol{X}}$ by $\hat{\boldsymbol{X}} = [\hat{\boldsymbol{x}}_0, \ldots, \hat{\boldsymbol{x}}_{N-1}]$, where $\hat{\boldsymbol{x}}_n \in \mathbb{R}^M$ and n is the time index. We assume sets of the observed signals, $\mathcal{C}_1$ and $\mathcal{C}_2$, where $\mathcal{C}_d$ contains the signals belonging to class d, $d \in \{1, 2\}$ is a class label, $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$, and $\emptyset$ is a set having no elements. CSP is defined the weight vector that minimize the intra-class variance in $\mathcal{C}_c$ under the normalization of samples, where c is a class label. More specifically, for c fixed, the weight vector is found by solving the following optimization problem [9], [10];

$$\begin{aligned}&\min_{\boldsymbol{w}} \quad E_{\boldsymbol{X}\in\mathcal{C}_c}\left[\frac{1}{N}\sum_{n=0}^{N-1}|\boldsymbol{w}^T(\hat{\boldsymbol{x}}_n-\boldsymbol{\mu})|^2\right],\\ &\text{subject to} \sum_{d=1,2}E_{\boldsymbol{X}\in\mathcal{C}_d}\left[\frac{1}{N}\sum_{n=0}^{N-1}|\boldsymbol{w}^T(\hat{\boldsymbol{x}}_n-\boldsymbol{\mu})|^2\right]=1,\end{aligned} \tag{1}$$

where $E_{\boldsymbol{X}\in\mathcal{C}_d}[\cdot]$ denotes the expectation over $\mathcal{C}_d$, $\boldsymbol{\mu}$ is the time average of $\boldsymbol{X}$ given by $\boldsymbol{\mu} = (1/N)\sum_{n=0}^{N-1}\hat{\boldsymbol{x}}_n$, $\cdot^T$ is the transpose of a vector or a matrix, and $|\cdot|$ is the absolute value of a scalar. The solution of (1) is given by the generalized eigenvector corresponding to the smallest generalized eigenvalue of the generalized eigenvalue problem described as

$$\boldsymbol{\Sigma}_c\boldsymbol{w} = \lambda(\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2)\boldsymbol{w}, \tag{2}$$

where $\boldsymbol{\Sigma}_d$, $d = 1, 2$, are defined as

$$\boldsymbol{\Sigma}_d = E_{\boldsymbol{X}\in\mathcal{C}_d}\left[\frac{1}{N}\sum_{n=0}^{N-1}(\hat{\boldsymbol{x}}_n-\boldsymbol{\mu})(\hat{\boldsymbol{x}}_n-\boldsymbol{\mu})^T\right]. \tag{3}$$

Although the solution of (1) is given by the eigenvector corresponding to the smallest eigenvalue in (2), we can use the other eigenvectors for classification [19]. The M eigenvectors can be obtained by solving (2) as $\hat{\boldsymbol{w}}_1, \ldots, \hat{\boldsymbol{w}}_M$, where $\hat{\boldsymbol{w}}_i$ is the eigenvector corresponding to the ith smallest eigenvalue of (2). We assume that the $2r$ eigenvectors are used for classification of unlabeled data, $\boldsymbol{X}$. Then we obtain the feature vector, $\boldsymbol{v} \in \mathbb{R}^{2r}$, from $\boldsymbol{X}$ defined as

$$\begin{aligned}\boldsymbol{v} =&[\sigma^2(\boldsymbol{X},\hat{\boldsymbol{w}}_1),\ldots,\sigma^2(\boldsymbol{X},\hat{\boldsymbol{w}}_r),\\ &\sigma^2(\boldsymbol{X},\hat{\boldsymbol{w}}_{M-r+1}),\ldots,\sigma^2(\boldsymbol{X},\hat{\boldsymbol{w}}_M)]^T.\end{aligned} \tag{4}$$

B. CSP with Spectrally Selected Covariance Matrices

In this section, the CSP method with the covariance matrices filtered in frequency domain is shown. First, we show that the covariance matrix defined in (3) can be represented as the sum of the covariance matrices calculated in narrow frequency bins. Introducing the filters, the CSP method with the filtered covariance matrices and the feature vector using the CSPs and the filters are defined.

