

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <http://orca.cf.ac.uk/119181/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Ai, Qingsong, Chen, Anqi, Chen, Kun, Liu, Quan, Zhou, Tichao, Xin, Sijin and Ji, Ze 2019. Feature extraction of four-class motor imagery EEG signals based on functional brain network. *Journal of Neural Engineering* 16 (2) , 026032. 10.1088/1741-2552/ab0328 file

Publishers page: <http://dx.doi.org/10.1088/1741-2552/ab0328> <<http://dx.doi.org/10.1088/1741-2552/ab0328>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



ACCEPTED MANUSCRIPT

Feature extraction of four-class motor imagery EEG signals based on functional brain network

To cite this article before publication: Qingsong Ai *et al* 2019 *J. Neural Eng.* in press <https://doi.org/10.1088/1741-2552/ab0328>

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an ‘Accepted Manuscript’ watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is © 2018 IOP Publishing Ltd.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.

As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence <https://creativecommons.org/licenses/by-nc-nd/3.0>

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the [article online](#) for updates and enhancements.

Feature extraction of four-class motor imagery EEG signals based on functional brain network

Qingsong Ai¹, Anqi Chen¹, Kun Chen^{1,2,*}, Quan Liu¹, Tichao Zhou¹, Sijin Xin^{1,*} and Ze Ji²

¹ School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China

² Cardiff School of Engineering, Cardiff University, Cardiff CF24 3AA, UK

E-mail: kunchen@whut.edu.cn; sijinxin@hotmail.com

Received xxxxxx

Accepted for publication xxxxxx

Published xxxxxx

Abstract

Objective. A motor-imagery-based brain-computer interface (MI-BCI) provides an alternative way for people to interface with the outside world. However, the classification accuracy of MI signals remains challenging, especially with an increased number of classes and the presence of high variations with data from multiple individual people. This work investigates electroencephalogram (EEG) signal processing techniques, aiming to enhance the classification performance of multiple MI tasks in terms of tackling the challenges caused by the vast variety of subjects.

Approach. This work introduces a novel method to extract discriminative features by combining the features of functional brain networks with two other feature extraction algorithms: common spatial pattern (CSP) and local characteristic-scale decomposition (LCD). After functional brain networks are established from the MI EEG signals of the subjects, the measures of degree in the binary networks are extracted as additional features and fused with features in the frequency and spatial domains extracted by the CSP and LCD algorithms. A real-time BCI robot control system is designed and implemented with the proposed method. Subjects can control the movement of the robot through four classes of MI tasks. Both the BCI competition IV dataset 2a and real-time data acquired in our designed system are used to validate the performance of the proposed method.

Main results. As for the offline data experiment results, the average classification accuracy of the proposed method reaches 79.7%, outperforming the majority of popular algorithms. Experimental results with real-time data also prove the proposed method to be highly promising in its real-time performance.

Significance. The experimental results show that our proposed method is robust in extracting discriminative brain activity features when performing different MI tasks, hence improving the classification accuracy in four-class MI tasks. The high classification accuracy and low computational demand show a considerable practicality for real-time rehabilitation systems.

Keywords: brain-computer interface, motor imagery, functional brain network, common spatial pattern, local characteristic-scale decomposition

1. Introduction

The idea of using brain signals to control a robot or prosthetic device without the involvement of the peripheral nerves and muscles began in 1929 when Berger discovered electroencephalogram (EEG) signals [1]. A brain-computer interface (BCI) provides an alternative method for natural communication between human brains and the outside world directly without relying on human nerves and muscle tissues [2]. The EEG-based system is one of the most widely used

techniques in BCI systems owing to its advantages of simple and high time resolution [3]. Among many brain-computer interaction control paradigms, motor imagery (MI)-based BCI is a very important brain-computer interaction strategy that realizes the control and exchange of information between the brain and the outside world by interpreting mental activities through recognizing EEG signals of different MI tasks [4].

Noninvasive BCI technology has matured and involved many areas, and the range of BCI applications has also been substantially enlarged. Research by the GRAZ-BCI team

focused on the pattern classification of the MI of different limb parts such as the left/right hands, feet, and tongue motor imagery [5]. At present, the MI BCI system they built is integrated to control wheelchairs, neural prostheses, and other devices in virtual and real environments [6-8]. Shi T. et al. realized a BCI system for unmanned aerial vehicle indoor navigation based on MI [9], and Müller-Putz G. et al. presented a hybrid BCI framework that was used in studies with nonimpaired as well as end users with motor impairments [10].

The effective extraction of discriminative features for identification of MI tasks from complex MI EEG signals is critical to the performance of BCI systems. However, acquired MI EEG signals are usually contaminated by strong artifacts and are highly nonstationary and nonlinear, posing a great challenge to the feature extraction from MI EEG. Researchers have proposed some classic feature extraction methods for MI EEG signals, including the adaptive auto regressive (AAR) model [11], wavelet transform (WT) [12], empirical mode decomposition (EMD) [13], and common spatial pattern (CSP) [14].

At present, CSP method and its inheritance are known to be the most effective feature extraction methods in MI EEG analyses [15]. CSP method can extract the spatial information from EEG signals and make a remarkable effect in two-class EEG signals classification. However, there are also deficiencies in CSP. First, it needs a multichannel signal to improve the classification effect. Second, it ignores the frequency domain characteristics of EEG signals. However, the frequency domain information is particularly important for MI task classification. Since MI EEG signals are nonlinear and nonstationary, the time domain analysis method does not reflect the frequency information, and the frequency domain analysis method includes the frequency information, but the time when the frequency information changes is unknown.

