Check for updates

Temporal Combination Pattern Optimization Based on Feature Selection Method for Motor Imagery BCIs

Jing Jiang[1], Chunhui Wang[1], Jinghan Wu[2], Wei Qin[3,4], Minpeng Xu[2] and Erwei Yin[3,4]*

[1] National Key Laboratory of Human Factors Engineering, China Astronaut Research and Training Center, Beijing, China, [2] Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin, China, [3] Unmanned Systems Research Center, National Innovation Institute of Defense Technology, Academy of Military Sciences China, Beijing, China, [4] Tianjin Artificial Intelligence Innovation Center (TAIIC), Tianjin, China

Common spatial pattern (CSP) method is widely used for spatial filtering and brain pattern extraction from electroencephalogram (EEG) signals in motor imagery (MI)-based brain-computer interfaces (BCIs). The participant-specific time window relative to the visual cue has a significant impact on the effectiveness of the CSP. However, the time window is usually selected experientially or manually. To solve this problem, we propose a novel feature selection approach for MI-based BCIs. Specifically, multiple time segments were obtained by decomposing each EEG sample of the MI task. Furthermore, the features were extracted by CSP from each time segment and were combined to form a new feature vector. Finally, the optimal temporal combination patterns for the new feature vector were selected based on four feature selection algorithms, i.e., mutual information, least absolute shrinkage and selection operator, principal component analysis and stepwise linear discriminant analysis (denoted as MUIN, LASSO, PCA, and SWLDA, respectively), and the classification algorithm was employed to evaluate the average classification accuracy. With three BCI competition datasets, the results of the four proposed algorithms were compared with traditional CSP algorithm in classification accuracy. Experimental results show that compared with traditional algorithm, the proposed methods significantly improve performance. Specifically, the LASSO achieved the highest accuracy (88.58%) among the proposed methods. Importantly, the average classification accuracies using the proposed approaches significantly improved 10.14% (MUIN), 11.40% (LASSO), 6.08% (PCA), and 10.25% (SWLDA) compared to that using CSP. These results indicate that the proposed approach is expected to be practical in MI-based BCIs.

Keywords: brain–computer interface (BCI), electroencephalogram (EEG), motor imagery (MI), common spatial pattern (CSP), feature selection, support vector machine (SVM)

INTRODUCTION

Brain Computer Interface (BCI) is a direct communication system between the brain and external devices, which does not rely on human peripheral nerves and muscles (Wolpaw et al., 2002; Nicolas-Alonso and Gomez-Gil, 2012). In the past few decades, BCI technology has achieved great progress and found a wide range of application scenarios in daily life, such as wheelchair control for disabled

patients, entertainment and smart home control for healthy users (Yin et al., 2013; Herweg et al., 2016; Xu et al., 2016). Most of the BCI systems are EEG-based due to its relatively low expense, high portability, better time resolution, and minimal risks to users as compared to other modalities (Arvaneh et al., 2011). At present, the most extensively used brain signals for BCI input are event-related potential (ERP) (Zhang et al., 2014; Jin et al., 2015, 2017; Yin et al., 2015, 2016; Xu et al., 2018), steady-state visual evoked potential (SSVEP) (Pan et al., 2013; Wang et al., 2016; Xing et al., 2018; Zhang et al., 2018), and motor-imagery (MI) (Zhang et al., 2015, 2017; Ang and Guan, 2017; Qiu et al., 2017; Meng et al., 2018; Lugo et al., 2019). MI-BCI works without outer stimuli, thus it is more intuitive for users (Yu et al., 2012; Velasco-Alvarez et al., 2013). Some MI-BCI systems depend on the well-known neurophysiological phenomenon of event-related synchronization (ERS) or event-related desynchronization (ERD), which is either enhancement or suppression of the EEG (Meng et al., 2013). Other MI-BCI systems use slow cortical potentials such as movement-related cortical potentials (Hinterberger et al., 2004; Ren et al., 2014).

The common spatial pattern (CSP) algorithm (Ramoser et al., 2000) is a method that can extract spatial feature for distinguishing the two types of MI tasks in EEG-based MI-BCI systems. The primary goal of the CSP is to compute spatial filters in a data driven manner, which maximizes the difference between the variance of two classes (Arvaneh et al., 2011). However, an important factor on which the effectiveness of the CSP algorithm depends is the specific time segment of the EEG used in the preprocessing phase (Blankertz et al., 2006). This parameter setting has great effect on subsequent CSP feature extraction and classification. Traditionally, a fixed single time period is used as a general setting for the majority of state-of-the-art MI-BCIs, such as a time period of 2.5–4.5 s after optic cue onset (He et al., 2012), 4–7 s or 0–3.5 s from the beginning of trial (Qiu et al., 2016). Although using subject-specific settings can improve the efficiency of CSP approach to a certain degree, these parameters are usually chosen experientially or manually (Blankertz et al., 2006) and may lead to bad experimental results due to the conflict from useless EEG signals.

In this study, a novel temporal combination pattern optimization method is proposed for MI-based BCIs. First, all the motor imagery trials for every type were decomposed to multiple time segments and were processed via band-pass filter on each time segment. Second, the features from multiple time segments of each sample were extracted by CSP spatial filter and were combined to form a new feature vector. Third, the optimal temporal combination patterns for the new feature vector were selected based on the four feature selection algorithms. Last, the classification algorithm support vector machine (SVM) was employed for identifying different categories, then the average classification accuracy was computed via the cross-validation approach. To our knowledge, no research on the incorporation of the feature selection method into temporal combination pattern optimization for motor imagery BCIs has been reported so far.

The remainder of this paper is organized as follows: section Methods describes the proposed methods, including multi-time segmenting and temporal band-pass filtering, EEG feature extraction based on CSP algorithm, feature selection, classification of optimal spatial-temporal features and event-related spectral perturbation; section Results shows description of the datasets and experimental results; section Discussion presents the discussion; and section Conclusion concludes the study.

METHODS

The proposed temporal combination pattern optimization method is shown in **Figure 1**. The method mainly includes four parts, which are used for EEG raw data preprocessing, feature extraction and classification (i.e., multi-time segmenting and temporal band-pass filtering, EEG feature extraction based on CSP algorithm, feature selection based on four methods, and classification of selected CSP features). In the training phase, the CSP spatial filters and discriminative CSP features for each time segment are computed using the training data labeled with the two-class MI action. Then the temporal combination patterns are selected, and the SVM classifier model is trained. In the evaluation phase, the class of each single-trial MI task is computed using the parameters obtained from the training phase.

Multi-Time Segmenting and Temporal Band-Pass Filtering

In the first stage, EEG raw signals are decomposed to a total of five time segments. The time segments are: $T_{si} = T_0(i-1) \sim 2 + T_0(i-1), i = 1, 2, ..., 5$, where the unit is second, $T_0 = 0.5$ for dataset 1, $T_0 = 0.25$ for dataset 2 and 3. When prompted via the optic cue, participants start performing motor imaging tasks. The indexes of these five time windows are represented by the numbers 1 to 5. A third-order Butterworth filter is applied to filter single-trial EEG data for each time segment with frequency range between 8 and 30 Hz. The reason we use these configurations of time segments and band-pass frequency range is that they contain most of empirically chosen parameters in the relevant literatures (Feng et al., 2018).