As well as Sec. II-A, we assume that the signal sets of $\mathcal{C}_1$ and $\mathcal{C}_2$. Let $\tilde{\boldsymbol{X}} \in \mathbb{R}^{M\times N}$ be the signal matrix defined as

$$[\tilde{\boldsymbol{X}}]_{m,n} = [\boldsymbol{x}_n]_m - [\boldsymbol{\mu}]_m, \quad n = 0, \ldots, N-1,\ m = 1, \ldots, M. \tag{5}$$

We define a mean vector of $\tilde{\boldsymbol{X}}$ as $\boldsymbol{\mu} = \frac{1}{N}\sum_{n=0}^{N-1}\boldsymbol{x}_n$. Let $\boldsymbol{F} \in \mathbb{C}^{N\times N}$ be a matrix that performs discrete Fourier transform defined as

$$[\boldsymbol{F}]_{k,l} = \frac{1}{\sqrt{N}}e^{-j2\pi(l-1)(k-1)/N}, \quad l, k = 1, \ldots, N, \tag{6}$$

then the equations;

$$\boldsymbol{F}^H\boldsymbol{F} = \boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{I}_N, \tag{7}$$

are given. The Fourier transform of $\boldsymbol{X}$ is obtained by

$$\boldsymbol{Y} = \boldsymbol{F}\boldsymbol{X}. \tag{8}$$

The kth element of each row of $\boldsymbol{Y}$ represented by $\boldsymbol{y}_k$ as $\boldsymbol{Y} = [\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{N-1}]$ is the kth coefficient of the discrete Fourier transform of $\boldsymbol{X}$ defined as

$$[\boldsymbol{y}_k]_m = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}[\boldsymbol{x}_n]_m e^{-j2\pi nk/N}, \tag{9}$$

for $k = 0, \ldots, N-1$ and $m = 1, \ldots, M$. By using the above equations, the covariance matrix of the signals in each set can be transformed as

$$\begin{aligned}\boldsymbol{S}_d &= \frac{1}{N}E_{\boldsymbol{X}\in\mathcal{C}_d}\left[\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T\right] = \frac{1}{N}E_{\boldsymbol{X}\in\mathcal{C}_d}\left[\tilde{\boldsymbol{X}}\boldsymbol{F}\boldsymbol{F}^H\tilde{\boldsymbol{X}}^T\right]\\ &= \frac{1}{N}\sum_{k=1}^{N-1}E_{\boldsymbol{X}\in\mathcal{C}_d}\left[\boldsymbol{y}_k\boldsymbol{y}_k^H\right]\\ &= \frac{1}{N}\sum_{k=1}^{K}E_{\boldsymbol{X}\in\mathcal{C}_d}\left[\boldsymbol{y}_k\boldsymbol{y}_k^H + \boldsymbol{y}_{N-k+1}\boldsymbol{y}_{N-k+1}^H\right]\\ &= \frac{1}{N}\sum_{k=1}^{K}E_{\boldsymbol{X}\in\mathcal{C}_d}\left[\boldsymbol{y}_k\boldsymbol{y}_k^H + \bar{\boldsymbol{y}}_k\bar{\boldsymbol{y}}_k^H\right]\\ &= \frac{1}{N}\sum_{k=1}^{K}E_{\boldsymbol{X}\in\mathcal{C}_d}\left[2\Re(\boldsymbol{y}_k\boldsymbol{y}_k^T)\right]\\ &= \frac{1}{N}\sum_{k=1}^{K}\boldsymbol{S}_d^{(k)},\end{aligned} \tag{10}$$

for $d = 1, 2$, where a matrix, $\boldsymbol{S}_d^{(k)} \in \mathbb{R}^{M\times M}$, is a covariance matrix at kth frequency bin defined as

$$\boldsymbol{S}_d^{(k)} = E_{\boldsymbol{X}\in\mathcal{C}_d}\left[2\Re(\boldsymbol{y}_k\boldsymbol{y}_k^T)\right], \tag{11}$$

the operator, $\Re(\cdot)$ takes the real value of an input, and $K = \lfloor N/2 \rfloor$. The covariance matrix of the filtered signals is

$$\hat{\boldsymbol{S}}_d(\boldsymbol{h}) = \frac{1}{N}\sum_{k=1}^{K}[\boldsymbol{h}]_k \boldsymbol{S}_d^{(k)}, \tag{12}$$

where $\boldsymbol{h}$ is a vector consisted of the filter coefficients that take either 0 or 1 as

$$[\boldsymbol{h}]_k \in \{0, 1\}, \quad k = 1, \dots, K. \tag{13}$$

We define the CSP method with the filters. Assume the set consisting of the N_f filters denoted by

$$\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_{N_f}]^T, \tag{14}$$

where $[\boldsymbol{h}_i]_j \in \{0, 1\}$, $i = 1, \dots, N_f$, $j = 1, \dots, K$, and N_f is the number of the filters. The CSP method with the filtered signals can be defined with the filtered covariance matrices by

$$\boldsymbol{w}_i = \arg\min_{\boldsymbol{w}} \frac{\boldsymbol{w}^T \hat{\boldsymbol{S}}_c(\boldsymbol{h}_i)\boldsymbol{w}}{\boldsymbol{w}^T \left(\hat{\boldsymbol{S}}_1(\boldsymbol{h}_i) + \hat{\boldsymbol{S}}_2(\boldsymbol{h}_i)\right)\boldsymbol{w}}, \tag{15}$$

for $i = 1, \dots, N_f$. The solution of (15) is the generalized eigenvector corresponding to the smallest generalized eigenvalue of the generalized eigenvalue problem:

$$\hat{\boldsymbol{S}}_c(\boldsymbol{h}_i)\boldsymbol{w} = \lambda\left(\hat{\boldsymbol{S}}_1(\boldsymbol{h}_i) + \hat{\boldsymbol{S}}_2(\boldsymbol{h}_i)\right)\boldsymbol{w}. \tag{16}$$