Therefore, combining the time domain method and the frequency domain method for analysis can more fully describe the characteristics of EEG signals. In 2012, based on the definition of the intrinsic scale component (ISC), Zhang Heng J. et al. proposed a new nonstationary signal analysis method named local characteristic-scale decomposition (LCD) [16]. LCD is known to be superior to the EMD algorithm in its endpoint effect, decomposition time, and iteration times, and is considered suitable for online analyses of EEG signals.

When users perform limb MI tasks, the corresponding motor sensation cortex of the brain is activated, and specific physiological phenomena such as event-related desynchronization (ERD) and event-related synchronization (ERS) will be generated simultaneously [17,18]. The aforementioned feature extraction methods are based on the ERD and ERS phenomena. However, about 15% to 30% of

users have the problem of “BCI illiteracy.” These users fail to produce signals with discriminative characteristics such as ERD/ERS; hence, the relevant rhythm signals cannot be measured [19,20]. In addition, owing to the individual differences in subjects, brain regions and the evoked characteristic signals activated by different subjects are not the same. These problems lead to rigorous screening of the subjects and a large amount of pretraining in BCI system experiments.

The brain can be considered a dynamic network that constantly organizes and reshapes its functional connections. EEG signals are recorded as time-series signals of brain activity, and studies show that such time-series signals captured at different brain regions reflect the brain activity synergy of their corresponding brain regions. Such EEG time-series signals acquired from multiple locations of the brain form a brain network [21,22], and cognitive activities can be analyzed by extracting different measures in the brain network in order to reflect differences between brain regions activated by different users. Finally, the classification accuracy can be improved.

Considering the above challenges, we combine the CSP and LCD algorithms to extract multiscale features of MI EEG signals. A functional brain network is constructed to characterize the interaction between each pair of electrode leads in order to extract measures as additional features of the BCI system. Finally, the above three types of features, namely, CSP, LCD, and brain networks, are fused for classifying MI tasks. The proposed method quantifies brain information with multidimensional and multiscaled features, aiming to minimize the effect caused by individual differences and to be effective and feasible for real-world BCI applications.

2. Methods

2.1 Feature extraction

2.1.1 Local characteristic-scale decomposition. LCD is a signal decomposition method that decomposes any complex signal $x(t)(t > 0)$ into the sum of n ISC component $c_i(t)(i = 1, 2, \dots, n)$ and a residue $u_n(t)$ (see equation 1). The ISC component must satisfy two conditions: 1) its local waveform is approximately a sine wave, and 2) the ISC of a single mode will not generate a negative frequency. The signal decomposition expression is as follows:

$$x(t) = \sum_{i=1}^n c_i(t) + u_n(t) \quad (1)$$

First, the baseline of the original signal is calculated using the cubic spline function, and then the baseline is subtracted from the original signal. If the residual signal satisfies two

conditions of the ISC component, the signal is an ISC component, otherwise, the signal is taken as the original signal and the above process is repeated. Meanwhile, after each ISC component is obtained, the standard deviation (SD) is calculated according to formula (2), and the iteration is terminated if the SD is less than 0.05[23].

$$SD = \sum_{t=0}^T \left[\frac{|h_{ip}(t) - h_{i(p-1)}(t)|^2}{h_{i(p-1)}^2(t)} \right] \quad (2)$$

where $h_{ip}(t)$ represents the i -th ISC component obtained after looping p times. Owing to the computation demand of the entire LCD calculation, only three channels (C3, C4, and Cz) that contribute most to the classification are selected for LCD decomposition [24]. Then, by conducting a lot of experiments and considering the computation time, the signals of these three channels are decomposed by the LCD and only the first three ISC components are taken, so 9 ISC components can be obtained from the C3, C4, and Cz channels in one experiment.

Perform Hilbert transform on each ISC component $c_i(t) (i = 1, 2, \dots, K)$, formulated as

$$y_i(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{c_i(\tau)}{t - \tau} d\tau \quad (3)$$

where K is the number of ISC components in each experiment, and K is set to 9 in our experiment. The parsing signal $z_i(t)$ is then constructed as follows:

$$z_i(t) = c_i(t) + jy_i(t) = a_i(t)e^{jf_i(t)} \quad (4)$$

where $a_i(t)$ and $f_i(t)$ represent the instantaneous amplitude and frequency of the i -th ISC component, respectively. Then, the instantaneous frequency is sorted, and part of the values are selected from the sorted instantaneous frequency $f'(t)$ at medium intervals as the frequency features $F_1 = [f_{11}, f_{12}, \dots, f_{1K}] \in R^{1 \times KP}$ of the MI EEG signals, where P is the number of features selected from the ISC component, and P is set to 20 through experimental experience. Each of the eigenvalues f_{11}, f_{12}, \dots in F_1 is a $1 \times P$ - dimensional vector.

2.1.2 Common spatial pattern. The CSP algorithm uses the theory of matrix simultaneous diagonalization in algebra to find a set of spatial filters in order to maximize the variance of one class of signals while minimizing the variance of the other class of signals.

Denote the original EEG signal of a trial as $E_{N \times T}$, where N is the number of electrode leads, and T is the number of data samples. Here, we take an example of a two-class experiment, where data are collected from two types of tasks named the left-hand and the right-hand MI tasks. The subjects are instructed to imagine the movement of their left hands or right hands, but without actual muscle activations in their hands.

The CSP feature is calculated with the following steps:

- 1) Calculate the covariance C of the two-class MI signals in each experiment, formulated as

$$C = \frac{EE^T}{\text{trace}(EE^T)} \quad (5)$$

where $\text{trace}(X)$ is the trace of matrix X , which is the sum of the diagonal elements of matrix X . The average covariance of all experiments is then calculated by summing up covariance matrices, in this case, the left-hand and right-hand MI data:

$$C_l = \sum_{i=1}^n C_{l,i} \quad C_r = \sum_{i=1}^n C_{r,i} \quad (6)$$

Then, the sum of the two types of covariance matrices is obtained:

$$C_c = C_l + C_r \quad (7)$$

- 2) Perform an eigenvalue decomposition of the mixed spatial covariance, formulated below:

$$C_c = U_c A_c U_c^T \quad (8)$$

where A_c is the eigenvalue diagonal matrix, and U_c is the corresponding eigenvector matrix.

