Medical & Biological Engineering & Computing (2022) 60:753–767
https://doi.org/10.1007/s11517-021-02488-7

ORIGINAL ARTICLE

Euler common spatial patterns for EEG classification

Jing Sun[1,2] · Mengting Wei[3] · Ning Luo[4] · Zhanli Li[5] · Haixian Wang[1,2]

Received: 27 October 2020 / Accepted: 15 November 2021 / Published online: 22 January 2022

Abstract
The technique of common spatial patterns (CSP) is a widely used method in the field of feature extraction of electroencephalogram (EEG) signals. Motivated by the fact that a cosine distance can enlarge the distance between samples of different classes, we propose the Euler CSP (e-CSP) for the feature extraction of EEG signals, and it is then used for EEG classification. The e-CSP is essentially the conventional CSP with the Euler representation. It includes the following two stages: each sample value is first mapped into a complex space by using the Euler representation, and then the conventional CSP is performed in the Euler space. Thus, the e-CSP is equivalent to applying the Euler representation as a kernel function to the input of the CSP. It is computationally as straightforward as the CSP. However, it extracts more discriminative features from the EEG signals. Extensive experimental results illustrate the discrimination ability of the e-CSP.

Graphical abstract

Keywords Euler representation · Common spatial patterns (CSP) · Brain-computer interface (BCI) · Electroencephalogram (EEG) · Feature extraction

1 Introduction

The automatic classification of movement-related EEG signals has received increasing attention in the field of brain-computer interfaces (BCIs). BCIs are a way to establish a direct connection between the human brain and external devices [1], and noninvasive EEG-based BCIs can be used for the rehabilitation of brain nerves [2]. EEG-based BCIs include P300, motor imagery (MI), and other types. In most cases, the P300 paradigm requires the subject's eyes to always look at the visual stimuli appearing on the screen, and at the same time, the EEG signal of this process is acquired. The P300 signal generally appears 300–400 ms after the event and is a positive potential signal [3]. The latency and amplitude of the P300 signal indirectly reflect the emotional and psychological changes of the subject when facing a stimulus or reflecting a potential intent [4]. The advantage of the P300 signal is that the subject can generate it without training because it is one of the internal signals in the brain [5]. However, when performing the analysis, it is necessary to overlay and average the signals many times to observe the changes in the EEG signal caused by the stimulation event. Furthermore, during the process of acquiring signals, the subject's eyes will be easily fatigued, and the practical effect is possibly reduced [6]. Motor imagery refers to a process of thinking that uses the brain to imagine the movement of its limbs, but there is actually no movement [7]. This thinking process is spontaneous by the brain and does not require external stimulation [8].

BCIs based on motor imagery have recently received increasing attention. To obtain a more effective EEG MI BCI, there are generally two optimization strategies: one

✉ Haixian Wang
hxwang@seu.edu.cn

1 Key Laboratory of Child Development and Learning Science of Ministry of Education, School of Biological Science & Medical Engineering, Southeast University, Nanjing 210096, Jiangsu, People's Republic of China

2 Institute of Artificial Intelligence of Hefei Comprehensive National Science Center, Hefei 230094, Anhui, People's Republic of China

3 Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

4 Institute of Software, Chinese Academy of Sciences, Beijing 100190, People's Republic of China

5 College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710054, Shanxi, People's Republic of China

is to maintain the high separability of the MI EEG features during feature extraction, and the other is to optimize the classifier based on the specific needs. In contrast, the use of the former as an optimization strategy has attracted increasing attention in recent years. In machine learning, many methods exist for EEG feature extraction, which can be divided into two categories according to the purpose of the processing: temporal filtering and spatial filtering. The temporal filter focuses on the temporal aspects of EEG signals, and its advantage lies in its low computational complexity while retaining an acceptable discrimination performance among simple MI tasks [9]. Spatial filtering reduces the influence of spatial noise. Typical spatial filters include common spatial pattern (CSP) and common average reference (CAR) filters [10]. These spatial filter methods are helpful to solve the problem of MI binary classification. In 1999, G. Pfurtscheller and F.H. Lopes da Silva [11] confirmed that when preparing and performing an exercise with one hand, the amplitude of the α band (8–12 Hz) and β band (13–30 Hz) in the contralateral sensorimotor cortical EEG signals decreased. This phenomenon is called event-related desynchronization (ERD), which is the reduced amplitude of activated cortical EEG signals. At the same time, the amplitude of the α band (8–12 Hz) and β band (13–30 Hz) in the cerebral cortex signaled on the same side as the exercising hand increased; this signal is called event-related synchronization (ERS). This outcome means that the corresponding cortex increased in amplitude in the resting state [12]. Some studies have also indicated that motor imagery tasks are associated with attenuation or increase in localized brain rhythm activity called ERD/ERS [13].

One of the most popular and efficient methods to extract ERD/ERS-related features is the CSP, which is widely used for motor imagery BCI designs [14, 15]. CSP aims to find an optimal spatial filter by maximizing the filtered variance of one class and minimizing the filtered variance of the other class simultaneously [16]. Generally, the improvement of CSP can be divided into two categories according to the direction: one is to optimize the objective function, and the other is to optimize the channel quality of EEG signals. Some studies adopt the second strategy to improve CSP; for instance, Dong et al. [17] established the CSP-CIM to improve the robustness of CSP with respect to outliers. The correntropy-induced metric (CIM) can saturate when two vectors are far apart from each other in the input space, and this property makes it appropriate for learning problems with outliers in the data and helps to build a robust objective function [17]. Of course, there are also studies that adopt the first strategy to improve CSP. Y. Park and W. Chung [18, 19] considered a local CSP generated from individual channels and their neighbors (termed "local regions" [20]) rather than a global CSP generated from all channels. To overcome some of the shortcomings of the global CSP approach, a local temporal CSP (LTCSP) was developed in [16]. Recently, Z. Yu et al. [21] proposed introducing a weight function based on the phase-locking value (PLV) [22] into the framework of LTCSP. B. Chakraborty et al. [23] also considered using the phase information and amplitude information of the EEG signal to model the CSP. The novelty of this method is that the formula for calculating phase information and amplitude information is completed with the help of complex space. A sparse CSP algorithm that incorporates the sparse technique and iterative search into the CSP was proposed by Fu et al. [24]. Geirnaert et al. [25] proposed a new attention decoding method by using the filter-bank common spatial pattern filters (FB-CSP). Compared to the traditional auditory attention decoding (AAD) algorithms, FB-CSP does not require access to clean source envelopes. More recently, an integrated framework, termed common time–frequency-spatial patterns (CTFSP), was proposed to extract sparse CSP features from multiple frequency bands in multiple time windows [26].

A new optimization method that can increase the difference between different classes of motor imagery tasks is desirable. Recent studies have indicated that the use of the Euler representation for feature extraction has greatly contributed to the improvement of image classification accuracy. By combining these studies [27–30], some advantages of the Euler representation have been confirmed as follows. (a) The Euler representation obtains a larger margin, which is helpful for image classification by using Euler representation to measure the distance between pixels in each image. (b) Moreover, compared with the measurement of images of the same category, the Euler representation can obtain a larger margin when measuring the distance between images of different categories. In other words, it can obtain a better separability between different class images.

Inspired by these previous research results, it is also possible to utilize the Euler representation for the extraction of EEG signal features. Therefore, in our proposed method, we apply the Euler representation to discriminate the characteristics of EEG data between the different MI tasks. In this paper, we present a novel feature extraction method for MI classification, which is called e-CSP. The method uses the Euler representation as a kernel function in the input of the CSP. Each sample value is mapped to a complex space by the formula of the Euler representation, and then the CSP is performed in Euler space. We use linear discriminant analysis (LDA) to observe the classification performance of our proposed method. Higher classification accuracy indicates higher class separability. A higher separability can be considered a better motor imagery discrimination performance [31]. The rest of the paper is organized as follows. Section 2 introduces the conventional CSP algorithm and our proposed method. The experiment and dataset description is presented in Sect. 3. Computational results are analyzed in Sect. 4. Finally, the conclusions are drawn in Sect. 5.