As well as CSP, we can use the multiple generalized eigenvectors corresponding to the largest and smallest eigenvalues as the spatial weights. The generalized eigenvector corresponding to the jth largest eigenvalue is denoted as $\boldsymbol{w}_i^{(j)}$. Then, the feature vector of the observed signal, $\boldsymbol{X}$, with $\boldsymbol{H}$ and the corresponding spatial weights is defined by

$$\begin{aligned}\boldsymbol{v} =& [\alpha_{\boldsymbol{X}}(\boldsymbol{w}_1^{(1)}, \boldsymbol{h}_1), \dots, \alpha_{\boldsymbol{X}}(\boldsymbol{w}_1^{(r)}, \boldsymbol{h}_1),\\ & \alpha_{\boldsymbol{X}}(\boldsymbol{w}_1^{(M-r+1)}, \boldsymbol{h}_1), \dots, \alpha_{\boldsymbol{X}}(\boldsymbol{w}_1^{(M)}, \boldsymbol{h}_1), \dots,\\ & \alpha_{\boldsymbol{X}}(\boldsymbol{w}_{N_f}^{(1)}, \boldsymbol{h}_{N_f}), \dots, \alpha_{\boldsymbol{X}}(\boldsymbol{w}_{N_f}^{(r)}, \boldsymbol{h}_{N_f}),\\ & \alpha_{\boldsymbol{X}}(\boldsymbol{w}_{N_f}^{(M-r+1)}, \boldsymbol{h}_{N_f}), \dots, \alpha_{\boldsymbol{X}}(\boldsymbol{w}_{N_f}^{(M)}, \boldsymbol{h}_{N_f})]^T,\end{aligned} \tag{17}$$

where the $2r$ spatial weight vectors for each filter are supposed to form the feature vector and $\alpha_{\boldsymbol{X}}(\boldsymbol{w}, \boldsymbol{h})$ is the logarithm of the variance of the signal extracted by $\boldsymbol{w}$ and $\boldsymbol{h}$ defined as

$$\alpha_{\boldsymbol{X}}(\boldsymbol{w}, \boldsymbol{h}) = \log\left[\boldsymbol{w}^T\left(\frac{1}{N}\sum_{k=1}^{K}[\boldsymbol{h}]_k 2\Re(\boldsymbol{y}_k \boldsymbol{y}_k^T)\right)\boldsymbol{w}\right]. \tag{18}$$

C. Search for Filters

In this section, we introduce an algorithm to determine the set of the filters, $\boldsymbol{H}$. For designing $\boldsymbol{H}$, we use the similarity of the covariance matrices over frequency bins that is evaluated by using the CSP method.

First, we find the CSP, $\tilde{\boldsymbol{w}}_k$, and the cost, $J_k(\tilde{\boldsymbol{w}}_k)$, in each frequency bin as

$$\tilde{\boldsymbol{w}}_k = \arg\min_{\boldsymbol{w}} \; J_k(\boldsymbol{w}), \tag{19}$$

and

$$J_k(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \boldsymbol{S}_c^{(k)} \boldsymbol{w}}{\boldsymbol{w}^T\left(\boldsymbol{S}_1^{(k)} + \boldsymbol{S}_2^{(k)}\right)\boldsymbol{w}}. \tag{20}$$

By using (19) and (20), we quantitatively evaluate the similarity of the covariance matrices over the frequency bins. The idea for the evaluation is that $\{\boldsymbol{S}_1^{(k')}, \boldsymbol{S}_2^{(k')}\}$ and $\{\boldsymbol{S}_1^{(k'')}, \boldsymbol{S}_2^{(k'')}\}$ are similar to each other under $\tilde{\boldsymbol{w}}_{k'}$, if the difference between $J_{k'}(\tilde{\boldsymbol{w}}_{k'})$ and $J_{k''}(\tilde{\boldsymbol{w}}_{k'})$ is small.