- 3) Construct the whitening transformation matrix first. Then, using the features of S_l, S_r with the same feature vector, decompose its eigenvalue:

$$P = A_c^{-1/2} U_c^T$$

$$S_l = PC_l P^T \quad S_r = PC_r P^T \quad (9)$$

$$S_l = BA_l B^T \quad S_r = BA_r B^T \quad (10)$$

The desired space filter $W = (B^T P)^T$ is obtained, and the filter is used to obtain $Z_{N \times T} = W_{N \times N} E_{N \times T}$.

4) Find the eigenvector f .

The dimension of f can be adjusted according to the quality of the EEG signals and the classifier requirements, but should not exceed the number of electrode leads N . Extract the first m rows and the last m rows of $Z (2m < N)$, take the p -th row of Z as Z_p . Then, f_p is calculated as

$$f_p = \log \left(\frac{\text{var}(Z_p)}{\sum_{n=1}^p \text{var}(Z_n)} \right) \quad p = 1 : 2m \quad (11)$$

where $\text{var}(X)$ represents the variance of the time-series signals.

Since there are four classes of MI tasks that need to be classified, it is necessary to expand the CSP to meet the technical requirements. There are two common methods to expand: one to one and one to the other. The one-to-one method is adopted to expand the CSP in this paper. For each experiment, six projection matrices are generated, and each projection matrix is concatenated to form a complete spatial feature vector F_2 .

$$F_2 = [f_{21}, f_{22}, f_{23}, f_{24}, f_{25}, f_{26}] \in R^{1 \times 6N_1} \quad (12)$$

where N_1 represents the number of channels in each experiment. The signal of one experiment here includes 22 channel data and 9 ISC component data, forming a total N_1 of 31. Each of the eigenvalues f_{21}, f_{22}, \dots in F_2 is a $1 \times N_1$ -dimensional vector.

2.1.3 Brain network. Researchers have found that biological networks generally have properties of small-world networks. This means the network has a large clustering coefficient and a short characteristic path length [25,26].

Considering the independence of EEG nodes and the synergistic effect between nodes, we can use the complex network theory to construct an EEG functional brain network. In a functional network based on EEG, each node corresponds to the brain regions detected by different leads, and the collected EEG signals constitute the time series of this node.

The definition of edges in functional networks is based on the functional connections, and the weights of the edges can be determined with various methods, which can be broadly categorized into two main branches: linear and nonlinear. The linear methods include the Pearson correlation, partial correlation, and partial coherence. The nonlinear methods include the synchronization likelihood, canonical correlation analysis, and mutual information.

Through the analysis of various connection methods, we choose the canonical correlation analysis (CCA) to calculate the nonlinear correlation between each pair of leads. This method can analyze the signal in the entire EEG frequency band as well as in a specific range of frequency spectra. This is ideal for analyzing instantaneous and unstable signals such as EEGs. CCA considers the linear combination of the two sets of variables and studies the correlation coefficient $\rho(u, v)$ between them. In all linear combinations, we find a linear combination with the largest correlation coefficient and use this maximum correlation coefficient to represent the correlation of the pair of variables. The correlation coefficient is expressed as

$$\rho_{uv} = \frac{\text{Cov}(U, V)}{\sqrt{\text{Var}(U)} \sqrt{\text{Var}(V)}} \quad (13)$$

Mathematically, the functional network obtained is a correlation matrix in which each element represents a correlation between two brain regions. After obtaining the correlation matrix, the next step is to binarize it by setting a threshold C_{thr} . If the value of an element in the matrix is greater than this threshold, it is considered that there is a functional connection between the two brain regions, and the value here is set to 1. Otherwise the value is set to 0, thereby establishing a complete binarized function network. The process of constructing functional brain networks based on EEG signals is shown in Figure 1.

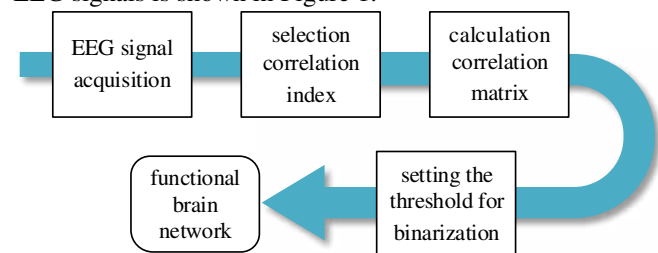


Figure 1. Diagram of building a functional network based on EEG signals.

In order to describe the topological structure of the established brain network, some common measures are provided, and the changes in the topological properties of the brain networks can be studied by analyzing these measures. We select five measures: degree, clustering coefficient, average shortest path length, local efficiency, and betweenness centrality. These complex network measures are described in detail below.

Degree

$$k_i^B = \sum_{j \in G} a_{ij} \text{ or } k_i^W = \sum_{j \in G} w_{ij} \quad (14)$$

where a_{ij} or w_{ij} is the corresponding element in the binary or weighted network matrix. The degree of a node is the number of edges of one node. A node with a higher degree is considered more important in the network.

Clustering coefficient

The clustering coefficients of the nodes are defined as follows [27]:

$$c_i^B = \frac{2R}{k_i^B(k_i^B - 1)} \quad \text{or} \quad (15)$$

$$c_i^W = \frac{2}{k_i^B(k_i^B - 1)} \sum_{j,k} (w_{ij}w_{jk}w_{ki})^{1/3}$$

where R represents the number of directly connected neighbors of node i . The clustering coefficient of a node is defined as the ratio between the actual number of edges existing between the neighbor nodes of the node and the maximum possible number of connected edges. The clustering coefficient reflects the local connectivity and measures the cluster characteristics and closeness within the functional brain network.