2 Methods

2.1 Common spatial pattern

The input EEG signals can be defined as a matrix of [N*C*T], where C is the number of EEG channels, N represents the number of time samples per trial, and T is the number of trials. It is assumed here that the data matrix of each trial is zero mean, i.e., each class of data is the center matrix. The sample covariance matrices of the two classes are computed as

$$R_1 = \frac{X_1 X_1^T}{tr(X_1 X_1^T)}, R_2 = \frac{X_2 X_2^T}{tr(X_2 X_2^T)}, \tag{1}$$

where X_1 and X_2 are sample data from two classes, while R_1 and R_2 are each of the respective covariance matrices. The final spatial covariance $\overline{R_1}$ and $\overline{R_2}$ are computed by averaging over the trials under each class. Then, the total covariance matrix R can be formed as

$$R = \overline{R_1} + \overline{R_2}. \tag{2}$$

Due to the symmetry of the covariance matrix R, the eigenvalue decomposition is given by

$$R = UDU^T,$$

where D is the diagonal matrix formed by the corresponding eigenvalues, and U is the eigenvector matrix of the matrix. Since the eigenvalues are arranged in ascending order [32, 33], descending order is performed first. Then, the whitening value matrix of R is computed as

$$P = \sqrt{D^{-1}}U^T \tag{3}$$

and transforms the mean covariance matrices $\overline{R_1}$ and $\overline{R_2}$ as

$$S_1 = P\overline{R_1}P^T, S_2 = P\overline{R_2}P^T.$$

Note that, by principal component decomposition, S_1 and S_2 can be formulated as

$$S_1 = B_1\Lambda_1 {B_1}^T, S_2 = B_2\Lambda_2 {B_2}^T,$$

where B_1 and B_2 are the eigenvector matrices of S_1 and S_2, and Λ_1 and Λ_2 are associated eigenvalue matrices with diagonal entries. It can be demonstrated that the eigenvectors of the matrix and the eigenvector matrices are equal [32, 34]; at the same time, the sum of the diagonal arrays and of the two eigenvalue matrices is an identity matrix. That is

$$B_1 = B_2 = B, \Lambda_1 + \Lambda_2 = I.$$

Thus, the eigenvector corresponding to the largest eigenvalue allows S_1 to S_2 to have the smallest eigenvalue and vice versa. Finally, the projection matrix is the corresponding spatial filter W, which is defined as

$$W = B^T P. \tag{4}$$

Using the projection matrix W, the EEG signals X_1 and X_2 are projected as

$$F_1 = W^T X_1 \text{ and } F_2 = W^T X_2,$$

where F_1 and F_2 (usually after a log transformation) are used for classification.

Combining the aforementioned analysis, we find that the spatial filter is obtained by comparing the filtered variance of one class to the filtered variance of the other class. Thus, the objective function of CSP can be written as

$$J(w) = \frac{tr(w^T\overline{R_1}w)}{tr(w^T\overline{R_2}w)}, \tag{5}$$

where w is the spatial filter W obtained by the above process. $\overline{R_1}$ and $\overline{R_2}$ are computed by averaging over the trials under each class. When the objective function reaches a maximum, the spatial filter W is optimal.

2.2 Euler representation for CSP

2.2.1 Motivation

To solve the optimization problem of the spatial filter in the CSP objective function, we consider using the kernel trick property to capture the nonlinear similarity of features to improve the discrimination ability of CSP. Different from the commonly used kernel function, which maps data into a higher-dimensional hidden space, the Euler representation maps data into an explicit space that has the same dimensionality as that of the original data space. In our proposed Euler representation for the CSP method, the Euler representation is used as a kernel function in the input of the algorithm, and it is expected to improve the CSP. To better explain the formulation of the Euler representation for CSP, we first introduce the definition of the cosine kernel function, i.e., cosine distance, and then analyze its advantages.

Definition: Given two arbitrary vectors x_j and $x_p \in R^m$ [35], the cosine distance metric between them is.

$$d(x_j, x_p) = \sum_{c=1}^{m} \{1 - \cos(\alpha\pi(x_j(c) - x_p(c)))\}, \tag{6}$$

where $x_j(c)$ is the c-th element of x_j.

In most feature extraction algorithms, the Euclidean distance is generally used as the measurement metric of discrimination between data. In fact, the cosine distance can also be used as a measurement metric. Two classes of motor imagery data are marked as A and B. Tables 1 and 2 list the distance between the same class and different classes of data under different

Table 1 Comparison of normalized dissimilarity measures between the same class

Subject	1	2	3	4	5
Dissimilarity	A-A	A-A	A-A	A-A	A-A
Euclidean	0.2674	0.2909	0.2245	0.3241	0.2635
Cosine-based	99.64	115.25	74.98	134.49	98.51

Table 2 Comparison of normalized dissimilarity measures between different classes

Subject	1	2	3	4	5
Dissimilarity	A-B	A-B	A-B	A-B	A-B
Euclidean	0.3266	0.3923	0.3491	0.5267	0.3699
Cosine-based	134.01	168.66	146.43	238.11	159.94

measurement metrics. By comparing the ratio of the between-class distance to the within-class distance, we find that under the cosine distance metric, this ratio is always greater than that of the Euclidean distance. The results illustrate that the cosine distance is more advantageous in determining the difference between classes, which will greatly contribute to improving the accuracy of classification. Thus, the cosine distance metric, which is defined in Eq. (6), is a more effective kernel function.

Combining the aforementioned analysis and the comparison with the Euclidean distance, we can obtain the following two characteristics of cosine distance:

(1) Using the cosine distance metric enlarges the distance between all sample value margins between different channels at the same time point. This enlargement demonstrates that the metric of the cosine distance can help obtain a large margin, making it advantageous for the classification of motor imagery.
(2) Compared to Euclidean distance, the cosine distance metric can enlarge the distance between samples of different classes.

Therefore, the cosine distance metric not only obtains a greater degree of discrimination between different channels but also increases the degree of discrimination between the different classes of motor imagery data. This outcome greatly improves classification accuracy. Combining the aforementioned analysis for the cosine distance metric, we build a robust formulation for CSP. By simple algebra, Eq. (6) can be rewritten as Eq. (7).

$$\begin{aligned} d(x_j,x_p) &= \sum_{c=1}^{m}\{1-\cos(\alpha\pi(x_j(c)-x_p(c)))\} \\ &= \left\|\frac{1}{\sqrt{2}}(e^{i\alpha\pi x_j}-e^{i\alpha\pi x_p})\right\|_2^2 \\ &= \|z_j-z_p\|_2^2 \end{aligned} \tag{7}$$

where

$$z_j=\frac{1}{\sqrt{2}}\begin{bmatrix} e^{i\alpha\pi x_j(1)} \\ \dots \\ e^{i\alpha\pi x_j(c)} \end{bmatrix}=\frac{1}{\sqrt{2}}e^{i\alpha\pi x_j} \tag{8}$$

is called the Euler representation of x_j.

Equation (7) illustrates that the sample values of different channels at two time points obtained by the cosine distance metric are equivalent to the ℓ_2-norm between the corresponding two vectors with the Euler representation. In addition, the ℓ_2-norm is a better function expression of the optimization problem than the cosine distance in practical applications. Therefore, our proposed method first maps the sample values x_j normalized in [0, 1] onto the complex representation $z_j \in C^m$ and then performs the complex CSP in Euler space. To more intuitively observe the contribution of the Euler representation for EEG data, we use its formulation, which is defined in Eq. (8) for EEG data. As shown in Fig. 1, the results indicate that the use of the Euler representation can obtain more consequential information. This outcome further confirms that our consideration is feasible and that the Euler representation is meaningful as an optimization strategy for CSP.

2.2.2 Euler CSP

The Euler representation is used to transform the two classes of motor imagery data, and for the EEG data of a single trial, $Z_{n1}=\frac{1}{\sqrt{2}}e^{i\alpha\pi X_{n1}}$, where X_{n1} is a vector consisting of the sampled values of each channel at a time point, then the EEG data for a single trial can be denoted as $Z=[Z_{n1},Z_{n2}...Z_N]$, and data for all trials are denoted as such.