Based on the similarity, we introduce the algorithm for finding the filters as follows. We sequentially obtain the filters forming $\boldsymbol{H}$. Therefore, we find $\boldsymbol{h}_i$ after obtaining $\boldsymbol{h}_1, \dots, \boldsymbol{h}_{i-1}$ in the following algorithm. A vector, $\boldsymbol{p} \in \mathbb{R}^K$, whose elements are the CSP cost is defined as

$$\boldsymbol{p} = [J_1(\tilde{\boldsymbol{w}}_1), \dots, J_K(\tilde{\boldsymbol{w}}_K)]^T. \tag{21}$$

The bin index corresponding to the minimum of $\boldsymbol{p}$ is defined as

$$k_{\min} = \arg\min_k \, [\boldsymbol{p}]_k \tag{22}$$

to decide the target covariance matrix that is used for evaluating its similarity. The frequency, η_k, $k = 1, \dots, K$, is selected by thresholding the similarities as

$$\eta_k = \begin{cases} 1, & J_k(\tilde{\boldsymbol{w}}_{k_{\min}}) \le T_s \\ 0, & \text{otherwise} \end{cases}, \tag{23}$$

where T_s is a threshold. Furthermore, we remove isolated bands including η_k, $k = 1, \dots, K$ and determine ith filter of $\boldsymbol{H}$ as

$$[\boldsymbol{h}_i]_k = \begin{cases} 1, & \displaystyle\sum_{k'=k-N_B+1}^{k} \eta_{k'} = T_b \text{ or } \sum_{k'=k}^{k+N_B-1} \eta_{k'} = T_b \\ 0, & \text{otherwise} \end{cases}, \tag{24}$$

for $k = 1, \dots, K$, where T_b is a threshold deciding the minimum bandwidth of the passbands. To prevent the next filter, $\boldsymbol{h}_{i+1}$ from having the same passband as $\boldsymbol{h}_1, \dots, \boldsymbol{h}_i$, we remove the possibility to select the same bins as the passbands of $\boldsymbol{h}_i$ for next $k_{\min}$ by updating $\boldsymbol{p}$ as

$$[\boldsymbol{p}]_k = \begin{cases} \infty, & [\boldsymbol{h}_i]_k = 1 \text{ or } k = k_{\min} \\ [\boldsymbol{p}]_k, & \text{otherwise} \end{cases}, \tag{25}$$

for $k = 1, \dots, K$. After updating $\boldsymbol{p}$, if all elements of $\boldsymbol{h}_i$ are zero, we return to (22) for finding $\boldsymbol{h}_i$ again, otherwise update $i \leftarrow i + 1$ and return to (22) for finding the next filter. The repetition is finished if $J_{k_{\min}}(\tilde{\boldsymbol{w}}_{k_{\min}}) > T_s$. After finishing the repetition, if $i = 1$, this means that any filters can not be decided. Therefore, we define a filter that is represented by a K-dimensional vector whose elements are all ones defined by $\boldsymbol{h}_1 = [1, \dots, 1]^T$, and set i to 2. Then the number of the filters is decided as $N_f = i - 1$. The pseudo-code of the procedure for finding the filters is shown in Algorithm 1.

Algorithm 1 Design of the filters by thresholding of the CSP cost

Input: $\{\boldsymbol{S}_1^{(k)}, \boldsymbol{S}_2^{(k)}\}_{k=1}^K$: the set of covariance matrices.
Parameters: T_s: the threshold of the similarity. T_b: the threshold of the bandwidth.
Output: $\boldsymbol{H}$: the set of the filters.

$i = 1$
Obtain $\boldsymbol{p}$ as (21) with (19) and (20)
repeat
  Choose $k_{\min}$ as (22)
  Obtain η_k by (23)
  Obtain $\boldsymbol{h}_i$ by (24)
  Update $\boldsymbol{p}$ by (25)
  if $\|\boldsymbol{h}_i\| \neq 0$ **then**
    $i \leftarrow i + 1$
  end if
until $J_{k_{\min}}(\tilde{\boldsymbol{w}}_{k_{\min}}) > T_s$
if $i = 1$ **then**
  $\boldsymbol{h}_1 = [1, \ldots, 1]^T$
  $i = 2$
end if
$N_f = i - 1$

TABLE I
DESCRIPTION OF THE DATASET.

Classes	right hand and right foot
Subject labels	*aa*, *al*, *av*, *aw*, *ay*
Number of channels	118
Signal length	3.5 secs
Sampling rate	100 Hz
Number of the trials per class	140

III. EXPERIMENT

We conducted an experiment of classification of EEG signals during motor imagery. We compare the proposed method with well-known conventional methods, CSP, common sparse spectral spatial pattern (CSSSP) [12], and filterbank CSP (FBCSP) [17] in the accuracy rate of the classification.

A. Data Description

We used dataset IVa from BCI competition III [20], which was provided by Fraunhofer FIRST (Intelligent Data Analysis Group) and Campus Benjamin Franklin of the Charité - University Medicine Berlin (Department of Neurology, Neurophysics Group) [21]. The condition for the dataset is shown in Table I. The signals in the provided datasets were recorded with the sampling rate of 1000 Hz.