Average shortest path length

The shortest path length is the smallest number of edges between two nodes. In other words, it is the minimum number of steps to travel through the network from node i to j . The average shortest path length is defined as the mean number of steps along the shortest paths between all possible pairs of network nodes. The definition is as follows:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in V, i \neq j} d_{ij} \quad (16)$$

where N represents the total number of network nodes. (The number of network nodes is the same as the number of electrode leads in our experiment), d_{ij} represents the distance between nodes i and j in the network.

Local efficiency

For a network G with N nodes, the global efficiency is calculated as shown in equation 17 [28]:

$$E_{glob}(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}} \quad (17)$$

The formula for calculating the local efficiency is as follows:

$$E_{loc}(G) = \frac{1}{N} \sum_{i \in G} E_{glob}(G_i) \quad (18)$$

where $E_{glob}(G_i)$ is the global efficiency of G_i , and G_i is a subgraph composed of the neighbors of node i . The global efficiency and local efficiency measure the information transmission ability of the network globally and locally.

Betweenness centrality

The betweenness centrality is defined as the number of shortest paths going through a node or edge [29]. The higher the betweenness centrality of a node, the greater the flow of information carried by the node, and the more significant the impact on the function of the functional brain network. The betweenness centrality is formulated in equation 19:

$$C_B(i) = \frac{2 \sum_{h \neq i, j \neq i, h < j \in V} g_{hj}(i)}{(N-1)(N-2)g_{hj}} \quad (19)$$

where g_{hj} is the number of all shortest paths between node $h \in V$ and $j \in V$, and V is the set of all nodes in the network.

After quantifying the relationship between nodes, it is necessary to select an appropriate threshold to binarize the adjacency matrix. Two principles need to be followed to establish the network: To ensure the integrity of the network, it should not contain any isolated node or isolated part; and the small-world characteristics of the network should be ensured [25,26]. According to the random model of Erdos-Renyi [30], if a graph with N nodes is to be fully connected, the

connection sparsity should be greater than $2 \ln N / N$. In addition, it should be ensured that its small-world attribute value σ is much greater than 1. Through a large number of experiments, the threshold was empirically set to 0.84, and the corresponding sparsity of the brain network was 0.35.

The offline experimental results of chapter 4 below show that the classification effect of the measure of degree in the binary network is better than for other measures. Thus, the measure of degree is used as the feature of the brain network for online experiments.

2.1.4 Feature fusion. After establishing a functional brain network based on MI EEG signals, the measures described in the previous section are extracted as the features of brain networks and are then fused with the multiscale features extracted from the CSP and LCD algorithms. There are two feature fusion strategies: parallel feature fusion and serial feature fusion. Compared to parallel feature fusion, one advantage of serial feature fusion is its simplicity in that it requires two steps: normalization and concatenation of multiple features. This effectively retains the discriminative information of various features for classification. For the above reason, the serial feature fusion strategy is adopted in this work.

With the features in the above spatial and frequency domains and the features of the brain network fused, the obtained feature vector of the EEG signals, denoted by $F \in R^{1 \times (KP+6N_1+MN)}$, is defined below:

$$F = [F_{11}, F_{22}, F_{33}] \quad (20)$$

$$F_{11} = \left[\frac{f_{11}}{\|f_{11}\|}, \frac{f_{12}}{\|f_{12}\|}, \dots, \frac{f_{1K}}{\|f_{1K}\|} \right] \in R^{1 \times KP}$$

$$F_{22} = \left[\frac{f_{21}}{\|f_{21}\|}, \frac{f_{22}}{\|f_{22}\|}, \dots, \frac{f_{26}}{\|f_{26}\|} \right] \in R^{1 \times 6N_1}$$

$$F_{33} = \left[\frac{f_{31}}{\|f_{31}\|}, \frac{f_{32}}{\|f_{32}\|}, \dots, \frac{f_{3M}}{\|f_{3M}\|} \right] \in R^{1 \times MN}$$

The three channels C3, C4, and Cz that contribute most to the classification are selected to perform LCD decomposition [24], which produces nine ISC components and then extracts the frequency domain feature F_1 . Next, the obtained nine ISC components are added to the 22 channels of original EEG signals, and the CSP algorithm is used to extract the spatial features from the 31-channel data as F_2 . The feature vector F_3 is extracted from the functional brain network, where $\|\cdot\|$ is the l_2 -norm in equation 20. A flowchart of the feature extraction algorithm is shown in Figure 2.

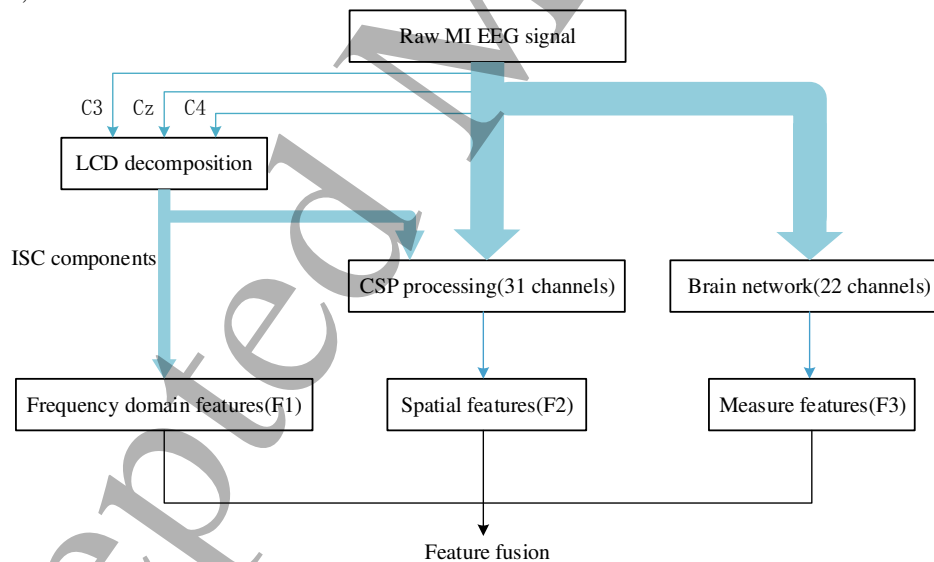


Figure 2. Flow of proposed feature extraction algorithm.