By simple algebra [16], the objective function of CSP (Eq. (5)) gives

$$\begin{aligned} J(w) &= \frac{tr(w^T\overline{R_1}w)}{tr(w^T\overline{R_2}w)} \\ &= \frac{tr(w^T(\frac{1}{T_1}\sum_{i=1}^{T_1}R_1(i))w)}{tr(w^T(\frac{1}{T_2}\sum_{i=1}^{T_2}R_2(i))w)} \\ &= \frac{\frac{1}{T_1}\sum_{i=1}^{T_1}\sum_{j=1}^{C}(\gamma_j^T X_1(i)X_1(i)^T/tr(X_1(i)X_1(i)^T))}{\frac{1}{T_2}\sum_{i=1}^{T_2}\sum_{j=1}^{C}(\gamma_j^T X_2(i)X_2(i)^T/tr(X_2(i)X_2(i)^T))} \end{aligned}, \tag{9}$$

where (i) denotes the i-th trial, γ_j is the j-th column of matrix w, and T_1 and T_2 correspond to the numbers of trials for X_1 and X_2, respectively. On account of the use of the Euler representation for X_1 and X_2, the numerator of the objective function becomes

$$\frac{1}{T_1}\sum_{i=1}^{T_1}\sum_{j=1}^{C}\left(\gamma_j^T Z_1(i)Z_1(i)^T/tr\left(Z_1(i)Z_1(i)^T\right)\right), \tag{10}$$

and the denominator is likewise formulated. The covariance matrices of X_1 and X_2 are calculated by

$$G_1=\frac{Z_1Z_1^T}{tr(Z_1Z_1^T)}, G_2=\frac{Z_2Z_2^T}{tr(Z_2Z_2^T)}, \tag{11}$$

where Z_1 and Z_2 are the Euler representations of X_1 and X_2, respectively. Then, the total covariance matrix G is calculated by

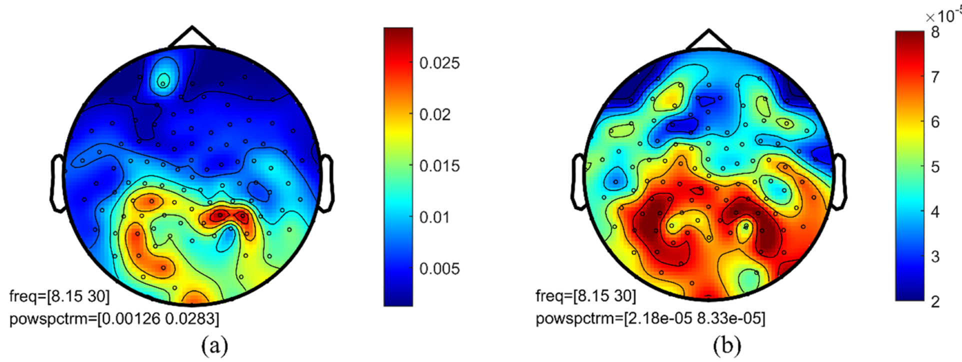


Fig. 1 Contribution of the Euler representation for EEG data. **a** The traditional data; **b** the data using the Euler representation

$$G = G_1 + G_2. \tag{12}$$

Hence, the objective function of e-CSP is rewritten as

$$J(Q) = \frac{tr(Q^T\overline{G_1}Q)}{tr(Q^T\overline{G_2}Q)}, \tag{13}$$

where Q is the new projection matrix of e-CSP, $\overline{G_1}$ is the spatial covariance matrix of the average trials for one class and $\overline{G_2}$ is the spatial covariance matrix of the average trials for the other class.

The procedure of solving the objective function, i.e., e-CSP, is summarized in the following algorithm.

Algorithm: The procedure of e-CSP

Input: EEG signals composed of multiple trials and multiple channels, which can be defined as a matrix of [N*C*T], where C is the number of EEG channels, N represents the number of time samples per trial and T is the number of trials.

Required: the data matrix of each trial is zero mean.

1. Calculate Z by Eq. (7) in X_1 and X_2.
2. Calculate the covariance matrix of X_1 and X_2 by Eq. (11) and the total covariance matrix by Eq. (12).
3. Use PCA to obtain the whitening eigenvalue matrix P.
4. Construct the spatial filter Q.
5. Feature extraction. The feature vector is denoted as $F = (f_1, f_2, ..., f_N)$. The features are log-transformed, i.e., $f_j \leftarrow \log(f_j / \sum f_j)$, where j is the j-th value of F.

Output: The e-CSP feature set F.

3 Data and experiment

In this section, we use two publicly available EEG data sets from BCI competitions and a new Motor Imagery dataset from Cho et al. [36] to evaluate the performance of the proposed algorithm and compare the results with the conventional CSP algorithm. In our experiments, linear discriminant analysis (LDA) and support vector machines (SVM) are used as classifiers. As shown in Fig. 2, the experiments mainly include three parts: preprocessing, feature extraction, and classification. To fully verify the performance of our proposed method, we design four experimental paradigms, which are described below.

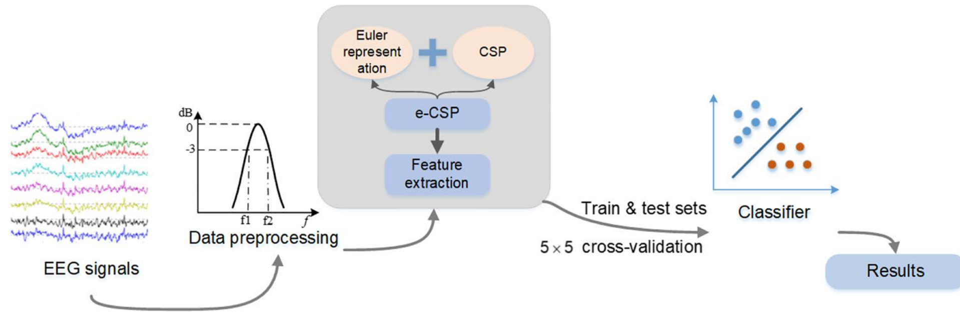


Fig. 2 The block diagram of the experiments

Table 3 BCI competition datasets information

Dataset	BCI competition III dataset IVa					BCI competition III dataset IIIa		
Subject	aa	al	av	aw	ay	s1	s2	s3
Number of training trials	168	224	84	56	28	90	60	60
Number of test trials	112	56	196	224	252	90	60	60
Total number of trials	280					180	120	120
Number of electrodes	118					60		
Sample rate	100 Hz					250 Hz		
MI task type	Right hand and right foot					Left hand and right hand		

(1) For experiment 1, we compare the classification accuracies of the e-CSP with other optimization methods of CSP in the BCI competition III dataset IVa and BCI competition III dataset IIIa publicly available datasets under the LDA classifier.
(2) For experiment 2, we empirically set the maximum number of projection vectors as the number of electrodes and set the parameter α in the e-CSP to 0.9, 1.2, and 0.5 in the BCI competition III dataset IVa, BCI competition III dataset IIIa, and Cho's dataset, respectively. In addition, to obtain the accuracies of the e-CSP and the CSP, we also calculated the F_1-score based on the accuracy to fully evaluate the performance of the e-CSP.
(3) For experiment 3, we set the parameter α in the e-CSP to 1.1 in the BCI competition III dataset IVa and BCI competition III dataset IIIa. We set the parameter α in the e-CSP to 0.5 in Cho's dataset. Then, we compare the classification accuracies of the e-CSP and CSP algorithms under the LDA and SVM classifiers.
(4) For experiment 4, we respectively set the number of projection vectors to 30, 48, and 24 in the BCI competition III dataset IIIa, BCI competition III dataset IVa, and Cho's dataset. We set the parameter α in the e-CSP from 0.1 to 1.9 to observe its influence on the accuracy.

All experiments are performed on the Windows 10 operating system, and the fivefold cross-validation was used for model evaluation. It includes three steps. First of all, the dataset was divided into five subsets, and then the four subsets of that were used to train the model while the rest subset was used to test and calculate the accuracy of the method; finally, this process was repeated five times such that each subset was used once as testing set. We recorded the final averaged accuracy. The rest of this section includes the description of the specific information of the dataset and the preprocessing of the data.