We furthermore applied to this dataset a Butterworth low-pass filter whose cutoff frequency is 50 Hz and the filter order is 4, and downsampled to 100 Hz. The dataset for each subject consisted of signals of 140 trials per class.

B. Results

For classification of the trial signals, the feature vectors were formed by each method as follows.

- **CSP:** We applied the Butterworth bandpass filter with the passband of 7–30 Hz, and minimized the variance cost of the right hand class in (1). The eigenvectors corresponding to the r largest and r smallest eigenvalues of the eigenvalue problem (2) as the spatial weights.
- **CSP-Exh:** This was an exhaustive search for the parameters using the CSP method. We obtained classification accuracy rates by 5×5 cross validation (CV) using various passbands for the bandpass filter used for preprocessing in CSP. The passband is represented as $[f_l, f_u]$ [Hz] for $f_l = 1, \ldots, 48$ and $f_u = f_l + 1, \ldots, 49$. The exhaustive search was performed to find f_l and f_u that give the best classification accuracy rate for each subject.
- **CSSSP:** The order of the FIR filter was fixed to 16 [12]. The bandpass filter between 7–30 Hz was applied as preprocessing Then, the bandpass filter designed by CSSSP was applied to the observed signals, and then we classified them by the same procedure as the CSP method.
- **FBCSP:** The filterbank comprising 9 bandpass filters covering 4–40 Hz was used. All filters were Chebyshev Type II filters with a bandwidth of 4 Hz each. The number of spatial weights, N_M, in each band was set to 8. These parameters were decided by referring [17]. We selected the r feature values used for classification by mutual information based best individual feature with a naïve Bayesian Parzen window (NBPW) classifier [17].
- **Proposed:** We assumed that the filter coefficients corresponding to bands except for 7–30 Hz were zero. After designing the filters with the thresholds of T_s and T_b, we extracted the feature vectors defined in (17).

In the column below the classification accuracy in Tables II, we show the parameters that we tuned to obtain the highest classification accuracy rate by using 5×5 CV for each method and subject. The parameters tested and decided in advance were as follows. The parameter for the dimension of the feature vector, r, was chosen out of $\{1, 2, \ldots, 20\}$. The regularization parameter, C, in CSSSP was chosen out of $\{0, 0.01, 0.1, 0.2, 0.5, 1, 2, 5\}$. The thresholds, T_s and T_b, in the proposed method was chosen out of $\{0.05, 0.1, \ldots, 0.5\}$ and $\{\lceil 350k/100 \rceil + 1\}_{k=1}^5$, respectively. Finally, the feature vector extracted by each method was projected into the 1-dimensional space determined by linear discriminant analysis [22] and was classified by a threshold that is the middle point of two class averages over the learning samples. The proposed method achieves the highest accuracy in three subjects and the accuracy averaged over all subjects.

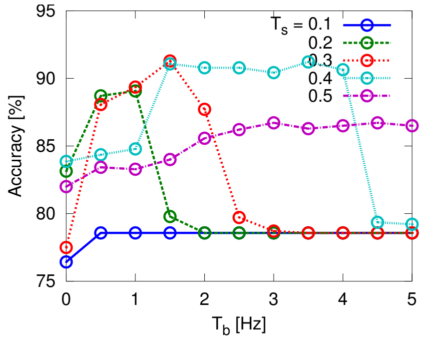
Figure 1 shows the relations between the classification accuracy rate and the parameters, T_b and T_s, in the proposed method. The parameter, r, is fixed to 1. As shown in the figure, the classification accuracy with the proposed method highly depends on the parameters. Moreover, the combinations of the parameters performing the high accuracy rates are different among the subjects.

Figures 2 and 3 show the examples of the filters designed by the proposed method. The same parameters as those shown

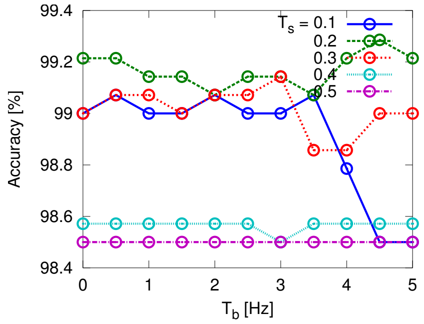
TABLE II

CLASSIFICATION ACCURACY [%] GIVEN BY 5×5 CV IN DATASET IVA FROM BCI COMPETITION III. THE FIGURE WITH ± REPRESENTS THE STANDARD DEVIATION (S.D.). THE VALUES BELOW THE COLUMNS SHOWN IN THE CLASSIFICATION ACCURACIES ARE THE PARAMETERS FOR EACH SUBJECT.

