

Research Article

Correlation Assisted Strong Uncorrelating Transform Complex Common Spatial Patterns for Spatially Distant Channel Data

Youngjoo Kim ¹, Jiwoo You,¹ Heejun Lee,¹ Seung Min Lee,² and Cheolsoo Park ¹

¹Department of Computer Engineering, Kwangwoon University, Seoul 01897, Republic of Korea

²School of Electrical Engineering, College of Creative Engineering, Kookmin University, Seoul 02707, Republic of Korea

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

Youngjoo Kim and Jiwoo You contributed equally to this work.

Received 28 September 2017; Revised 26 February 2018; Accepted 1 April 2018; Published 15 May 2018

Academic Editor: Toshihisa Tanaka

Copyright © 2018 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.

The Strong Uncorrelating Transform Complex Common Spatial Patterns (SUTCCSP) algorithm, designed for multichannel data analysis, has a limitation on keeping the correlation information among channels during the simultaneous diagonalization process of the covariance and pseudocovariance matrices. This paper focuses on the importance of preserving the correlation information among multichannel data and proposes the correlation assisted SUTCCSP (CASUT) algorithm to address this issue. The performance of the proposed algorithm was demonstrated by classifying the motor imagery electroencephalogram (EEG) dataset. The features were first extracted using CSP algorithms including the proposed method, and then the random forest classifier was utilized for the classification. Experiments using CASUT yielded an average classification accuracy of 78.10 (%), which significantly outperformed those of original CSP, Complex Common Spatial Patterns (CCSP), and SUTCCSP with p -values less than 0.01, tested by the Wilcoxon signed rank test.

1. Introduction

Noninvasive measurements of physiological signals including electroencephalogram (EEG), electrocardiogram (ECG), and electromyogram (EMG) have become widely used throughout the biomedical industry [1–5]. Out of the various feature engineering methods, researchers have shown that the common spatial patterns (CSP) algorithm is a strong feature extraction algorithm for multichannel EEG data, yielding high performance for classification problems [6, 7]. CSP is a mathematical methodology to decompose spatial subcomponents of multivariate signals, whose variance difference between two classes is maximized [8]. CSP designs spatial filters for the multichannel EEG signals based on the spatial distribution of neural activities in the cortex areas [6, 7] and adopts a supervised learning approach, while the other spatial filter algorithms such as principal component analysis (PCA) and independent component analysis (ICA) are designed in an unsupervised manner [9, 10].

Furthermore, a complex version of CSP, termed CCSP, uses the covariance matrix that maintains the power sum information of the real and imaginary parts of the complex-valued data [11]. Another complex-valued CSP algorithm, termed analytic signal-based CSP (ACSP), was proposed by Falzon et al. to discriminate different mental tasks [12, 13]. However, given that the Hilbert transformed analytic signals could only produce circular signals (rotation invariant probability distribution) and that physiological signals are improper (mismatch of power between different channel data), the augmented complex CSP was introduced to fully exploit the second-order statistics of noncircular complex vectors [11, 14].

Strong Uncorrelating Transform CCSP (SUTCCSP), which is an advanced version of the augmented complex CSP, was applied to the two-class classification problem of motor imagery EEG and produced a minimum of 4% improvement over the conventional CSP, ACSP, and augmented CSP [11]. This is due to the power difference information preserved in the pseudocovariance matrix, accompanied with the sum

of power maintained in the covariance matrix. However, during the simultaneous diagonalization process of the covariance and pseudocovariance matrices, the correlation term vanishes owing to the process of applying the strong uncorrelating transform [11, 15, 16]. Such effort to preserve correlation has not been made so far for the CSP algorithm, and the correlation assisted version of SUTCCSP is newly proposed in this paper.

The basic terminologies and procedure of SUTCCSP and the proposed method are explained in Section 2, followed by extensive simulation results on the benchmark motor imagery dataset of 105 subjects in Section 3. Finally, the concluding remarks are given in Section 4 with additional discussions in terms of the performance difference of distinct channel pairs that have less correlation compared with results of Section 3.

2. Proposed Method

Here we explain SUT based on the terminologies used in [9, 14] and show how the correlation information is utilized with CSP algorithms [11, 16].

Let x be a complex-valued random vector such as

$$x = x_r + jx_i, \quad (1)$$

where j is $\sqrt{-1}$, x_r is the real part, and x_i is the imaginary part of a complex random vector. \mathbf{X}_k is a zero-mean complex-valued matrix consisting of values with the form of (1), where k denotes the two different classes, $k \in \{1, 2\}$. \mathbf{X}_k has the dimension of the number of channels by the number of samples. Then the covariance (\mathbf{C}) and pseudocovariance (\mathbf{P}) matrices are defined as follows:

$$\begin{aligned} \mathbf{C}_k &= E[\mathbf{X}_k \mathbf{X}_k^H], \\ \mathbf{P}_k &= E[\mathbf{X}_k \mathbf{X}_k^T], \end{aligned} \quad (2)$$

where $E(\cdot)$ is the statistical expected value operator and $(\cdot)^H$ is the conjugate transpose. Then, we can define the composite covariance (\mathbf{C}_c) and pseudocovariance (\mathbf{P}_c) matrices as follows:

$$\begin{aligned} \mathbf{C}_c &= \sum_k \mathbf{C}_k = E[\mathbf{X}_1 \mathbf{X}_1^H] + E[\mathbf{X}_2 \mathbf{X}_2^H] \\ \mathbf{P}_c &= \sum_k \mathbf{P}_k = E[\mathbf{X}_1 \mathbf{X}_1^T] + E[\mathbf{X}_2 \mathbf{X}_2^T]. \end{aligned} \quad (3)$$

