

Research Article

Motor Imagery Classification Using Mu and Beta Rhythms of EEG with Strong Uncorrelating Transform Based Complex Common Spatial Patterns

Youngjoo Kim,¹ Jiwoo Ryu,¹ Ko Keun Kim,² Clive C. Took,³
Danilo P. Mandic,⁴ and Cheolsoo Park¹

¹Department of Computer Engineering, Kwangwoon University, 20 Gwangun Rd, Nowon-gu, Seoul 01897, Republic of Korea

²LG, 38 Baumoe-ro, Seocho-gu, Seoul 137724, Republic of Korea

³Department of Computing, University of Surrey, Guildford, Surrey GU27XH, UK

⁴Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London SW72AZ, UK

Correspondence should be addressed to Cheolsoo Park; parkcheolsoo@kw.ac.kr

Received 15 April 2016; Revised 25 August 2016; Accepted 5 September 2016

Academic Editor: Stefan Haufe

Copyright © 2016 Youngjoo Kim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recent studies have demonstrated the disassociation between the mu and beta rhythms of electroencephalogram (EEG) during motor imagery tasks. The proposed algorithm in this paper uses a fully data-driven multivariate empirical mode decomposition (MEMD) in order to obtain the mu and beta rhythms from the nonlinear EEG signals. Then, the strong uncorrelating transform complex common spatial patterns (SUTCCSP) algorithm is applied to the rhythms so that the complex data, constructed with the mu and beta rhythms, becomes uncorrelated and its pseudocovariance provides supplementary power difference information between the two rhythms. The extracted features using SUTCCSP that maximize the interclass variances are classified using various classification algorithms for the separation of the left- and right-hand motor imagery EEG acquired from the Physionet database. This paper shows that the supplementary information of the power difference between mu and beta rhythms obtained using SUTCCSP provides an important feature for the classification of the left- and right-hand motor imagery tasks. In addition, MEMD is proved to be a preferred preprocessing method for the nonlinear and nonstationary EEG signals compared to the conventional IIR filtering. Finally, the random forest classifier yielded a high performance for the classification of the motor imagery tasks.

1. Introduction

The development of the brain computer interface (BCI) system allows one to control and communicate with the surrounding environment [1]. This includes applications ranging from education and entertainment to environmental control and communication through a noninvasive measurement of electroencephalogram (EEG) [2–4]. In particular, measuring EEG during different motor imagery tasks (e.g., left- and right-hand motor imagery) is a widely used paradigm to implement a BCI system. The EEG data acquired during these motor imagery tasks is then classified in order to control the BCI system.

The cornerstone to generate control signals that will facilitate the classification of different mental tasks is to

extract the appropriate features from the obtained EEG data. A number of different methods were used to perform feature extraction for various signals [5, 6], such as principal component analysis (PCA), which was used to classify additional forward tasks and relaxation [7] as well as left- and right-hand motor imagery tasks [8]. Independent component analysis (ICA) was also applied to motor imagery tasks (e.g., left- and right-hand [9] or finger lifting tasks [10]), which produced a high classification accuracy. Common spatial patterns (CSP) algorithm is another widely used feature extraction method, which is based on the fact that such neural activities are spatially distributed on the cortex areas [11, 12]. Recently, a complex version of the CSP algorithm has been proposed to analyze two different EEG signals simultaneously in a complex form, which provides features containing the

power sum information of the real and imaginary parts. The complex CSP algorithm was also combined with the strong uncorrelating transform (SUT) in order to cater for general complex signals considering the noncircularity (noncircular signals are not invariant to rotations, which may have power difference between real and imaginary parts of the complex form) of the data and maximizing the variance-ratio between two motor imagery tasks [13].

In addition, the power sum of the mu and beta rhythms gained from the complex version of CSP is closely related to the event-related phenomena, indicating the changes of the frequency components in the EEG data. Due to the synchronization in neuronal activities, the phenomena of a decrease (event-related desynchronization, ERD) [14, 15] or an increase (event-related synchronization, ERS) [16] in power of frequency components were found [17]. The brain lateralization of ERD [18] of the EEG activity during motor imagery tasks of the left- and right-hand is also well known [13, 19].

There have been further attempts to analyze EEG signals by investigating the different frequency band components of EEG separately, such as the mu (8–13 Hz) and beta (13–25 Hz) rhythms. This is based on the fact that the beta rhythm has distinct topographies and responses to the limb movements, compared to the mu rhythm, and thus the mu and beta rhythms should be individually considered [20, 21]. Brinkman et al. showed that the oscillatory power of the mu rhythm in the sensorimotor cortex ipsilateral to the tasks increased, while that of the beta rhythm in the contralateral sensorimotor cortex decreased simultaneously [22]. However, many of the previous studies on motor imagery responses analyzed the data considering all frequency components as a whole, which ignores the difference between the mu and beta rhythms [11, 13, 23]. In order to utilize the different information from the mu and beta rhythms for improved performance of motor imagery classification, we propose the application of SUTCCSP by constructing a complex formed data of these two rhythms.

Furthermore, considering the multichannel, nonlinear, and nonstationary property of EEG signals, recent studies have proved that the empirical mode decomposition (EMD) based algorithm is more effective than the conventional Fourier analysis in preprocessing physiological signals including motor imagery EEG signals [23–26]. In addition to the preprocessing method, various nonlinear classifiers have been applied to the classification of motor imagery tasks during the last decade [18]. The most commonly used classification method used in this area of the BCI research is the linear discriminant analysis (LDA) [27]. However, recent studies suggest that nonlinear classifiers are more practical by taking into account nonlinear relationships between data and the robustness against noise and outliers [28, 29]. Recent studies show that the well known machine learning based nonlinear classifier, random forest (RF) produces a high classification rate in the application of motor imagery classification [28, 30].

In this paper, we propose to classify the motor imagery EEG signals by analyzing the mu and beta rhythms separately

in a complex form of data using SUTCCSP. EEG signals from the Physionet database [31, 32] are preprocessed using MEMD in order to decompose the signals into mu and beta rhythms. The two distinct signals are then utilized to produce a complex data, which is composed of mu and beta rhythms in its real and imaginary terms for the analysis of SUTCCSP. Taking into account the functional disassociation of mu and beta rhythms, the extracted features using this method contain valuable information of the power difference between these two distinct rhythms. In addition, random forest (RF) is used in order to classify the left- and right-hand motor imagery tasks. As a result, the SUTCCSP algorithm results in a higher classification accuracy (80.05%) compared to the conventional CSP algorithm (78.04%) that does not account for the difference between the two distinct rhythms.

2. Methods

2.1. Motor Imagery EEG Datasets. The proposed algorithms were applied to the EEG data from the Physiobank Motor/Mental Imagery (MMI) database [31–33]. The database consists of a total of 109 subjects who performed the left- and right-hand motor imagery tasks. Each subject performed a total average of 46.62 ± 0.96 trials for the left- and right-hand motor imagery tasks. The average numbers of trials were 23.62 ± 0.61 and 23.00 ± 0.62 for the left- and right-hand motor imagery data, respectively. The EEG data was sampled at 160 Hz for all subjects yielding 640 samples for each single trial. However, in this study, we excluded the data of 4 subjects including S088, S092, S100, S104, since these subjects had damaged recordings (S088, S092, and S100) and too little samples (S104) in their left- and right-hand motor imagery datasets [34]. Therefore, a total of 105 pieces of subject data out of 109 were used for the experiment.

Out of the 64 channels of EEG data recorded with the 10-10 system, 14 channels were chosen as shown in Figure 1 for the feasible implementation of BCI with small number of channels [35]. The channels were selected so that it could cover all the regions of the scalp, including the frontal, central, parietal, and occipital lobe. Then, signals were decomposed using the MEMD in order to extract the mu and beta rhythms. After the preprocessing procedure, SUTCCSP was applied to complex formed data consisting of mu and beta rhythms, in order to extract the features for the classification of the left- and right-hand motor imagery tasks.