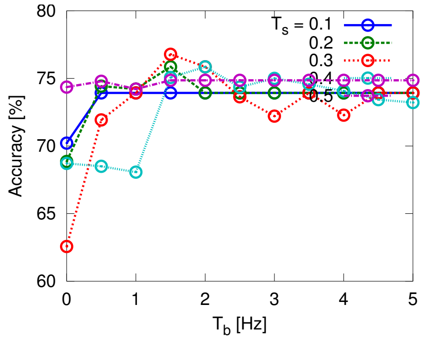
Method	Subject					Ave.
	aa	*al*	*av*	*aw*	*ay*	
CSP	81.5±4.0	98.8±1.5	74.14±4.1	97.1±2.5	93.1±4.2	88.9
(r)	4	8	3	4	3	
CSP-Exh	92.7±3.0	99.6±0.7	76.9±6.2	99.6±0.8	94.3±3.8	92.6
(r, f_l–f_u)	1, 11–16	1, 12–20	2, 10–12	5, 9–15	3, 9–22	
CSSSP	91.9±3.0	99.2±1.5	74.9±4.1	99.3±1.0	93.7±3.7	91.8
(r, C)	1, 0	10, 0.01	3, 2	5, 0.01	3, 0	
FBCSP	92.0±4.0	99.1±1.5	72.4±5.3	98.5±1.5	90.5±4.0	90.5
(r)	1	1	2	9	10	
Proposed	91.3±3.5	99.4±1.4	78.1±5.9	99.4±1.1	95.9±2.5	92.8
(r, T_b, T_s)	1, 2, 0.4	5, 5, 0.2	2, 4, 0.45	3, 0, 0.1	3, 4, 0.5	



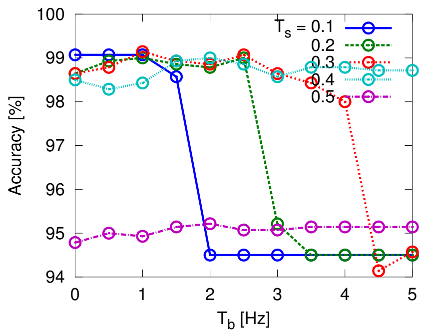
(a) *aa*



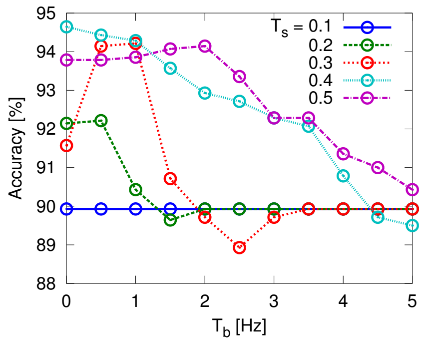
(b) *al*



(c) *av*



(d) *aw*



(e) *ay*

Fig. 1. Relations between the classification accuracy rate and the parameters, the thresholds for the similarity, T_s, and the bandwidth, T_b. r is fixed to 1.

in Table II are used for each subject. We can observe that the different filters are designed for each subject. However, the band in between 10–15 Hz is selected for the passband in all subjects. An oscillation in the band of 10–15 Hz is called the mu rhythm known as a rhythmic component associated to imagery tasks of body movement.

IV. CONCLUSIONS

This paper has proposed the new method to design the set of the filters with multiple passbands. The ideas of the proposed method are to evaluate the similarities between the covariance matrices in each frequency bin by using the CSP method and to decide the passbands based on the similarities. Although the experimental results suggest that the proposed method can outperform some conventional methods in the classification accuracy rate in the motor-imagery based BMIs, the classification accuracy depends on the parameters. How to choose the parameters has still been an open problem in this paper.

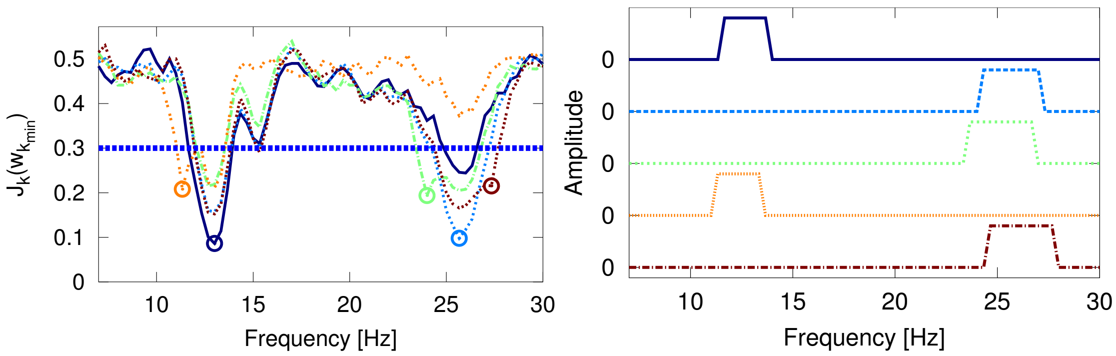
REFERENCES

[1] G. Dornhege, J. d. R. Millan, T. Hinterberger, D. McFarland, and K.-R. Muller, Eds., *Toward Brain-Computer Interfacing*. The MIT Press, 2007.

