Relevant Feature Integration and Extraction for Single-Trial Motor Imagery Classification

Lili Li[1], Guanghua Xu[1,2]*, Feng Zhang[1], Jun Xie[1] and Min Li[1]

[1] School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China, [2] State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an, China

Brain computer interfaces provide a novel channel for the communication between brain and output devices. The effectiveness of the brain computer interface is based on the classification accuracy of single trial brain signals. The common spatial pattern (CSP) algorithm is believed to be an effective algorithm for the classification of single trial brain signals. As the amplitude feature for spatial projection applied by this algorithm is based on a broad frequency bandpass filter (mainly 5–30 Hz) in which the frequency band is often selected by experience, the CSP is sensitive to noise and the influence of other irrelevant information in the selected broad frequency band. In this paper, to improve the CSP, a novel relevant feature integration and extraction algorithm is proposed. Before projecting, we integrated the motor relevant information to suppress the interference of noise and irrelevant information, as well as to improve the spatial difference for projection. The algorithm was evaluated with public datasets. It showed significantly better classification performance with single trial electroencephalography (EEG) data, increasing by 6.8% compared with the CSP.

Keywords: classification, motor imagery, brain computer interface, single trial, feature extraction

INTRODUCTION

Brain-computer interface (BCI) is a way of communication that aims to provide a communication path between humans and computers. It directly translates brain activity into a series of control commands. Accordingly, it provides a non-muscular output channel for the brain and communicates with devices directly (Yu et al., 2014). This interface may offer disabled people a great prospect by solely translating their intentions that are reflected in their brain signals into actual instructions (Lemm et al., 2005). In addition, BCI can also be used as a neuro-rehabilitation tool to improve motor and/or cognitive performance of people after neurological diseases, such as stroke (van Dokkum et al., 2015) and tetraplegia (Vuckovic et al., 2015). In the BCI system, several modalities have been used for brain signal acquisition, such as electrocorticographic (ECoG) (Leuthardt et al., 2004), electroencephalography (EEG) (Bennet et al., 2016), magnetoencephalography (MEG) (Sardouie and Shamsollahi, 2012), functional magnetic resonance imaging (fMRI) (Ruiz et al., 2014), functional near-infrared spectroscopy (fNIRS) (Naseer and Hong, 2013, 2015; Hong et al., 2015; Naseer et al., 2016a,b), and intracortical neuronal spikes (Gupta et al., 2016). Among them, because of the real-time, low-cost, portable and noninvasive properties of EEG, it is one of the most convenient means to measure neurophysiological activity in the practical BCI application (Mihajlovic et al., 2015).

Electroencephalography (EEG) modulated by motor imagery (MI) is one of the most studied types of EEG signals of the BCI systems for the similarities of motor-related area involvement with motor execution (Pfurtscheller et al., 1997). MI can be revealed on brain activity patterns of the imagination of a motor action, but without its physical movement. During an MI task, the EEG activity is accompanied by an increase or decrease in the EEG magnitude which is known as an event-related synchronization or desynchronization (ERS/ERD). The ERD and ERS are non-phase-locked modulations of the EEG power, usually confined to a specific frequency band. ERD and ERS have been suggested to reflect the cortical activation and cortical deactivation (Hu et al., 2015). In particular, ERD of μ-rhythm (8–12 Hz) is usually associated with MI (Neuper and Pfurtscheller, 2001; ter Horst et al., 2013). BCI based on MI is an efficient path of rehabilitation, and it achieves excellent findings on complex movement (Qiu et al., 2017).

A big challenge for BCI based on motor imagery is to correctly and efficiently identify and extract subject-specific features from the blurred scalp EEG and translate those features into device commands (Wu et al., 2008). Based on topographic patterns, the Common Spatial Pattern (CSP) has been shown to be very efficient in the establishment of subject-specific discriminative spatial filters (Dornhege et al., 2006). The CSP algorithm decomposes multi-channel EEG from two classes into spatial patterns and enhances the separability between the two classes by diagonalizing the covariance matrix at the same time (Park et al., 2014). However, the conventional CSP algorithm selects multi-channel magnitude features on frequency band, which is selected by experience (Dornhege et al., 2006). As a result, it is sensitive to noise and the influence of other irrelevant information in the selected broad frequency band. Therefore, method for the optimization of the characteristics is urgently needed.

A noteworthy attempt, namely the Common Spatio-Spectral Pattern (CSSP) algorithm has been reported in Lemm et al. (2005). In the CSSP algorithm, the filter is constructed by the method of time-delay embedding. However, the CSSP algorithm limits the flexibility of the filters. The Common Sparse Spectral-Spatial Pattern (CSSSP) performs simultaneous optimization of an arbitrary Finite Impulse Response (FIR) filter (Dornhege et al., 2006). The spectral weighted common spatial pattern (SPEC-CSP) (Tomioka et al., 2006) optimizes the filter in the frequency domain and the spatial filter is an iterative procedure. But, this method is computationally expensive. The Filter Bank Common Spatial Patterns (FBCSP) (Ang et al., 2012) uses mutual information to select the optimal frequency band and time range. Xu applies particle swarm method to optimize frequency band and time interval (Xu et al., 2014). Local temporal correlation common spatial patterns employs local temporal information to estimate covariance matrices instead of Euclidean distance method of CSP (Zhang et al., 2013). The Regularizing Common Spatial Patterns (RCSP) adds a regularization algorithm to the CSP algorithm by a priori knowledge (Lotte and Guan, 2011). However, it does not consider the multivariable nature of the EEG signals, and thus it limits the feasibility of this method.

In this paper, an algorithm designated Spectral Component Common Spatial Pattern (SCCSP) is proposed. It provides a new approach to further improve the classification performance of the motor-imagery-based BCIs. To feature optimize, it focuses on the changes of the amplitude spectrum during motor imagery, and utilizes Independent Components Analysis (ICA) to extract the components from multi-channel amplitude spectrum with the aim of separating motor-relevant and irrelevant information from obscure EEG amplitude features applied by CSP. Accordingly, SCCSP could increase the classification accuracy of single-trial motor imagery EEG by improving the spatial difference of projecting.

DATA ACQUISITION AND CONFIGURATION

Two publically available datasets from BCI competitions were collected for the evaluation of the proposed algorithm for motor imagery. For the classification algorithm of CSP is the binary-class classification algorithm, two classes of motor imagery EEG data are collected from the two public datasets. The first public dataset recording the imagination left and right hands movement is collected from the publically available dataset BCI competition IV, dataset IIa (http://bbci.de/competition/iv/), including all 9 subjects. This dataset records EEG with twenty-two electrodes with a sampling rate of 250 Hz. Each trial (experiment) lasts 7.5 s. The subjects imaged movements from $t = 3$ s to $t = 6$ s in trials. Before this period, it is the period for preparation. The second public dataset is the dataset IIIa from the BCI competition III using a 60-channel amplifier with a sampling rate of 250 Hz, including all 3 subjects. The subjects imaged left and right hand movements from $t = 3$ s to $t = 7$ s in trials. Before this period, it is the period for preparation. Both of datasets were online filtered by a bandpass filter and a 50 Hz notchfilter to remove artifacts. A summary of the two datasets is presented in **Table 1**. The electrodes locations of two datasets are shown in **Figure 1**.

NEUROPHYSIOLOGICAL AND MATHEMATICAL METHODOLOGY

Feature Extraction and Integration

For each part of the human body, there exists a respective region in the primary motor cortex and somatosensory area

TABLE 1 | Summary of the datasets.