2.2. Multivariate Empirical Mode Decomposition (MEMD). Multivariate empirical mode decomposition provides an accurate data-driven time-frequency analysis for multichannel, nonlinear, and nonstationary signals, and thus MEMD was proved to be more effective in preprocessing the motor imagery EEG signals in terms of the baseline filtering [23, 36]. With the help of MEMD, the multivariate signals were decomposed into a linear combination of multiple common oscillatory modes called intrinsic mode functions (IMFs) [37], and the baseline filtering was done by eliminating

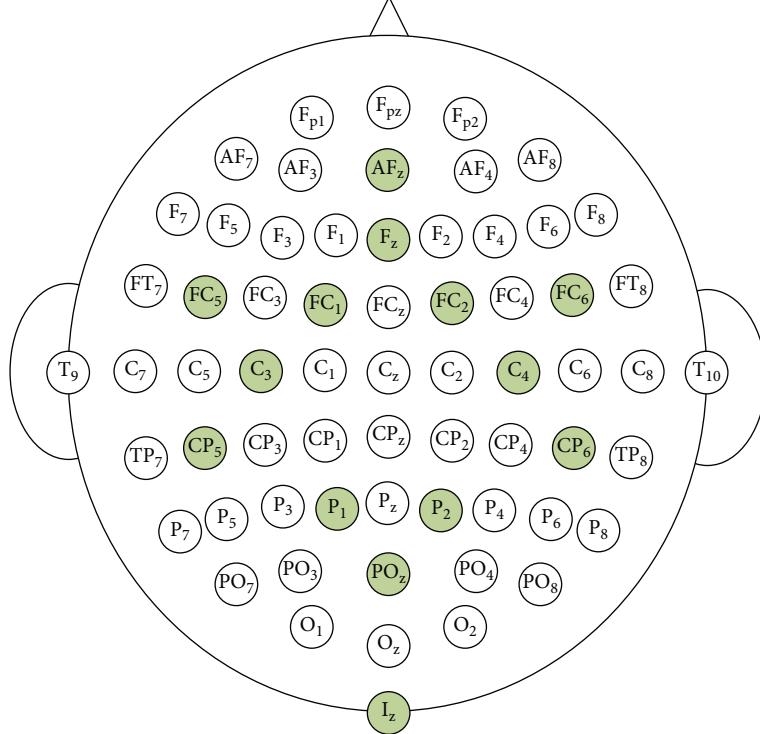


FIGURE 1: Channel selection. EEG montage of the selected 14 channels from the 64 channels provided from the Physionet database.

irrelevant IMFs. The original multivariate signal, $\mathbf{S}(t)$, is represented with the decomposed IMFs as follows:

$$\mathbf{S}(t) = \sum_{k=1}^n i_k(t), \quad (1)$$

where $i_k(t)$ represents k th IMF. When the original empirical mode decomposition (EMD) algorithm was applied to each channel of the data, the resulting IMFs of the same order from different channels did not have similar frequency characteristics (mode mixing problem) [38]. MEMD resolves this problem by using the mean envelope of the projected signals on a multidimensional projection space.

In this paper, the noise-assisted MEMD was applied to the motor imagery EEG data in order to further reduce the mode mixing problem by using an additional channel with white Gaussian noise [23]. Therefore, we used the noise-assisted MEMD to extract reliable frequency components. The IMFs that corresponded to the mu and beta rhythms were selected by investigating the power spectra of the IMFs calculated using the periodogram (Bartlett window) [39]. Bartlett window was mainly used since it was easy to implement and the aliasing problem was not as critical as the much simpler rectangular window. The averaged spectra of the 105 pieces of subject data using all trials were investigated and Figure 2 displays the averaged power spectra of the first 6 IMFs out of all 11 IMFs (from $i_1(t)$ to $i_6(t)$). The parameters for FFT points and window size were both set to 640. As shown in Figure 2, $i_2(t)$, $i_3(t)$, and $i_4(t)$ cover the frequency bands of the mu and beta rhythms, and thus $i_4(t)$ and $i_3(t) + i_2(t)$

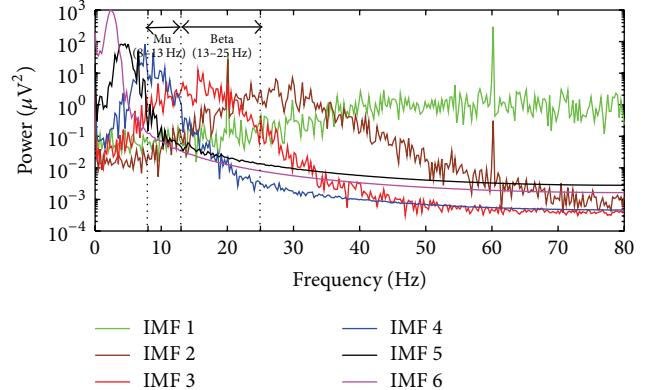


FIGURE 2: Average power spectra of all trials of 105 subjects. The power spectra show the first six IMFs, 1st IMF to the 6th IMF out of a total of 11 IMFs, which were decomposed in order using MEMD. $i_2(t)$ and $i_3(t) + i_4(t)$ correspond to the mu (8–13 Hz) and beta rhythms (13–25 Hz), respectively. The high peak at 60 Hz, indicating the power noise, was not to be considered.

of each trial were used as the real and imaginary parts of the constructed complex data, respectively.

2.3. Strong Uncorrelating Transform Complex Common Spatial Pattern. Falzon et al. first proposed the complex version of common spatial patterns in order to discriminate EEG responses to mental tasks using analytic signal-based CSP (ACSP) with Hilbert transform [40]. However, Hilbert transform could only be applied to narrowband signals, and

thus empirical mode decomposition was used to produce more accurate narrowband signals compared to the Fourier analysis [41, 42].

In addition to ACSP, there have been approaches that consider the noncircularity of complex signals [13]. When a complex random variable, $z = z_\mu + jz_\beta$ is defined, covariance (\mathbf{C}) and pseudocovariance (\mathbf{P}) are derived as follows [43]:

$$\begin{aligned}\mathbf{C} &= E[zz^H] = E[(z_\mu + jz_\beta)(z_\mu - jz_\beta)^T] \\ &= E[z_\mu z_\mu^T + z_\beta z_\beta^T] + jE[z_\beta z_\mu^T - z_\mu z_\beta^T] \\ &= E[z_\mu^2 + z_\beta^2] \geq 0,\end{aligned}\quad (2)$$

$$\begin{aligned}\mathbf{P} &= E[zz^T] = E[(z_\mu + jz_\beta)(z_\mu + jz_\beta)] \\ &= E[z_\mu z_\mu^T] - E[z_\beta z_\beta^T] + jE[z_\beta z_\mu^T + z_\mu z_\beta^T] \\ &= E[z_\mu^2 - z_\beta^2] + 2jE[z_\mu z_\beta],\end{aligned}\quad (3)$$

where $E[\cdot]$ indicates the statistical expectation operator and $(\cdot)^H$ and $(\cdot)^T$ denote the Hermitian and transpose of a vector, respectively. Equation (2) shows that the covariance contains the sum of power information of the mu and beta rhythms, whereas (3) shows that the pseudocovariance includes the power difference and the correlation information of the mu and beta rhythms. When the given data is circular, (3) is equivalent to zero, since z_μ and z_β have the same variance ($E[z_\mu z_\mu^T] - E[z_\beta z_\beta^T] = 0$) and are uncorrelated ($E[z_\beta z_\mu^T + z_\mu z_\beta^T] = 0$). However, most of the biological signals are noncircular in the real world [13]. Therefore, the complex form preserves the pseudocovariance information [44] and the augmented form of complex CSP (ACCSP) that holds both the covariance and pseudocovariance information can be applied.

Then, the strong uncorrelating transform combined with ACCSP is used in order to diagonalize the covariance and pseudocovariance matrices simultaneously, assuring that the multichannel complex data can no longer be correlated [13]. The diagonalization process leaves the sum of power and the power difference information of mu and beta rhythms from the augmented covariance and pseudocovariance matrix. The steps for SUTCCSP is described below:

When given the number of channels (N) and sample size (S), \mathbf{A}_L and \mathbf{A}_R are both $N \times S$ matrices that hold the zero-mean data of the two classes, left-hand (L) and right-hand (R) motor imagery tasks, respectively. In this paper, obtained $i_4(t)$ and $i_3(t) + i_2(t)$ that cover the mu and beta rhythms of the EEG data recorded during motor imagery tasks are used as the real and imagery terms shown as \mathbf{A}_L or $\mathbf{A}_R = i_4(t) + j(i_3(t) + i_2(t))$. (original CSP algorithm uses \mathbf{A}_L or $\mathbf{A}_R = i_2(t) + i_3(t) + i_4(t)$.)