Here \mathbf{C}_c can then be decomposed into Θ_c and Λ_c as follows:

$$\mathbf{C}_c = \Theta_c \Lambda_c \Theta_c^H = \Theta_c \Lambda_c^{1/2} \Lambda_c^{1/2} \Theta_c^H, \quad (4)$$

where Θ_c has eigenvectors in each column for the corresponding diagonal eigenvalues of Λ_c . Note that Θ_c and Λ_c consist of real elements and the nondiagonal elements of Λ_c are zero. This allows \mathbf{C}_c to be whitened by the whitening matrix $\Phi = \Lambda_c^{-1/2} \Theta_c^H$ in the original CCSP algorithm, resulting in $\Phi \mathbf{C}_c \Phi^H = \mathbf{I}$, where \mathbf{I} denotes the identity matrix [11].

Using the whitening matrix $\Phi = \Lambda_c^{-1/2} \Theta_c^H$ from the original CCSP algorithm [11], the pseudocovariance matrix can also be decomposed using Takagi's factorization as shown in the following equation [17]:

$$\Phi \mathbf{P}_c \Phi^T = \Delta \Lambda \Delta^T, \quad (5)$$

where Δ and Λ are yielded by symmetric matrices. This leads to a derivation of the strong uncorrelating transform matrix \mathbf{S} as follows:

$$\mathbf{S} = \Delta^H \Phi. \quad (6)$$

Using the matrix \mathbf{S} , it is now possible to diagonalize the covariance and pseudocovariance matrices simultaneously. Firstly, the composite covariance matrix can be diagonalized as follows:

$$\begin{aligned} \mathbf{S} \mathbf{C}_c \mathbf{S}^H &= \mathbf{S} \mathbf{C}_1 \mathbf{S}^H + \mathbf{S} \mathbf{C}_2 \mathbf{S}^H = \mathbf{M}_1 + \mathbf{M}_2 \\ &= \mathbf{I} (\because \mathbf{M}_k = \mathbf{S} \mathbf{C}_k \mathbf{S}^H, k \in \{1, 2\}) \end{aligned} \quad (7)$$

$$\mathbf{Y}^{-1} \mathbf{M}_k \mathbf{Y} = \Lambda_k, \quad \sum_k \Lambda_k = \mathbf{I},$$

where \mathbf{Y} and Λ are the estimations of eigenvectors and eigenvalues of \mathbf{M} , respectively. Next, the composite pseudocovariance can also be diagonalized as follows:

$$\begin{aligned} \mathbf{S} \mathbf{P}_c \mathbf{S}^H &= \mathbf{S} \mathbf{P}_1 \mathbf{S}^H + \mathbf{S} \mathbf{P}_2 \mathbf{S}^H = \Lambda \\ \hat{\mathbf{S}} &= \Lambda^{-1/2} \Delta^H \Phi \\ \hat{\mathbf{M}}_k &= \hat{\mathbf{S}} \mathbf{P}_k \hat{\mathbf{S}}^T \end{aligned} \quad (8)$$

$$\hat{\mathbf{S}} \mathbf{P}_c \hat{\mathbf{S}}^T = \sum_k \hat{\mathbf{M}}_k = \mathbf{I}$$

$$\hat{\mathbf{Y}}^{-1} \hat{\mathbf{M}}_k \hat{\mathbf{Y}} = \hat{\Lambda}_k, \quad \sum_k \hat{\Lambda}_k = \mathbf{I},$$

where $\hat{\mathbf{S}}$ is the strong uncorrelating transform matrix for the pseudocovariance and $\hat{\mathbf{Y}}$ and $\hat{\Lambda}$ are the estimations of the eigenvectors and eigenvalues of $\hat{\mathbf{M}}$, respectively. Therefore, the two spatial filters \mathbf{W} and $\hat{\mathbf{W}}$ can be designed as follows:

$$\begin{aligned} \mathbf{W} &= \mathbf{Y}^{-1} \mathbf{S} \\ \hat{\mathbf{W}} &= \hat{\mathbf{Y}}^{-1} \hat{\mathbf{S}}. \end{aligned} \quad (9)$$