2.2. Feature selection and classification

The multicluster feature selection (MCFS) algorithm was applied to sort features [31]. The basic principle of MCFS is first to construct a p -nearest neighbor graph according to (21):

$$W_{ij} = \begin{cases} 1 & \text{if } x_i \in N(x_j) \text{ or } x_j \in N(x_i) \\ 0 & \text{others} \end{cases} \quad (21)$$

where x_i or x_j corresponds to the extracted feature, and $N(x_i)$ represents the nearest neighbor of x_i . Define a diagonal matrix D , $D_{ii} = \sum_j W_{ij}$. We can compute the graph Laplacian $L = D - W$ and solve the generalized eigenvalue problem in equation 22 to obtain the feature vector y_i corresponding to the minimum eigenvalue.

$$Ly = \lambda Dy \quad (22)$$

Then, the sparse coefficient vector a_i is obtained by solving the normalized regression problem, as shown in equation 23:

$$\min_{a_i} \|y_i - X^T a_i\|^2 \quad (23)$$

where X is the input data matrix, and a_i is the M -dimensional vector that contains the combination coefficient for different features. For every feature j , we define the MCFS score for the feature as

$$MCFS(j) = \max_i |a_{i,j}| \quad (24)$$

where $a_{i,j}$ is the j -th element of vector a_i . We then sort all features according to their MCFS scores and select a number of features.

This work employs the spectral regression discriminant analysis (SRDA) classifier [32]. The SRDA algorithm combines spectral analysis and linear regression. It effectively solves the feature decomposition problem in the LDA algorithm, and saves a considerable amount of classification time and storage space. When extended to multiclassification problems, the SRDA algorithm first uses the regression model to reduce the dimension, and then by using spectrum analysis, feature data can be classified by simply solving a series of regular least squares problems. [33].

3. Experimental setup

3.1 BCI competition IV dataset 2a

The performance of our proposed method is evaluated using the BCI competition IV dataset 2a [34], which is widely used and publicly available. This dataset provided by Graz University was recorded from nine healthy subjects. According to the international 10/20 system [35], 22 Ag/AgCl electrodes were placed. The subject was required to perform the following four classes of MI tasks in each trial: left hand, right hand, both feet, and tongue. This means the subject imagined the movement of his limb without actual muscle activation. During the experiment, the subjects sat in front of

the computer and performed corresponding actions according to the screen prompts. A detailed dataset description can be found in [34].

The EEG signal was recorded simultaneously at a sampling frequency of 250 Hz and processed by a band-pass filter with 0.5 Hz–100 Hz to remove interference from other frequency bands. An embedded notch filter of 50 Hz eliminated power line noises. Each subject's dataset consisted of a training set and a test set, and each set contained 288 experiments.

In order to remove the artifacts and enhance the signal-to-noise ratio of the signals, the EEG signals need to be effectively preprocessed before extracting the features. First, according to the characteristics of ERD and ERS, the data are band-pass filtered between 8 Hz and 30 Hz [36] by a five-order Butterworth band-pass filter. Then, various artifacts in the EEG signals are removed effectively by wavelets. The 'sym4' wavelet is selected to decompose the signal, and then the threshold function is used to set a critical threshold. If the wavelet coefficient is less than the threshold, it is considered that the coefficient is mainly caused by noise, and the coefficient is removed. If the wavelet coefficient is larger than the threshold, the coefficient is considered to be mainly caused by the signal, and the coefficient is retained. Finally, inverse-transform is performed on the retained wavelet coefficients to reconstruct the denoised signal.

3.2. Self-designed BCI system

3.2.1 Subjects and experimental setup. The EEG data were recorded using a UE-16B EEG amplifier at a sampling rate of 1000 Hz. All 16 channels were selected (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, and T6). Since the frequency of the MI EEG signal is concentrated at 8–30 Hz, and the cutoff frequency of the low-pass filter was set to 100 Hz (with a 50-Hz notch filter enabled). The left and right ear electrodes, A1 and A2, were used as reference electrodes, and the forehead was used as ground electrode G. The controlled object in the online BCI system was a humanoid robot, the NAO robot, produced by Aldebaran Robotics.

The subjects were eight graduate students aged from 23 to 26, including two females and six males. All of them were right-handed and had no neurological history. To minimize environmental effects such as light stimulations, the experiments were carried out in a quiet environment with dimmed lighting. The subjects were seated in comfortable backrest chairs to reduce muscle strain that could interfere with the experimental results. Before the experiments, all subjects were instructed and trained about the experimental procedure. The study was conducted with approval from the Wuhan University of Technology.

3.2.2 Data acquisition. The MI EEG data acquisition included training data collection and real-time data collection. The training data collection session consisted of four runs with 2-min breaks between two consecutive runs. Each run had 25 MI trials. The MI tasks to be performed in the four runs were left hand, right hand, both foot, and tongue movements. Each trial consisted of a 2-s ready period, a 4-s motor imagery period, and a 2-s break period. When the prompt picture was displayed on the screen, the subject needed to perform the corresponding MI task until the prompt image disappeared. The subject then rested for 2 s and waited for the next experiment to begin. The experimental process for training EEG data acquisition is shown in Figure 3.