Then, the covariance and pseudocovariance matrices of \mathbf{A}_L and \mathbf{A}_R can be generated as follows:

$$\mathbf{C}_L = \text{cov}(\mathbf{A}_L) = E[\mathbf{A}_L \mathbf{A}_L^H],$$

$$\mathbf{C}_R = \text{cov}(\mathbf{A}_R) = E[\mathbf{A}_R \mathbf{A}_R^H],$$

$$\mathbf{P}_L = p \text{cov}(\mathbf{A}_L) = E[\mathbf{A}_L \mathbf{A}_L^T],$$

$$\mathbf{P}_R = p \text{cov}(\mathbf{A}_R) = E[\mathbf{A}_R \mathbf{A}_R^T].$$

(4)

Then, a composite spatial covariance and pseudocovariance matrices are calculated as

$$\mathbf{C}_c = \mathbf{C}_L + \mathbf{C}_R = E[\mathbf{A}_L \mathbf{A}_L^H] + E[\mathbf{A}_R \mathbf{A}_R^H],$$

(5)

$$\mathbf{P}_c = \mathbf{P}_L + \mathbf{P}_R = E[\mathbf{A}_L \mathbf{A}_L^T] + E[\mathbf{A}_R \mathbf{A}_R^T].$$

Using the eigen decomposition, there exists a factorization form of

$$\mathbf{C}_c = \mathbf{U}_c \Lambda_c \mathbf{U}_c^H,$$

(6)

so that \mathbf{C}_c can be whitened by applying whitening matrix $\mathbf{G} = \Lambda_c^{-1/2} \mathbf{U}_c^H$, $\mathbf{I} = \mathbf{G} \mathbf{C}_c \mathbf{G}^H$, and the pseudocovariance matrix is then decomposed using Takagi's factorization as follows [13]:

$$\bar{\mathbf{P}}_c = \mathbf{G} \mathbf{P}_c \mathbf{G}^T = \mathbf{Y} \Lambda \mathbf{Y}^T.$$

(7)

The SUT transform matrix \mathbf{Q} is defined as

$$\mathbf{Q} = \mathbf{Y}^H \mathbf{G}.$$

(8)

Therefore, SUTCCSP is a transform of the whitened factorization form of the covariance and pseudocovariance matrices. The process above allows both the covariance and pseudocovariance matrices to be diagonalized simultaneously

$$\begin{aligned}\mathbf{Q} \mathbf{C}_c \mathbf{Q}^H &= \mathbf{Q} \mathbf{C}_L \mathbf{Q}^H + \mathbf{Q} \mathbf{C}_R \mathbf{Q}^H = \mathbf{I}, \\ \mathbf{Q} \mathbf{P}_c \mathbf{Q}^H &= \mathbf{Q} \mathbf{P}_L \mathbf{Q}^T + \mathbf{Q} \mathbf{P}_R \mathbf{Q}^T = \Lambda.\end{aligned}\quad (9)$$

(9)

Assuming $\mathbf{S}_L = \mathbf{Q} \mathbf{C}_L \mathbf{Q}^H$ and $\mathbf{S}_R = \mathbf{Q} \mathbf{C}_R \mathbf{Q}^H$, the SUT transform yields an estimation of the eigenvectors from the covariance matrix so that

$$\begin{aligned}\Lambda_L &= \mathbf{B}^{-1} \mathbf{S}_L \mathbf{B}, \\ \Lambda_R &= \mathbf{B}^{-1} \mathbf{S}_R \mathbf{B}.\end{aligned}\quad (10)$$

(10)

The estimation of the eigenvectors from the pseudocovariance matrix is also obtained as follows:

$$\widehat{\mathbf{Q}} = \Lambda^{-1/2} \mathbf{Y}^H \mathbf{G},$$

(11)

$$\widehat{\mathbf{S}}_L = \widehat{\mathbf{Q}} \mathbf{P}_L \widehat{\mathbf{Q}}^T,$$

(12)

$$\widehat{\mathbf{S}}_R = \widehat{\mathbf{Q}} \mathbf{P}_R \widehat{\mathbf{Q}}^T,$$

$$\widehat{\mathbf{Q}} \mathbf{P}_c \widehat{\mathbf{Q}}^T = \widehat{\mathbf{S}}_L + \widehat{\mathbf{S}}_R = \mathbf{I},$$

(13)

$$\widehat{\mathbf{B}}^{-1} \widehat{\mathbf{S}}_L \widehat{\mathbf{B}} = \widehat{\Lambda}_L,$$

(14)

$$\widehat{\mathbf{B}}^{-1} \widehat{\mathbf{S}}_R \widehat{\mathbf{B}} = \widehat{\Lambda}_R,$$

(14)

where $\widehat{\mathbf{B}}$ and $\widehat{\Lambda}$ indicate the eigenvectors and eigenvalues of $\widehat{\mathbf{S}}$, respectively. The derived equations (10) and (14) lead to

$\Lambda_L + \Lambda_R = \mathbf{I}$ and $\widehat{\Lambda}_L + \widehat{\Lambda}_R = \mathbf{I}$, which is equivalent to $\Lambda_L = \mathbf{I} - \Lambda_R$ and $\widehat{\Lambda}_L = \mathbf{I} - \widehat{\Lambda}_R$. Therefore, if the values of Λ_L and $\widehat{\Lambda}_L$ are in descending order, values of Λ_R and $\widehat{\Lambda}_R$ will be in the ascending order. This is the main property of the CSP algorithm, which illustrates that the variance of one class is maximized, while the variance of the other class is minimized, when applying the following covariance and pseudocovariance spatial filter:

$$\begin{aligned}\mathbf{W} &= \mathbf{B}^{-1} \mathbf{G}, \\ \widehat{\mathbf{W}} &= \widehat{\mathbf{B}}^{-1} \widehat{\mathbf{G}}.\end{aligned}\tag{15}$$

The final step to extract the features is to apply spatial filter \mathbf{W} to given data \mathbf{A} as

$$\begin{aligned}\mathbf{V} &= \mathbf{WA}, \\ \widehat{\mathbf{V}} &= \widehat{\mathbf{WA}},\end{aligned}\tag{16}$$

where \mathbf{V} and $\widehat{\mathbf{V}}$ are the covariance and pseudocovariance spatial filtered data, respectively, and their power values are used as features. These obtained features are then separated using classifiers to classify the left- and right-hand motor imagery tasks. To demonstrate the advantage using the pseudocovariance information, we show the complex version of CSP (CCSP), using the similar procedure to SUTCCSP, only using the covariance information in \mathbf{W} , except $\widehat{\mathbf{W}}$. Therefore, unlike SUTCCSP preserving both the power sum and difference information, CCSP only preserves the power sum information of the real and imaginary terms of the complex variable. In this way, we show how features containing the power difference information affect the classification accuracy of the motor imagery tasks.

2.4. Classifiers. In this study, five machine learning algorithms, including the random forest (RF), logistic model tree (LMT), model tree (MT), k -nearest neighbor (KNN), and logitboost (LB) [30, 45–49], were implemented using WEKA and utilized for the benchmark test of classifiers [27, 28]. These five classifiers have frequently been used to classify various motor imagery tasks [45, 50–53]. These classifiers are mainly based on bagging or boosting (random forest, logistic model tree, model tree, and logitboost) [30, 46, 47, 49]. In addition to the classifiers based on bagging and boosting, the well known k -nearest neighbor algorithm, which uses the neighboring distance measures of features, was also used for the comparison of the classification methods [48]. Both bagging and boosting are based on an ensemble method, using a combination of multiple learning methods to produce a better prediction. Particularly, random forest, an extended version of bagging, has been proven to be an effective classification algorithm for the classification of motor imagery tasks and emotional dimensions using EEG signals [28, 45]. This is due to the characteristics of random forest, the robustness against outliers and noise, and the useful internal estimates of the error, correlation, and variable importance [30]. Breiman introduces the definition of the random forest as follows [30]:

Definition 1. Random forest is an ensemble of tree-structured classifiers $h(x, \theta_k)$ ($k \in \mathbb{N}$), where random vectors $\{\theta_k\}$ generated at k th tree are independent and identically distributed and each tree votes for the most popular class ($c \in \{1, -1\}$), given input x from the training set.

In addition, all parameters of each classifier including the random forest were set with the default parameters of WEKA. The number of trees for random forest was set as 100 and the maximum depth of the trees was set as unlimited. For model tree, which uses a regression model for every class value [47], the minimum number of instances per leaf was set as 4. Logitboost performs an additive logistic regression [49] and the percentage of the weight mass used for base training was set as 100 with a total of 10 iterations. Logistic model tree is also based on the linear logistic regression models. However, it uses logitboost and regression functions as base learners [46] and the number of iterations for early heuristic stopping was set as 50. Finally, index k from k -nearest neighbor algorithm was set as 1.

3. Results

3.1. Classification of the Left- and Right-Hand Motor Imagery EEG Data. The classification performance using the original CSP, CCSP, and SUTCCSP of significant subjects was compared using five machine learning algorithms implemented using WEKA. Machine learning algorithms include random forest, logistic model tree, model tree, k -nearest neighbor, and logitboost. The classification performances of all classifiers were calculated using a five-cross validation (30 iterations for different random selected training sets for each subject) and the similar analyses from [13] were applied. Specifically, significant subjects were chosen when their classification accuracies were above a certain percentage, which was set with a confidence limit of 95% (cf. [54]). The average percentage limit for 45 trials of the motor imagery tasks was approximately 64%, and thus subjects with classification rates over 64% were categorized as significant subjects. The rationale of using only significant subjects was to exclude the subjects with too low classification accuracies. Since Ahn and Jun claimed that the subjects who had performed poorly with tasks had little brain activity across the different regions of the brain or less brain network, these subjects were excluded from the evaluation [55]. The bar chart of the number of significant subjects for CSP, CCSP, and SUTCCSP is shown in Figure 3. Note that the number of significant subjects of CCSP and SUTCCSP are bigger than those of CSP.