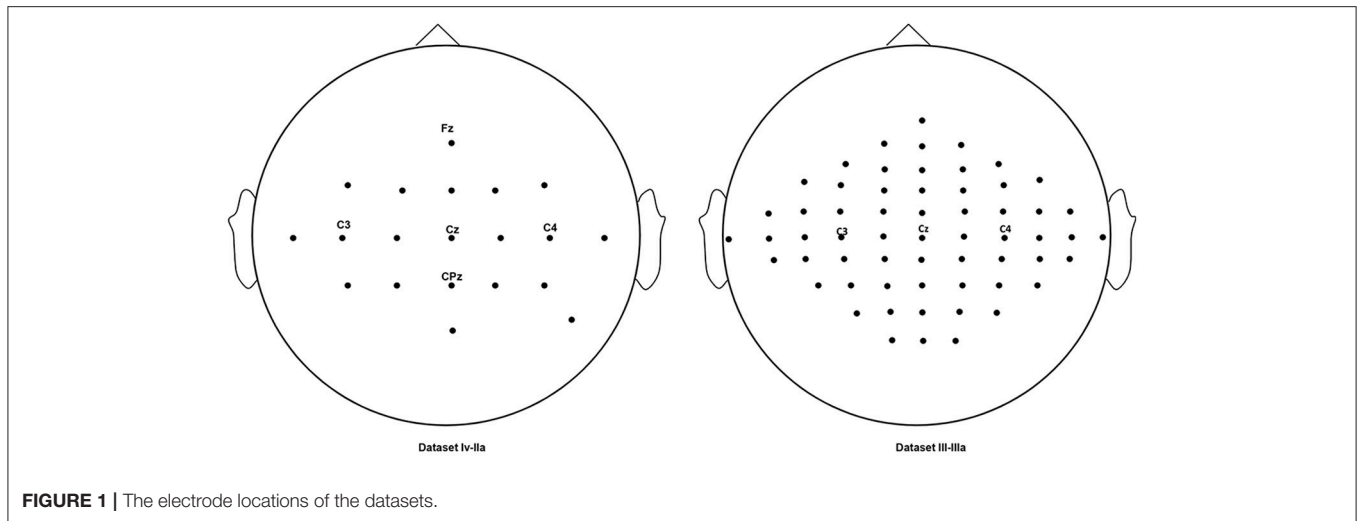| Dataset	Channels	MI type	Subjects	Subject number	Trials
IIa	22	Left vs. right hand	A01	1	138
			A02	2	136
			A03	3	137
			A04	4	129
			A05	5	129
			A06	6	113
			A07	7	133
			A08	8	132
			A09	9	116
IIIa	60	Left vs. right hand	k3b	10	90
			k6b	11	58
			l1b	12	60

FIGURE 1 | The electrode locations of the datasets.

of the neocortex (Chainay et al., 2004; Blankertz et al., 2008). The imaged part is surrounded by the other regions which represent other parts of the human body. Previous studies (Pei et al., 2005; Byblow et al., 2007) indicated that there was a parallel functional process between the lateral somatosensory area and the mid-central area during activation, indicating the independence of the hand and feet/leg areas during imagery. The inhibition mechanism was independent of the excitation mechanism on the somatosensory area (Ikeda et al., 2000). Accordingly, it is hypothesized that the area which represents the part of the imaged human body is independent of other areas which represent the parts of the un-imaged human body in the neocortex. However, the effect of volume conduction, EEG modulated by MI should be the combination of several independent components. Thus, it is urgent to source separation.

Independent Components Analysis (ICA) is a blind source separation method under the temporal information. It has emerged as a valuable signal processing method for the analysis of multivariate channel data (Woods et al., 2015). Let the time-varying observed signals be $X = [x_1(t), ..., x_m(t)]^T$, and the $S = [s_1(t), s_2(t)..., s_n(t)]^T$ $t = t_0, ..., T$, is matrix that contains unknown pure components, m and n indicate the channels of the observed signals and components, respectively. ICA assumes that the signal X is an instantaneous linear mixture of independent sources:

$$X = ES \tag{1}$$

where the matrix E of size $m \times n$ is the mixing matrix, whose component represents the linear memoryless mixing channels. To recover all the independent components (ICs) of the observed signals, ICA aims to obtain a de-mixing matrix W with minimal knowledge of E and S. The recovered signals $U = (u_1, u_2,..., u_n)^T$ are given by Equation 2 (Monakhova et al., 2015).

$$U = WX \tag{2}$$

Therefore, the ICA problem can be restated as the problem of finding W such that the sources of U are maximally independent.

We focus on the improvement of the classification accuracy based on the oscillatory feature (ERD/ERS). Motor imageries are accompanied by the ERD in specific frequency band (Pfurtscheller and Neuper, 1997), indicating an obvious sinking of the amplitude spectrum. To maximize the separability between classes, the feature extraction and integration algorithm is designed by integrating motor-related information. To suppress the influence of imagination irrelevant information and noise, we want to extract relevant information from blurred feature and integrate imagination related information on multi-channel dimensions into a single dimension. In this paper, we extended the conventional ICA algorithm to the frequency domain, and named it as Spectral Independent Components Analysis (SICA). It is hypothesized that the independent component, which is relevant to the imagination contains most of motor imagination information under the information theory. Information maximization algorithm of ICA (Hansen et al., 2001) was applied, and two independent components, imagination relevant information and imagination irrelevant information, were extracted with the SICA. Accordingly, in SICA, the Equation 2 is reconstructed as below:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \hat{W} X_f = \hat{W} \left[\left| \sum_{t=t_0}^{T} x_1 e^{-j\omega t} \right|, \left| \sum_{t=t_0}^{T} x_2 e^{-j\omega t} \right|, \bullet \bullet \bullet, \right.$$
$$\left. \left| \sum_{t=t_0}^{T} x_m e^{-j\omega t} \right| \right]^T \tag{3}$$

Where, the matrix of size $2 \times m$ is the separation matrix. The matrix X_f of size $m \times k$ is the amplitude spectrum matrix of multi-channel. k is the length of amplitude spectrum.

According to neurophysiological observations, when subjects engage in the unilateral limb imagination, large populations of neurons in the contralateral cortex will be excited, and the scalp EEG rhythm around 10 Hz (μ-rhythm) is significantly

suppressed. Namely, the cortex is activated (ERD) (Pfurtscheller and Neuper, 1997). This is a reliable feature of brain activity for BCIs based on motor imagery. For evaluating the effectiveness of SICA algorithm, the 6 channels simulation data of the adult MI EEG without any kinds of mental disease and damage were applied on the same hemisphere. Practically, the ERD often appears on several channels. To imitate this phenomenon, the μ-rhythm on two channels (the 5 and 6th channels) of the simulation data was suppressed. The simulation data of every channel was the sum of the sinusoidal signals with the frequency range from 0 to 20 Hz. The amplitude of the frequencies obtained a greater one on the low frequencies (simulation of real EEG nature), and the sum of maximum and minimum was under 12 uV. The frequency spectrum of simulation data on 6 channels is shown in the **Figure 2** after Fast Fourier Transform (FFT). SICA based on information maximization algorithm was used to extract the independent components from the frequency spectrum information of simulation data and the results are illustrated in **Figure 2** (component 1 and component 2). The results of the simulation data indicated that the μ-rhythm suppression or activation should be the criterion for the separation of independent components, and the μ-rhythm suppression information was integrated effectively and clearly. Further, SICA is an effective tool in the amplitude spectrum for feature extraction and integration of MI.