EEG Feature Extraction Based on CSP Algorithm

As a spatial filtering approach, CSP has been extensively applied in MI-BCI systems. Its main goal is to minimize the variance of another type of data while maximizing the variance of one type of data, thereby obtaining a set of filters that can extract spatial features (Wang et al., 1999). The variance difference of two-class signals after band-pass filtering is maximized by using CSP projection matrix W. The eigenvectors corresponding to m minimum and maximum eigenvalues are used to form a new final filter $W_{2m} \in \Re^{c \times 2m}$. m is fixed to 2 in this paper. The features from a trial of EEG signal E are computed as below:

$$f = \log(\text{var}(W_{2m}^T E)) \tag{1}$$

where $f \in \Re^{1 \times 2m}$. Merging of feature matrices from N_s time segments forms a new feature matrix for the ith trial as follows:
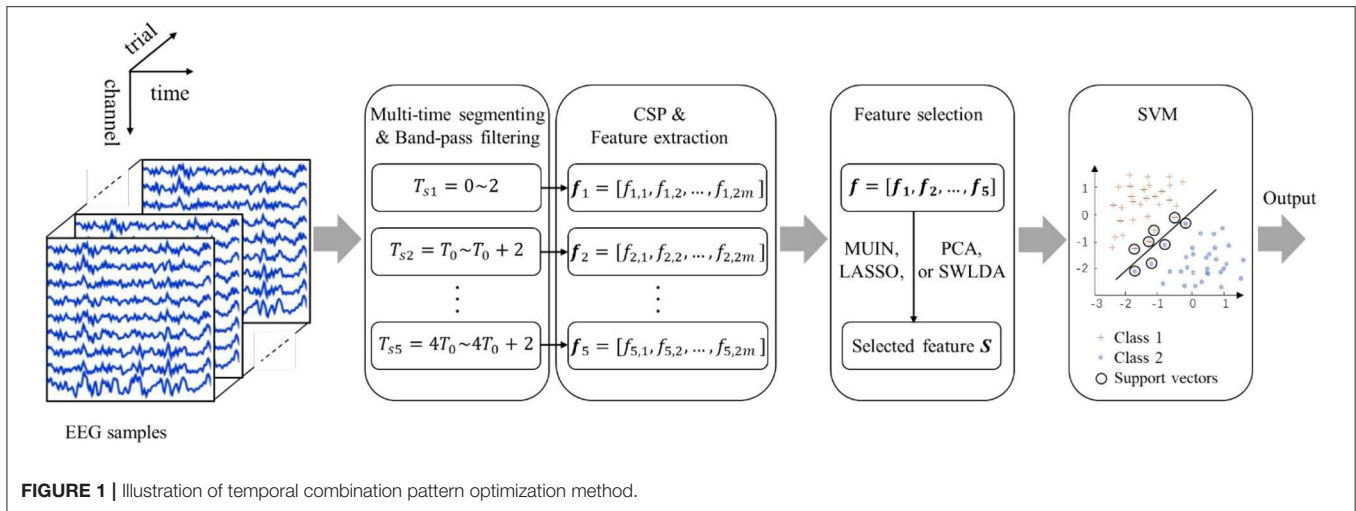
FIGURE 1 | Illustration of temporal combination pattern optimization method.

$$f_i = [f_{1,i}, f_{2,i}, ..., f_{N_s,i}] \tag{2}$$

where $f_i \in \Re^{1 \times (N_s * 2m)}$. The number of time segments $N_s = 5$ in this paper.

To distinguish from the evaluation data, the feature vectors and true class labels from the training data are denoted as $\bar{f}$ and $\bar{y}$, respectively,

$$\bar{f} = [\bar{f}_1, \bar{f}_2, ..., \bar{f}_{n_t}]^T \tag{3}$$

$$\bar{y} = [\bar{y}_1, \bar{y}_2, ..., \bar{y}_{n_t}]^T \tag{4}$$

where $\bar{f} \in \Re^{n_t \times (N_s * 2m)}$, $\bar{y} \in \Re^{n_t \times 1}$, $\bar{y}_i$ and $\bar{f}_i$ denote true class mark and the feature matrix for the ith training trial, respectively, $i = 1,2,..., n_t$; n_t denotes the total number of training trials.

Feature Selection

In the third stage, various feature selection algorithms are employed to choose distinguishing feature matrix $\bar{f}$ extracted from training data. Based on the studies performed on the BCI Competition IV Dataset I, III Dataset IVa and III Dataset IIIa, the four feature selection-based algorithms (i.e., mutual information, least absolute shrinkage and selection operator, principal component analysis and stepwise linear discriminant analysis) produced better performance than other feature selection methods (Meier et al., 2008; Ang et al., 2012), so these algorithms are applied in this paper.

The Mutual Information-Based Feature Selection Algorithm

Mutual information is a useful measure of information in information theory. It can be seen as the amount of information about a random variable contained in another random variable. The mutual information-based feature selection algorithm is described as follows:

Step 1: Initialize the features F and the selected features S.

Initialize $S = \varnothing$, $F = \bar{f} = [v_1^T, v_2^T, ..., v_{n_s \times 2m}^T]$ from the training data, where v_j^T is the jth column vector of $\bar{f}$.

Step 2: Calculate the mutual information $I(v_j; \omega)$ of every feature v_j using the type tag $\omega = \{1, 2\}$ (Kwak and Choi, 2002).

$$
\begin{aligned}
I(v_j; \omega) &= H(\omega) - H(\omega|v_j) \\
&= -\sum_{\omega=1}^{2} p(\omega)\log_2 p(\omega) - (-\sum_{\omega=1}^{2} p(\omega|v_j)\log_2 p(\omega|v_j))
\end{aligned} \tag{5}
$$

where $p(\omega|v_j)\log_2 p(\omega|v_j) = \sum_{i=1}^{n_t} p(\omega|v_{j,i})\log_2 p(\omega|v_{j,i})$, $v_{j,i}$ denotes the jth feature of the ith trial from v_j; n_t is the number of training trials. The probability $p(\omega|v_{j,i})$ is calculated via the Bayes theory as follows:

$$p(\omega|v_{j,i}) = \frac{p(v_{j,i}|\omega)p(\omega)}{\sum_{\omega=1}^{2} p(v_{j,i}|\omega)p(\omega)} \tag{6}$$

where $p(\omega)$ denotes the prior probability of class ω, $p(\omega|v_{j,i})$ is the conditional probability of type ω given $v_{j,i}$, $p(v_{j,i}|\omega)$ is the conditional probability of feature $v_{j,i}$ given type ω. $p(v_{j,i}|\omega)$ can be estimated as follows:

$$\hat{p}(v_{j,i}|\omega) = \frac{1}{n_\omega} \sum_{r \in I_\omega} \phi(v_{j,i} - v_{j,r}, h) \tag{7}$$

where n_ω represents the number of trials belonging to category ω among all training trials, I_ω represents the set of trials belonging to category ω among all training trials, $v_{j,r}$ denotes the feature values from the rth trial of v_j, ϕ represents a smooth Gaussian kernel function, and its smoothing parameter is h. The univariate Gaussian kernel is employed as:

$$\phi(y, h) = \frac{1}{\sqrt{2\pi}} e^{-(y^2/(2h^2))} \tag{8}$$

$$h = \left(\frac{4}{3n_\omega}\right)^{1/5} \sigma \tag{9}$$

where σ denotes the standard deviation of y in formula (10).

Step 3: Sort all the mutual information $I(v_j; \omega)$ of features in descending order in step 2. Then the first $k/2$ and the corresponding pair of features are selected.

LASSO-Based Feature Selection Algorithm

LASSO was first proposed by Robert Tibshirani in 1996 (Tibshirani, 1996), and its full name is Least absolute shrinkage and selection operator. This method is a kind of compression estimation. It obtains a more refined model by constructing a penalty function, so that it compresses some regression coefficients, that is, the sum of the absolute values of the mandatory coefficients is less than a fixed value; at the same time, some regression coefficients are set to zero. Therefore, the advantage of subset shrinkage is retained, and it is a method for processing biased estimates with complex collinearity data (Nigham and Aggarwal, 2005; Meier et al., 2008). The calculation process of the LASSO-based method is as bellow:

Step 1: Initialize the features F and the selected features S as the step 1 of Section 2.3.1.
Step 2: Calculate the contribution degree (CD) of every feature v_j with the type tag.