All subjects, who were marked significant subjects using either CSP, CCSP, or SUTCCSP, were included in the significant subject pool in order to calculate the average classification rate across the subjects. As a result, a total of 24 subjects were chosen as significant subjects, and the classification rates calculated using the five classifiers are shown in Table 1. Overall, SUTCCSP produced the highest classification rate among the CSP algorithms when classified using random forest as shown in Table 1. Additionally, error bars for these results are displayed in Figure 4, which show that SUTCCSP yields higher classification rates among the CSP algorithms.

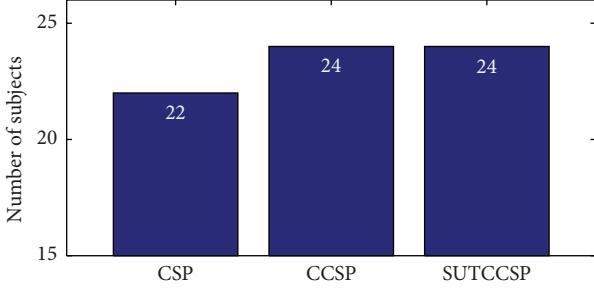


FIGURE 3: Bar chart indicating the number of significant subjects. Note that CCSP and SUTCCSP produce two more significant subjects compared to CSP.

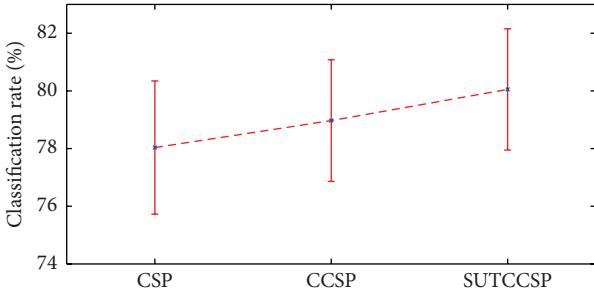


FIGURE 4: Errorbar of classification rates of the significant subjects for CSP, CCSP, and SUTCCSP. Note that SUTCCSP produces higher classification rates compared to CSP and CCSP, confirmed by Student's *t*-test in Table 3.

TABLE 1: Comparison of classification accuracies among different classification algorithms with significant subjects.

Classifiers	CSP (%)	CCSP (%)	SUTCCSP (%)
RF	78.04 ± 2.31	78.97 ± 2.11	80.05 ± 2.10
LMT	77.16 ± 2.33	77.62 ± 1.97	77.75 ± 2.02
MT	76.89 ± 2.18	76.07 ± 2.04	76.02 ± 2.13
KNN	77.36 ± 2.35	77.49 ± 2.22	77.02 ± 2.28
LB	76.88 ± 2.17	75.98 ± 2.11	76.25 ± 2.25

The classification accuracies for the insignificant subjects were also shown in Table 2 to compare with those for the significant subjects in Table 1. Table 2 shows that the performances of the insignificant subjects were close to 50%, which were consistent with what random chance might produce. For this, subjects yielding low performance were not appropriate for the evaluation, and, thus, the insignificant subjects were excluded in the main analysis [55].

In addition, the scatterplots of the classification rates of the significant subjects are displayed in Figure 5 to compare the results of SUTCCSP with CSP (a) and CCSP (b). The diagonal lines in Figure 5 represent the cases where the classification rates of CSP or CCSP and SUTCCSP are the same. An additional study of the one-way analysis of the variance, Student's *t*-test, was conducted, where classification accuracies of SUTCCSP were compared with CSP (p_1) and CCSP (p_2). Note that most dots in both scatterplots lie above

TABLE 2: Comparison of classification accuracies among different classification algorithms with insignificant subjects.

Classifiers	CSP (%)	CCSP (%)	SUTCCSP (%)
RF	51.70 ± 0.78	50.87 ± 0.62	51.25 ± 0.71
LMT	49.72 ± 0.81	49.50 ± 0.79	50.59 ± 0.80
MT	49.71 ± 0.83	49.87 ± 0.78	50.70 ± 0.74
KNN	51.56 ± 0.57	51.03 ± 0.53	51.74 ± 0.55
LB	51.32 ± 0.63	50.79 ± 0.64	51.49 ± 0.67

TABLE 3: Student's *t*-test results, which compare classification accuracies of SUTCCSP with CSP (p_1) and CCSP (p_2).

Classifiers	p_1	p_2
RF	0.0075	0.0037
LMT	0.5055	0.7437
MT	0.2539	0.8798
KNN	0.6405	0.3620
LB	0.4736	0.5437

the line, meaning that SUTCCSP outperforms both CSP and CCSP ($p_1 < 0.01$ and $p_2 < 0.005$).

Table 3 shows the results of Student's *t*-test performed for the other classifiers including the random forest, which compares the classification rates of SUTCCSP with CSP (p_1) and CCSP (p_2). In detail, significant p values of random forest ($p_1 < 0.01$ and $p_2 < 0.005$) reassure that SUTCCSP outperforms CSP and CCSP. The other classification algorithms except the random forest gave no significance in terms of p values in Table 3 ($p_1, p_2 > 0.05$), and their classification rates did not vary across CSP, CCSP, and SUTCCSP in Table 1.

It is noted that the results of SUTCCSP outperformed those of CCSP, whose features include only the information of the power sum of the mu and beta rhythms. This suggests that the power difference information of them preserved by SUTCCSP can be considered as an important factor for the classification between the left- and right-hand motor imagery tasks. Therefore, results in Tables 1 and 3 prove that SUTCCSP outperforms CSP and CCSP, with random forest, a preferred classification method with CSP based feature extraction algorithms, particularly SUTCCSP.

In order to marginalize the performance difference that comes from the classifiers, the average across the performance of different classifiers was calculated for CSP (77.27%), CCSP (77.22%), and SUTCCSP (77.41%). As a result, SUTCCSP resulted in a slightly higher performance compared to the other CSP algorithms by 0.15 (%).

In addition, an additional experiment using both the CSP features and features of the power difference information of mu and beta rhythms (CSP + PD) was conducted to show that the performance improvement is coming from SUTCCSP and not solely from the random forest classifier. If our best result is yielded only from the classifying technique of the random forest classifier and not SUTCCSP, the result of CSP + PD should be comparable to that of SUTCCSP when feeding the classifier with the same amount of power difference information. However, despite using the same classifier, CSP

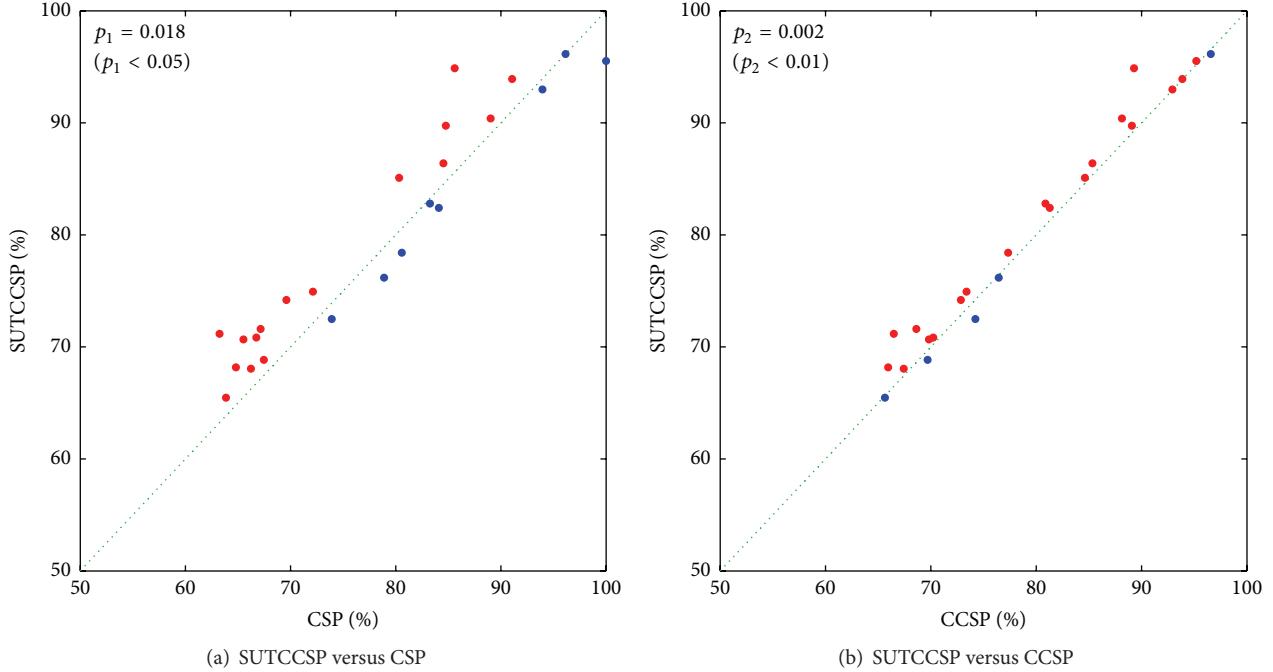


FIGURE 5: Scatterplots of the classification rates comparing SUTCCSP with CSP (a) and CCSP (b). The dots are marked red where the classification rates of SUTCCSP are larger than those of CSP or CCSP and marked blue in opposite cases. The enhanced performance using SUTCCSP is confirmed by p value of Student's t -test.