Projecting

The aim of CSP is the maximization of the difference between signals of two classes after feature extraction and integration in this study. $Y_k = [y_1(t), y_2(t), ..., y_p(t)]^T$ is defined as the kth time

domain feature after feature extraction and integration, where p is the number of ICs. The normalization covariance matrices C_1 and C_2 of the two classes are calculated using Equation 4. The covariance space $C = C_1 + C_2$ consists of the covariance matrices of the two classes. Whiten the matrix C and receive a matrix P as shown in Equation 5.

$$C_{1/2} = \frac{(Y_{1/2} \times Y_{1/2}^T)}{\text{trace}(Y_{1/2} \times Y_{1/2}^T)} \quad (4)$$

$$PCP^T = I \quad (5)$$

S_1 and S_2 are defined as $S_1 = PC_1P^T$ and $S_2 = PC_2P^T$, and then calculate the orthogonal matrix R and the diagonal matrix D by singular value decomposition.

$$S_i = RD_iR^T \quad (6)$$

Where, i = 1, 2, as $I = S_1 + S_2$, $D_2 = I - D_1$. Therefore, when the eigenvalue of S_i (i = 1, 2) is closer to I, the eigenvalue of the other S_i (i = 2, 1) is closer to **0**. The difference of the two classes is maximization. The filter is constructed by Equation 7.
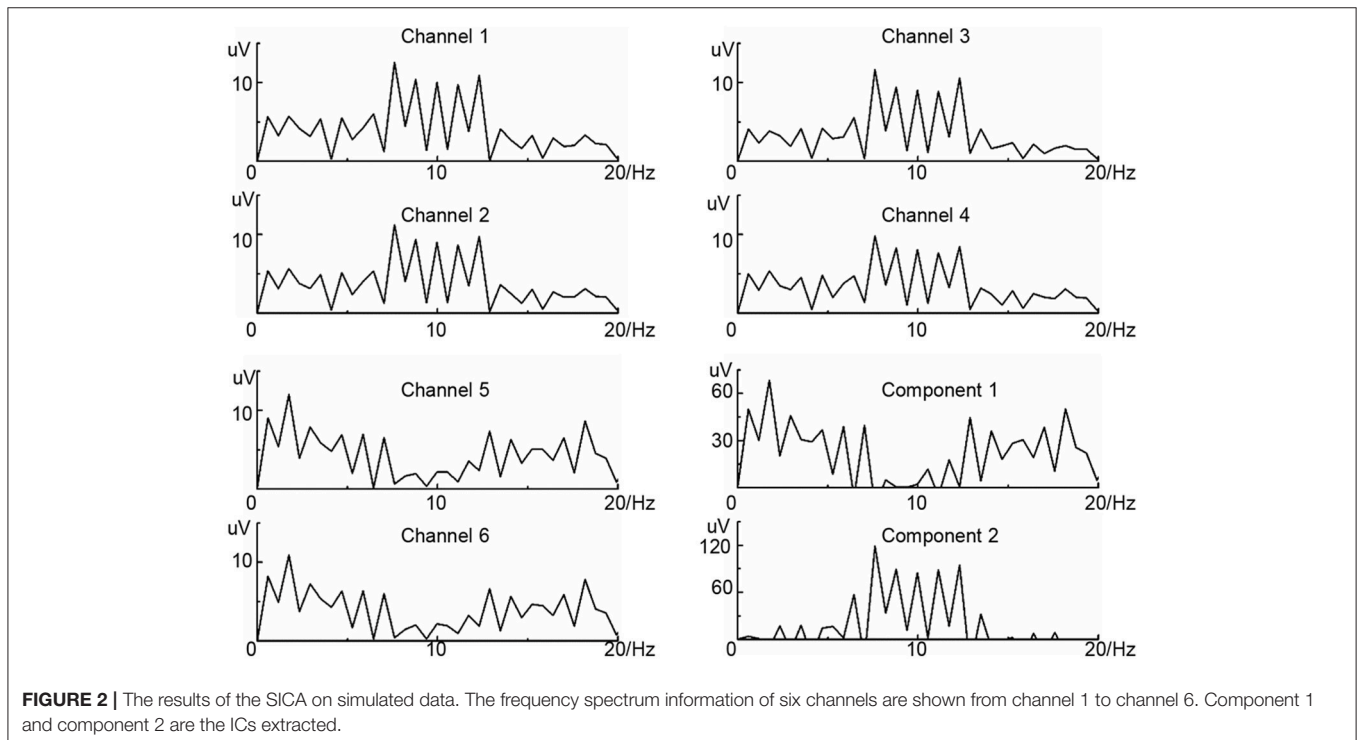
$$K = R^T P \quad (7)$$

$$Z = KY_k \quad (8)$$

The characteristic for the classifier is calculated by Equation 9.

$$f = \frac{\text{var}(Z_1)}{\text{var}(Z_1) + \text{var}(Z_2)} \quad (9)$$

Where, Z_1 and Z_2 are the projection of Y_k by the filters of two classes.

FIGURE 2 | The results of the SICA on simulated data. The frequency spectrum information of six channels are shown from channel 1 to channel 6. Component 1 and component 2 are the ICs extracted.

Data Processing

All trials were extracted from the two datasets with a bandpass filter of 5–30 Hz by a fourth-order Butterworth filter before analysis. $A_k = [a_1(t), a_2(t),..., a_g(t)]^T$ $t = t_0, ..., T$ was the kth EEG record, where g is the number of electrodes. To suppress the mutual interference of the hemispheres, and to extract and integrate the imagination relevant information by SICA; the EEG data were separated by hemisphere and named as A_k^l and A_k^r in every trial. After fast Fourier transform as illustrated by Equation 10, H_k^l and H_k^r were analyzed by SICA for the feature extraction and integration. Two independent components $U_1^{l(r)}$ and $U_2^{l(r)}$ which contained imagination relevant or irrelevant information were extracted over each hemisphere. In other words, the imagination relevant information was separated from irrelevant information and integrated together on each hemisphere. After inverse Fourier transform, four temporal components Y_1^l, Y_2^l, Y_1^r, and Y_2^r were rearranged as feature matrix according to hemisphere, and the component matrix $Y_k = [Y_1^l, Y_2^l, Y_1^r, Y_2^r]^T$ was projected. The flow chart of proposed method is illustrated in **Figure 3**.

$$H_k^{l(r)}(e^{-j\omega t}) = \left| \sum_{t=t_0}^{T} A_k^{l(r)}(t)e^{-j\omega t} \right| \tag{10}$$

Furthermore, whether the proposed method could provide better single trial classification capability than conventional CSP which relied on the bandpass filter was verified by the classification accuracy on twelve subjects of the public datasets. The conventional CSP only applied a bandpass filter from 5 to 30 Hz before projecting. Additionally, the results of the competing feature extraction and integration method, temporal ICA was also reported for comparison. The method named as ICA-CSP which extracted imagination relevant and irrelevant information by conventional temporal ICA before the components were projected. Similarly, ICA-CSP extracted four temporal components and rearranged them according to the hemisphere as SCCSP. The Analytic Common Spatial Patterns (ACSP), CSSSP and the Bilinear Common Spatial Pattern (BCSP) (Yu et al., 2013) and FBCSP were also studied for comparison. The parameters of the FBCSP were the same as the previously reported (Ang et al., 2012). After projecting, a classifier was adopted by LIBSVM (Chang and Lin, 2011) with Radial Basis Function (RBF) by the algorithms. The training and test trials did not overlap on every subject. The numbers of the training

and testing trials were half of the whole trials for every subject. The classification performance was evaluated by classification accuracy which is the ratio between the correct number after the classifier and the sum of trials. K-fold cross-validation was applied as cross-validation. The number of K was half of trials in every subject to make sure that every data could be used as the training data and testing data once. K was higher than 10 in all subject. The Lilliefors test was used to evaluate results if they obeyed normal distribution. One-way Analysis of Variance (ANOVA) with repeated measures was applied for statistical analysis of results, and pair t-test and least significant difference were used as a *post-hoc* test methods. All calculations were performed in MATLAB.