Finally, the spatially filtered vector, \mathbf{Z} , is calculated as follows:

$$\mathbf{Z} = \mathbf{W} \mathbf{X}. \quad (10)$$

Let N be the number of data channels, and z_p the p th row vector in \mathbf{Z} ;

$$\mathbf{Z}' = \begin{bmatrix} z_1 \\ \vdots \\ z_m \\ z_{N-m+1} \\ \vdots \\ z_N \end{bmatrix} = \begin{bmatrix} z'_1 \\ \vdots \\ z'_{2m} \end{bmatrix}, \quad (11)$$

where z'_p corresponds to each row of the new matrix \mathbf{Z}' . Now the final subfeatures, f_p and f'_p , by SUTCCSP are calculated as follows:

$$f_p = \log \left(\frac{\text{var}(\Re[z'_p])}{\sum_p \text{var}(\Re[z'_p])} \right) \quad (12)$$

$$f'_p = \log \left(\frac{\text{var}(\Im[z'_p])}{\sum_p \text{var}(\Im[z'_p])} \right),$$

where p varies between 1 and $2m$ and $\text{var}(\cdot)$ is the variance of (\cdot) . Here, selecting one pair of filter is equivalent to choosing the first and last rows in each real and imaginary part of the covariance and pseudocovariance matrices, separately. The number of filter pairs was chosen to maximize the performance for each subject. Such consideration of selecting the appropriate number of filter pairs could be important in real time applications. Next, Pearson's correlation coefficient for x_r and x_i is calculated as follows [17]:

$$\rho_{x_r, x_i} = \frac{E[(x_r - \mu_{x_r})(x_i - \mu_{x_i})]}{\text{std}(x_r) \text{std}(x_i)}, \quad (13)$$

where $\text{std}(\cdot)$ is the standard deviation of (\cdot) and μ_x is the mean of x . The maximum number of correlation coefficients between the real and imaginary parts of (1) is equal to the number of channel pairs due to the multichannel attribute of the data. The high dimension of the number of channel pairs should be reduced owing to the curse of dimensionality. PCA is applied to reduce the high dimension in this paper, due to its simple implementation and fast speed [18, 19].

Let $\mathbf{\Gamma}$ be the matrix containing ρ_{x_r, x_i} for $N(N-1)/2$ channel pairs, where N is the number of channels. By applying PCA to the correlation coefficient matrices, the principal component coefficients, known as loadings, are estimated [20]. Here we will define $\mathbf{\Psi}$ as an N -by- L matrix of loadings, where L is the reduced number of dimensions. An additional subfeature f''_q containing the correlation information of two data channels is calculated as follows:

$$f''_q = \mathbf{\Gamma} \mathbf{\Psi} \quad (q = 1, \dots, L). \quad (14)$$

The final feature matrices for two different classes are f_p , f'_p , and f''_q for each class. In this paper, the covariance matrix information from the original CSP is added to the

feature matrices of CCSP, SUTCCSP, and CASUT, which could provide a fair test to compare CSP with these three algorithms. Accordingly, the feature matrices of CASUT were designed to contain the information of variance, power sum, and difference, as well as the correlation information lost due to the strong uncorrelating transform.

3. Experiments

3.1. Data Acquisition. As Park et al. used the Physiobank Motor Mental Imagery (MMI) database to test the performance of SUTCCSP, this study uses the same dataset in order to compare the proposed CASUT with the former CSP algorithms including SUTCCSP [11, 21–23]. Out of the 109 subjects who conducted the left- and right-hand motor imagery tasks, three subjects (S088, S092, and S100) had damaged recordings, and one subject (S104) had an insufficient amount of data [15, 24]. For these reasons, 105 subjects were used to examine the classification accuracy of CASUT. All subject data consist of 45 trials of performing the left- and right-hand tasks, which were recorded using 64 channel electrodes with the 10-10 EEG system and sampled by 160 Hz [25].

In order to verify the performance of CASUT in preserving the correlation information, the channel pairs that yield high correlation coefficients were selected (values over 0.9 and less than or equal to 1). All trials for the left-hand motor imagery task of 105 subjects were combined into one single trial set, and the correlation coefficients of the all possible distinct 2016 pairs among the 64 channels were calculated. Then the average of the correlation coefficient values over all trials of the left-hand task was calculated, in order to determine which channel pair has a high correlation coefficient. The same calculation was conducted on the trials of the right-hand motor imagery task. The channel pairs, whose correlations were in the range of the following equation, were denoted as

$$r_t = \{(x, y) \mid t \cdot 10^{-1} < \rho_{x_r, x_i} \leq (t+1) \cdot 10^{-1}\}, \quad (15)$$

where (x, y) is a pair of two distinct channels x and y , ρ_{x_r, x_i} are the correlation coefficients between x and y , and t is a number in the range of $0 \leq t \leq 9$.

The EEG recordings were preprocessed using the fifth-order Butterworth IIR bandpass filter extracting the frequency components into 8–25 Hz [6, 26, 27]. Such preprocessing techniques were identical to the preprocessing techniques used by Park et al. [11].

3.2. Classification Results

3.2.1. Analysis of 105 Subjects. The average classification accuracies over all 105 subjects were calculated in order to compare the proposed algorithm with CSP, CCSP, and SUTCCSP. Table 1 shows the average classification rates with the standard deviations for each algorithm. Note that the classification rate of CASUT outperforms those of CSP, CCSP, and SUTCCSP.