We segmented the data in this duration into several epochs to conduct a series of experiments statistically, and found that the data between 2.5 s and 3.5 s achieved the best performance. Therefore, this 1-s data epoch was selected for feature extraction and classification.

The real-time data acquisition session of the robot control continuously collected data, and the data of the 1-s time period was collected and sent to the signal processing module, where the data of the next second was collected at the same time.

3.2.3 System framework and experiment design. This subsection introduces the BCI system framework designed for evaluating the proposed algorithm. The system includes four functional components: signal acquisition, signal processing, human-computer interaction, and robot control. The signal acquisition module is responsible for the data acquisition,

filtering, and amplification of EEG signals, which are then sent to the human-computer interaction module in real time. Next, the human-computer interaction module stores the received EEG data, and the signal processing module starts to process the data that are eventually converted into a set of control commands. The commands are sent to the robot control module through socket communication. The NAO robot interprets the commands and starts executing the corresponding movement behaviors according to the received commands. A block diagram of the overall system is shown in Figure 4.

The mapping between the MI tasks and the corresponding movement control commands of the robot are listed in Table 1 together with their labels used for classification. As can be seen, the robot has four basic behaviors moving in four directions: forward, left, right, and backward. Correspondingly, the subjects are asked to execute the mental activities of moving four body parts, both feet, left hand, right hand, and tongue.

Table 1. Mapping between motor imagery tasks and movement control commands of robot.

Classification Labels	Motor Imagery	Robot Movement
0	Both feet	Forward
1	Left hand	Left
2	Right hand	Right
3	Tongue	Backward

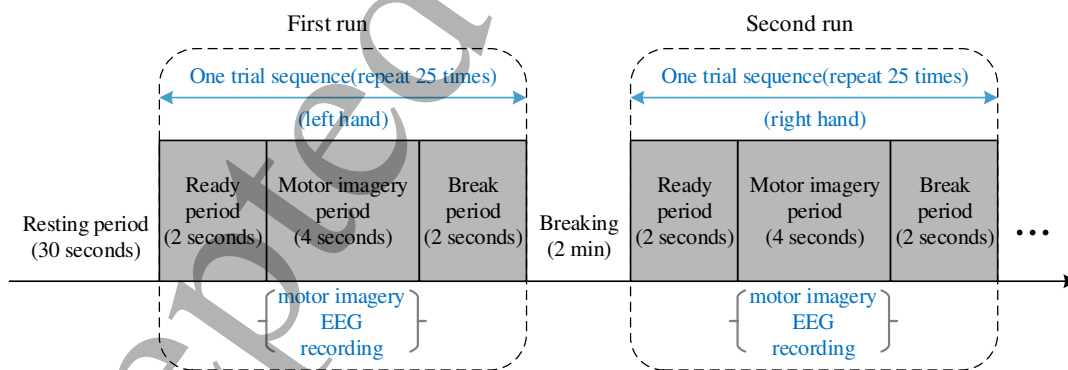


Figure 3. Training EEG data acquisition experiment design of real-time brain-computer interface system.

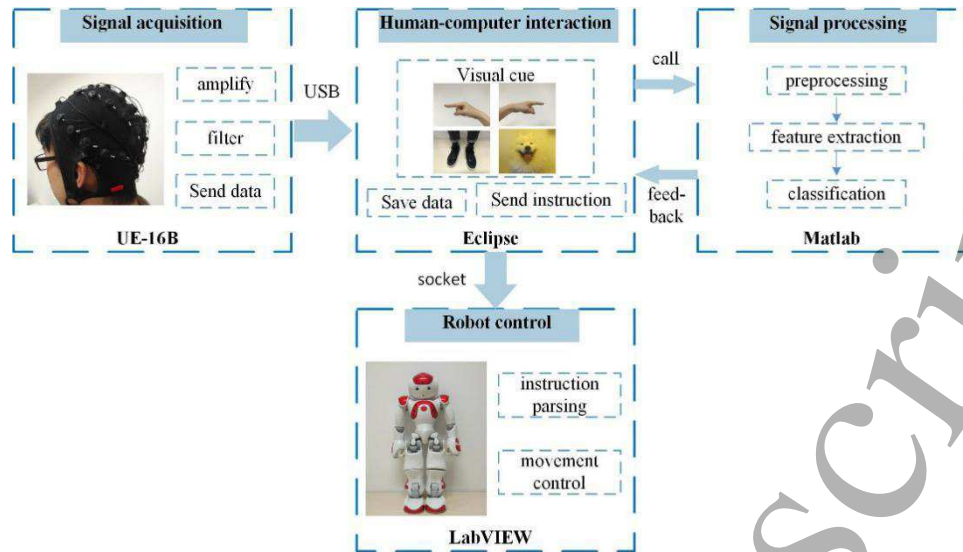


Figure 4. Structure of real-time brain-computer interface system.

In terms of interacting with the robot, two control strategies were adopted for the motion control of the NAO robot: synchronous control and asynchronous control.

1) Synchronous control

The principle of the NAO robot synchronously controlled by the MI signals is as follows. The robot needs to perform a predefined sequence of motion behaviors. A motion behavior corresponds to a control instruction. In this work, we test with a total of 17 instructions. For example, [2 2 0 0 2 2 0 0 1 0 0 1 1 3 3 3 1] is the sequence of labels (as seen in Table 1) of the control commands for testing in this work. The subject then performs the corresponding MI tasks with the given instructions displayed in order. During the experiment, the robot will carry out the corresponding motion only if the classification result of the MI signal matches the preset command. The experiment is completed when the robot completes all instructions.