Given a random variable of the type tag $y = [y_1, y_2, ..., y_{n_t}]^T, y_i = \{1, 2\}$ (Meier et al., 2008; Meinshausen and Yu, 2009), a standard linear regression function can be expressed as:

$$y = F\beta + \varepsilon \tag{10}$$

where y is a $n_t \times 1$vector, $F = \overline{f} = [v_1^T, v_2^T, ..., v_{n_s \times 2m}^T]$ denotes a $n_t \times (2mn_s)$ matrix, ε denotes a noise vector whose mean is 0 and variance is constant. With reference to (Wang et al., 1999), the estimated value of LASSO can be expressed as:

$$\hat{\beta} = \arg\min_{\beta} (||y - F\beta||_2^2 + \lambda ||\beta||_1) \tag{11}$$

where $|| \cdot ||_1, || \cdot ||_2$ represent the l_1-norm and l_2-norm respectively. λ denotes a compensation factor which can encourage a sparse solution LASSO estimate $\hat{\beta}$ (i.e., many entries in $\hat{\beta}$ are equal to zero). By using quadratic programming (Schittkowski, 1985), the solution $\hat{\beta}$of the optimization problem depicted by Eq. (13) can be computed. The entries β_j in the LASSO estimator $\hat{\beta} = [\beta_1, \beta_2, ..., \beta_{2mn_s}]^T$ between the class label y and the feature F imply the contribution degree (CD) of the jth feature v_j. Since $\hat{\beta}$ is sparse

to some degree, meaning that most of the values in $\hat{\beta}$ are 0, we can classify the CDs of different features in the feature set F to the class label y. The variable CDs of diverse features is defined as: $CD_j = |\beta_j|$.

Step 3: Sort the CDs of features which meet $CD_j > 0$ in descending order in step 2. Then we choose the first k features. If the number of selected features is less than k, the penalty parameter λ will be adjusted.

Principal Component Analysis-Based (PCA) Feature Selection Algorithm

PCA is a statistical approach that is very effective in linear dimensionality reduction and feature selection. A set of variables that may be correlated is transformed into a set of linearly uncorrelated variables by orthogonal transformation. The transformed set of variables is called the principal component. The PCA-based algorithm is described as follows:

Step 1: Initialization as the step 1 of Section 2.3.1.
Step 2: Calculate the linear transformation W of the feature F given the type tag$y = [y_1, y_2, ..., y_{n_t}]^T, y_i = \{1, 2\}$.

The PCA can maximize the retention variance through a linear transformation$y = WF$. In other words, it finds the W by minimizing the reestablishment error. Each row vector of W points to the normalized orthogonal eigenvector computed from the signal covariance matrix. We can use singular value decomposition (SVD) as one simple approach to PCA.

Step 3: For dimensionality reduction, the column vectors of $U_F$corresponding to the k largest eigenvalues are chosen to form a final required transformation matrix W.

Stepwise Linear Discriminant Analysis-Based (SWLDA) Feature Selection Algorithm

The SWLDA is a commonly used algorithm in pattern recognition. Its main idea is to minimize the distance within the class and maximize the distance between the classes, to obtain the optimal projection direction to produce the best classification results. SWLDA mainly includes two parts, stepwise forward and backward analysis, and weighting the input variables using least squares regression to achieve classification of the type tag $\omega = \{1, 2\}$. The main flow of SWLDA-based method is as bellow:

Step 1: Initialize the features F and the selected features S as the step 1 of Section 2.3.1.

Create an initial function that can select features. This function calculates the significance (i.e., p-value) via the F-test to get the most significant feature that can predict the type tag. The condition that the feature is selected to enter S is set to $p < 0.1$, and the condition to remove from S is set to $p > 0.15$ (Krusienski et al., 2006).
Step 2: A new feature $v_j$is interpolated into the process. By performing the forward stepwise analysis, the largest statistically significant feature will be added into F if $p < 0.1$. By performing the backward stepwise analysis, the least statistically significant feature will be removed from S if $p > 0.15$.
Step 3: Repeat step 2 until the number of selected features in S reaches a predefined number or until the entry/removal

condition is not satisfied for all features in the model. We set the maximum number of significant features to k.

Classification of Optimal Spatial-Temporal Features

The fourth part introduces the model building and pattern recognition of the selected feature vectors. It is worth noting that because the selected CSP spatial filters appear in pairs, when using the proposed algorithm for feature selection, if a feature is chosen, its corresponding feature is also chosen. After performing feature selection in the section Feature selection, the selected feature vectors for training EEG signals are denoted as $F \in \Re^{N_t \times k}$ where n_t denotes the total number of training trials and k is set to 4. $k = 4$ means two pairs of feature vectors are selected, that is, the selected features appear in pairs. We set $k=4$ in this work because the number of features is also 4 (i.e., $2m = 4$) for the algorithm without feature selection, which is compared with the four feature selection-based algorithms in the Section Experimental Results.

The support vector machine (SVM) is used in this work since it has broad applications in classification. Plenty of BCI researches reported outstanding performance using SVM for classification. SVM finds a normal vector and a side-play amount of a discrimination hyperplane to separate the data from two classes by maximizing margins between two classes (Meng et al., 2013). LIBSVM is used as the classification tool in the current study (Chang and Lin, 2011).

Event-Related Spectral Perturbation

The event-related spectral perturbation (ERSP) is a common method that can be used to examine the spectral power changing law of the electroencephalogram from the view of time-frequency domain, which could show ERD/ERS phenomenon of diverse MI tasks. Variable ERSP with frequency and time is defined as follows:

$$ERSP(f, t) = \frac{1}{N} \sum_{i=1}^{N} (G_i(f, t)^2) \qquad (12)$$

where N denotes the total number of trials, and $G_i(f, t)$ denotes the spectral power for ith trial at a specific time (t) and frequency (f) (Delorme and Makeig, 2004). For each motor imagery trial, we calculate the average spectrum power from -3 to 5 s and between 1 and 35 Hz, where $t = 0$ denotes the time of cue onset. In this article, a two-dimensional plot of ERSP over time and frequency for the three main channels (i.e., C3, CZ, C4) is displayed and used for discussion.

In addition, topographical distribution was displayed to show the distribution of ERD/ERS in different regions of the brain during the execution of different motor imaging tasks. The averaged ERSP value was calculated within alpha band and imagination period (4 s). Since 60 EEG channels (except HEO, VEO, CB1 and CB2) were recorded in the dataset 3, the time-frequency and topographical figures were plotted using the EEG data from participant "l1," which achieved a relatively high classification accuracy.

RESULTS

Description of the Datasets

Dataset 1 (BCI Competition IV Dataset I) (Ang et al., 2012): The dataset comprised calibration and evaluation data from seven participants, including four healthy individuals (named "a," "b," "f," "g") and three artificially generated "participants" (named "c," "d," "e"). We only used the calibration data from each participant consisting of two runs, totaling 200 trials for two types of MI tasks where these two types of tasks come from left hand, right hand or foot. **Figure 2A** shows the timeline of every trial. First, a fixed cross for 2 s appears on the display to prompt the participants to prepare. Then, a 4 s arrow would appear on the display to prompt the participant to start the MI task. The directions of the arrows are left, right, or down, respectively, indicating left, right and foot imagination tasks. Last, the screen was all black for 2 s, indicating that the trial was over. In this process, a fixed cross appeared for a total of 6 s, as shown in the **Figure 2A**. EEG signals for 59 channels were recorded at 1,000 Hz sampling rate. These data were then bandpass filtered (0.05–200 Hz) and downsampled to 100 Hz. The details of the competition including ethical approval, and the data download website is as follows: http://www.bbci.de/competition/iv/.