The normality was tested to determine whether to use the parametric or nonparametric version of a statistical test

TABLE 1: Average classification accuracies of CSP, CCSP, SUTCCSP, and CASUT across 105 subjects.

| CSP method | CSP | CCSP | SUTCCSP | CASUT |
|-----------------------------|--------------|--------------|--------------|--------------|
| Classification accuracy (%) | 70.62 ± 1.35 | 70.60 ± 1.41 | 73.05 ± 1.32 | 73.69 ± 1.30 |

TABLE 2: The resulting p -values of the KS test for each CSP algorithm for 105 subjects.

| CSP method | CSP | CCSP | SUTCCSP | CASUT |
|-------------|--------|--------|---------|--------|
| p -values | 0.1784 | 0.1568 | 0.0777 | 0.2533 |

TABLE 3: Results of the Wilcoxon signed rank test conducted on performance accuracies of CASUT compared with those of CSP, CCSP, and SUTCCSP using 105 subjects.

| | p_1 | p_2 | p_3 |
|------------|------------|-------------|---------|
| p -value | $<10^{-7}$ | $<10^{-10}$ | <0.05 |

such as ANOVA. Accordingly, the resulting p -values of the Kolmogorov-Smirnov goodness-of-fit hypothesis test (KS test) in Table 2 show that the classification accuracies of CSP algorithms could not always satisfy the normality assumption [28]. Therefore, the nonparametric Friedman test was used instead of the parametric ANOVA, to compare three or more matched groups regardless of their normality [29, 30].

The p -value for the Friedman test, which was less than 10^{-15} , indicates that it is safe to perform the post hoc test. Instead of the parametric paired Student's t -test, the Wilcoxon signed rank test, which can be used regardless of the normality, was conducted as the post hoc test [28]. Although the average classification accuracy difference between CASUT and SUTCCSP looked small, the Wilcoxon signed rank test performed on the accuracies of the two algorithms yielded significant p -values (<0.05), as shown in Table 3. The p -values, p_1 , p_2 , and p_3 , indicate the results of the Wilcoxon signed rank test conducted on the classification accuracies of CASUT compared with those of original CSP, CCSP, and SUTCCSP, respectively.

3.2.2. Analysis of Significant Subjects. For a thorough validation of the classification performances of the CSP algorithms, an additional analysis that was conducted by Park et al. was adopted by selecting the significant subjects prior to any further analysis [11]. This is crucial due to the possibility of little brain network information in the recorded EEG and activities of poorly performed subjects, based on the study of Ahn and Jun [31]. For these reasons, the subjects were categorized as significant, when the performance of each subject exceeded the minimum classification accuracy of 64%, defined using the confidence limit of 95% [32]. Figure 1 shows the number of significant subjects for each CSP algorithm. It can be observed that the number of significant subjects using CASUT was the highest out of all four CSP algorithms. The results throughout this chapter were based on the histograms of Figure 1, from which the data of the significant subjects were chosen for further analysis.

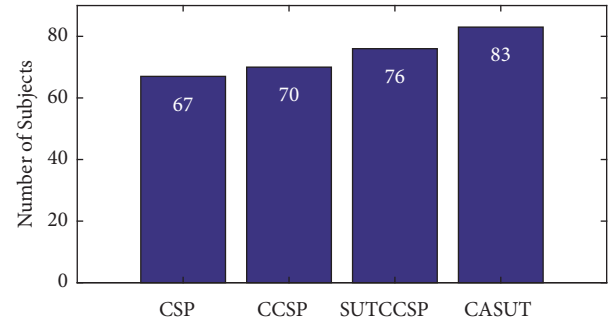


FIGURE 1: Number of significant subjects of CSP, CCSP, SUTCCSP, and CASUT. Note that the number of subjects for CASUT is the highest out of the four CSP algorithms.

Table 4 lists the average classification accuracies over the significant subjects and their standard deviations for CSP algorithms. It can be also noted that the average classification rate of CASUT outperformed those of CSP, CCSP, and SUTCCSP. The KS test was also performed for the significant subjects. However, the results from Table 5 indicated that the classification accuracies of the CSP algorithm do not follow the normal distribution. Accordingly, the Friedman test, which can be used regardless of the normality, was conducted. The p -value from the Friedman test yielded a value less than 10^{-12} , and thus the post hoc test was conducted and shown in Table 6. Note that the low p -values (<0.01) by the Wilcoxon signed rank test demonstrate the enhanced performance of CASUT.

Additional plots of the error bar and whisker diagram of the classification accuracies of CSP, CCSP, and CASUT were illustrated in Figures 2 and 3, respectively. The blue crosses in Figure 2 were identical to the average classification rates shown in Table 4. The red lines in Figure 3 indicate the median classification rates, and it can be observed that the median of CASUT outperforms those of the other three CSP algorithms. The superiority of CASUT over the other CSP algorithms was also confirmed by the Wilcoxon signed rank test results in Table 6.