2) Asynchronous control

Asynchronous control differs in that it does not require a pre-known sequence of control commands. The NAO robot starts performing behaviors purely based on the classification results. The task is to control the robot to move toward a predefined goal position in the room (Figure 9). The decisions of robot motions are not constrained and are only based on the classification results of the subject's mental activities. However, false classifications are unavoidable. The expected motions from the subject's mental activities cannot always be correctly recognized, resulting in incorrect motion behaviors executed by the NAO robot.

To avoid or minimize the number of false alarms, an error control mechanism is deployed in the experiment of the asynchronous control robot. It is assumed that the subject's mental activities will generally remain temporally consistent, meaning the subject will not change his or her task very quickly. Therefore, the false alarm detector used in the work compares the classification results of two consecutive detections by comparing the current classification result with the previous one. If the two classification results are the same, it is determined that the movement is correct, and the corresponding control instruction is sent to the robot, the sent instruction is also compared with the next classification result. However, if the two are inconsistent, a false alarm will be triggered, and the control command is not sent until the sending condition is met again. This control mechanism greatly increases the classification success rate of MI tasks that the subject needs to perform, and it effectively reduces the chances of performing incorrect movements by the NAO robot.

4. Results and discussion

4.1 Offline data analysis

To quantify the validity of the proposed method, an offline study is carried out using BCI competition IV dataset 2a. We first study the effectiveness of the brain network features in order to choose the optimal feature sets. According to the forenamed complex network theory, a brain network can be used as a weighted network or it can be transformed into a binary network. Briefly, the functional brain network of each subject is first established by CCA, which is a weighted network, and is converted into a binary network according to the set threshold. Then, five brain-network measures are

1
2
3 extracted from the weighted network and the binary network:
4 node degree, betweenness centrality, clustering coefficient,
5 average shortest path length, and local efficiency. The
6 connections between functional brain networks can be
7 analyzed by graph theory.
8

9 Figure 5 shows the average adjacency matrices of weighted
10 brain networks of the nine subjects under the four classes of
11 MI tasks. The dimension of the matrices is 22×22 , and the
12 horizontal and vertical axes represent the signal channels. The
13 elements in the matrices represent the correlation coefficients
14 between all pairs of lead signals in the entire brain region. The
15 correlation coefficients are normalized and range from 0 to 1.
16 A coefficient closer to 1 indicates a higher degree of
17 correlation between the two corresponding leads.
18

19 Next, the weighted brain network is binarized. If the
20 correlation coefficient is greater than the set threshold, it is
21 considered that a connection edge is established between the
22 two leads and the corresponding element in the adjacency
23 matrix is set to 1. For the opposite situation, the element is set
24 to 0. There is no edge in the general node set that is directly
25 connected to itself without passing through other nodes, so the
26 diagonal element in the adjacency matrix is 0. This is the
27 average adjacency matrix of the binary brain network of the
28 nine subjects shown in Figure 6.
29

30 As presented in Figures 5 and 6, the brain functional
31 connectivity of all four kinds of MI signals is overall quite
32 large and evenly distributed. Among them, the connection of
33 the right-hand MI signal is stronger than that of the other three
34 kinds of MI signals, and the connection strength of the tongue
35 MI signal is the weakest. This is clearly shown in Figure 6. In
36 terms of the electrode locations from Figure 5, the connection
37 of the right-hand MI signal is significantly concentrated on the
38 8th-12th leads, corresponding to the central region of the
39 brain. This is consistent with previous studies of MI brain
40 activities [24].
41

42 After building the weighted brain network and the binary
43 brain network, five measures can be extracted from the two
44 networks and combined with multiscale features respectively
45 to form all of the feature inputs. The classification
46 performance is first evaluated with each of the five measures
47 combined with multiscale features. The classification results
48 of the nine subjects are listed in Table 2.
49

50 As can be seen from Table 2, for weighted networks, the
51 data in each row of the table vary considerably, and this may
52 be caused by the subject's adaptability problem. By comparing
53 the average classification accuracy of the nine subjects,
54 features extracted by the CSP and LCD algorithms combined
55 with the measure of degree give the best results. The average
56 classification accuracy of four classes of MI reaches 79.69%.
57 The lowest classification accuracy for all five measures is the
58
59

clustering coefficient, which is 76.05%. As for the binary
network, the results of the five measures are very close, and
the average classification accuracies obtained by the measures
of degree and average shortest path length are both 79.5%,
which is very close to the best result. Compared to the
weighted networks, the results show that the difference
between the five measures extracted from the binary networks
is not significant.

A comparison of the average classification accuracy of the
five measures in the weighted network and the binary network
is illustrated in Figure 7. It can be seen that in addition to the
measure of degree, the other four measures extracted from the
binary network have better results than the weighted network.
The result with the measure of clustering coefficients in
weighted networks is the worst of all the features, while the
average classification accuracy obtained by the measure of
degree in the weighted network is the highest.

Previous studies of brain networks in disease states showed
that the clustering coefficient of patients with Alzheimer's
disease is significantly lower than that of normal people [37].
Our experimental results of the weighted brain network show
that the clustering coefficient does not accurately represent the
brain differences between different MI movements, but at
present, there are few research studies on the brain networks
of healthy people under different activities.

As a whole, the classification performance by measures
extracted from the binary networks is better overall. It also
shows that the measure of degree with the weighted networks
is very effective in classification despite the relatively poor
performance of the other four measures. In Figure 8, the
difference between the measures of degree of the binary
networks in the four classes of MI tasks can be clearly seen.
Based on the data of subject 8, the mean value of the degree
in each class of MI data is calculated, and the distribution of
the value for 22 channels can be observed.