Dataset 2 (BCI Competition III Dataset IVa) (Blankertz et al., 2006): The dataset included EEG signal from five healthy participants. Each participant was instructed to complete 280 trials. In each trial, the participants were asked to perform 3.5 s right-hand and foot motor imagery missions. Then participant
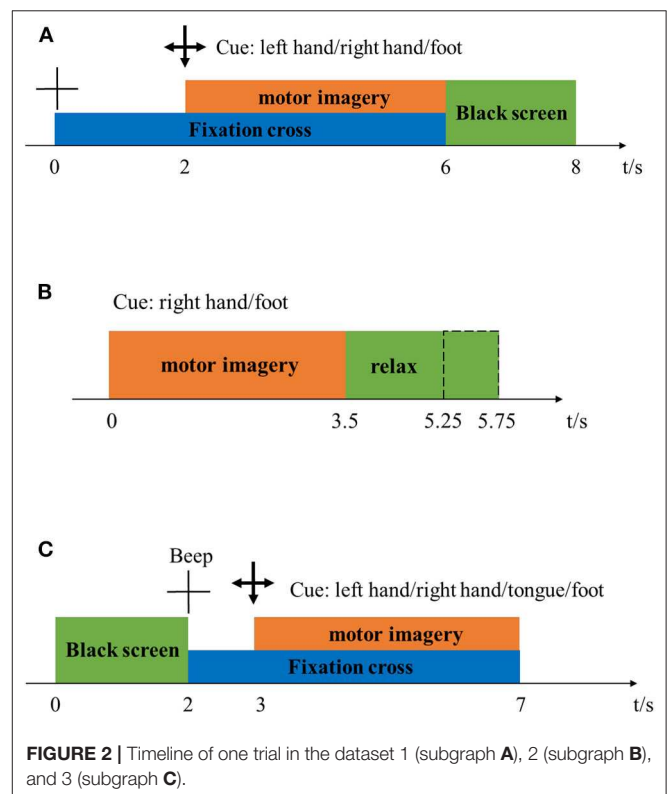

FIGURE 2 | Timeline of one trial in the dataset 1 (subgraph **A**), 2 (subgraph **B**), and 3 (subgraph **C**).

TABLE 1 | Accuracy (%) and significance comparisons of different methods applied on dataset 1, 2, and 3.

| Participant	Methods				
	CSP	MUIN	LASSO	PCA	SWLDA
a	55.5	87.5	86.5	78.0	84.5
b	66.0	82.5	83.0	78.0	82.0
c	77.5	92.0	92.0	66.0	87.0
d	90.5	96.5	98.0	93.5	97.5
e	92.5	100.0	100.0	98.0	98.5
f	85.5	91.5	91.0	90.5	91.5
g	54.5	82.0	79.5	76.0	79.0
Mean ± std	74.6 ± 14.8	90.3 ± 6.3	90.0 ± 7.0	82.9 ± 10.6	88.6 ± 7.0
aa	80.7	81.8	83.6	80.0	83.2
al	97.5	95.7	98.9	98.9	98.9
av	68.2	68.6	70.4	66.1	72.5
aw	95.7	96.8	97.1	96.4	96.8
ay	92.1	92.1	96.4	93.2	96.8
Mean ± std	86.9 ± 11.0	87.0 ± 10.6	89.3 ± 10.9	86.9 ± 12.3	89.6 ± 10.2
k3	85.6	93.3	93.9	92.2	91.7
k6	60.8	57.5	61.7	62.5	60.8
l1	90.0	95.8	96.7	95.8	94.2
Mean ± std	78.8 ± 12.8	82.2 ± 17.5	84.1 ± 15.9	83.5 ± 14.9	82.2 ± 15.2
p-value	–	0.0097	0.0016	0.048	0.002
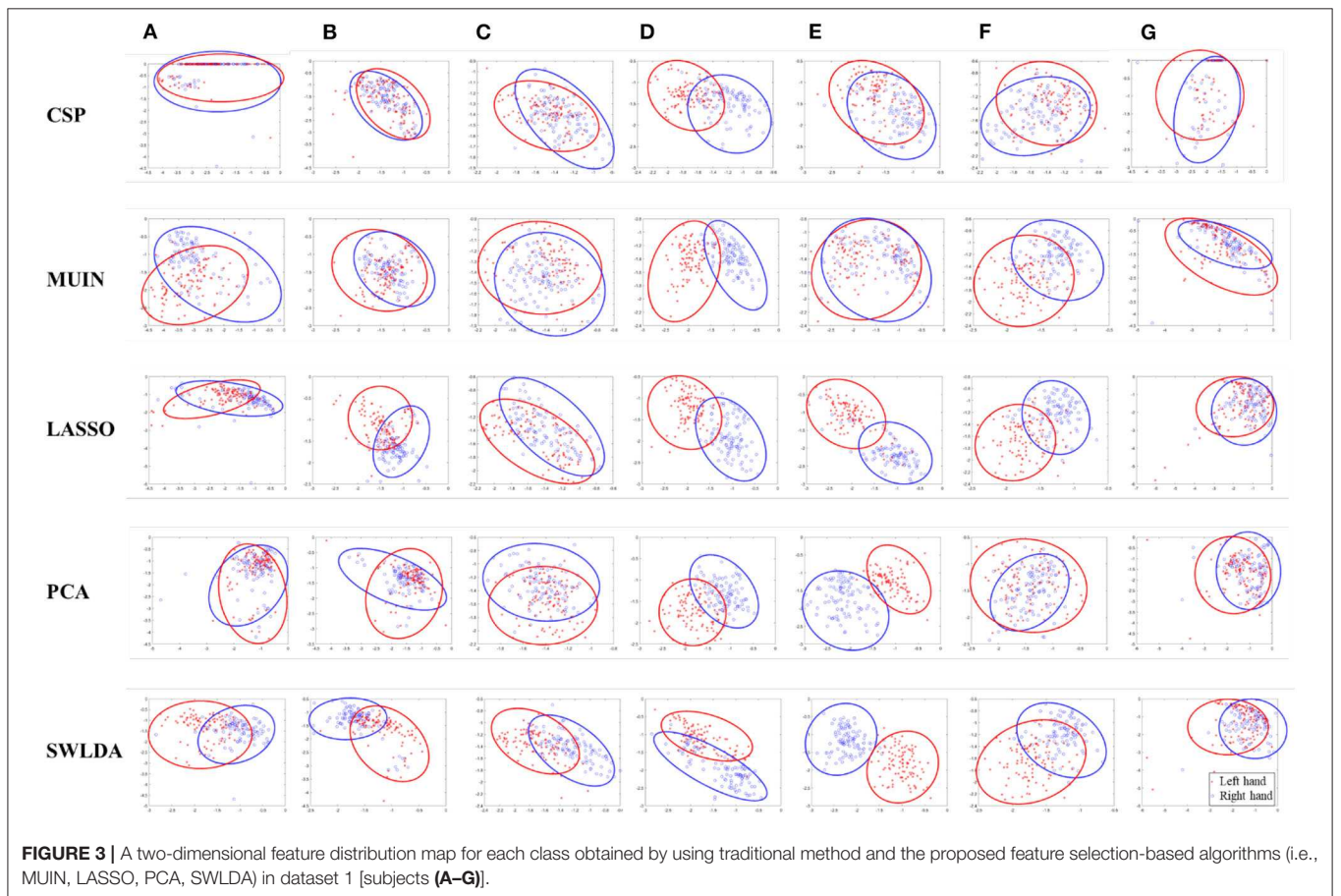
FIGURE 3 | A two-dimensional feature distribution map for each class obtained by using traditional method and the proposed feature selection-based algorithms (i.e., MUIN, LASSO, PCA, SWLDA) in dataset 1 [subjects **(A–G)**].

was instructed to relax for a period of changing length (see **Figure 2B**). EEG signals for 118 channels were recorded at 1,000 Hz sampling rate. These data were then bandpass filtered (0.05–200 Hz) and downsampled to 100 Hz. The details of the competition including ethical approval, and the data download website is as follows: http://www.bbci.de/competition/iii/.