Table 4 Cho's dataset information

Dataset	Cho's dataset									
Subject	S01	S02	S03	S06	S14	S23	S35	S41	S43	S50
Number of training trials	150	66	150	150	100	150	150	134	100	150
Number of test trials	50	134	50	50	100	50	50	66	100	50
Total number of trials	200									
Number of electrodes	64									
Sample rate	128 Hz									
MI task type	Left hand and right hand									

Table 5 Classification accuracies of the proposed method and different methods in the BCI competition III dataset IVa. The bold content in each row of the table represents the highest accuracy achieved in that row

Subject	e-CSP	CSP phase + amp	LTCSP	p-LTCSP
aa	90.00 ± 1.8	**96.4 ± 1.22**	73.21 ± 1.76	77.68 ± 1.16
al	96.78 ± 1.16	66.96 ± 1.01	**98.00 ± 1.18**	97.10 ± 1.28
av	**93.10 ± 1.86**	57.95 ± 0.87	71.43 ± 1.77	71.94 ± 1.07
aw	**96.10 ± 1.86**	75.36 ± 1.13	89.72 ± 1.41	92.41 ± 1.38
ay	**95.60 ± 2.86**	64.76 ± 0.97	74.60 ± 1.91	72.41 ± 1.10
Mean	**94.31 ± 2.07**	72.29 ± 1.08	81.39 ± 1.33	82.63 ± 1.24

Table 6 Classification accuracies of the proposed method and different methods in the BCI competition III dataset IIIa. The bold content in each row of the table represents the highest accuracy achieved in that row

Subject	e-CSP	CSP phase + amp	LTCSP	p-LTCSP
s1	94.10 ± 2.35	76.48 ± 1.22	94.83 ± 1.89	**96.6 ± 1.88**
s2	**92.90 ± 2.87**	86.48 ± 1.66	73.30 ± 1.22	71.67 ± 1.28
s3	**97.50 ± 1.46**	80.87 ± 0.97	96.40 ± 0.97	97.50 ± 0.98
Mean	**94.83 ± 2.66**	81.27 ± 0.98	88.51 ± 1.77	88.59 ± 1.32

3.1 Data description

3.1.1 BCI competition datasets

The two publicly available datasets used to evaluate the performance of our proposed method are the BCI competition III dataset IVa and the BCI competition III dataset IIIa. The detailed information is described in Table 3.

3.1.2 Cho's dataset

This dataset conducted a BCI experiment for motor imagery movement (MI movement) of the left and right hands with 52 subjects (19 females, mean age ± SD age = 24.8 ± 3.86 years). Each subject took part in the same experiment, and subject ID was denoted and indexed as S01, S02, …, S52. In this paper, we randomly select the data of 10 subjects to evaluate the performance of the proposed method. The detailed information is described in Table 4.

For the MI performance, some brief descriptions were provided [36]. The performance of subjects "S02" and "S23" are worse than that of the other subjects, as will be seen in the experiments.

3.2 Data preprocessing

According to the background knowledge of neurophysiology, it can be known that during imagination, the neuronal cluster discharge of the sensorimotor cortex of the brain can cause ERS or ERD [37]. The characteristic frequency of these phenomena mainly includes an α-rhythm of 8–12 Hz and a β-rhythm of 13–30 Hz. Thus, in the preprocessing phase, a fifth-order Butterworth filter was first used for 8–30 Hz bandpass filtering. As in [38], the cut-off speed will increase when the order of the filter increases. For MI signals, the selection of the 5th order filter usually leads to the optimal cut-off speed. For the signal data for feature extraction, EEG data between 0.5 and 2.5 s after the visual cue of the motor imagery task were selected because this is the best time period for the EEG to detect ERD/ERS [38].

4 Results and discussion

In this section, the results of the three experiments will be detailed in turn. The classification accuracy is used as an index to evaluate the performance of the algorithm in our study. Higher accuracy indicates higher class separability. A higher separability can be considered a better motor imagery discrimination performance. Moreover, the F_1-score is also used as an index to compare the performance of the e-CSP and the CSP. The results include three parts: the performance comparison of e-CSP and CSP based on other optimization strategies, the performance comparison of e-CSP and CSP, and the impact of the parameter α in the Euler representation.

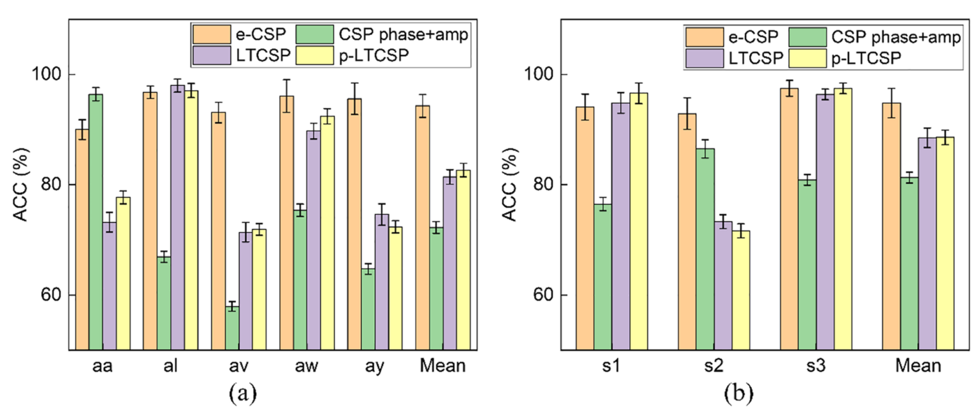
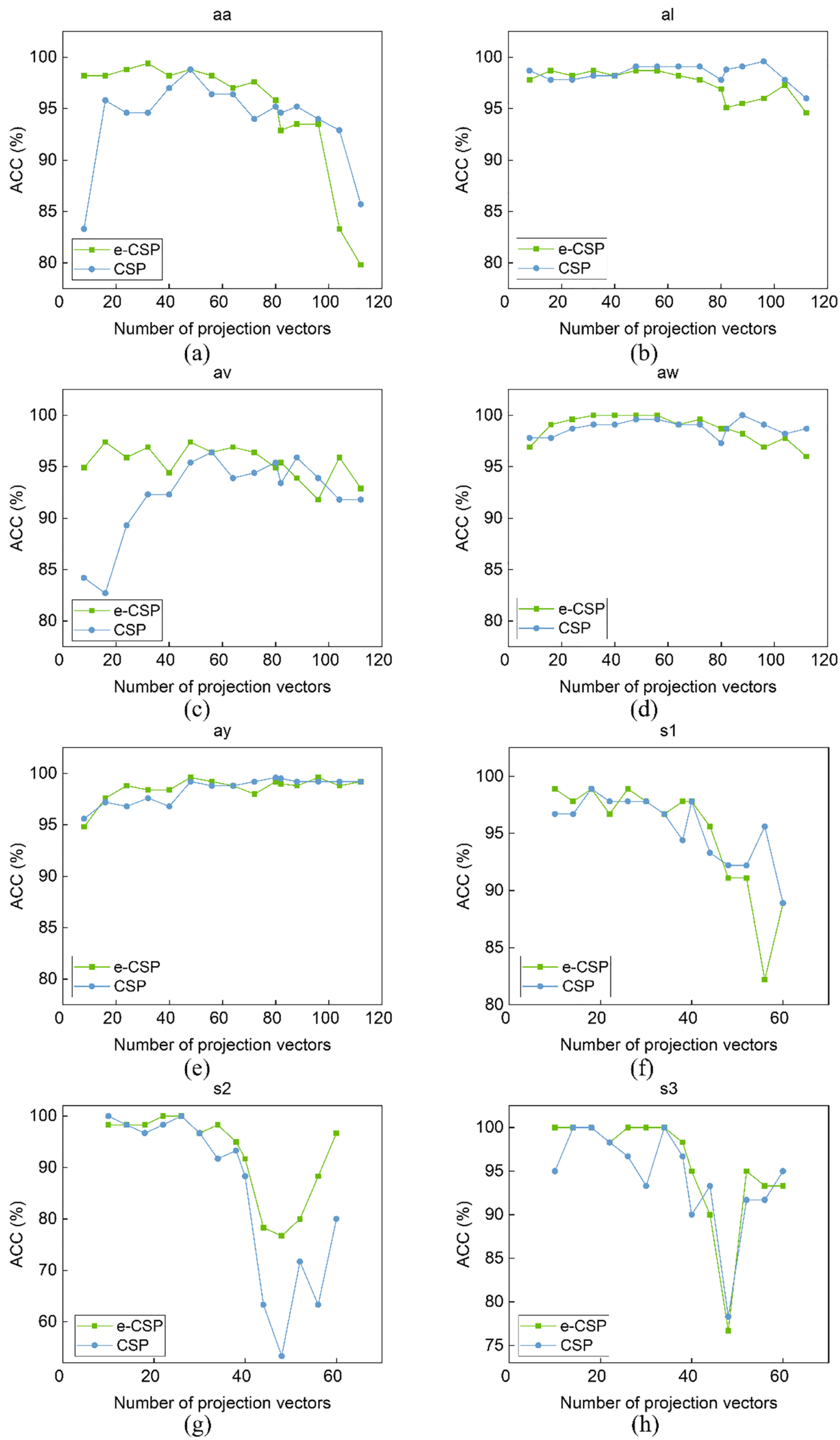