TABLE 4: Classification rates of significant subjects.

Preprocessing method	CSP (%)	CCSP (%)	SUTCCSP (%)
IIR	75.92 ± 2.04	76.06 ± 1.98	76.98 ± 1.90
MEMD	78.04 ± 2.31	78.97 ± 2.11	80.05 ± 2.10

TABLE 5: Student's t -test results for IIR and MEMD, which compare classification accuracies of SUTCCSP with CSP (p_1) and CCSP (p_2).

Preprocessing method	p_1	p_2
IIR	0.1115	0.0229
MEMD	0.0075	0.0037

TABLE 6: Student's t -test results for MEMD versus IIR filter for all 105 subjects using the different CSP algorithms.

Feature extraction	$p < 0.00001$
CSP	0.000003
CCSP	0.0000005
SUTCCSP	0.000008

+ PD achieved a 78.63 ± 1.93 (%) classification accuracy, which is slightly higher than that of CSP but still below the best classification accuracy using SUTCCSP (80.05%) by 1.42 (%). This states that the performance improvement of the best classification accuracy is coming from the preservation of unique power difference information using SUTCCSP and does not solely come from the random forest classifier.

Additional studies comparing MEMD with IIR filtered results are presented in Tables 4–6 in order to prove that MEMD is more effective for the dataset of motor imagery EEG signals. Table 4 shows the classification results of the preprocessed data using MEMD, which outperforms those using the 5th-order Butterworth IIR filter by approximately 1.1%, 2.9%, and 3.0% for CSP, CCSP, and SUTCCSP features, respectively. Note the largest difference between the performance of IIR and MEMD for SUTCCSP features. The classification accuracies using MEMD and the IIR filter were calculated with the average classification rate of the significant subjects among the 105 subjects.

Student's t -test was also utilized to compare the variances between the classification performances of MEMD and IIR prefiltered data of the significant subjects in Table 5. Results display t -test performed for two cases: t -test for the original CSP versus SUTCCSP (p_1) and CCSP versus SUTCCSP (p_2). Results show that the IIR filter gave no significant p value for p_1 (> 0.05) and a relatively high value for p_2 (< 0.05), whereas MEMD gave significantly low p values for both p_1 (< 0.01) and p_2 (< 0.005).

Table 6 also shows the results of Student's t -test performed for MEMD versus IIR filter for all 105 subjects using the different CSP algorithms. All CSP methods show significant p values (< 0.00001), which suggest that MEMD is more competent than the IIR filter for the preprocessing of EEG. As Park et al. have shown that MEMD is effective in preprocessing motor imagery EEG signals due to the nonlinear and nonstationary characteristics of the data [23], Tables 4–6 demonstrate that MEMD is more effective than

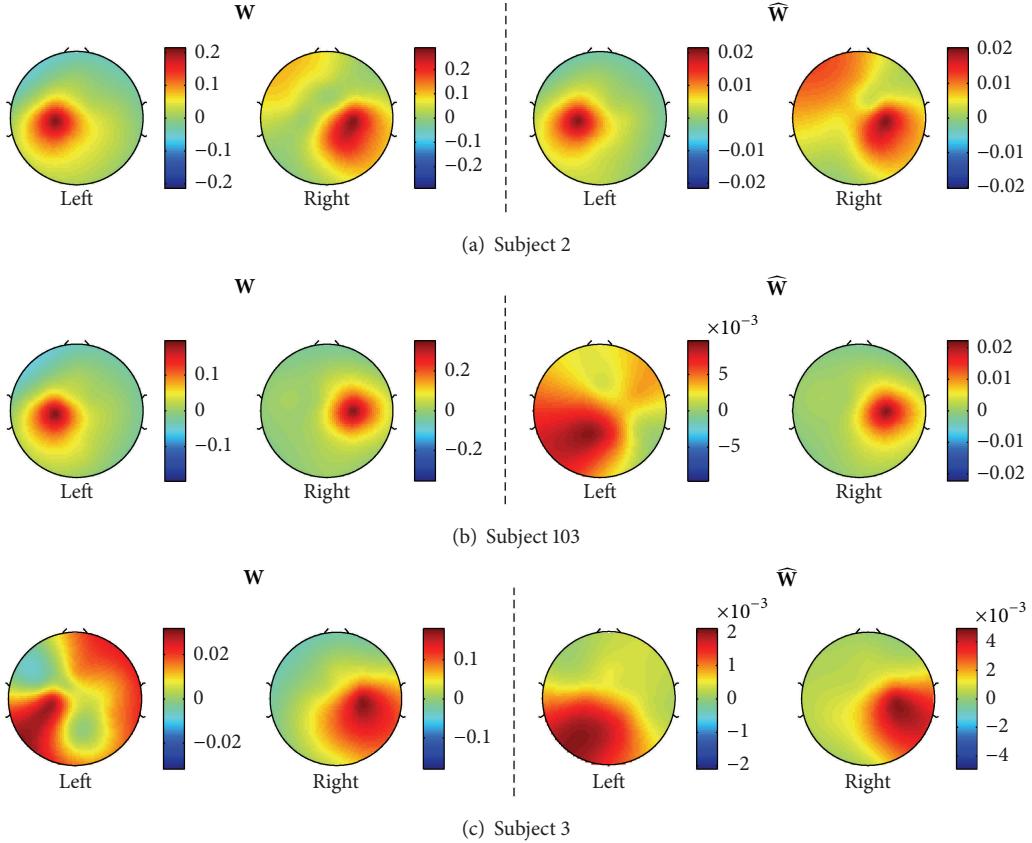


FIGURE 6: Spatial patterns of the top three subjects in descending order, who had the best classification rates: subject 34 (96.13%), 72 (95.51%), and 7 (94.89%). The left two topographies show the spatial patterns for the covariance matrices (\mathbf{W}) of the left and right-hand tasks, respectively, while the right two topographies show the spatial patterns for the pseudocovariance matrices ($\widehat{\mathbf{W}}$). Note that the patterns of \mathbf{W} and $\widehat{\mathbf{W}}$ are prominently different, meaning $\widehat{\mathbf{W}}$ can produce additional information to \mathbf{W} .

the IIR filter for this left- and right-hand motor imagery EEG dataset.

3.2. Spatial Pattern Topographies. Figure 6 illustrates the spatial pattern topographies for the top three subjects in descending order who had the best classification rates out of the 24 significant subjects: subject 34 (96.13%), 72 (95.51%), and 7 (94.89%). The left and right two topographies of Figure 6 correspond to the spatial patterns of covariance (\mathbf{W}) and pseudocovariance ($\widehat{\mathbf{W}}$) matrices, respectively.

Figure 7 also shows the spatial pattern topographies of the three subjects, who gave the worst classification rates: subject 2 (68.18%), 103 (68.04%), and 33 (65.47%). Overall, the typical spatial patterns during the motor imagery tasks, the synchronization on the ipsilateral hemisphere, and desynchronization on the contralateral hemisphere [19] can be noted for the covariance spatial patterns. A prominent ipsilateral power difference is also shown for the spatial patterns of the pseudocovariance of subject 33. This is also seen in Figure 8, illustrating spatial patterns of average covariance (\mathbf{W}) and pseudocovariance ($\widehat{\mathbf{W}}$) matrices of all 24 significant subjects.

Furthermore, the difference between the covariance and pseudocovariance spatial patterns suggests that the pseudocovariance spatial filters provide additional information

about the power difference between the mu and beta rhythms of EEG recorded during motor imagery tasks.