Dataset 3 (BCI Competition III Dataset IIIa) (Blankertz et al., 2006): The dataset included EEG signals from three subjects labeled "k3," "k6," and "l1," who were instructed to complete 90, 60 and 60 trials, respectively. In each trial, the screen was first completely black for 2 s, during which subjects could rest and relax. Then a "beep" sound was issued, and a cross "+" appeared on the display to remind the participants to prepare for this trial. This process lasted for 1 s. From $t = 3$, a 4 s-long arrow appeared on the screen, prompting the subject to start performing motor imaging tasks. The directions of the arrows were left, right, up, or down, which represented the left, right, tongue, and foot motor imaging tasks (see **Figure 2C**). EEG signals for 60 channels were recorded at 250 Hz sampling rate. For the purpose of evaluating the feature selection methods of optimal temporal combinations, the EEG signals of left-hand and right-hand motor imagery missions were selected in the current study. The details of the competition including ethical approval, and the data download website is as follows: http://www.bbci.de/competition/iii/.

Experimental Results

For the above competition data in section Description of the datasets, we made a comparison from multiple angles on the performance among the proposed feature selection of optimal temporal combination patterns algorithms and the traditional CSP algorithm (denoted CSP). The fixed time windows (2–6 s for the dataset 1 and 3–6 s for the dataset 2 and 3 as shown in **Figure 2**) were employed for analyses. For the proposed algorithms, we used four feature selection-based methods: mutual information, least absolute shrinkage and selection operator, principal component analysis and stepwise linear discriminant analysis (denoted as MUIN, LASSO, PCA and SWLDA, respectively).

The experiment was performed by applying 10×10-fold cross-validation to avoid overfitting on the datasets 1–3. In each fold of this procedure, we optimized the temporal combination pattern and trained CSP spatial filters and classification model on the 90% training samples. The remain 10% samples were used as testing data to evaluate the performance of different algorithms. After finishing this procedure for all 10-folds, all data had been involved in the test set.

Table 1 shows the accuracy and significance comparisons of different methods (i.e., CSP, MUIN, LASSO, PCA and SWLDA algorithms) applied on dataset 1, 2, and 3. Experimental results show that classification accuracy is significantly enhanced by using the proposed algorithms compared with the traditional CSP method. Specifically, for three datasets, the proposed feature selection-based methods achieve higher classification accuracies. For seven subjects of dataset 1, the average accuracies are 74.6% (CSP), 90.3% (MUIN), 90.0% (LASSO), 82.9% (PCA), and 88.6% (SWLDA), respectively. For five subjects of dataset 2,

the average accuracies are 86.9% (CSP), 87.0% (MUIN), 89.3% (LASSO), 86.9% (PCA), and 89.6% (SWLDA), respectively. For three subjects of dataset 3, the average accuracies are 78.8% (CSP), 82.2% (MUIN), 84.1% (LASSO), 83.5% (PCA), and 82.2% (SWLDA), respectively. The results from three datasets confirm the superiority of the proposed algorithms over the CSP algorithm (paired t-tests, $p < 0.05$). Among the four feature selection-based methods, the results show that the LASSO yield best average classification accuracy (88.6%) across all the participants compared to MUIN, PCA, and SWLDA (87.6, 84.3,
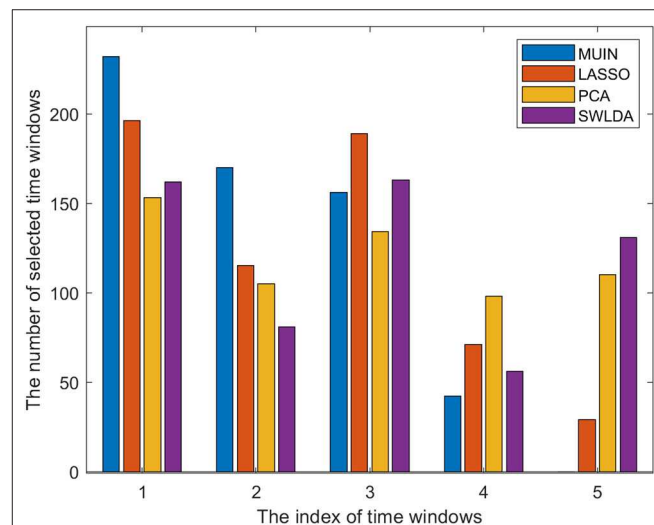

FIGURE 4 | The bar chart represents the total number of selected time windows for the proposed algorithms in datasets 1–3. (The meaning of time window indexes can be found in Section Multi-time segmenting and temporal band-pass filtering).

TABLE 2 | Ratio comparison of samples with selected features from different time windows to the total samples.

Participant	Methods			
	MUIN	**LASSO**	**PCA**	**SWLDA**
a	0.7	0.6	0.6	0.8
b	0.2	0.5	0.9	0.3
c	0.9	1	0.5	1
d	0	0.7	0.5	0.9
e	0.3	1	0.8	0.8
f	1	0	0.8	1
g	0.7	0.7	0.9	0.8
aa	0.7	1	0.9	1
al	1	0.8	1	0.5
av	0.9	1	0.7	0.5
aw	0.6	0.7	0.7	0.4
ay	1	1	0.7	1
k3	1	1	1	0.8
k6	0.7	1	0.5	0.7
l1	0.5	1	0.9	0.7
Mean ± std	0.68 ± 0.3	0.8 ± 0.27	0.76 ± 0. 17	0.75 ± 0.22

and 87.7%). By applying a paired sample t-test on the datasets 1–3, we find that this superiority is significant ($p = 0.03$, 0.04, and 0.046). But in these three methods (MUIN, PCA, and SWLDA), the classification accuracy between each pair is not significant ($p > 0.05$). The results from the three competition datasets confirm that the proposed algorithms are significantly better than the CSP algorithm in terms of classification accuracy. Using a paired t-test, the resulting $p < 0.05$.

To explain the superiority of the proposed algorithms, we compared the two-dimensional feature distribution between the methods made with and without feature selection. As shown in **Figure 3**, we depicted the two-feature distribution of each class in dataset 1 for seven participants "a"-"g." The red asterisks and blue circles represented the left and right feature types, respectively. For every subject, the sub-graph in the first row indicates that it was drawn directly using CSP, and the sub-graph in the bottom indicates that it was obtained using the proposed algorithms (i.e., MUIN, LASSO, PCA, SWLDA see section Methods). The results from **Figure 3** demonstrate that the features selected by the proposed algorithms are easier to distinguish and perform the pattern recognition, compared with the features obtained by traditional method.

We also calculated the total numbers of selected time windows based on the proposed methods for datasets 1–3 to determine whether the selected time windows varied in different methods (see **Figure 4**). The results indicate that the numbers of selected time windows for the proposed methods are different. Interestingly, the number of features selected by the four methods in the time window indexes 1–3 is more than that in the time window indexes 4–5, which indicates that the early period of the motor imagery task contributes more to the classification accuracy.

We further calculated the proportion of samples with selected features from different time windows to the total samples in datasets 1–3 via 10-fold cross-validation. The results of applying different approaches on the data of each participant are shown in the **Table 2**. The mean ratios are all more than 50%, which shows that the proportion of features selected from the combination of different time windows exceeds the proportion from the same time window for the four methods. More specifically, the ratio for method LASSO is the highest, which explains to some extent that the method LASSO achieves a relatively high average classification accuracy (88.6%). Note that both ratios of participant d in MUIN and participant f in LASSO are zero. It may be because the features from the same time window are sufficient for both cases to provide all the information needed for classification.

DISCUSSION

As we know, feature extraction plays an important role in motor imagery (MI)-based BCI studies (Park et al., 2013; Kevric and Subasi, 2017). As an effective spatial feature extraction algorithm, the CSP algorithm has been widely applied in MI-BCI related research fields and has achieved admirable results. Recently, some researches have improved and optimized the traditional CSP algorithm to solve the issue of the parameter setting in the time

segment of the EEG used. These solutions could be divided into the following two categories.

(1) Automatic selection for parameters. Time segment selection method based on correlation (Feng et al., 2018) is proposed to choose a subject-specific time window for CSP correlation analysis. Efficient wrapper-based methodology (German et al., 2013) is proposed for automatic selection of features computed in different time segments.