In Figure 4, the scatterplots comparing classification rates of CASUT with CSP, CCSP, and SUTCCSP were displayed. The red dots above the dotted green lines indicate that classification rates were higher by CASUT than the other CSP algorithm. The black dots mean that CASUT and the compared CSP algorithm have the same classification rates,

TABLE 4: Average classification accuracies across the significant subjects of CSP, CCSP, SUTCCSP, and CASUT.

| CSP method | CSP | CCSP | SUTCCSP | CASUT |
|-----------------------------|--------------|--------------|--------------|--------------|
| Classification accuracy (%) | 74.68 ± 1.33 | 75.06 ± 1.36 | 77.20 ± 1.27 | 78.10 ± 1.18 |

TABLE 5: The resulting p -values of the KS test for each CSP algorithm for significant subjects.

| CSP method | CSP | CCSP | SUTCCSP | CASUT |
|-------------|--------|--------|---------|--------|
| p -values | 0.2087 | 0.0359 | 0.0282 | 0.0418 |

TABLE 6: Results of the Wilcoxon signed rank test conducted on the classification accuracies of CASUT compared with those of CSP, CCSP, and SUTCCSP for significant subjects.

| | p_1 | p_2 | p_3 |
|------------|------------|------------|---------|
| p -value | $<10^{-7}$ | $<10^{-8}$ | <0.01 |

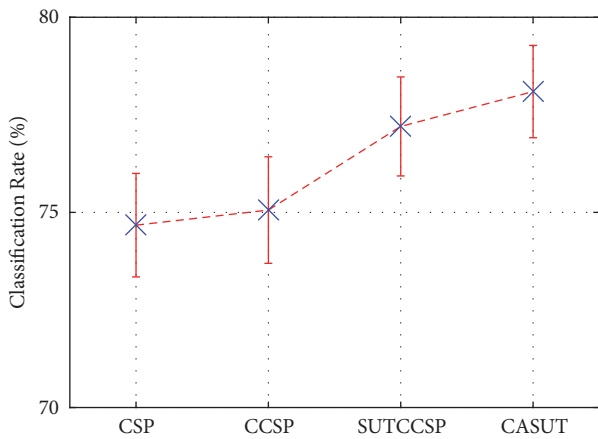


FIGURE 2: Error bar of the classification accuracies of CSP, CCSP, SUTCCSP, and CASUT. Note that CASUT produces higher classification rates compared with those of the other CSP algorithms, which is confirmed by the Wilcoxon signed rank test results of Table 6.

and blue means that the performances of CASUT are lower than those of the compared CSP algorithm. This demonstrates the fact that the majority of classification accuracies by CASUT were higher than those of the other CSP algorithms. Additionally, when two or more subjects yielded the same classification accuracies by two of the different algorithms, the dots for the subjects in these figures were duplicated. Therefore, the number of selected subjects in Figure 1 and the number of dots in Figure 4 may differ.

Lastly, the number of subjects, classified significantly using CASUT and classified insignificantly using the other CSP algorithms, was counted and shown in Figure 5. The bar chart indicates the number of subjects that were classified as significant by CASUT, but not by CSP, CCSP, and SUT, respectively.

On the other hand, there was only one subject whose data was classified as insignificant by CASUT, while the other CSP algorithms classified it as significant. These results also demonstrate the superiority of CASUT over the other conventional CSP algorithms.

TABLE 7: Average classification accuracies across 105 subjects of CSP, CASUT, CCSP, CACCSP, SUTCCSP, and CASUT.

| CSP method | Classification accuracy (%) |
|---------------|-----------------------------|
| CSP | 74.06 ± 1.30 |
| CACSP | 75.11 ± 1.33 |
| CCSP | 74.18 ± 1.38 |
| CACCSP | 74.83 ± 1.30 |
| SUTCCSP | 76.45 ± 1.27 |
| CASUT | 77.36 ± 1.19 |

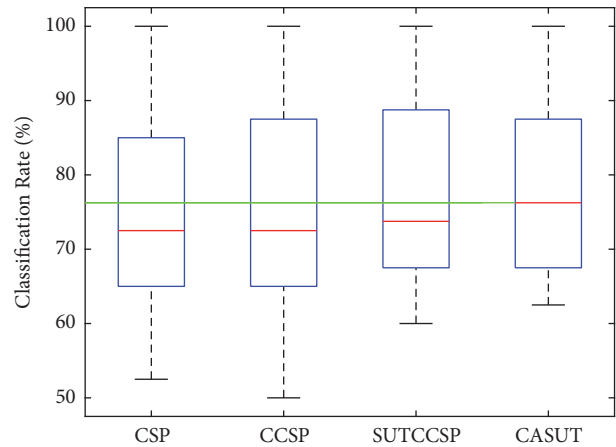


FIGURE 3: Whisker diagram of the classification accuracies of CSP, CCSP, SUTCCSP, and CASUT. The median of CASUT is highest compared with CSP, CCSP, and SUTCCSP.

3.2.3. *Analysis of Correlation Assisted CSP.* The various versions of CSP algorithms were additionally investigated for further interpretation of the effects of correlation information on the features of motor imagery tasks. To this end, correlation assisted CSP (CACSP) is defined as a CSP algorithm containing the correlation information, whereas correlation assisted CCSP (CACCSP) is defined as CCSP including the correlation information. The benchmark tests including CSP, CACSP, CCSP, CACCSP, SUTCCSP, and CASUT could provide an exact interpretation of the effects of correlation information on the features of the motor imagery tasks.