RESULTS

Figure 4 shows the ERD/ERS maps at 5–15 Hz of the fifth subject from the dataset IIa during the left hand MI. It indicated that the μ-suppression appeared on several contralateral electrodes. The classification accuracies of the six methods are presented after cross validation in **Table 2**. They showed that the SCCSP outperformed CSP, ICA-CSP, CSSSP, BCSP and ACSP, achieving 6.8, 3.5, 11.5, 26, and 15.5% higher average classification accuracy than these algorithms, respectively. Among the 12 subjects, SCCSP showed better performance than CSP in 10 subjects. The Lilliefors test showed that the classification accuracy from six algorithms obeyed the normal distribution. The probabilities were 0.1852, 0.5, 0.3136, 0.2141, 0.5, and 0.3909 for the CSP, SCCSP, ICA-CSP, CSSSP, BCSP, and ACSP, respectively. ANOVA indicated that there was significant difference among the six algorithms [$F_{(1, 72)} = 8.53$, $p < 0.001$]. Moreover, the paired t-test showed that the better performances of SCCSP over CSP ($p < 0.05$), ICA-CSP ($p < 0.05$), CSSSP ($p < 0.05$), BCSP ($p < 0.001$) and ACSP ($p < 0.001$) were significant. Least significant difference, used as *post-hoc* test, showed that the better performances of SCCSP over CSSSP, BCSP, and ACSP were significant at 0.05 level. Additionally, ICA-CSP achieved 3.3% higher average classification accuracy than CSP. The kappa value was also applied to evaluate the consistency of classification performance.
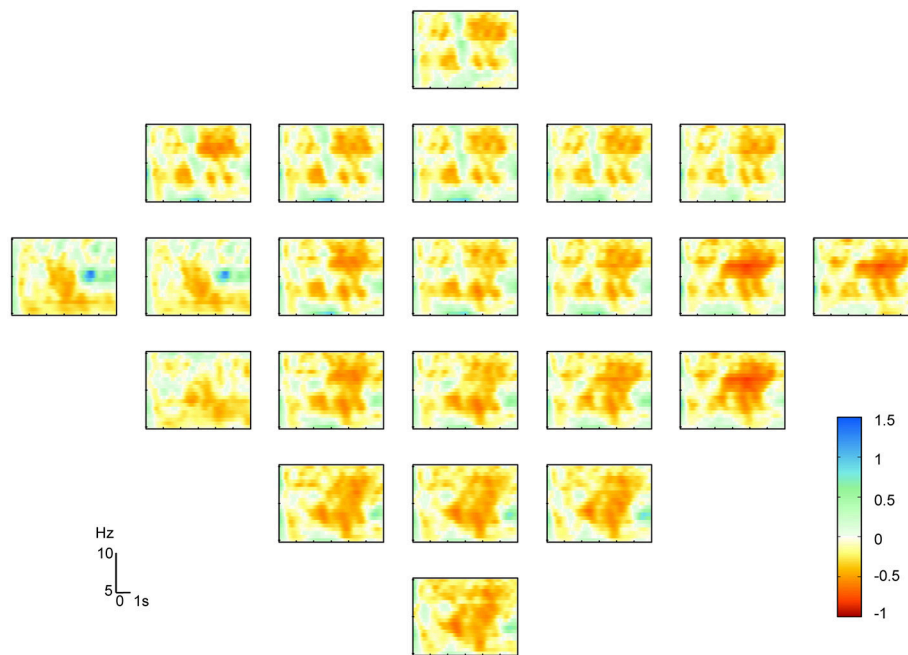
$$kappa = \frac{p_o - p_e}{1 - p_e} \tag{11}$$

FIGURE 3 | The flow chart for data processing.
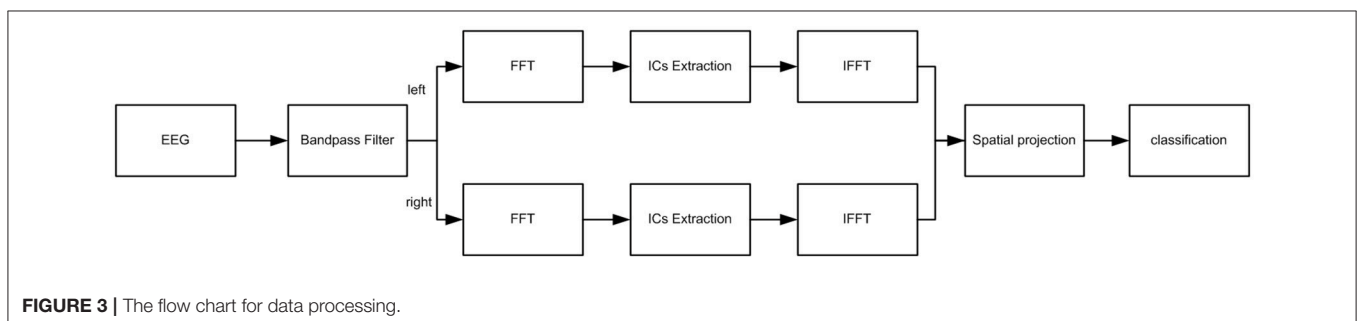
FIGURE 4 | The ERD/ERS maps of subject five on left hand motor imagery. X-axis and Y-axis represent time and frequency, respectively.

TABLE 2 | Classification accuracies (%) of subjects.

Dataset	Subjects	CSP Mean (std)	SCCSP Mean (std)	ICA-CSP Mean (std)	CSSSP (Yu et al., 2013)	BCSP (Yu et al., 2013)	ACSP (Yu et al., 2013)
IIa	A01	84.3 (1.58)	87.1 (2.63)	85.2 (3.05)	86.1	70.8	90.2
	A02	79.4 (3.27)	86.8 (3.62)	81.0 (2.63)	52.0	50	52.0
	A03	82.4 (2.89)	89.7 (1.39)	89.5 (3.56)	86.1	61.8	95.1
	A04	82.3 (3.15)	83.9 (1.94)	81.1 (2.73)	65.9	55.5	69.4
	A05	89.1 (2.34)	90.6 (2.75)	86.8 (1.99)	68.0	49.3	56.9
	A06	83.9 (2.78)	83.9 (3.11)	91.9 (4.53)	66.6	56.2	70.1
	A07	74.2 (2.92)	86.4 (2.84)	80.5 (3.50)	75.0	57.6	78.4
	A08	83.3 (2.59)	89.4 (2.70)	89.0 (3.15)	95.1	63.1	97.2
	A09	74.1 (2.22)	96.3 (8.24)	76.8 (4.99)	93.0	76.3	91.6
IIIa	k3b	93.5 (2.76)	91.3 (2.01)	87.5 (2.50)	95.5	78.8	76.6
	k6b	80.0 (5.96)	90.0 (2.97)	80.0 (4.39)	55.1	63.7	56.8
	l1b	83.3 (2.98)	96.7 (6.05)	100.0 (2.57)	95.0	76.6	51.6
Average		82.5 (3.0)	89.3 (3.4)	85.8 (3.3)	77.8 (16.0)	63.3 (10.3)	73.8 (17.0)

Where, p_o is the classification accuracy; p_e denotes the probability of expected agreement. The results of the kappa values are listed in **Table 3**. The SCCSP outperformed CSP, ICA-CSP and FBCSP, achieving 0.247, 0.094, and 0.109 higher average kappa value than these algorithms, respectively. The Lilliefors test showed that the kappa value from these algorithms followed the normal distribution. The probabilities were 0.5, 0.3573, 0.5, and 0.076 for the CSP, SCCSP, ICA-CSP, and FBCSP, respectively. The ANOVA indicated that there was significant difference among these algorithms [$F_{(1, 36)} = 5.99, p < 0.05$] in the kappa value. The paired t-test showed that the better performances of the SCCSP

over CSP ($p < 0.001$) and ICA-CSP ($p < 0.05$) were significant. The better performance of the ICA-CSP over CSP ($p < 0.001$) was significant. Moreover, the probability of SCCSP performance over FBCSP was 0.08. Least significant difference, used as *post-hoc* test, showed that the better performance of SCCSP over CSP was significant at 0.05 level.