(2) Feature selection using information measure. For example, Fisher's common spatial pattern (FCSP) (Fattahi et al., 2013) uses the Fisher's criterion as an optimal function for estimating the spatial and spectral filters. Mutual information-based method (Ang et al., 2012) automatically optimizes the time windows and frequency ranges, by calculating the MUIN variable between the spatial and temporal features reflected by the EEG data and the activity of the corresponding micro-neurons. CSP-tangent space mapping (TSM) algorithm (Kumar et al., 2017) is proposed by utilizing Riemannian tangent space information for extracting features.

Although all the solutions improve the performance of the traditional CSP algorithm to varying degrees, none of them consider the temporal combination patterns during the MI task. From **Table 2** we can find that the feature vectors are selected from diverse time windows for most participants. Although the time during which the subject performs the task is known, the time when the brain activity is associated with the task is unknown and it may even be intermittent. Although we know the time when the subjects perform the motor imagery task, the time when ERD / ERS phenomenon occurs is unknown. Therefore, using a fixed single time period for data interception and pattern recognition does not obtain the best classification performance. Although some studies have begun to focus on the optimization of time segment (Ince et al., 2009; Higashi and Tanaka, 2013; Qiu et al., 2016; Feng et al., 2018), the combination of time windows is not considered.

In this study, the proposed feature selection-based method considers the temporal combination patterns among participants' data during motor imagery mission, and applies the feature selection algorithms to optimize the combination patterns of time windows for every participant. We then verified the effectiveness of the proposed methods using competition datasets and obtained significantly higher classification accuracy, compared with the CSP method (see **Table 1**).

As shown in **Figure 1**, the band-pass filtering plays an important role in the proposed algorithm and has a significant impact on the final result. Because of the heavy load of computation, it is difficult to employ multiple frequency bands for filtering the EEG measurements in optimization of the frequency-temporal feature. Therefore, a fixed range of band-pass frequencies is used to filter the EEG data. Theoretically, the interaction effect between the time window and the frequency range will affect the classification effect. Therefore, without optimizing the frequency range, optimizing only the time window may not get the best classification results (Xu et al., 2014). Thus, a fixed band-pass frequency range may not be the optimal band-pass filter setting to some extent. In subsequent research work, to solve this problem, we will try to incorporate the multiple frequency bands in the temporal

optimization to find optimal combination of frequency band and time segment. Maybe, different subjects have different optimal time segments. In the future, we will try to use other multi-time segments and optimize the multi-time segments for each subject to further improve the performance of the proposed method.

The time-frequency plots of different motor imaging tasks are shown in **Figure 5**, where blue areas indicate the ERD phenomenon. We used data from the C3, CZ, and C3 channels of the subject "l1". Left-hand and right-hand MI tasks are represented by LH and RH, respectively. As shown in **Figure 5**,

after the subjects perform the MI tasks for 500 ms, a clear and persistent ERD phenomenon appeared in the alpha band. Under two kinds of motor imagery tasks, the ERD phenomenon appears in the 8–9 Hz frequency band at the C4 channel, and the phenomenon between two tasks is not quite different. In contrast, the ERD phenomenon under the two tasks at the C3 channel is significantly different. Compared with right hand MI task, the ERD feature band of left-hand MI task is broader in alpha rhythm (9–11 Hz). In addition, the ERD in alpha rhythm is intermittent and instable (i.e., sometimes ERD is strong but sometimes is weak). This explains the
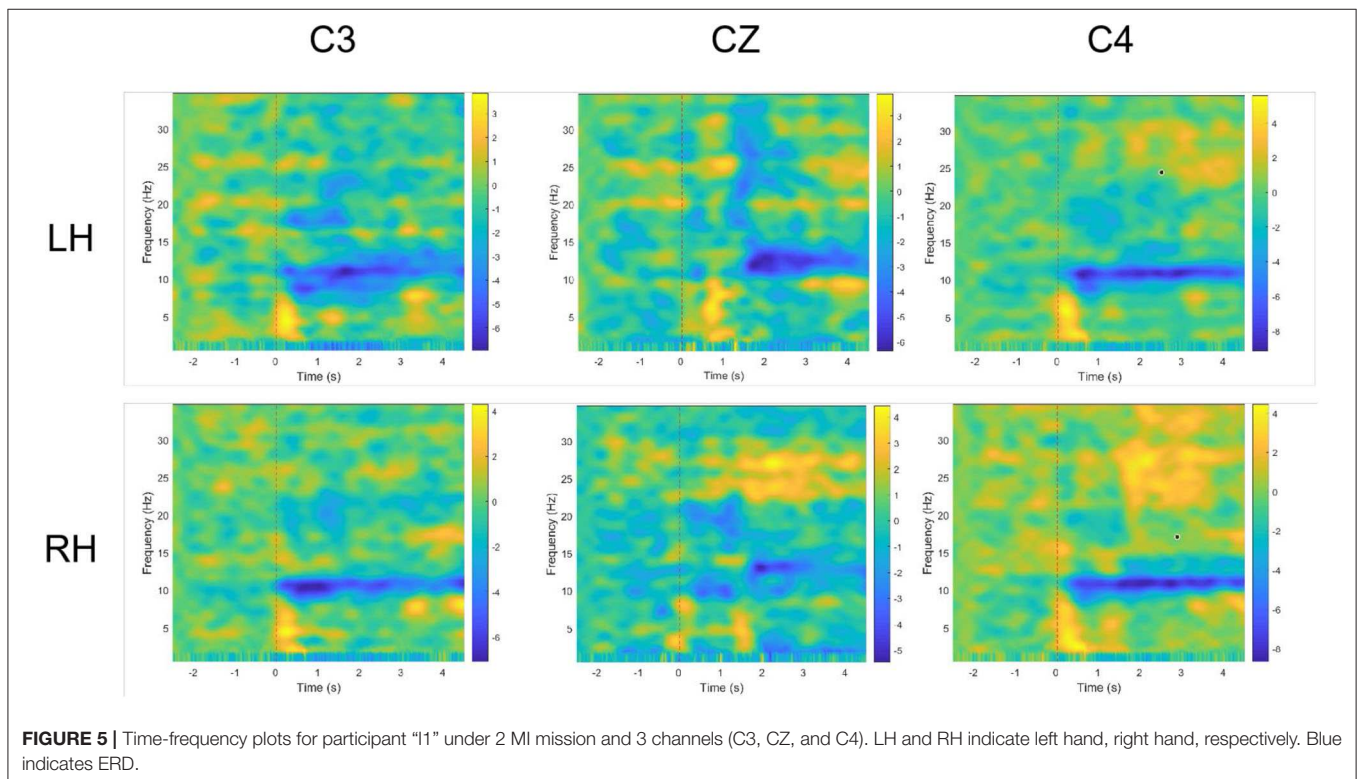

FIGURE 5 | Time-frequency plots for participant "l1" under 2 MI mission and 3 channels (C3, CZ, and C4). LH and RH indicate left hand, right hand, respectively. Blue indicates ERD.
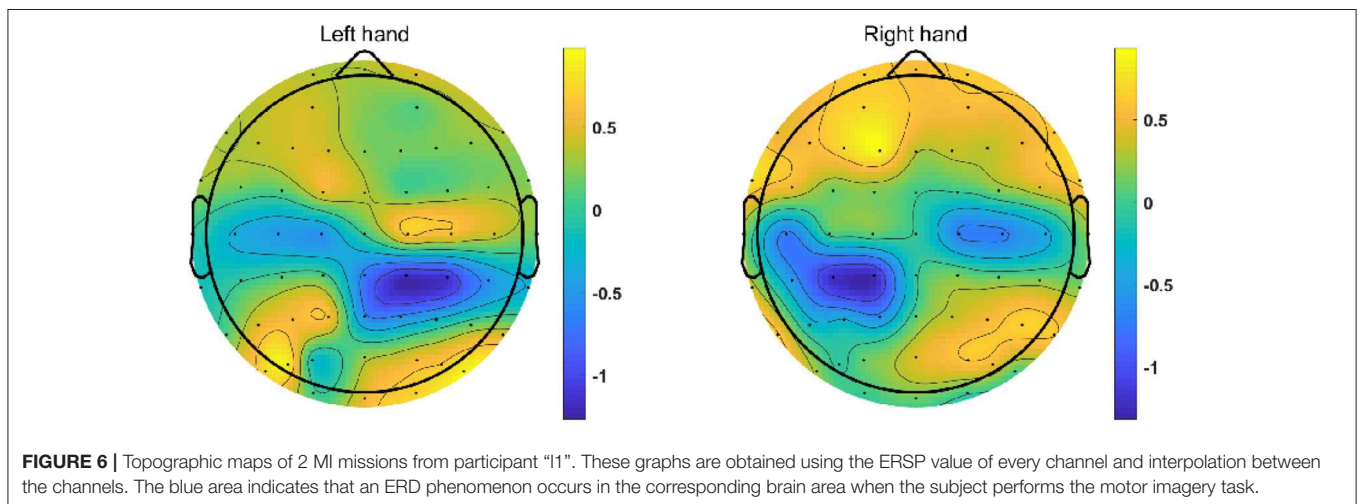