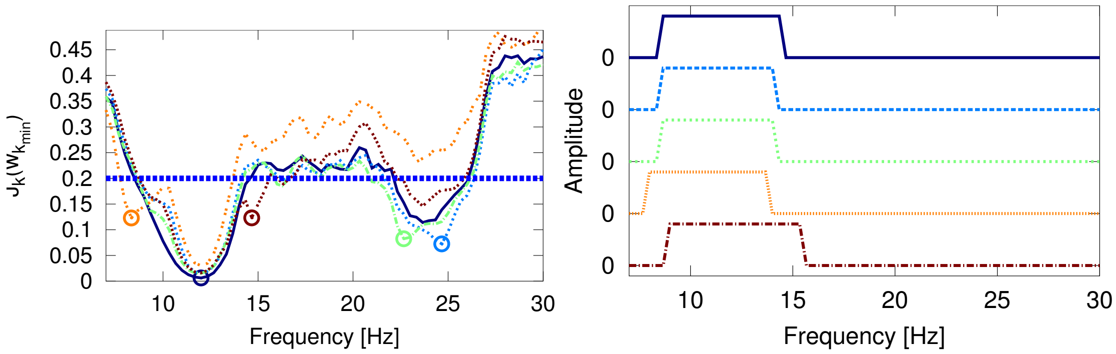
[2] S. Sanei and J. Chambers, *EEG Signal Processing*. Wiley-Interscience, 2007.

[3] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

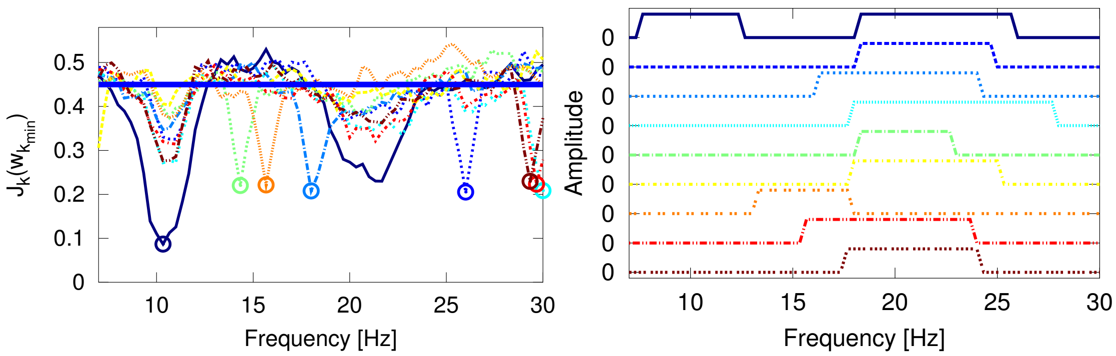
[4] D. J. McFarland and J. R. Wolpaw, "Brain-computer interface operation of robotic and prosthetic devices," *Computer*, vol. 41, no. 10, pp. 52–56, 2008.



(a) *aa*



(b) *al*



(c) *av*

Fig. 2. Examples of $J_k(\tilde{\boldsymbol{w}}_{k_{\min}})$ (left) and the corresponding filters (right) designed by the proposed method for the subjects *aa*, *al*, and *av*. The proposed method is applied with the same parameters as those shown in Table II for each subject. The filters, $\boldsymbol{h}_1, \dots, \boldsymbol{h}_{N_f}$ are shown in order from top to bottom. The circles in the figures in the right panels represent $k_{\min}$. The boldest lines in the figures in the right panel represent T_s.

[5] C. Zhang, Y. Kimura, H. Higashi, and T. Tanaka, "A simple platform of brain-controlled mobile robot and its implementation by SSVEP," in *Proceedings of The 2012 International Joint Conference on Neural Networks (IJCNN)*, 2012, pp. 1–7.

[6] D. J. McFarland and J. R. Wolpaw, "Brain-computer interfaces for communication and control," *Communications of the ACM*, vol. 54, no. 5, pp. 60–66, 2011.

[7] J. R. Wolpaw and D. J. McFarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans," *The National Academy of Sciences*, vol. 101, no. 51, pp. 17 849–17 854, 2004.

[8] J. J. Daly and J. R. Wolpaw, "Braincomputer interfaces in neurological rehabilitation," *The Lancet Neurology*, vol. 7, no. 11, pp. 1032–1043, 2008.

[9] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," *Clinical Neurophysiology*, vol. 110, no. 5, pp. 787–798, 1999.