3.3. Asymmetries of the Power Difference and Sum of Mu and Beta Rhythms. Figures 6–8 have shown that the spatial patterns of the left- and right-hand motor imagery tasks for each individual subject are distinguishable. Additional calculations of the asymmetry of the power difference and sum from the symmetric channels (FC5-FC6, FCI-FC2, C3-C4, CP5-CP6, and P1-P2) would give a clear explanation for these results. Figure 9 displays the asymmetries of the average power difference (a) and sum (b) of the mu and beta rhythms for all trials of the significant subjects. Outliers were excluded with criteria of ± 5 standard deviation cutoff [56]. The asymmetries of all the symmetric channels were calculated using the following equation:

$$\text{Asymmetry} = \frac{\text{CH}_L - \text{CH}_R}{\text{CH}_L + \text{CH}_R}, \quad (17)$$

where CH_L and CH_R indicate the symmetric channels of the left and right hemispheres, respectively. Therefore, when the power difference or sum from the left hemisphere is greater than the right, the resulting asymmetry will be positive and vice versa. Figures 9(a) and 9(b) use the power difference

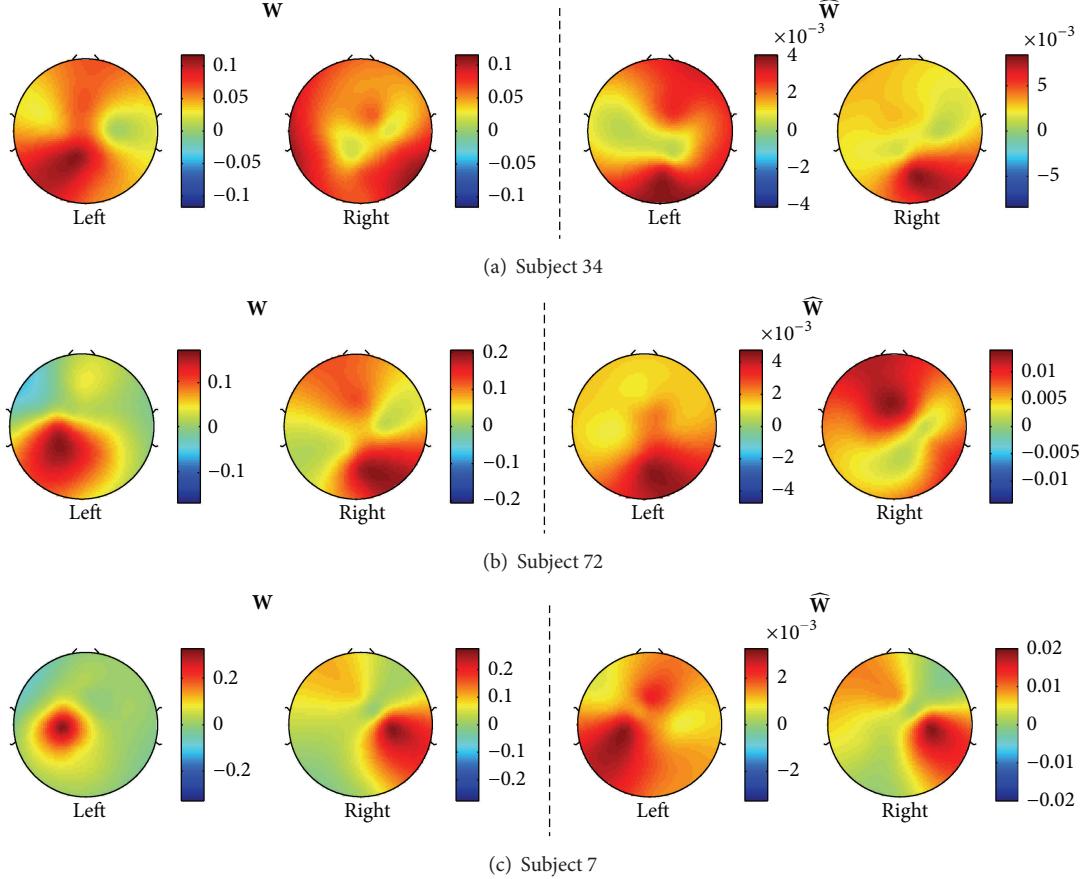


FIGURE 7: Spatial patterns of the worst three subjects in descending order, who gave the worst classification rates: subject 2 (68.18%), 103 (68.04%), and 33 (65.47%). The left two topographies show the spatial patterns for the covariance matrices (\mathbf{W}) of the left- and right-hand tasks, respectively, while the right two topographies show the spatial patterns for the pseudocovariance matrices ($\widehat{\mathbf{W}}$). Note that the patterns of the left- and right-hand motor imagery are prominently different even for the subjects who had poor classification rates.

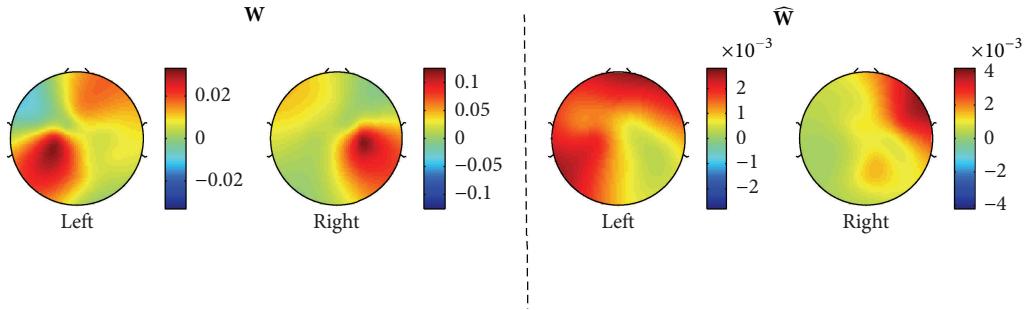


FIGURE 8: Spatial patterns of the average covariance and pseudocovariance matrices of all 24 significant subjects are shown.

and sum of the mu and beta rhythms, respectively, from the symmetric channels. Figure 9(a) shows a marked difference for the channels of the central area compared to Figure 9(b). Since the motor cortex responsible for all voluntary movements is located in the central region of the human brain, these asymmetry results can explain why the spatial pattern topographies of the left- and right-hand motor imagery tasks are prominently different in Figures 6–8. In particular, in channels C3 and C4, the power difference asymmetries of the

left- and right-hand tasks have different signs, meaning that the power difference of the mu and beta rhythms is greater in C3 during the left-hand task, whereas the power difference is greater in C4 during the right-hand task. Since this difference is only shown in Figure 9(a) and not in Figure 9(b), this could explain why the features containing the power difference between the mu and beta rhythms in the pseudocovariance matrix result in a better classification performance than those without the power difference information.

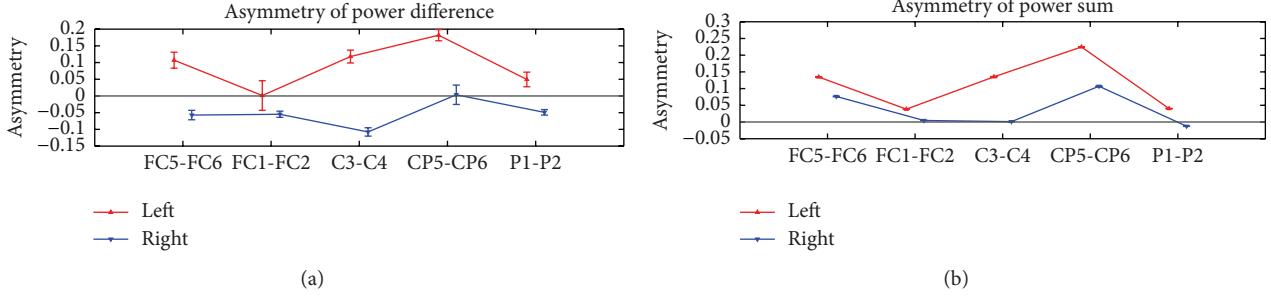


FIGURE 9: Asymmetries of the power difference and sum of symmetric channel EEG of motor imagery tasks. (a) Asymmetry of the power difference between mu and beta rhythms from the symmetric channels. (b) Asymmetry of the power sum of mu and beta rhythms from the symmetric channels. Note that the asymmetry of the power difference shows more prominent difference between the left- and right-hand motor imagery tasks with the different signs of values, compared to the asymmetry of the power sum.

TABLE 7: Comparison of the classification accuracies of CSP algorithms using only motor cortex channels.

CSP (%)	CCSP (%)	SUTCCSP (%)
75.82 ± 2.24	76.62 ± 2.09	77.70 ± 2.09

TABLE 8: Student's *t*-test results, which compare classification accuracies of SUTCCSP with CSP (p_1) and CCSP (p_2) using only motor cortex channels.

p_1	p_2
0.0144	0.0090

4. Discussion

4.1. Motor Cortex Channels. In this study, the channel data that covers all the brain regions was used. Therefore, an additional experiment using only the motor cortex channels was conducted, since the motor cortex is originally known to be responsible for the motor movement or imagery tasks [57]. The same data analysis methods from the original study were used. As shown in Table 7, the performances of all CSP algorithms were lower than those of Table 1. A reasonable explanation could be that supplementary information was kept in the occipital and the nearby parietal region, due to the perception of the visual stimulus guidance of the motor imagery tasks. Other studies using the positron emission tomography (PET) and the functional magnetic resonance imaging (fMRI) also showed that other areas including the parietal region, anterior cingulate gyrus, and the cerebellum were activated [58, 59]. Additional pairwise Student's *t*-tests to calculate p_1 and p_2 were conducted as shown in Table 8. This shows that SUTCCSP results in a higher performance even when using only motor cortex channels ($p_1 < 0.05$ and $p_2 < 0.01$, resp.).