FIGURE 6 | Topographic maps of 2 MI missions from participant "l1". These graphs are obtained using the ERSP value of every channel and interpolation between the channels. The blue area indicates that an ERD phenomenon occurs in the corresponding brain area when the subject performs the motor imagery task.

rationality of temporal combination pattern optimization to some extent.

Figure 6 presents the topographical distributions of two mental tasks obtained from participant "l1". We can find a clear ERD phenomenon on near all C3 and C4 channels during both MI tasks, which means the ERD is mainly distributed in the sensorimotor region of the cerebral cortex corresponding to human limbs. Additionally, the ERD of left-hand MI task is stronger in right hemisphere compared to that in left hemisphere, and vice versa.

CONCLUSION

In this study, a novel feature selection-based method of optimal temporal combination patterns is proposed for MI-BCI systems. In our method, multiple time segments were obtained from each MI sample. After that, the features were extracted by CSP from each time segment and were combined to form a new feature vector. Finally, the four feature selection algorithms and the classification were applied to evaluate the effectiveness of the proposed method. The results from three competition datasets suggested that the proposed algorithms (i.e., MUIN, LASSO, PCA and SWLDA) could improve the performance compared to traditional feature extraction approach (i.e., CSP). Experimental results showed that the LASSO achieved the highest accuracy (88.58%) among the proposed methods. More specifically, the average classification accuracies using the proposed approaches significantly improved 10.14% (MUIN), 11.40% (LASSO), 6.08% (PCA), and 10.25% (SWLDA) compared to using CSP directly on the datasets 1–3 ($p < 0.05$). The proposed algorithms hold promise for practical applications in MI-based BCIs.

DATA AVAILABILITY STATEMENT

Information for existing publicly accessible datasets are contained within the article.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Conceptualization: JJ and CW. Methodology, data curation, and writing—original draft preparation: JJ. Software, formal analysis, and visualization: JW. Validation: JJ, CW, and JW. Investigation: WQ and MX. Resources: EY. Writing— review and editing: JJ, EY, and JW. Supervision: MX. Project administration: JJ and EY. Funding acquisition: EY. All authors contributed to the article and approved the submitted version.

FUNDING

REFERENCES

Ang, K. K., Chin, Z. Y., Zhang, H. H., and Guan, C. T. (2012). Mutual information of optimal spatial–temporal patterns for single-trial EEG-based BCIs. *Pattern Recognit.* 45, 2137–2144. doi: 10.1016/j.patcog.2011.04.018

Ang, K. K., and Guan, C. T. (2017). EEG-based strategies to detect motor imagery for control and rehabilitation. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 392–401. doi: 10.1109/TNSRE.2016.2646763

Arvaneh, M., Guan, C., Ang, K. K., and Quek, C. (2011). Optimizing the channel selection and classification accuracy in EEG-based BCI. *IEEE Trans. Biomed. Eng.* 58, 1865–1873. doi: 10.1109/TBME.2011.2131142

Blankertz, B., Muller, K. R., Krusienski, D. J., Schalk, G., Wolpaw, J. R., Schlogl, A., et al. (2006). The BCI competition III: validating alternative approaches to actual BCI problems. *IEEE Trans. Neural Syst. Rehabil. Eng.* 14, 153–159. doi: 10.1109/TNSRE.2006.875642

Chang, C. C., and Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. doi: 10.1145/1961189.1961199

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Fattahi, D., Nasihatkon, B., and Boostani, R. (2013). A general framework to estimate spatial and spatio-spectral filters for EEG signal classification. *Neurocomputing* 119, 165–174. doi: 10.1016/j.neucom.2013.03.044

Feng, J., Yin, E., Jin, J., Saab, R., Daly, I., Wang, X., et al. (2018). Towards correlation-based time window selection method for motor imagery BCIs. *Neural Netw.* 102, 87–95. doi: 10.1016/j.neunet.2018.02.011

German, R. B., Pedro, J. G., and Joaquín, R. D. (2013). Efficient automatic selection and combination of EEG features in least squares classifiers for motor imagery brain-computer interfaces. *Int. J. Neural Syst.* 23:1350015. doi: 10.1142/S0129065713500159

He, W., Wei, P., Wang, L., and Zou, Y. (2012). A novel EMD-based common spatial pattern for motor imagery brain-computer interface. In *IEEE-EMBS International Conference on Biomedical and Health Informatics* (Hong Kong; Shenzhen: IEEE), 216–219.

Herweg, A., Gutzeit, J., Kleih, S., and Kübler, A. (2016). Wheelchair control by elderly participants in a virtual environment with a brain-computer interface (BCI) and tactile stimulation. *Biol. Psychol* 121, 117–124. doi: 10.1016/j.biopsycho.2016.10.006

Higashi, H., and Tanaka, T. (2013). Common spatio–time–frequency patterns for motor imagery-based brain machine interfaces. *Comput. Intell. Neurosci.* 4:537218. doi: 10.1155/2013/537218

Hinterberger, T., Schmidt, S., Neumann, N., Mellinger, J., Blankertz, B., Curio, G., et al. (2004). Brain-computer communication and slow cortical potentials. *IEEE Trans. Biomed. Eng.* 51, 1011–1018. doi: 10.1109/TBME.2004.827067

Ince, N. F., Goksu, F., Tewfik, A. H., and Arica, S. (2009). Adapting subject specific motor imagery EEG patterns in space–time–frequency for a brain computer interface. *Biomed. Signal Process. Control* 4, 236–246. doi: 10.1016/j.bspc.2009.03.005

Jin, J., Sellers, E. W., Zhou, S., Zhang, Y., Wang, X., and Cichocki, A. (2015). A P300 brain–computer interface based on a modification of the mismatch negativity paradigm. *Int. J. Neural Syst.* 25:1550011. doi: 10.1142/S0129065715500112

Jin, J., Zhang, H., Daly, I., Wang, X., and Cichocki, A. (2017). An improved P300 pattern in BCI to catch user's attention. *J. Neural. Eng.* 14:036001. doi: 10.1088/1741-2552/aa6213

Kevric, J., and Subasi, A. (2017). Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system. *Biomed. Signal Process. Control* 31, 398–406. doi: 10.1016/j.bspc.2016.09.007

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., et al. (2006). A comparison of classification techniques for the P300 Speller. *J. Neural Eng.* 3, 299–305. doi: 10.1088/1741-2560/3/4/007

Kumar, S., Mamun, K., and Sharma, A. (2017). CSP-TSM: optimizing the performance of Riemannian tangent space mapping using common spatial pattern for MI-BCI. *Comput. Biol. Med.* 91, 231–242. doi: 10.1016/j.compbiomed.2017.10.025

Kwak, N., and Choi, C. H. (2002). Input feature selection by mutual information based on parzen window. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 1667–1671. doi: 10.1109/TPAMI.2002.1114861