The feature extraction and integration result of subject 5 is presented in **Figure 5**. The result presented in **Figure 5A** shows the topographical view of the average time-frequency representation of ERD/ERS values in μ-rhythm during hand imagery. The result in **Figure 5B** shows the filtered result by

the bandpass filter in CSP. **Figure 5C** reveals the result obtained by feature extraction and integration algorithm proposed where the components were converted by matrix *W*. The result in **Figure 5A** was consistent with the fact that the EEG suppressions were contralateral to the imagined hand movement (Pfurtscheller and da Silva, 1999).

To study the stability of the SCCSP, the number of the trials for training the classifier was varied from 2 to 50 with about
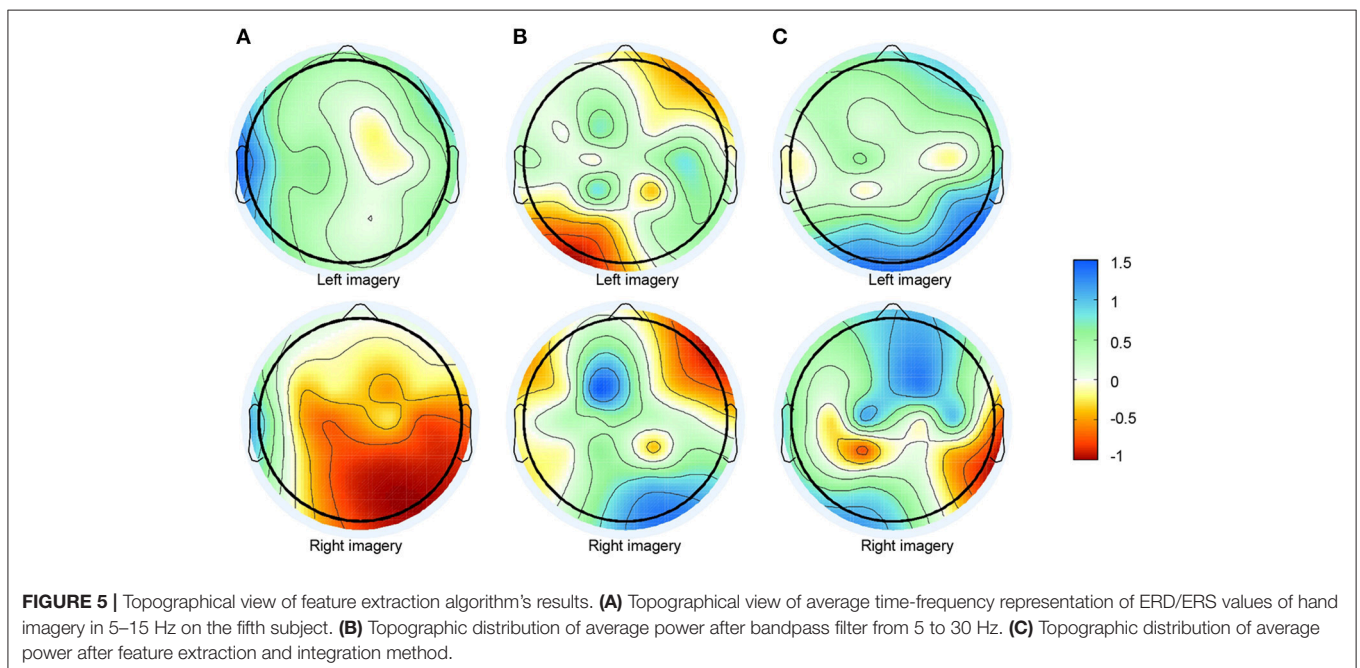
10 steps. The results of classification accuracy with error bar for every step are presented in **Figure 6** after cross validation. The average classification accuracy and standard deviation of accuracy of the SCCSP and CSP were calculated and are shown in **Table 4**. The SCCSP achieved 12.1% higher average accuracy than CSP. The Lilliefors test shows that the average classification accuracy from these algorithms obeyed the normal distribution. The probabilities were 0.5 and 0.5 for CSP and SCCSP, respectively. The ANOVA results indicated that there was significant difference between these algorithms [$F_{(1, 24)} = 35.97$, $p < 0.001$] for classification accuracy. The paired *t*-test showed that the better performance of SCCSP over CSP was significant ($p < 0.001$). Furthermore, SCCSP had a smaller average standard deviation of classification accuracy than CSP. A classification of the *f* in Equation 9 of subject 5 is shown in **Figure 7** for visualization. The statistical results of the *f* under the SCCSP and CSP are shown in **Figure 8**. A paired *t*-test analysis showed that the SCCSP achieved a higher difference between the two classes than CSP ($p < 0.05$). For quantitative analysis, the within-class distance *B* and between-class distance *D* were applied.

TABLE 3 | Kappa scores of BCI competition IV dataset IIa.

Dataset	Subjects	CSP Mean (std)	SCCSP Mean (std)	ICA-CSP Mean (std)	FBCSP (Ang et al., 2011)
IIa	A01	0.556 (0.0316)	0.664 (0.0526)	0.687 (0.0579)	0.747
	A02	0.599 (0.0654)	0.776 (0.0724)	0.689 (0.0359)	0.416
	A03	0.539 (0.0579)	0.776 (0.0277)	0.560 (0.0712)	0.824
	A04	0.419 (0.0629)	0.732 (0.0388)	0.771 (0.0546)	0.400
	A05	0.656 (0.0469)	0.838 (0.0549)	0.742 (0.0399)	0.608
	A06	0.490 (0.0556)	0.701 (0.0622)	0.607 (0.0906)	0.309
	A07	0.430 (0.0585)	0.758 (0.0568)	0.668 (0.0700)	0.849
	A08	0.411 (0.0518)	0.735 (0.0540)	0.615 (0.0631)	0.787
	A09	0.372 (0.0455)	0.717 (0.1648)	0.507 (0.0998)	0.772
Average		0.497 (0.0528)	0.744 (0.0649)	0.650 (0.0648)	0.635 (0.208)

$$B = \frac{1}{M} \sum_{k \in C_1/C_2} \|d_k\| \tag{12}$$

$$D(C_1 C_2) = \frac{1}{N_1 N_2} \sum_{i \in C_1} \sum_{j \in C_2} \|d_{ij}\| \tag{13}$$

$$\lambda = \frac{B_1 + B_2}{D} \tag{14}$$

Where, $\| d_k \|$ denotes the Euclidean distance between the f_k and the gravity in C_1 or C_2. $\| d_{ij} \|$ is the Euclidean distance between the f_i in C_1 and the f_j in C_2. To evaluate the difference between the two classes, the ratio of within-class distance and