Table 7 lists the average classification rates calculated using CSP, CACSP, CCSP, CACCSP, SUTCCSP, and CASUT

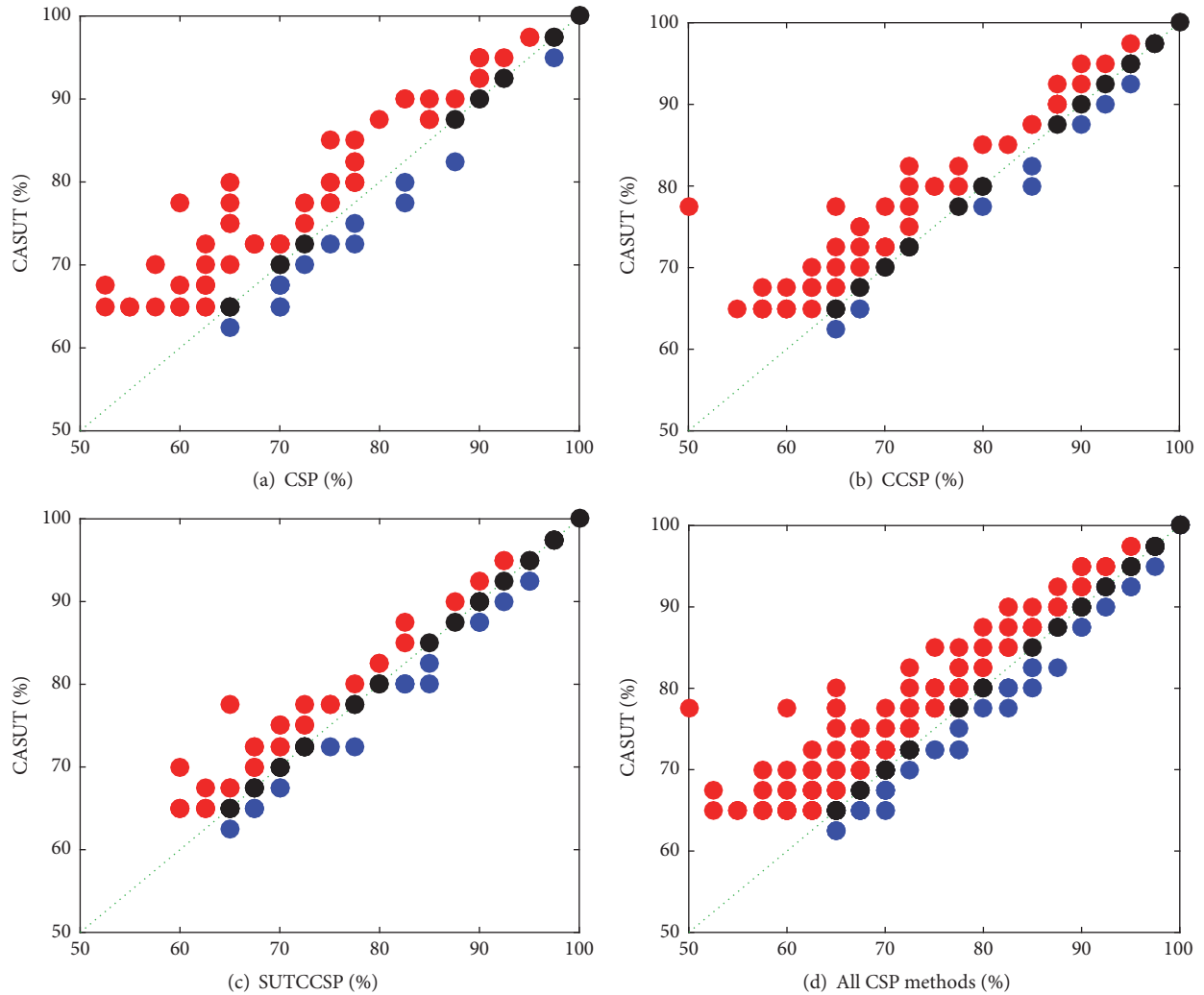


FIGURE 4: Scatterplot of classification rates of CASUT with (a) CSP, (b) CCSP, (c) SUTCCSP, and (d) the overlapping results of (a), (b), and (c). Note that most of the dots are located above the dotted green line, which indicates higher performance of CASUT.

TABLE 8: Results of the Wilcoxon signed rank test conducted on performance accuracies of CASUT compared with those of CSP, CACSP, CCSP, CACCSP, and SUTCCSP.

| | p -value |
|------------------------------|------------|
| p_1 | $<10^{-7}$ |
| $\mathbf{P}_{\text{CACSP}}$ | $<10^{-5}$ |
| p_2 | $<10^{-8}$ |
| $\mathbf{P}_{\text{CACCSP}}$ | $<10^{-7}$ |
| p_3 | <0.01 |

in the same conditions set in Table 4. Friedman test was conducted and a p -value less than 10^{-15} was confirmed. In Table 8, the Wilcoxon signed rank test was performed on CASUT with the other CSP algorithms, including CACSP and CACCSP. Results in bold show the results of the additional implementations of CSP and CCSP, that is, CACSP and CACCSP, respectively. Note that all p -values are significant,