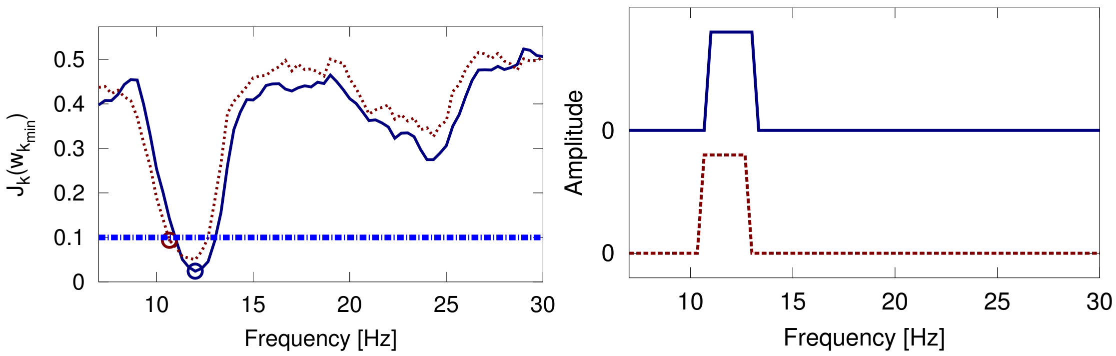
[10] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.

[11] S. Lemm, B. Blankertz, G. Curio, and K.-R. Müller, "Spatio-spectral filters for improving the classification of single trial EEG," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 9, pp. 1541–1548, 2005.

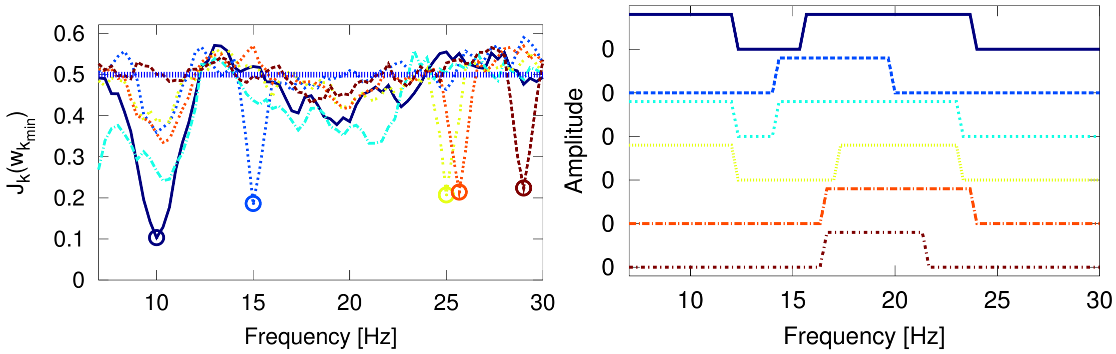
[12] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K.-R. Müller, "Combined optimization of spatial and temporal filters for improving brain-computer interfacing," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2274–2281, 2006.

[13] R. Tomioka, G. Dornhege, G. Nolte, B. Blankertz, K. Aihara, and K. R. Müller, "Spectrally weighted common spatial pattern algorithm for single trial eeg classification," *Dept. Math. Eng., Univ. Tokyo, Tokyo, Japan, Tech. Rep*, vol. 40, 2006.

[14] H. Higashi and T. Tanaka, "Classification by weighting for spatio-frequency components of EEG signal during motor imagery," in *Proceedings of 2011 IEEE International Conference on Acoustics, Speech*



(a) *aw*



(b) *ay*

Fig. 3. Examples of $J_k(\tilde{\boldsymbol{w}}_{k_{\min}})$ (left) and the corresponding filters (right) designed by the proposed method for the subjects *aw* and *ay*. The proposed method is applied with the same parameters as those shown in Table II for each subject. The filters, $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_{N_f}$ are shown in order from top to bottom. The circles in the figures in the right panels represent $k_{\min}$. The boldest lines in the figures in the right panel represent T_s.

and Signal Processing (ICASSP), 2011, pp. 585–588.

[15] ——, "Simultaneous design of FIR filter banks and spatial patterns for EEG signal classification," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 1100–1110, 2013.

[16] N. Tomida, H. Higashi, and T. Tanaka, "A joint tensor diagonalization approach to active data selection for EEG classification," in *Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 983–987.

[17] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in *Proceedings of The 2008 International Joint Conference on Neural Networks (IJCNN)*, 2008, pp. 2390–2397.

[18] ——, "Mutual information-based selection of optimal spatialtemporal patterns for single-trial EEG-based BCIs," *Pattern Recognition*, vol. 45, no. 6, pp. 2137–2144, 2012.

[19] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.

[20] B. Blankertz, K.-R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. Millan, M. Schroder, and N. Birbaumer, "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.

[21] G. Dornhege, B. Blankertz, G. Curio, and K.-R. Müller, "Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 993–1002, 2004.

[22] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.