4.2. Proposed Algorithm. In this study, MEMD was applied to the multichannel data, where channels were selected to cover all the brain regions. The nonlinear property of such real-world EEG data makes it difficult for the conventional frequency analysis methods to decompose the signals into

the natural oscillations, such as Fourier transform based on the fixed sinusoidal functions. However, the data-driven MEMD method obtains the frequency components without any basis functions, and, thus, for the real-world physiological data, such as electromyogram (EMG) or electrocardiogram (ECG), MEMD has the potential to provide a highly accurate frequency analysis.

In addition, it is known that the left- and right-hand imagery is associated with bilateral desynchronization of mu rhythms, greater on the contralateral side, and the mu rhythm has prominent hemispheric asymmetry with the right-hand imagery, while the beta rhythm is more prominent with the left-hand imagery [20, 60]. Recent studies have shown the dissociation of mu and beta rhythms, which has not been identified in the previous studies [22]. Brinkman et al. showed an increase in the oscillatory power of the mu rhythm in the sensorimotor cortex ipsilateral to the tasks, whereas that of the beta rhythm decreased in the contralateral sensorimotor cortex simultaneously. The disassociation between mu and beta rhythms can explain why the complementary information of the power difference between mu and beta rhythms from the pseudocovariance could provide crucial information to classify the left- and right-hand motor imagery tasks. Also, the ipsilateral difference of the power difference of the left- and right-hand motor imagery EEG, shown in Figure 9, demonstrates that the complementary information of the power difference between mu and beta rhythms is an important feature for the classification of the left- and right-hand motor imagery tasks.

4.3. Processing Speed of CSP Algorithms. In order to calculate the processing speed of the system as in real-time applications, the computational complexity in terms of Big-O notation and the actual processing time of the three CSP algorithms were compared, since these algorithms were directly applied to the data segments of the test set data. The software used for calculation was MATLAB R2016a, since all codes were implemented using this software. The hardware specifications include a Windows 10 OS and an x64-based Intel(R) Core(TM) i7-6700HQ CPU (2.60 GHz) processor with 16.0 G of RAM.

TABLE 9: Actual processing time for CSP, CCSP, and SUTCCSP.

CSP (ms)	CCSP (ms)	SUTCCSP (ms)
11.227	9.894	30.272

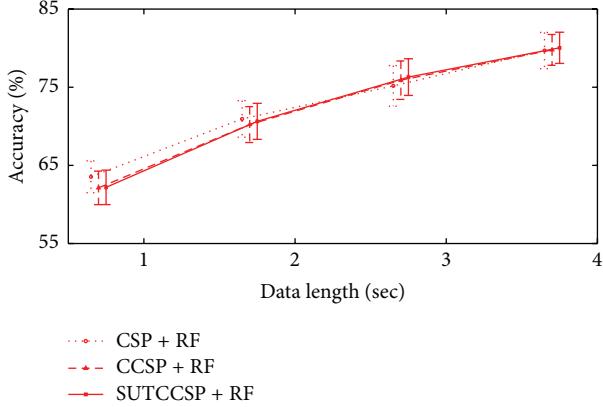


FIGURE 10: The errorbar of RF conducted with features from CSP, CCSP, and SUTCCSP when varying the data length. Note that the performance decreases as the data length reduces.

When the computational complexity of the three CSP algorithms was calculated using Big-O notation, multiplication was calculated as $O(n^3)$ since MATLAB uses the blocked-matrix multiplication method. The computational complexity of each algorithm was $O(n^4)$ for all CSP algorithms. Although it may seem that there were no differences among the three methods, the actual processing time would be different, since the coefficients for SUTCCSP were much larger than those of CSP and CCSP due to more multiplication during processing a large number of features.

When calculating the actual computational time processed with the specifications mentioned above, approximately six trials were used for the test set since our study conducts a fivefold cross validation. The actual processing time of the three methods is shown in Table 9. Note that the unit of milliseconds (ms) would be of negligible difference among the different approaches for real-time processing.

4.4. Variation of Data Length. In order to search for performance differences when varying the data length, an additional experiment was conducted that shows the classification accuracies using data samples varying from 1 second (160 samples) to 4 seconds (640 samples). For benchmark testing, the best classification method, along with the significant subject pool from the CSP algorithms, was selected, since it yields the best results. Then, all experiments were processed from the preprocessing step with MEMD, feature extraction with CSP methods, and then classification. Results are displayed in Figure 10.

Figure 10 clearly shows that the data using a full data length of 640 samples increases the classification accuracy for all CSP algorithms. Therefore, it proves that the performance decreases when reducing the data length.

5. Conclusion

In this paper, we have used SUTCCSP to extract the different responses of mu and beta rhythms of EEG to the motor imagery tasks. Results showed improved classification performance using SUTCCSP with consideration of the power difference between mu and beta rhythms, compared to the original CSP algorithm. The functional disassociation between the mu and beta rhythms can explain the contribution of the supplementary information of the power difference to the motor imagery classification. Finally, our investigation of preprocessing and classification methods for the motor imagery EEG analysis has confirmed that MEMD and the random forest classifier are the optimal algorithms for this purpose.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

Youngjoo Kim and Jiwoo Ryu participated in the design of the study, carried out the main EEG data analysis using SUTCCSP, and drafted the manuscript. Cheolsoo Park supervised experiments and the statistical analysis throughout the study. Youngjoo Kim, Jiwoo Ryu, Ko Keun Kim, Clive C. Took, Danilo P. Mandic, and Cheolsoo Park helped the drafting and revision of the manuscript. All authors read and approved the final manuscript. Youngjoo Kim and Jiwoo Ryu are equal contributor.

Acknowledgments

The present research has been conducted by the Research Grant of Kwangwoon University in 2016, Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2014R1A1A2059483) and Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korean Government (MISP) (no. B0184-15-1003, The Development of oneM2M Conformance Testing Tool and QoS Technology).

References

- [1] B. Graimann, B. Z. Allison, and G. Pfurtscheller, *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*, The Frontiers Collection, Springer, Berlin, Germany, 2010.
- [2] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Müller, and G. Curio, “The non-invasive Berlin Brain-Computer Interface: fast acquisition of effective performance in untrained subjects,” *NeuroImage*, vol. 37, no. 2, pp. 539–550, 2007.
- [3] G. Pfurtscheller, G. R. Müller-Putz, A. Schlogl et al., “15 Years of BCI research at Graz University of Technology: current projects,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 205–210, 2006.
- [4] S. Sanei and J. A. Chambers, *EEG Signal Processing*, John Wiley & Sons, New York, NY, USA, 2007.