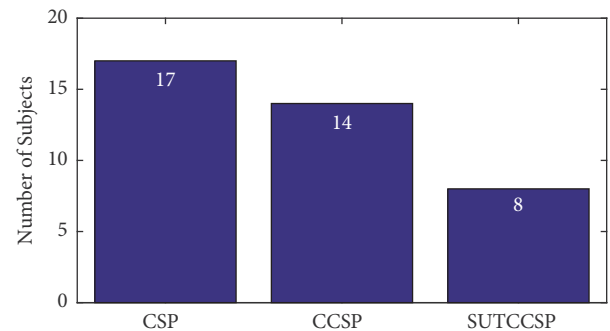


FIGURE 5: Number of subjects that were classified as significant with CASUT, but not with CSP, CCSP, and SUT, respectively.

indicating the enhanced performance of CASUT over the others. Since CCSP contains the power sum information, additional to the CSP features, and SUTCCSP preserves the power difference information, supplementary to CCSP, the

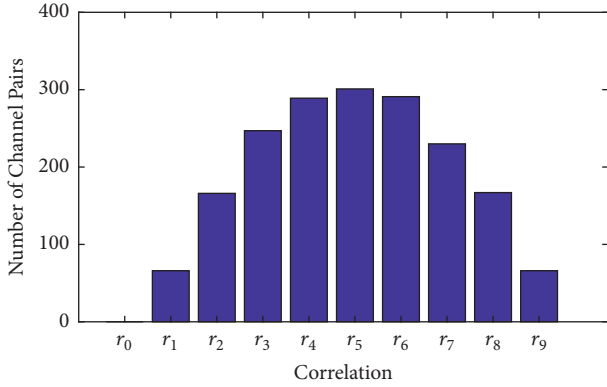


FIGURE 6: Number of channel pairs for each correlation range (r_0 to r_9).

gradually increasing classification rates could be expected as shown in Table 4.

Similarly, the performances of CSP and CCSP increase as the correlation information is added to their original features. Additionally, the highest classification accuracy in these benchmark tests was yielded using CASUT, indicating that CASUT outperforms all former CSP algorithms introduced so far.

4. Discussion and Conclusion

The correlation range chosen to evaluate the performance of CASUT was r_9 , based on (15). As shown in Figure 6, the number of channel pairs for each correlation range (r_0 to r_9) differs from zero to 301 channel pairs. In order to examine the effects of the correlation information on the CSP algorithms, the average classification accuracies over 105 subjects across different correlation ranges were calculated based on the same analysis in Section 3. Results demonstrate that the performance of CASUT gradually decreases as the correlation information is degraded as shown in Figure 7. Additionally, Figure 8 illustrates the resulting p -values estimated using the Wilcoxon signed rank test on CASUT compared with SUTCCSP, indicating less significance with small correlation coefficients. This proves that CASUT is the most effective feature extraction approach, when sufficient correlation information exists among the multichannel data.

The limitations of SUTCCSP have been addressed in this study due to the loss of the correlation information during the simultaneous diagonalization process of the covariance and pseudocovariance matrices. To that end, the correlation assisted version of SUTCCSP, denoted by CASUT, has been proposed for the first time by preserving the correlation information among multichannel data. The proposed algorithm was tested on the two-class motor imagery classification problem, and its classification accuracies obtained using the channel pairs with high correlation were significantly improved by CASUT compared with those of CSP, CCSP, and SUTCCSP, with p -values less than 0.01. Additional experiments on the various ranges of correlation prove that the correlation information is crucial to the classification of

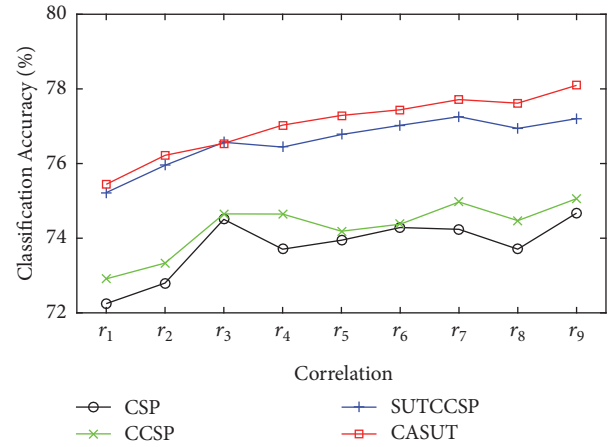


FIGURE 7: Classification accuracies for different correlation ranges (r_1 to r_9) of CSP, CCSP, SUTCCSP, and CASUT.