FIGURE 5 | Topographical view of feature extraction algorithm's results. **(A)** Topographical view of average time-frequency representation of ERD/ERS values of hand imagery in 5–15 Hz on the fifth subject. **(B)** Topographic distribution of average power after bandpass filter from 5 to 30 Hz. **(C)** Topographic distribution of average power after feature extraction and integration method.
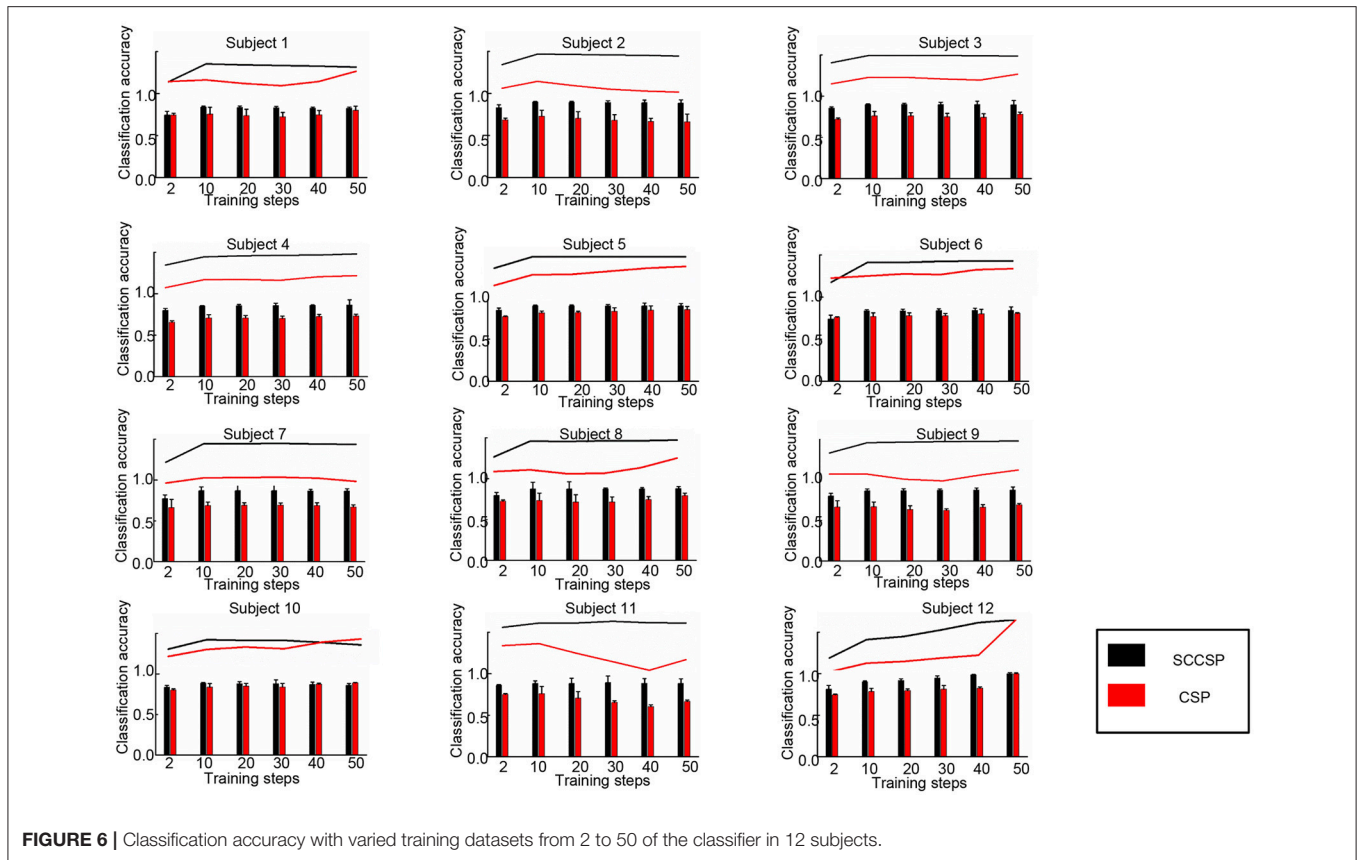
FIGURE 6 | Classification accuracy with varied training datasets from 2 to 50 of the classifier in 12 subjects.

TABLE 4 | Average classification accuracies (%) and standard deviation of accuracy (%) of the datasets IIa and IIIa in different steps.
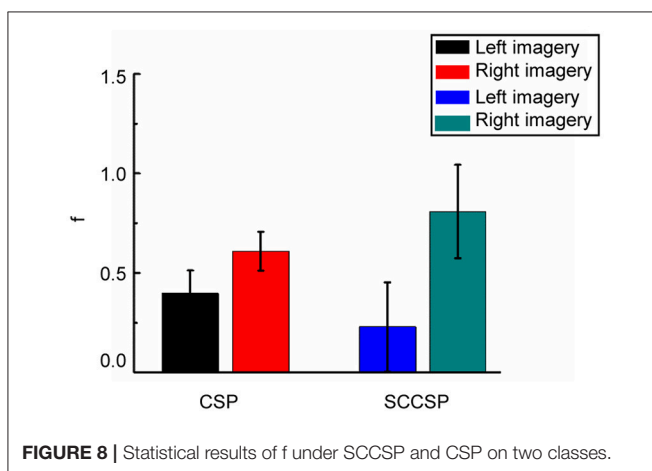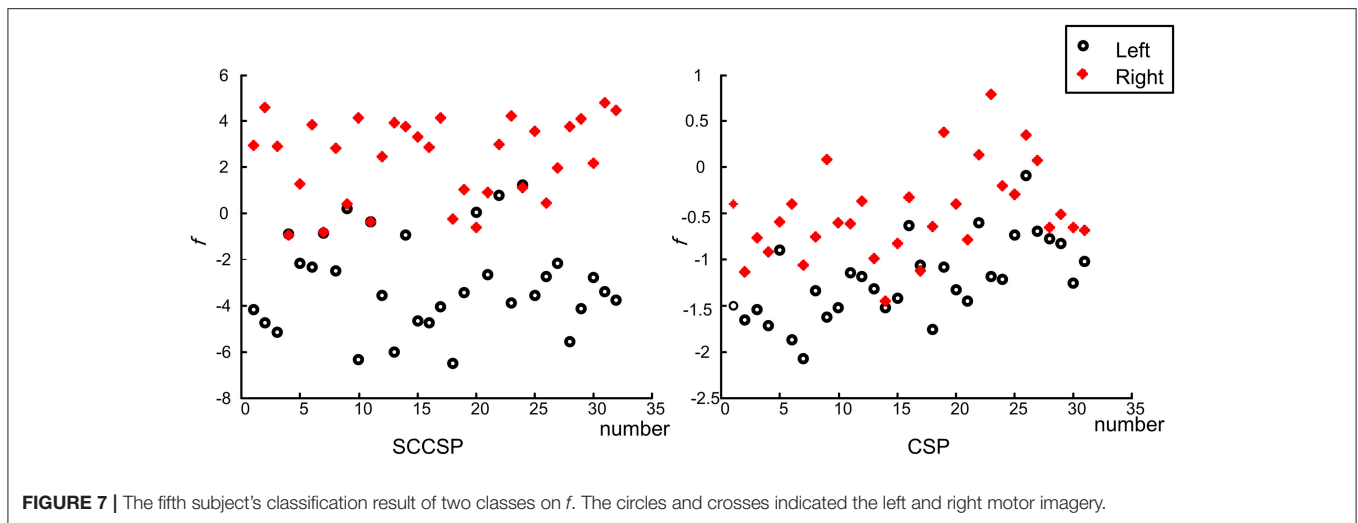
Dataset	Subjects	CSP Mean (std)	SCCSP Mean (std)
IIa	A01	75.2 (2.7)	81.7 (3.6)
	A02	69.0 (2.5)	88.3 (2.5)
	A03	75.3 (2.0)	89.1 (1.8)
	A04	70.48 (2.7)	84.8 (2.5)
	A05	82.1 (3.0)	88.8 (2.1)
	A06	78.3 (1.7)	82.3 (4.1)
	A07	68.3 (1.3)	85.2 (3.9)
	A08	73.8 (2.9)	86.4 (3.3)
	A09	65.9 (2.4)	86.0 (2.9)
IIIa	k3b	82.8 (8.9)	92.9 (6.8)
	k6b	69.0 (6.0)	88.2 (1.2)
	l1b	84.8 (3.0)	86.9 (1.8)
Average		74.6 (3.3)	86.7 (3.0)

between-class distance λ is derived using Equation 14. A lower λ indicated a greater separability between classes. As a result, the SCCSP achieved nearly twenty times reduction of λ compared to the conventional CSP algorithm, on average.

DISCUSSION AND CONCLUSION

Before the onset of motor imagery and execution, somatosensory area which is a part of the posterior parietal lobe needed some information, such as location, which comes from proprioception and visual area, etc. The prefrontal lobe and posterior parietal lobe determine and control movements. The axons of the prefrontal lobe and posterior parietal lobe concentrate on the Brodmann area 6 of which including the Supplementary Motor Area (SMA) and the Premotor Area (PMA). Most of the corticospinal tracts connect with the efferent fibers of the Brodmann area 6 which encodes the movement and primary motor cortex. The independence discussion of the inhibition mechanism and excitation mechanism on different motor-function area of somatosensory area (Ikeda et al., 2000; Pei et al., 2005) provided a great possibility of activation independence on function areas which represent different parts of body in the primary motor cortex. This reveals that EEG of one-task mental motor imagery should be the combination of time and spatial independent sources on motor-related areas.