Fig. 3 Average classification accuracies of the proposed method and different methods on the two datasets. **a** The BCI competition III dataset IVa, and **b** the BCI competition III dataset IIIa


aa
ACC (%)
Number of projection vectors
e-CSP
CSP
(a)
al
ACC (%)
Number of projection vectors
e-CSP
CSP
(b)
av
ACC (%)
Number of projection vectors
e-CSP
CSP
(c)
aw
ACC (%)
Number of projection vectors
e-CSP
CSP
(d)
ay
ACC (%)
Number of projection vectors
e-CSP
CSP
(e)
s1
ACC (%)
Number of projection vectors
e-CSP
CSP
(f)
s2
ACC (%)
Number of projection vectors
e-CSP
CSP
(g)
s3
ACC (%)
Number of projection vectors
e-CSP
CSP
(h)

◂**Fig. 4** Average classification accuracies versus the number of projection vectors in the BCI competition III dataset IVa and BCI competition III dataset IIIa. The horizontal axis in the figure represents the number of projection vectors, and the vertical axis represents the classification accuracy

4.1 Performance comparison of the e-CSP with other methods on BCI competition datasets

To evaluate the performance of the proposed method, we also compare the classification accuracies of several other methods in two public competition datasets. We obtained the accuracies of the other three methods in the two competition datasets from previous studies [16, 21, 23]. The results are presented in Table 5, Table 6, and Fig. 3. Table 5 lists the classification accuracies of the proposed e-CSP and the other three methods in the BCI competition III dataset IVa. The e-CSP obtains higher classification accuracies for the subjects “av” and “aw” than the other three methods. The mean classification accuracy of the e-CSP is better than that of the other methods. The e-CSP classification accuracy is 22.02%, 12.92%, and 11.68% higher than those of the CSP phase + amp, LTCSP, and p-LTCSP, respectively. As seen in Table 6, the performance of the CSP phase + amp [23] is very unstable in comparison with the e-CSP. The LTCSP [16] and p-LTCSP [21] are both CSPs based on temporally local manifolds, and they both show better classification performance for subject “al” than the e-CSP. However, the mean classification accuracy of e-CSP is higher than that of the LTCSP and p-LTCSP.

We also evaluate the performance of the e-CSP in the BCI competition III dataset IIIa. Table 6 shows the classification accuracies of the e-CSP, CSP phase + amp, LTCSP, and p-LTCSP. It is evident that the e-CSP also achieves the greatest mean classification accuracy. The mean classification accuracy of the e-CSP is better than that of the other methods, with accuracies that are 13.56%, 6.32%, and 6.24% higher than those of the CSP phase + amp, LTCSP, and p-LTCSP, respectively.

The classification accuracies of the four methods on the two datasets are presented in Fig. 3. By contrast, the e-CSP has relatively superior classification accuracies across the eight subjects. A *T*-test was performed for the e-CSP and one of the other methods one by one. The results also reveal that the accuracies of the e-CSP are significantly higher than other methods ($P < 0.05$). The results demonstrate that the Euler representation is indeed a meaningful optimization strategy based on CSP.

4.2 Performance comparison of e-CSP with CSP

In addition, to verify the effectiveness of the e-CSP on BCI competition datasets, we also performed experiments using Cho’s dataset. The average classification accuracies across all subjects of the conventional CSP and the proposed e-CSP with the number of projection vectors can be seen in Fig. 4 and Table 7. As shown in Fig. 4, some results are revealed in the curve of the classification accuracies across the 8 subjects.

(1) When the number of projection vectors is small, the classification accuracy of the e-CSP is higher than that of the CSP, and when it reaches a certain threshold, the performance of the conventional CSP will be slightly higher than that of the e-CSP. For different subjects, the threshold points are different, but the overall performance trend is the same. As shown in Fig. 4(b), when the number of projection vectors is set as 40, the classification accuracy of the CSP and e-CSP is the same. When the number of projection vectors is less than 40, the classification accuracy obtained by the e-CSP is higher than that obtained by the CSP. However, when the number of projection vectors is greater than 40, the classification accuracy obtained by the e-CSP is lower than that of the CSP. The results illustrate that to achieve the same classification accuracy: the e-CSP requires fewer features than those of the CSP. In other words, the e-CSP is more effective.

(2) In addition, principal component analysis (PCA) is used to select the features with a variance of 95% from the extracted features. During the evaluation process, it is found that the proportion of the effective features among all the features extracted by the e-CSP is always higher than that of the CSP when extracting the same number of features. This result further validates the effectiveness of the e-CSP.

(3) As shown in Fig. 4(a), (b), and (d), although the accuracy obtained by the e-CSP is greater than 80% regardless of the number of the projection vectors, the accuracies show a decreasing trend. In previous works, the number of the pair of filters (i.e., projection vectors) was empirically set to 3 [37]. In this paper, in order to compare the impact of the number of projection vectors, we set it from the maximum value to the minimum value. The trend shown in the figure demonstrates that a few leading projection vectors contain more discriminative features.

The above results indicate that, compared with conventional CSP, the e-CSP shows better classification performance when the number of projection vectors is relatively small. When the number of features extracted by these two algorithms is the same, the features extracted by the e-CSP contain more effective features. These results reveal that the e-CSP has better feature discrimination performance, which provides a more efficient proposal for

Table 7 Classification accuracies versus the number of projection vectors for each subject in the Cho's dataset

Number of projection vectors	Method	Subject									
		S01	S02	S03	S06	S14	S23	S35	S41	S43	S50
4	CSP	46	46	52	56	77	52	90	66	66	46
	e-CSP	62	53	82	76	79	60	90	61	69	64
8	CSP	60	52	58	72	78	58	90	63	68	46
	e-CSP	66	60	84	78	80	58	90	72	74	60
12	CSP	68	57	84	70	79	56	86	69	77	64
	e-CSP	60	58	84	78	80	60	88	72	70	62
16	CSP	68	56	78	74	80	64	86	61	78	60
	e-CSP	60	64	82	74	81	58	92	73	75	62
20	CSP	64	53	72	72	82	68	86	63	76	68
	e-CSP	58	66	80	74	80	54	88	73	76	72
24	CSP	62	53	80	74	83	64	76	64	81	80
	e-CSP	58	62	80	74	81	66	90	73	76	74
28	CSP	58	51	78	78	78	72	82	66	81	76
	e-CSP	62	59	80	78	83	60	86	75	74	76
32	CSP	60	48	76	80	81	68	82	69	84	78
	e-CSP	58	63	74	78	80	64	86	70	75	76
36	CSP	68	49	72	66	84	72	84	69	81	78
	e-CSP	62	61	76	74	80	66	84	64	72	74
40	CSP	66	50	76	66	77	70	80	64	82	78
	e-CSP	58	58	82	78	78	68	86	64	76	76
44	CSP	68	53	66	62	82	64	80	66	81	76
	e-CSP	60	58	80	78	77	68	84	64	74	78
48	CSP	68	56	66	64	78	64	78	61	82	72
	e-CSP	56	57	80	78	78	68	78	66	75	74
52	CSP	66	51	62	66	71	56	80	61	81	70
	e-CSP	62	59	76	78	77	62	74	69	73	70
56	CSP	62	46	62	66	75	56	74	63	84	66
	e-CSP	62	59	74	76	77	64	78	66	74	70
60	CSP	54	51	68	68	72	54	78	66	81	68
	e-CSP	66	59	72	78	76	64	76	69	75	70
64	CSP	60	47	70	68	73	56	76	66	80	70
	e-CSP	66	54	76	76	76	60	76	73	74	72

feature extraction of EEG data. The good performance of the e-CSP can also be confirmed in the curve of classification accuracy of the left- and right-hand motor imagery tasks. In Fig. 4(g), it can be observed that under various projection vector numbers, the classification accuracy of the e-CSP is higher than that of the conventional CSP from a macro perspective. Moreover, the average accuracy of the e-CSP is significantly higher than that of CSP by performing a T-test ($P < 0.05$). In summary, by comparing the experimental results above, it is obvious that a substantial improvement of the CSP is obtained by the e-CSP in the majority of cases.