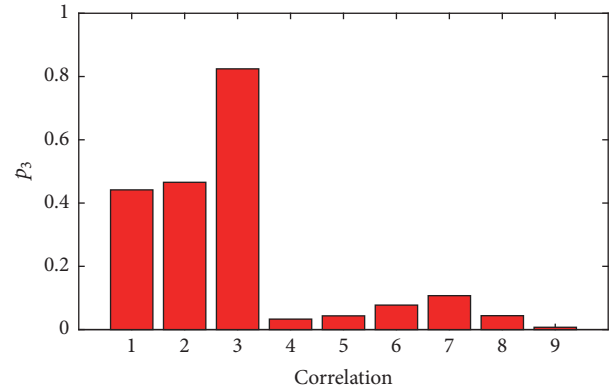


FIGURE 8: Resulting p -values of Wilcoxon signed rank tests conducted on CASUT with SUTCCSP for different correlation ranges (r_1 to r_9).

the two-class motor imagery tasks and that CASUT yields the highest classification accuracies compared with the other CSP algorithms.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Youngjoo Kim and Jiwoo You participated in the design of the study, carried out the key experiments and analyses, and drafted the manuscript. Youngjoo Kim and Jiwoo You are equal contributors. Heejun Lee and Seung Min Lee helped in drafting and revising the manuscript. Cheolsoo Park supervised the experiments and analyses. Youngjoo Kim, Jiwoo You, Heejun Lee, Seung Min Lee, and Cheolsoo Park all read and approved the final manuscript.

Acknowledgments

The present research was supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (no. 2017-0-00167, Development of Human Implicit/Explicit Intention Recognition Technologies for Autonomous Human-Things Interaction) and the Research Grant of Kwangwoon University in 2018.

References

- [1] A. B. Usakli, "Improvement of EEG signal acquisition: an electrical aspect for state of the Art of front end," *Computational Intelligence and Neuroscience*, vol. 2010, Article ID 630649, 2010.
- [2] P.-E. Aguera, K. Jerbi, A. Caclin, and O. Bertrand, "ELAN: a software package for analysis and visualization of MEG, EEG, and LFP signals," *Computational Intelligence and Neuroscience*, vol. 2011, Article ID 158970, 11 pages, 2011.
- [3] Y. Choi, "Data-driven Complexity Measure of an EEG with Application to Brain Injury and Recovery," *IEIE Transactions on Smart Processing & Computing*, vol. 6, no. 5, pp. 334–340, 2017.
- [4] 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, no. 1, pp. 17–21, 2013.
- [5] C. Kim, H. Kim, S. Kim, H. Park, and J. Lee, "A novel non-contact heart rate estimation algorithm and system with user identification," *IEIE Transactions on Smart Processing & Computing*, vol. 5, pp. 395–402, 2016.
- [6] 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.
- [7] 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.
- [8] Z. J. Koles, M. S. Lazar, and S. Z. Zhou, "Spatial patterns underlying population differences in the background EEG," *Brain Topography*, vol. 2, no. 4, pp. 275–284, 1990.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] O. Falzon, K. P. Camilleri, and J. Muscat, "The analytic common spatial patterns method for EEG-based BCI data," *Journal of Neural Engineering*, vol. 9, no. 4, Article ID 045009, 2012.
- [14] J. Navarro-Moreno, M. D. Estudillo-Martínez, R. M. Fernández-Alcala, 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.
- [15] Y. Kim, J. Ryu, K. K. Kim, C. C. Took, D. P. Mandic, and C. Park, "Motor Imagery Classification Using Mu and Beta Rhythms of EEG with Strong Uncorrelating Transform Based Complex Common Spatial Patterns," *Computational Intelligence and Neuroscience*, vol. 2016, Article ID 1489692, 2016.
- [16] C. C. Took, S. C. Douglas, and D. P. Mandic, "Maintaining the integrity of sources in complex learning systems: Intraference and the correlation preserving transform," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 3, pp. 500–509, 2015.
- [17] S. C. Douglas, J. Eriksson, and V. Koivunen, "Adaptive estimation of the strong uncorrelating transform with applications to subspace tracking," in *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006*, pp. 941–944, May 2006.
- [18] A. Rana and S. Arora, "Comparative Analysis of Medical Image Fusion," *International Journal of Computer Applications*, vol. 73, no. 9, pp. 10–13, 2013.
- [19] E. Mostacci, C. Truntzer, H. Cardot, and P. Ducoroy, "Multivariate denoising methods combining wavelets and principal component analysis for mass spectrometry data," *Proteomics*, vol. 10, no. 14, pp. 2564–2572, 2010.
- [20] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1–3, pp. 37–52, 1987.
- [21] 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.
- [22] A. L. Goldberger, L. A. 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.
- [23] "General-Purpose Software System for Brain-Computer Interface (BCI)," 2016, <http://www.bci2000.org>.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] B. D. Spurr and W. W. Daniel, "Applied Nonparametric Statistics," *Biometrics*, vol. 34, no. 4, p. 721, 1978.
- [29] M. Friedman, "The use of ranks to avoid the assumption of normality implicit in the analysis of variance," *Journal of the American Statistical Association*, vol. 32, no. 200, pp. 675–701, 1937.

- [30] M. Friedman, "A comparison of alternative tests of significance for the problem of m rankings," *The Annals of Mathematical Statistics*, vol. 11, no. 1, pp. 86–92, 1940.
- [31] 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.
- [32] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, G. Pfurtscheller, and G. R. Müller-Putz, "Better than random? A closer look on BCI results," *International Journal of Bioelectromagnetism*, vol. 10, no. 1, pp. 52–55, 2008.