In this paper, we extended the temporal ICA to amplitude spectrum analysis. A novel SCCSP algorithm for motor imagery classification based on SICA was proposed. This SCCSP method provided greater classification accuracy than CSP, ICA-CSP, CSSSP, BCSP, and ACSP. The kappa results also exhibited a better performance than CSP, ICS-CSP, and FBCSP. SICA is the extension of blind source separation. Therefore, the

FIGURE 7 | The fifth subject's classification result of two classes on *f*. The circles and crosses indicated the left and right motor imagery.

FIGURE 8 | Statistical results of f under SCCSP and CSP on two classes.

better classification performances of SCCSP and ICA-CSP may indicate that the time-frequency independence nature of motor-related sources in this experiment. Moreover, the greater average classification accuracy of SCCSP than ICA-CSP may show a possibility of greater separability or independence on frequency domain. In practice, the channels which reveal μ-suppression varied with trials. For the volume conduction, the suppression appeared in a wide region. This was a challenge to improve the spatial separability of the features. However, the algorithms for projecting were sensitive to the arrangement of feature, spatial distribution. Under SCCSP, a feature extraction and integration method based on SICA was applied. This method can extract the relevant imagination information into one component. That is, the integration of the feature algorithm could separate the motor relevant information from blurred data on multi-channel, concentrate relevant feature, suppress the influence of other region which represent other un-imaged parts of the body, and noise, and enlarge the spatial distribution separability of the features. The pure bandpass filter applied by CSP only suppressed the interference of other frequencies, while the information

of other irrelevant function areas and noise remained in the frequency band selected. The results presented in **Figures 5B,C** illustrate that the proposed algorithm obtained a greater spatial separability, while the information extracted by bandpass filter was obscure. The greater spatial separability extracted by the feature extraction and integration algorithm was favorable for improving the classification accuracy. Therefore, this SCCSP can reduce the interferences both in the other frequency bands and in the frequency band selected to improve classification accuracies. Moreover, the results of SCCSP and spatio-spectral filter selection method by cognitive fuzzy inference system (SCIF) (Das et al., 2016) indicated that there was 3.3% accuracy increase of SCCSP over SCIF on dataset IIa. Though there were different datasets, it was comparability. SCCSP achieved 5.18% higher average accuracy than the dynamic frequency feature selection method mentioned in Luo et al. (2016). Therefore, SCCSP achieve a better performance on MI classification.

SCCSP provided a lower average standard deviation of classification accuracy than almost all other methods. The statistical results of classification accuracy and kappa values indicated that the feature extraction and integration of SCCSP should be individual variability and adaptability. That is, SCCSP can decrease the individual difference. In **Table 3**, ICA-CSP and SCCSP both achieved significantly higher performances regarding the kappa value. Therefore, ICA is an efficient feature extraction algorithm to improve the spatial separability of features. The results presented in **Figure 6** and **Table 4** illustrated that SCCSP achieves greater average classification accuracy and a smaller standard deviation compared with CSP, simultaneously. The curve of the classification accuracy on SCCSP was steadier than CSP, and it obtained greater classification accuracy under the small training dataset. Therefore, SCCSP was less affected by the number of training datasets. This is very important for BCI application. **Figure 7**, **8** illustrated that the SCCSP algorithm obtained a greater separability between classes after classifier. The statistical results of λ indicated that SCCSP can improve the classification accuracy by improving the separability of classes. In BCI applications, there existed multi-class classification problem.

The algorithms by spatial projection applied multiple binary-class classification to achieve multi-class classification, such as CSP. Thus, SCCSP can obey this way to classify multi-classes. In this way, one class can be seperated from other classes. Moreover, SCCSP applied SICA and spatial projection to obtain the spatial filter, and furter, the method may also be extended to other higher time resolution signal modalities analysis, such as fNIRS.

In conclusion, in this study, SCCSP has been introduced to the CSP family. This algorithm naturally integrates the relevant information and suppresses the influence of irrelevant information. The accuracy merits of SCCSP as supplemental to the broadband CSP filtering have been attentively validated on the public datasets of motor imagery EEG signals. The quantitative comparisons suggest superior discrimination and stable capability of the proposed method over the conventional CSP. Moreover, the test with varied training datasets shows excellent performance in small training datasets, and this is important in practical application. However, SCCSP spends longer time than CSP. This algorithm needed to be optimized. The SCCSP is affected by the μ-rhythm oscillation on the homolateral hemisphere. For improving the classification accuracy, the suppressed method of the homolateral hemisphere influence should be studied in further studies. In the future, we plan to study the SCCSP for multi class classification.

AUTHOR CONTRIBUTIONS

Analyzed the data: LL and FZ. Wrote the paper: LL. Contributed materials and analysis tools: LL and GX. Language correction: JX and ML.

FUNDING

REFERENCES

Ang, K. K., Chin, Z. Y., Zhang, H., and Guan, C. (2011). "Filter Bank Common Spatial Pattern (FBCSP) algorithm using online adaptive and semi-supervised learning," in *2011 International Joint Conference on Neural Networks* (San Jose, CA: IEEE), 392–396. doi: 10.1109/IJCNN.2011.6033248

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

Bennet, L., Fyfe, K. L., Yiallourou, S. R., Merk, H., Wong, F. Y., and Horne, R. S. C. (2016). Discrimination of sleep states using continuous cerebral bedside monitoring (amplitude-integrated electroencephalography) compared to polysomnography in infants. *Acta Paediat.* 105, E582–E587. doi: 10.1111/apa.13602

Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., and Mueller, K.-R. (2008). Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Process. Mag.* 25, 41–56. doi: 10.1109/MSP.2008.4408441

Byblow, W. D., Coxon, J. P., Stinear, C. M., Fleming, M. K., Williams, G., Mueller, J. F. M., et al. (2007). Functional connectivity between secondary and primary motor areas underlying hand-foot coordination. *J. Neurophysiol.* 98, 414–422. doi: 10.1152/jn.00325.2007

Chainay, H., Krainik, A., Tanguy, M. L., Gerardin, E., Le Bihan, D., and Lehericy, S. (2004). Foot, face and hand representation in the human supplementary motor area. *Neuroreport* 15, 765–769. doi: 10.1097/00001756-200404090-00005

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

Das, A. K., Suresh, S., and Sundararajan, N. (2016). A discriminative subject-specific spatio-spectral filter selection approach for EEG based motor-imagery task classification. *Expert Syst. Appl.* 64, 375–384. doi: 10.1016/j.eswa.2016.08.007

Dornhege, G., Blankertz, B., Krauledat, M., Losch, F., Curio, G., and Mueller, K.-R. (2006). Combined optimization of spatial and temporal filters for improving brain-computer interfacing. *IEEE Trans. Biomed. Eng.* 53, 2274–2281. doi: 10.1109/TBME.2006.883649

Gupta, I., Serb, A., Khiat, A., Zeitler, R., Vassanelli, S., and Prodromakis, T. (2016). Real-time encoding and compression of neuronal spikes by metal-oxide memristors. *Nat. Commun.* 7:12805. doi: 10.1038/ncomms12805

Hansen, L. K., Larsen, J., and Kolenda, T. (2001). "Blind detection of independent dynamic components," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing* (Salt Lake City, UT: IEEE), 3197–3200.