Table 8 Average F_1-score of the CSP and the e-CSP for each subject in three datasets

Method	Subject								
	aa	al	av	aw	ay	s1	s2	s3	S01
CSP	0.5	0.98	0.52	0.71	0.47	1	1	1	0.63
e-CSP	0.52	0.98	0.54	0.7	0.6	1	1	1	0.65
	S02	S03	S06	S14	S23	S35	S41	S43	S50
CSP	0.57	0.69	0.65	0.68	0.6	0.82	0.62	0.71	0.72
e-CSP	0.64	0.77	0.69	0.72	0.69	0.87	0.75	0.83	0.78

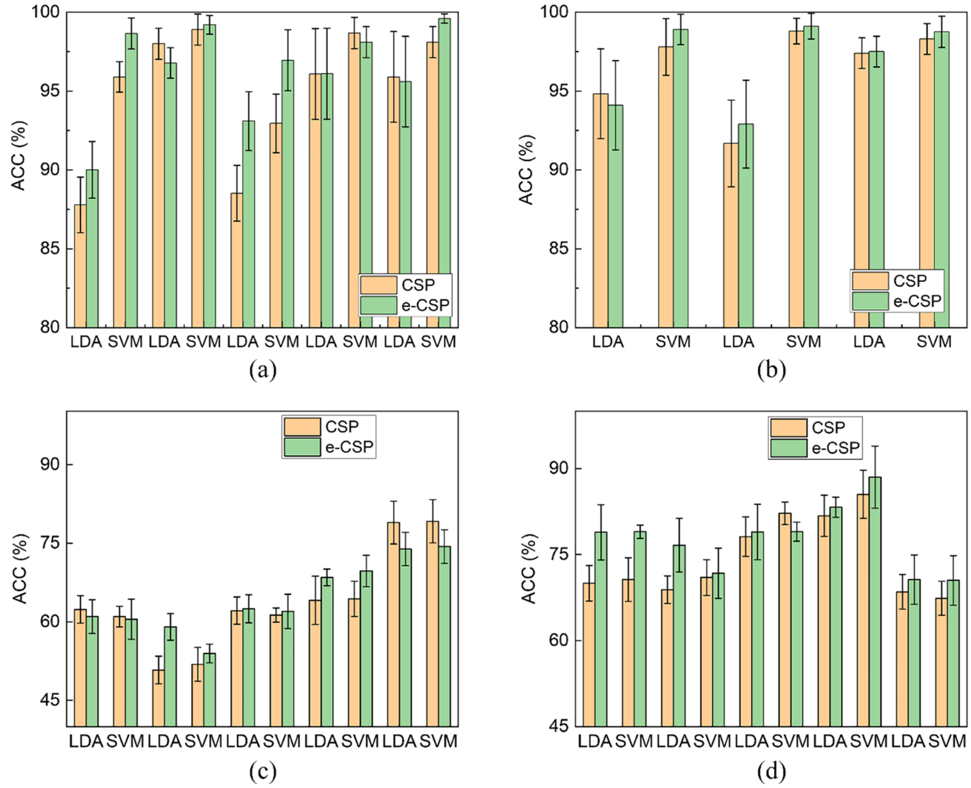


Fig. 5 Average classification accuracies versus the different methods under two classifiers on two datasets. **a** The BCI competition III dataset IVa; **b** the BCI competition III dataset IIIa; **c** the subjects S01, S02, S23, S41, and S43 in the Cho's dataset; and **d** the subjects S03, S06, S14, S35, and S50 in the Cho's dataset

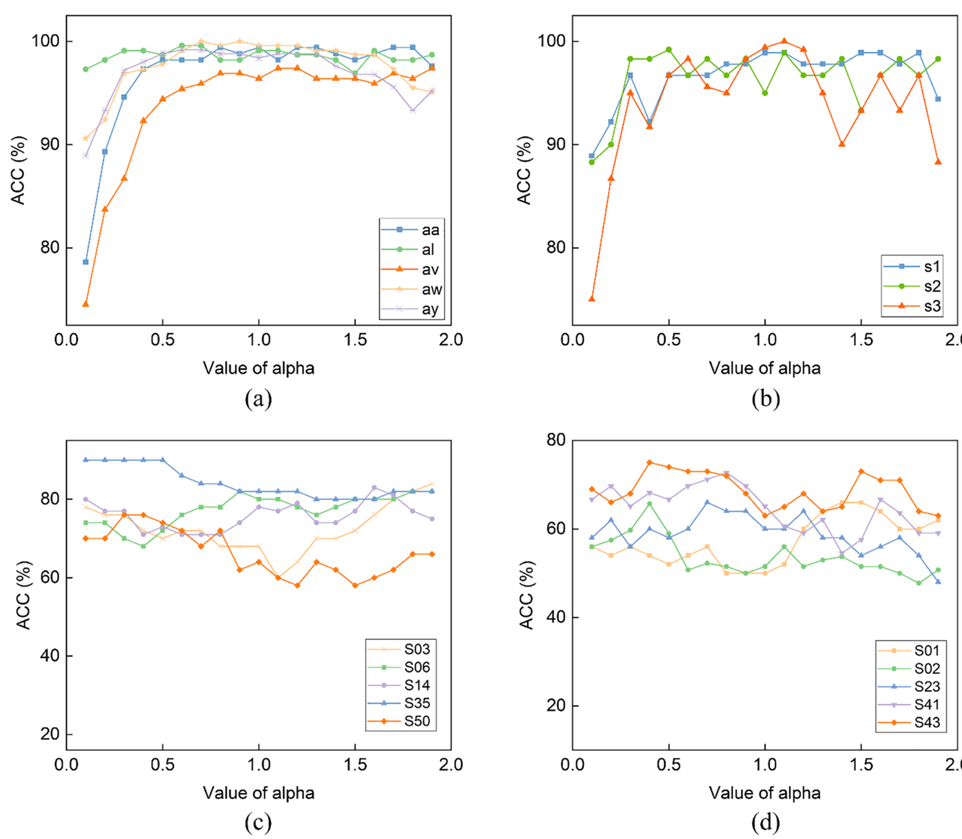


Fig. 6 Average classification accuracies versus the α value of different datasets. **a** The BCI competition III dataset IVa; **b** the BCI competition III dataset IIIa; **c** the subjects S03, S06, S14, S35, and S50 in Cho's dataset; and **d** the subjects S01, S02, S23, S41, and S43 in the Cho's dataset. The horizontal axis in the figure represents the value of the parameter α, and the vertical axis represents the classification accuracy

Table 9 Values of α for the best classification of the proposed method in the two datasets

Dataset	BCI competition III dataset IVa				BCI competition III dataset IIIa			
Subject	aa	al	av	aw	ay	s1	s2	s3
The best α	0.8	0.9	1.2	1.0	1.3	1.1	1.1	1.2

Table 10 Values of α for the best classification of the proposed method in the Cho's dataset

Dataset	Cho's dataset									
Subject	S01	S02	S03	S06	S14	S23	S35	S41	S43	S50
The best α	1.4	0.4	1.9	0.9	1.6	0.7	0.5	0.8	0.4	0.4

As shown in Table 7, some results similar to those on the two public datasets can also be observed in Cho's dataset. Although the classification accuracy of the e-CSP is not the same when the number of projection vectors is different, the same is that the classification accuracy of the e-CSP is higher than the CSP in the majority of cases. Furthermore, this behavior becomes more obvious when the number of projection vectors is small. Some other phenomena are also deserved to be discussed. When the number of projection vectors was set as 48 and 52, the classification accuracies of subjects "S01" and "S43" obtained by the e-CSP are not greater than that by the CSP. It may be due to the impact of the parameter α. Thus, we also inspect the fluctuation of classification accuracy of the e-CSP with the change of the parameter α in the following section. In summary, the performance of Cho's dataset has further verified the effectiveness of the e-CSP.