- [5] W. Liu, D. P. Mandic, and A. Cichocki, "Analysis and online realization of the CCA approach for blind source separation," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1505–1510, 2007.
- [6] S. Ge, M. Han, and X. Hong, "A fully automatic ocular artifact removal from EEG based on fourth-order tensor method," *Biomedical Engineering Letters*, vol. 4, no. 1, pp. 55–63, 2014.
- [7] C. R. Hema, M. P. Paulraj, S. Yaacob, A. H. Adom, and R. Nagarajan, "Motor imagery signal classification for a four state brain machine interface," *International Journal of Computer, Information, Systems and Control Engineering*, vol. 1, no. 5, pp. 1375–1380, 2007.
- [8] A. Vallabhaneni and B. He, "Motor imagery task classification for brain computer interface applications using spatiotemporal principle component analysis," *Neurological Research*, vol. 26, no. 3, pp. 282–287, 2004.
- [9] X. Guo and X. Wu, "Motor imagery EEG classification based on dynamic ICA mixing matrix," in *Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE '10)*, Chengdu, China, June 2010.
- [10] C.-I. Hung, P.-L. Lee, Y.-T. Wu, L.-F. Chen, T.-C. Yeh, and J.-C. Hsieh, "Recognition of motor imagery electroencephalography using independent component analysis and machine classifiers," *Annals of Biomedical Engineering*, vol. 33, no. 8, pp. 1053–1070, 2005.
- [11] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," *Clinical Neurophysiology*, vol. 110, no. 5, pp. 787–798, 1999.
- [12] P. Li, P. Xu, R. Zhang, L. Guo, and D. Yao, "L1 norm based common spatial patterns decomposition for scalp EEG BCI," *BioMedical Engineering Online*, vol. 12, no. 1, article 77, 2013.
- [13] C. Park, C. C. Cheong-Took, and D. P. Mandic, "Augmented complex common spatial patterns for classification of noncircular EEG from motor imagery tasks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 1–10, 2014.
- [14] G. Pfurtscheller, "Graphical display and statistical evaluation of event-related desynchronization (ERD)," *Electroencephalography and Clinical Neurophysiology*, vol. 43, no. 5, pp. 757–760, 1977.
- [15] H. Yuan, A. Doud, A. Gururajan, and B. He, "Cortical imaging of event-related (de)synchronization during online control of brain-computer interface using minimum-norm estimates in frequency domain," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 5, pp. 425–431, 2008.
- [16] G. Pfurtscheller, "Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest," *Electroencephalography and Clinical Neurophysiology*, vol. 83, no. 1, pp. 62–69, 1992.
- [17] G. Pfurtscheller and F. H. Lopes Da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [18] G. Pfurtscheller and C. Neuper, "Motor imagery activates primary sensorimotor area in humans," *Neuroscience Letters*, vol. 239, no. 2–3, pp. 65–68, 1997.
- [19] Y. Chung, J. Kang, and S. Kim, "Analysis of correlated EEG activity during motor imagery for brain-computer interfaces," in *Proceedings of the Control Automatic System*, pp. 337–341, 2011.
- [20] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements," *Brain Topography*, vol. 12, no. 3, pp. 177–186, 2000.
- [21] G. Pfurtscheller, M. Pregenzer, and C. Neuper, "Visualization of sensorimotor areas involved in preparation for hand movement based on classification of μ and central β rhythms in single EEG trials in man," *Neuroscience Letters*, vol. 181, no. 1–2, pp. 43–46, 1994.
- [22] L. Brinkman, A. Stolk, H. C. Dijkerman, F. P. De Lange, and I. Toni, "Distinct roles for alpha- and beta-band oscillations during mental simulation of goal-directed actions," *The Journal of Neuroscience*, vol. 34, no. 44, pp. 14783–14792, 2014.
- [23] C. Park, D. Looney, N. Ur Rehman, A. Ahrabian, and D. P. Mandic, "Classification of motor imagery BCI using multivariate empirical mode decomposition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 1, pp. 10–22, 2013.
- [24] J. Ryu and C. Park, "Time-frequency analysis of electrohysterogram for classification of term and preterm birth," *IEIE Transactions on Smart Processing and Computing*, vol. 4, no. 2, pp. 103–109, 2015.
- [25] V. Bajaj and R. B. Pachori, "Epileptic seizure detection based on the instantaneous area of analytic intrinsic mode functions of EEG signals," *Biomedical Engineering Letters*, vol. 3, article 1, 2013.
- [26] S. Lahmiri and M. Boukadoum, "A weighted bio-signal denoising approach using empirical mode decomposition," *Biomedical Engineering Letters*, vol. 5, no. 2, pp. 131–139, 2015.
- [27] F. Lotte, M. Congedo, A. Lécuyer, and F. Lamarche, "A review of classification algorithms for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, pp. 1–13, 2007.
- [28] D. Steyrl, R. Scherer, J. Faller, and G. R. Müller-Putz, "Random forests in non-invasive sensorimotor rhythm brain-computer interfaces: a practical and convenient non-linear classifier," *Biomedical Engineering*, vol. 61, no. 1, pp. 77–86, 2016.
- [29] X. Xu, Z. Wang, X. Zhang, W. Yan, W. Deng, and L. Lu, "Human face recognition using multi-class projection extreme learning machine," *IEIE Transactions on Smart Processing and Computing*, vol. 2, article 6, 2013.
- [30] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [31] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.
- [32] A. L. Goldberger, L. A. N. Amaral, L. Glass et al., "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–E220, 2000.
- [33] General-Purpose Software System for Brain-Computer Interface (BCI), 2016, <http://www.bci2000.org>.
- [34] A. Loboda, A. Margineanu, G. Rotariu, and A. M. Lazar, "Discrimination of EEG-based motor imagery tasks by means of a simple phase information method," *International Journal of Advanced Research in Artificial Intelligence*, vol. 3, no. 10, 2014.
- [35] H. Shan, H. Xu, S. Zhu, and B. He, "A novel channel selection method for optimal classification in different motor imagery BCI paradigms," *BioMedical Engineering Online*, vol. 14, article 93, 2015.
- [36] Y. Kim and C. Park, "Strong uncorrelated transform applied to spatially distant channel EEG data," *IEIE Transactions on Smart Processing and Computing*, vol. 4, no. 2, pp. 97–102, 2015.

- [37] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proceedings of the Royal Society A*, vol. 466, pp. 1291–1302, 2010.
- [38] N. ur Rehman and D. P. Mandic, "Filter bank property of multivariate empirical mode decomposition," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2421–2426, 2011.
- [39] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*, Prentice Hall, Upper Saddle River, NJ, USA, 1997.
- [40] O. Falzon, K. P. Camilleri, and J. Muscat, "Complex-valued spatial filters for task discrimination," in *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, Buenos Aires, Argentina, August–September 2010.
- [41] S. R. Long, N. R. Huang, C. C. Tung et al., "The Hilbert techniques: an alternate approach for non-steady time series analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, pp. 6–11, 1995.
- [42] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [43] J. Navarro-Moreno, M. D. Estudillo-Martínez, R. M. Fernández-Alcalá, and J. C. Ruiz-Molina, "Estimation of improper complex-valued random signals in colored noise by using the Hilbert space theory," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2859–2867, 2009.
- [44] D. P. Mandic and V. S. L. Goh, "Complex valued adaptive filters," in *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, chapter 6, John Wiley & Sons, Chichester, UK, 2009.
- [45] A. Heraz and C. Frasson, "Predicting learner answers correctness through brainwaves assessment and emotional dimensions," in *Artificial Intelligence in Education*, vol. 200 of *Frontiers in Artificial Intelligence and Applications*, pp. 49–56, 2009.
- [46] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Machine Learning*, vol. 59, no. 1–2, pp. 161–205, 2005.
- [47] E. Frank, Y. Wang, S. Inglis, G. Holmes, and I. H. Witten, "Using model trees for classification," *Machine Learning*, vol. 32, no. 1, pp. 63–76, 1998.
- [48] D. W. Aha, D. Kibler, and M. K. Albert, "Instance-based learning algorithms," *Machine Learning*, vol. 6, no. 1, pp. 37–66, 1991.
- [49] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *The Annals of Statistics*, vol. 28, no. 2, pp. 337–407, 2000.
- [50] M. E. Abdalsalam, M. Z. Yusoff, N. Kamel, A. Malik, and M. Meselhy, "Mental task motor imagery classifications for noninvasive brain computer interface," in *Proceedings of the IEEE 5th International Conference on Intelligent and Advanced Systems (ICIAS '14)*, pp. 1–5, June 2014.
- [51] A. Schlögl, F. Lee, H. Bischof, and G. Pfurtscheller, "Characterization of four-class motor imagery EEG data for the BCI-competition 2005," *Journal of Neural Engineering*, vol. 2, no. 4, pp. L14–L22, 2005.
- [52] N. Naseer and K.-S. Hong, "Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain-computer interface," *Neuroscience Letters*, vol. 553, pp. 84–89, 2013.
- [53] M. Mesbah, A. Khorshidtalab, H. Baali, and A. Al-Ani, "Motor imagery task classification using a signal-dependent orthogonal transform based feature extraction," in *Neural Information Processing*, vol. 9490, pp. 1–9, Springer, 2015.
- [54] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random? A closer look on BCI results," *International Journal of Bioelectromagnetism*, vol. 10, no. 1, pp. 52–55, 2008.
- [55] M. Ahn and S. C. Jun, "Performance variation in motor imagery brain-computer interface: a brief review," *Journal of Neuroscience Methods*, vol. 243, pp. 103–110, 2015.
- [56] M. Strand, M. W. Oram, and Å. Hammar, "Emotional information processing in major depression remission and partial remission: faces come first," *Applied Neuropsychology*, vol. 20, no. 2, pp. 110–119, 2013.
- [57] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.
- [58] J. Decety, D. Perani, M. Jeannerod et al., "Mapping motor representations with positron emission tomography," *Nature*, vol. 371, no. 6498, pp. 600–602, 1994.
- [59] S. M. Rao, J. R. Binder, P. A. Bandettini et al., "Functional magnetic resonance imaging of complex human movements," *Neurology*, vol. 43, no. 11, pp. 2311–2318, 1993.
- [60] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "EEG-based discrimination between imagination of right and left hand movement," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 6, pp. 642–651, 1997.