Hong, K.-S., Naseer, N., and Kim, Y.-H. (2015). Classification of prefrontal and motor cortex signals for three-class fNIRS-BCI. *Neurosci. Lett.* 587, 87–92. doi: 10.1016/j.neulet.2014.12.029

Hu, L., Zhang, Z. G., Mouraux, A., and Iannetti, G. D. (2015). Multiple linear regression to estimate time-frequency electrophysiological responses in single trials. *Neuroimage* 111, 442–453. doi: 10.1016/j.neuroimage.2015.01.062

Ikeda, A., Ohara, S., Matsumoto, R., Kunieda, T., Nagamine, T., Miyamoto, S., et al. (2000). Role of primary sensorimotor cortices in generating inhibitory motor response in humans. *Brain* 123, 1710–1721. doi: 10.1093/brain/123.8.1710

Lemm, S., Blankertz, B., Curio, G., and Muller, K. R. (2005). Spatio-spectral filters for improving the classification of single trial EEG. *IEEE Trans. Biomed. Eng.* 52, 1541–1548. doi: 10.1109/TBME.2005.851521

Leuthardt, E. C., Schalk, G., Wolpaw, J. R., Ojemann, J. G., and Moran, D. W. (2004). A brain-computer interface using electrocorticographic signals in humans. *J. Neural Eng.* 1, 63–71. doi: 10.1088/1741-2560/1/2/001

Lotte, F., and Guan, C. (2011). Regularizing common spatial patterns to improve bci designs: unified theory and new algorithms. *IEEE T. Biomed. Eng.* 58, 355–362. doi: 10.1109/TBME.2010.2082539

Luo, J., Feng, Z., Zhang, J., and Lu, N. (2016). Dynamic frequency feature selection based approach for classification of motor imageries. *Comput. Biol. Med.* 75, 45. doi: 10.1016/j.compbiomed.2016.03.004

Mihajlovic, V., Grundlehner, B., Vullers, R., and Penders, J. (2015). Wearable, wireless eeg solutions in daily life applications: what are we missing? *IEEE J. Biomed. Health Inform.* 19, 6–21. doi: 10.1109/JBHI.2014.2328317

Monakhova, Y. B., Tsikin, A. M., Mushtakova, S. P., and Mecozzi, M. (2015). Independent component analysis and multivariate curve resolution to improve spectral interpretation of complex spectroscopic data sets: application to infrared spectra of marine organic matter aggregates. *Microchem. J.* 118, 211–222. doi: 10.1016/j.microc.2014.10.001

Naseer, N., and Hong, K.-S. (2013). Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain-computer interface. *Neurosci. Lett.* 553, 84–89. doi: 10.1016/j.neulet.2013.08.021

Naseer, N., and Hong, K.-S. (2015). fNIRS-based brain-computer interfaces: a review. *Front. Hum. Neurosci.* 9:3. doi: 10.3389/fnhum.2015.00003

Naseer, N., Noori, F. M., Qureshi, N. K., and Hong, K.-S. (2016a). Determining optimal feature-combination for LDA classification of functional near-infrared spectroscopy signals in brain-computer interface application. *Front. Hum. Neurosci.* 10:237. doi: 10.3389/fnhum.2016.00237

Naseer, N., Qureshi, N. K., Noori, F. M., and Hong, K.-S. (2016b). Analysis of different classification techniques for two-class functional near-infrared spectroscopy-based brain-computer interface. *Comput. Intell. Neurosci.* 2016. 11. doi: 10.1155/2016/5480760

Neuper, C., and Pfurtscheller, G. (2001). Evidence for distinct beta resonance frequencies in human EEG related to specific sensorimotor cortical areas. *Clin. Neurophysiol.* 112, 2084–2097. doi: 10.1016/S1388-2457(01)00661-7

Park, C., Took, C. C., and Mandic, D. P. (2014). Augmented complex common spatial patterns for classification of noncircular EEG from motor imagery tasks. *IEEE Trans. Neural. Syst. Rehabil. Eng.* 22, 1–10. doi: 10.1109/TNSRE.2013.2294903

Pei, X. M., Xu, J., Zheng, C. X., and Bin, G. Y. (2005). "Selective two-channel linear descriptors for studying dynamic interaction of brain regions," in *Advances in Natural Computation* (Changsha), 646–653.

Pfurtscheller, G., and da Silva, F. H. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857. doi: 10.1016/S1388-2457(99)00141-8

Pfurtscheller, G., and Neuper, C. (1997). Motor imagery activates primary sensorimotor area in humans. *Neurosci. Lett.* 239, 65–68. doi: 10.1016/S0304-3940(97)00889-6

Pfurtscheller, G., Neuper, C., Flotzinger, D., and Pregenzer, M. (1997). EEG-based discrimination between imagination of right and left hand movement. *Electroencephalogr. Clin. Neuro.* 103, 642–651. doi: 10.1016/S0013-4694(97)00080-1

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.* 99, 1–1. doi: 10.1109/TNSRE.2017.2655542

Ruiz, S., Buyukturkoglu, K., Rana, M., Birbaumer, N., and Sitaram, R. (2014). Real-time fMRI brain computer interfaces: self-regulation of single brain regions to networks. *Biol. Psychol.* 95, 4–20. doi: 10.1016/j.biopsycho.2013.04.010

Sardouie, S. H., and Shamsollahi, M. B. (2012). Selection of efficient features for discrimination of hand movements from MEG using a BCI competition IV data set. *Front. Neurosci.* 6:42. doi: 10.3389/fnins.2012.00042

ter Horst, A. C., van Lier, R., and Steenbergen, B. (2013). Mental rotation strategies reflected in event-related (de)synchronization of alpha and mu power. *Psychophysiology* 50, 858–863. doi: 10.1111/psyp.12076

Tomioka, R., Dornhege, G., Nolte, G., Blankertz, B., Aihara, K., and Müller, K.-R. (2006). *Spectrally Weighted Common Spatial Pattern Algorithm for Single Trial EEG Classification.* Technical Reports, Department of Mathematics, Engineering University of Tokyo, Japan, 40.

van Dokkum, L. E. H., Ward, T., and Laffont, I. (2015). Brain computer interfaces for neurorehabilitation-its current status as a rehabilitation strategy post-stroke. *Ann. Phys. Rehabil. Med.* 58, 3–8. doi: 10.1016/j.rehab.2014.09.016

Vuckovic, A., Wallace, L., and Allan, D. B. (2015). Hybrid brain-computer interface and functional electrical stimulation for sensorimotor training in participants with tetraplegia: a proof-of-concept study. *J. Neurol. Phys. Ther.* 39, 3–14. doi: 10.1097/NPT.00000000000 00063

Woods, R. P., Hansen, L. K., and Strother, S. (2015). How many separable sources? Model selection in independent components analysis. *PLoS ONE* 10:e0118877. doi: 10.1371/journal.pone.0118877

Wu, W., Gao, X., Hong, B., and Gao, S. (2008). Classifying single-trial EEG during motor imagery by iterative spatio-spectral patterns learning (ISSPL). *IEEE Trans. Biomed. Eng.* 55, 1733–1743. doi: 10.1109/TBME.2008. 919125

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 Proces.* 10, 289–295. doi: 10.1016/j.bspc.2013. 08.012

Yu, K., Wang, Y., Shen, K., and Li, X. (2013). The Synergy between complex channel-specific FIR filter and spatial filter for single-trial eeg classification. *PLoS ONE* 8:e76923. doi: 10.1371/journal.pone.0076923

Yu, X., Chum, P., and Sim, K.-B. (2014). Analysis the effect of PCA for feature reduction in non-stationary EEG based motor imagery of BCI system. *Optik* 125, 1498–1502. doi: 10.1016/j.ijleo.2013.09.013

Zhang, R., Xu, P., Liu, T., Zhang, Y., Guo, L., Li, P., et al. (2013). Local temporal correlation common spatial patterns for single trial EEG classification during motor imagery. *Comput. Math. Method Med.* 2013:591216. doi: 10.1155/2013/591216