Based on the classification accuracy, we calculated the F_1-score to further compare the performance of the two methods. F_1-score is defined as the harmonic mean of precision and recall, whose range is [0, 1] [39]. The results are shown in Table 8. It is demonstrated that the F_1-score of the e-CSP is higher than that of the CSP. It is worth mentioning that better accuracy does not mean a higher F_1-score, but the higher F_1-score represents the robustness of the method.

In Fig. 5, the average classification accuracy of the conventional CSP and e-CSP algorithms across all subjects under both LDA and SVM classifiers are obtained. In the experiment, the linear kernel function was set for the SVM classifier, and the LDA classifier was followed the Fisher criterion. In the majority of cases, the classification accuracy of the e-CSP is higher than that of the CSP regardless of whether LDA or SVM is used. Statistical significance of accuracy was estimated by performing a T-test with a confidence level of 0.05, which verified the result mentioned above. It is evident that the results obtained by the SVM are in exceptionally good agreement with those obtained by the LDA. This outcome proves that our proposed algorithm has no classifier dependency, and the classifier has no effect on the essence of the results.

4.3 Performance comparison of e-CSP based on parameter α

The change in the average classification accuracies across all subjects of parameter α in the Euler representation of the two public datasets and Cho's dataset is reported in Fig. 6. Tables 9 and 10, respectively, list the value of α when the e-CSP obtains the best classification accuracy. The parameter α is set from 0.1 to 1.9 with a step size of 0.1 to observe the change in classification accuracy. Similar behavior is observed in the BCI competition III dataset IVa and BCI competition III dataset III a two publicly datasets; although the values of the best classification accuracy of each subject are not the same, they are concentrated in the range of 0.8–1.3. However, some different results between the two datasets can also be noticed. As shown in Fig. 6(a), the parameter α has a small influence on the performance of e-CSP, while it is seen that in another dataset (Fig. 6(b)), with the change in parameter α, the classification accuracy displays large fluctuations.

However, in Cho's dataset, it can be observed that as the parameters change, the performance of the e-CSP does not

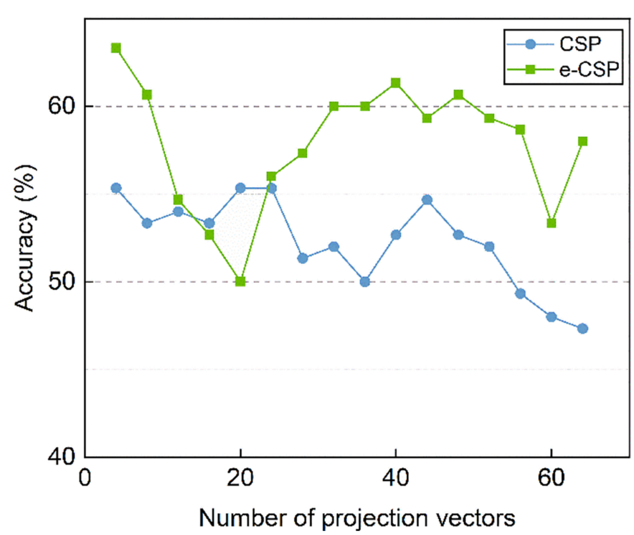


Fig. 7 Average classification accuracies of the proposed method and CSP in the lower limb MI data

show similar behavior as in the two public datasets. As shown in Fig. 6(c), for the subjects "S35" and "S50", the e-CSP can achieve relatively high accuracies when the values of the parameter α are concentrated in the range of 0.1–0.5; moreover, as the parameter α increases, the accuracy decreases. As shown in Fig. 6(d), the curve of the classification accuracy for the subject "S01" shows a rising trend with the increase of the parameter α. For the most of subjects, when the parameter α is in [0.3 0.8] interval, the e-CSP overall has good performance. Thus, we set the parameter α as 0.5 in experiment 1 and experiment 2 on Cho's dataset.

Moreover, we compared the classification accuracies of the CSP and the e-CSP in lower limb MI data. The dataset was collected by our laboratory, which was based on the MI task of the right foot and left foot. As shown in Fig. 7, the average classification accuracy of the e-CSP is greater than that of the CSP.

5 Conclusion

A novel feature extraction method is proposed for EEG classification, which is called the e-CSP. Different from other methods proposed to optimize the objective function of the CSP, e-CSP maps the sample values into the Euler space by explicit Euler representation and then performs the complex CSP in Euler space. Compared to the conventional CSP, e-CSP uses the cosine distance metric to measure the within-class and between-class scatters in the criterion function. The effectiveness of the e-CSP is illustrated by comparing with the CSP in the BCI competition III dataset IVa, the BCI competition III dataset IIIa and the Cho's dataset. From the experimental results, the e-CSP performs better than the CSP under both the LDA and SVM classifiers. Moreover, the e-CSP enhances EEG classification and outperforms some other recent related methods in the publicly available BCI competition datasets.

Acknowledgements This work was supported by the National Natural Science Foundation of China under Grant 62176054 and the University Synergy Innovation Program of Anhui Province under Grant GXXT-2020-015. The authors would like to thank the anonymous reviewers for their thoughtful comments and suggestions.

Funding National Natural Science Foundation of China, 62176054, Haixian Wang, the University Synergy Innovation Program of Anhui Province, GXXT-2020–015, Haixian Wang

References

1. Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM (2002) Brain-computer interfaces for communication and control. Clin Neurophysiol 113(6):767–791. https://doi.org/10.1016/S1388-2457(02)00057-3
2. Daly JJ, Wolpaw JR (2008) Brain-computer interfaces in neurological rehabilitation. The Lancet Neurol 7(11):1032–1043. https://doi.org/10.1016/s1474-4422(08)70223-0
3. Kathner I, Wriessnegger SC, Müller-Putz GR, Kübler A, Halder S (2014) Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface. Biol Psychol 102:118–129. https://doi.org/10.1016/j.biopsycho.2014.07.014
4. Reza MF, Ikoma K, Ito T, Ogawa T, Mano Y (2007) N200 latency and P300 amplitude in depressed mood post-traumatic brain injury patients. Neuropsychol Rehabil 17(6):723–734. https://doi.org/10.1080/09602010601082441
5. Zhang N, Tang J, Liu Y, Zhou Z (2016) A novel P300-BCI with virtual stimulus continuous motion. The 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), pp. 415–417. https://doi.org/10.1109/IHMSC.2016.222
6. Wang L, Liu X, Liang Z, Yang Z, Hu X (2019) Analysis and classification of hybrid BCI based on motor imagery and speech imagery. Measurement 147:106842. https://doi.org/10.1016/j.measurement.2019.07.070
7. Hernán D, Fernando M, Elizabeth F, Córdova F (2018) Intra and inter-hemispheric correlations of the order/chaos fluctuation in the brain activity during a motor imagination task. Procedia Computer Science 139:456–463. https://doi.org/10.1016/j.procs.2018.10.250
8. Florin P, Fazli S, Badower Y, Blankertz B, Müller K-R (2007) Single trial classification of motor imagination using 6 dry EEG electrodes. PLoS One 2(7):e637. https://doi.org/10.1371/journal.pone.0000637
9. Lee S-B, Kim H-J, Kim H, Jeong J-H, Lee S-W, Kim D-J (2019) Comparative analysis of features extracted from EEG spatial, spectral and temporal domains for binary and multiclass motor imagery classification. Inf Sci 502:190–200. https://doi.org/10.1016/j.ins.2019.06.008
10. Sözer AT, Fidan CB (2018) Novel spatial filter for SSVEP-based BCI: a generated reference filter approach. Comput Biol Med 96:98–105. https://doi.org/10.1016/j.compbiomed.2018.02.019
11. Pfurtscheller G, Lopes da Silva FH (1999) Event related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 110(11):1842–1857. https://doi.org/10.1016/s1388-2457(99)00141-8
12. Hatamikia S, Nasrabadi AM (2015) Subject transfer BCI based on composite local temporal correlation common spatial pattern. Comput Biol Med 64:1–11. https://doi.org/10.1016/j.compbiomed.2015.06.001
13. Blankertz B, Tomioka R, Lemm S, Kawanabe M, Müller K-R (2008) Optimizing spatial filters for robust EEG single-trial analysis. IEEE Signal Process Mag 25(1):41–56. https://doi.org/10.1109/msp.2008.4408441
14. Vuckovic A, Sepulveda F (2008) A four-class BCI based on motor imagination of the right and the left-hand wrist. International Symposium on Applied Sciences on Biomedical and Communication Technologies, pp. 1–4. https://doi.org/10.1109/ISABEL.2008.4712628
15. Hong K-S, Khan MJ (2017) Hybrid brain–computer interface techniques for improved classification accuracy and increased number of commands: a review. Front Neurorobotics 11:35. https://doi.org/10.3389/fnbot.2017.00035
16. Wang H, Zheng W (2008) Local temporal common spatial patterns for robust single-trial EEG classification. IEEE Trans Neural Syst Rehabil Eng 16(2):131–139. https://doi.org/10.1109/TNSRE.2007.914468
17. Dong J, Chen B, Lu N, Zheng N, Wang H (2017) Correntropy induced metric based common spatial pattern. The 27th International Workshop on Machine Learning for Signal Processing (MLSP), pp.1–6. https://doi.org/10.1109/MLSP.2017.8168132