Lugo, Z. R., Pokorny, C., Pellas, F., Noirhomme, Q., Laureys, S., Müller-Putz, G., et al. (2019). Mental imagery for brain–computer interface control and communication in non-responsive individuals. *Ann. Phys. Rehabil. Med.* 1252, 1–7. doi: 10.1016/j.rehab.2019.02.005

Meier, L., Geer, S., and Bühlmann, P. (2008). The group lasso for logistic regression. *J. R. Stat. Soc. B* 70, 53–71. doi: 10.1111/j.1467-9868.2007.00627.x

Meinshausen, N., and Yu, B. (2009). Lasso-type recovery of sparse representations for high dimensional data. *Ann. Stat.* 37, 246–270. doi: 10.1214/07-AOS582

Meng, J., Huang, G., Zhang, D., and Zhu, X. (2013). Optimizing spatial spectral patterns jointly with channel configuration for brain–computer interface. *Neurocomputing* 104, 115–126. doi: 10.1016/j.neucom.2012.11.004

Meng, J., Streitz, T., Gulachek, N., Suma, D., and He, B. (2018). Three-dimensional brain–computer interface control through simultaneous overt spatial attentional and motor imagery tasks. *IEEE Trans. Biomed. Eng.* 65, 2417–2427. doi: 10.1109/TBME.2018.2872855

Nicolas-Alonso, L. F., and Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors* 12, 1211–1279. doi: 10.3390/s120201211

Nigham, A., and Aggarwal, V. (2005). *The LPASSO Method for Regression Regularization.* Boston, MA: Technical Report, MIT.

Pan, J. H., Li, Y. Q., Zhang, R., Gu, Z. H., and Li, F. (2013). Discrimination between control and idle states in asynchronous SSVEP-based brain switches: a pseudo-key-based approach. *IEEE Trans. Neural Syst. Rehabil. Eng.* 21, 435–443. doi: 10.1109/TNSRE.2013.2253801

Park, S. A., Hwang, H. J., Lim, J. H., Choi, J. H., Jung, H. K., and Im, C. H. (2013). Evaluation of feature extraction methods for EEG-based brain–computer interfaces in terms of robustness to slight changes in electrode locations. *Med. Biol. Eng. Comput.* 51, 571–579. doi: 10.1007/s11517-012-1026-1

Qiu, Z., Allison, B. Z., Jin, J., Zhang, Y., Wang, X., Li, W., et al. (2017). Optimized motor imagery paradigm based on imagining Chinese characters writing movement. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 1009–1017. doi: 10.1109/TNSRE.2017.2655542

Qiu, Z., Jin, J., Lam, H. K., Zhang, Y., Wang, X., and Cichocki, A. (2016). Improved SFFS method for channel selection in motor imagery based BCI. *Neurocomputing* 207, 519–527. doi: 10.1016/j.neucom.2016.05.035

Ramoser, H., Muller-Gerking, J., and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* 8, 441–446. doi: 10.1109/86.895946

Ren, X., Ning, J., Chuang, L., Mrachacz-Kersting, N., Dremstrup, K., and Farina, D. (2014). Enhanced low-latency detection of motor intention from EEG for closed-loop brain-computer interface applications. *IEEE Trans. Biomed. Eng.* 61, 288–296. doi: 10.1109/TBME.2013.2294203

Schittkowski, K. (1985). NLQPL: a FORTRAN subroutine solving constrained nonlinear programming problems. *Ann. Oper. Res.* 11, 485–500. doi: 10.1007/BF02022087

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc.* 58, 267–288. doi: 10.1111/j.2517-6161.1996.tb02080.x

Velasco-Alvarez, F., Ron-Angevin, R., da Silva-Sauer, L., and Sancha-Ros, S. (2013). Audio-cued motor imagery-based brain-computer interface: navigation through virtual and real environments. *Neurocomputing* 121, 89–98. doi: 10.1016/j.neucom.2012.11.038

Wang, H., Zhang, Y., Waytowich, N. R., Krusienski, D. J., Zhou, G., Jin, J., et al. (2016). Discriminative feature extraction via multivariate linear regression

for SSVEP-based BCI. *IEEE Trans. Neural Syst. Rehabil. Eng.* 24, 532–541. doi: 10.1109/TNSRE.2016.2519350

Wang, Y., Berg, P., and Scherg, M. (1999). Common spatial subspace decomposition applied to analysis of brain responses under multiple task conditions: a simulation study. *Electroenceph. Clin. Neurophysiol.* 110, 604–614. doi: 10.1016/S1388-2457(98)00056-X

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clin. Neurophysiol.* 113, 767–791. doi: 10.1016/S1388-2457(02)00057-3

Xing, X., Wang, Y. J., Pei, W. H., Guo, X. H., Liu, Z. D., Wang, F., et al. (2018). A high-speed SSVEP-based BCI using dry EEG electrodes. *Sci. Rep.* 8:14708. doi: 10.1038/s41598-018-32283-8

Xu, M. P., Jia, Y. H., Qi, H. Z., Hu, Y., He, F., Zhao, X., et al. (2016). Use of a steady-state baseline to address evoked vs. oscillation models of visual evoked potential origin. *NeuroImage* 134, 204–212. doi: 10.1016/j.neuroimage.2016.03.073

Xu, M. P., Xiao, X. L., Wang, Y. J., Qi, H. Z., Jung, T. P., and Ming, D. (2018). A brain computer interface based on miniature-event-related potentials induced by very small lateral visual stimuli. *IEEE Trans. Biomed. Eng.* 65, 1166–1175. doi: 10.1109/TBME.2018.2799661

Xu, P., Liu, T., Zhang, R., Zhang, Y., and Yao, D. (2014). Using particle swarm to select frequency band and time interval for feature extraction of EEG based BCI. *Biomed. Signal Process. Control* 10, 289–295. doi: 10.1016/j.bspc.2013.08.012

Yin, E., Zeyl, T., Saab, R., Chau, T., Hu, D., and Zhou, Z. (2015). A hybrid brain–computer interface based on the fusion of P300 and SSVEP scores. *IEEE Trans. Neural Syst. Rehabil. Eng.* 23, 693–701. doi: 10.1109/TNSRE.2015.2403270

Yin, E., Zeyl, T., Saab, R., Hu, D., Zhou, Z., and Chau, T. (2016). An auditory-tactile visual saccade-independent P300 brain–computer interface. *Int. J. Neural. Syst.* 26:1650001. doi: 10.1142/S0129065716500015

Yin, E. W., Zhou, Z. T., Jiang, J., Chen, F. L., Liu, Y. D., and Hu, D. W. (2013). A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm. *J. Neural. Eng.* 10:026012. doi: 10.1088/1741-2560/10/2/026012

Yu, T. Y., Li, Y. Q., Long, J. Y., and Gu, Z. H. (2012). Surfing the internet with a BCI mouse. *J. Neural Eng.* 9:036012. doi: 10.1088/1741-2560/9/3/036012

Zhang, Y., Wang, Y., Jin, J., and Wang, X. Y. (2017). Sparse Bayesian learning for obtaining sparsity of EEG frequency bands based feature vectors in motor imagery classification. *Int. J. Neural Syst.* 27, 537–552. doi: 10.1142/S0129065716500325

Zhang, Y., Zhou, G., Jin, J., Wang, X., and Cichocki, A. (2015). Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface. *J. Neurosci. Methods* 255, 85–91. doi: 10.1016/j.jneumeth.2015.08.004

Zhang, Y. S., Xu, P., Cheng, K. W., and Yao, D. Z. (2014). Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface. *J. Neurosci. Methods* 221, 32–40. doi: 10.1016/j.jneumeth.2013.07.018

Zhang, Y. S., Yin, E. W., Li, F. L., Zhang, Y., Tanaka, T., Zhao, Q. B., et al. (2018). Two-stage frequency recognition method based on correlated component analysis for SSVEP-based BCI. *IEEE Trans. Neural Syst. Rehabil. Eng.* 26, 1314–1323. doi: 10.1109/TNSRE.2018.2848222