18. Park Y, Chung W (2019) Frequency-optimized local region common spatial pattern approach for motor imagery classification. IEEE Trans Neural Syst Rehabil Eng 27(7):1378–1388. https://doi.org/10.1109/TNSRE.2019.2922713
19. Park Y, Chung W (2020) Optimal channel selection using correlation coefficient for CSP based EEG classification. IEEE Access 8:111514–111521. https://doi.org/10.1109/ACCESS.2020.3003056
20. Park Y, Chung W (2018) BCI classification using locally generated CSP features. The 6th International Conference on Brain–Computer Interface (BCI), pp.1–4. https://doi.org/10.1109/IWW-BCI.2018.8311492
21. Yu Z, Ma T, Fang N, Wang H, Li Z, Fan H (2020) Local temporal common spatial patterns modulated with phase locking value. Biomed Signal Process Control 59:101882. https://doi.org/10.1016/j.bspc.2020.101882
22. Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. Hum Brain Mapp 8(4):194–208. https://doi.org/10.1002/(SICI)1097-0193(1999)8:4%3c194::AID-HBM4%3e3.0.CO;2-C
23. Chakraborty B, Ghosal S, Ghosh L, Konar A, Nagar AK (2019) Phase-sensitive common spatial pattern for EEG classification. IEEE International Conference on Systems, Man and Cybernetics (SMC), pp. 3654–3659. https://doi.org/10.1109/SMC.2019.8914070
24. Geirnaert S, Francart T, Bertrand A (2021) Fast EEG-based decoding of the directional focus of auditory attention using common spatial patterns. IEEE Trans Biomed Eng 68(5):1557–1568. https://doi.org/10.1109/TBME.2020.3033446
25. Miao Y, Jin J, Daly I, Zuo C, Wang X, Cichocki A, Jung T (2021) Learning common time-frequency-spatial patterns for motor imagery classification. IEEE Trans Neural Syst Rehabil Eng 29:699–707. https://doi.org/10.1109/TNSRE.2021.3071140
26. Fu R, Han M, Tian Y, Shi P (2020) Improvement motor imagery EEG classification based on sparse common spatial pattern and regularized discriminant analysis. J Neurosci Methods 343:108833. https://doi.org/10.1016/j.jneumeth.2020.108833
27. Tan H, Zhang X, Guan N, Tao D, Huang X, Luo Z (2015) Two-dimensional Euler PCA for face recognition. International Conference on Multimedia Modeling, pp. 548–559. https://doi.org/10.1007/978-3-319-14442-9_59
28. Liwicki S, Tzimiropoulos G, Zafeiriou S, Pantic M (2013) Euler principal component analysis. Int J Comput Vis 101(3):498–518. https://doi.org/10.1007/s11263-012-0558-z
29. Liao S, Gao Q, Yang Z, Chen F, Nie F, Han J (2018) Discriminant analysis via joint Euler transform and L2,1-norm. IEEE Trans Image Process 27(11):5668–5682. https://doi.org/10.1109/TIP.2018.2859589
30. Liu Y, Gao Q, Han J, Wang S (2018) Euler sparse representation for image classification. The 32th AAAI Conference on Artificial Intelligence (AAAI-18), pp. 3691–3697. https://aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16524
31. Talukdar U, Hazarika SM, Gan JQ (2020) Adaptation of common spatial patterns based on mental fatigue for motor-imagery BCI. Biomed Signal Process Control 58:101829. https://doi.org/10.1016/j.bspc.2019.101829
32. Ghaheri H, Ahmadyfard AR (2013) Extracting common spatial patterns from EEG time segments for classifying motor imagery classes in a brain computer interface (BCI). Sci Iran 20(6):2061–2072
33. Koren Y, Carmel L (2004) Robust linear dimensionality reduction. IEEE Trans Vis Comput Graph 10(4):459–470. https://doi.org/10.1109/TVCG.2004.17
34. Tu W, Sun S (2012) A subject transfer framework for EEG classification. Neurocomputing 82:109–116. https://doi.org/10.1016/j.neucom.2011.10.024
35. Fitch AJ, Kadyrov A, Christmas WJ, Kittler J (2005) Fast robust correlation. IEEE Trans Image Process 14(8):1063–1073. https://doi.org/10.1109/tip.2005.849767
36. Cho H, Ahn M, Ahn S, Kwon M, Jun SC (2017) EEG datasets for motor imagery brain–computer interface. GigaScience 6(7):1–8. https://doi.org/10.1093/gigascience/gix034
37. Fang N (2018) Lp-norm-based local temporal common spatial patterns. Southeast University, Master's Thesis
38. Lotte F, Guan C (2011) Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. IEEE Trans Biomed Eng 58(2):355–362. https://doi.org/10.1109/TBME.2010.2082539
39. Chicco D, Jurman G (2020) The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21:6. https://doi.org/10.1186/s12864-019-6413-7

Jing Sun was born in Henan, China, in 1998. She received B.S. degree in electrical engineering from South-central Minzu University, Wuhan, Hubei, China, in 2019. She is currently working toward her M.S. degree at the Key Laboratory of Child Development and Learning Science of Ministry of Education, School of Biological Science & Medical Engineering, Southeast University, Nanjing, Jiangsu, China. Her research interest focuses on EEG signal processing.

Mengting Wei achieved her M.S. degree from University of Hertfordshire in 2012. Her research interests include brain-computer interfaces and pattern recognition.

Ning Luo was born in Sichuan, China. He received the B.S. and M.S. degrees of Science in Tsinghua University and the Ph.D. degree in Economics from University of International Business and Economics, Beijing, China. He was a visiting scholar at the University of Michigan, USA. He is currently working at the Institute of Software of the Chinese Academy of Sciences. He has long been engaged in research on big data analysis and pattern analysis.

Zhanli Li was born in August 1964. He received the Ph.D. degree in Xi'an Jiaotong University in 1997. He is currently a professor at the School of Computer Science and Technology at Xi'an University of Science and Technology. His research interests include computer vision and pattern recognition.

Haixian Wang received the B.S. and M.S. degrees in statistics and the Ph.D. degree in computer science from Anhui University, Hefei, Anhui, China, in 1999, 2002, and 2005, respectively. Currently, he is with the Key Laboratory of Child Development and Learning Science of Ministry of Education, School of Biological Science & Medical Engineering, Southeast University, Nanjing, Jiangsu, China. His research interests focus on biomedical signal processing, brain-computer interfaces, and machine learning.