The results show that the value of degree is generally high
in the left-hand MI task, while it is obviously low in the tongue
MI task, and the value of degree of the other two kinds of MI
tasks is very close. It is shown that the measure of degree can
well distinguish the left-hand and tongue MI tasks. In addition,
it can be seen from the values of degree under different
channels that the degree of channels 9-11 is generally higher
than that of other channels in each kind of MI task. These
channels correspond to the central region of the brain, which
is consistent with the conclusion in Figure 5.

The feature extraction algorithm proposed in this work
combines three algorithms: CSP, LCD, and brain network. In
order to see the performance of each subset of features
independently to show the contribution of each method, the
classification effect of the three methods is tested separately.

Table 3 shows the classification accuracy and kappa coefficient of the three methods. As can be seen from Table 3, the contribution of the CSP algorithm is the largest, the average classification accuracy is 72.9%, where the kappa score is 0.64, and the average classification accuracy of the LCD algorithm is the lowest at only 28.4%, where the kappa score is only 0.04. The classification accuracy of the combined algorithm reaches 79.7%, which is about 7% higher than the CSP algorithm. In addition to subject 7, the proposed method achieved the best classification effect on the other subjects. The results show that the classification effect after combining the three algorithms is greatly improved compared with the single algorithm.

In order to further verify the validity of the proposed method, the measures of degree in the binary networks with better results in both the weighted networks and binary networks are extracted as the brain network features. This is compared with some other popular feature extraction methods: support vector machine (SVM) [38], tangent space linear discriminant analysis (TSLDA) [39], and CSP

combined with LCD method [40]. A 10-fold cross-validation procedure is applied here. The results of the classification accuracy and the kappa score of the different feature extraction algorithms are listed in Table 4.

As can be seen from Table 4, except for subjects 5, 6, and 7, the proposed algorithm outperforms the other methods with all other subjects. Among them, subject 9 achieves the highest classification accuracy of 89.7%, and the kappa score is 86%. The average classification accuracy obtained by the proposed algorithm is 79.7%, which is 6% higher than that of the CSP-LCD algorithm, and the kappa score is 0.73. The proposed feature extraction algorithm combined with functional brain networks has the advantage of containing both the frequency and spatial features extracted from the MI EEG signals, and at the same time, tackles the difficulties caused by the differences between different subjects by utilizing brain network information. It is obvious that the accuracy and robustness of classification in the four classes of MI tasks are considerably improved by embedding features from three sources: CSP, LCD, and extra functional brain networks.

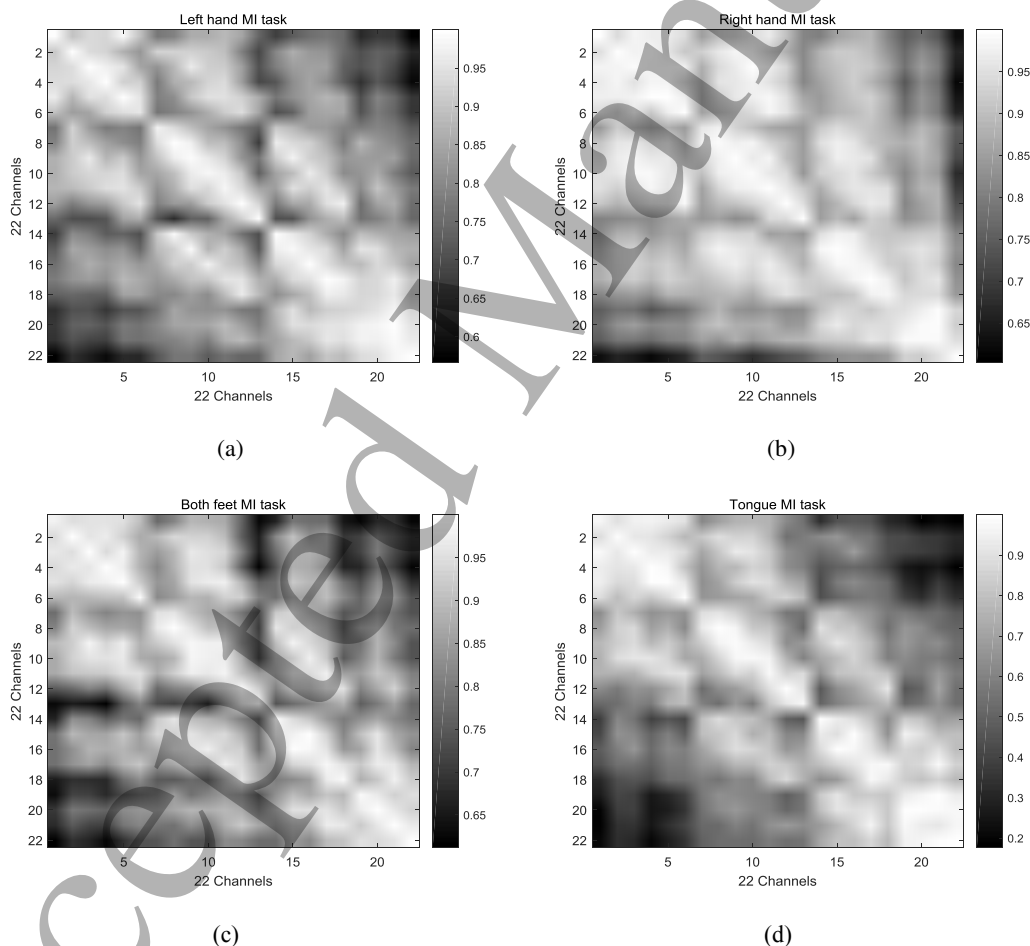


Figure 5. Adjacency matrix of weighted brain networks under four kinds of MI tasks: (a) left hand, (b) right hand, (c) both feet, and (d) tongue.

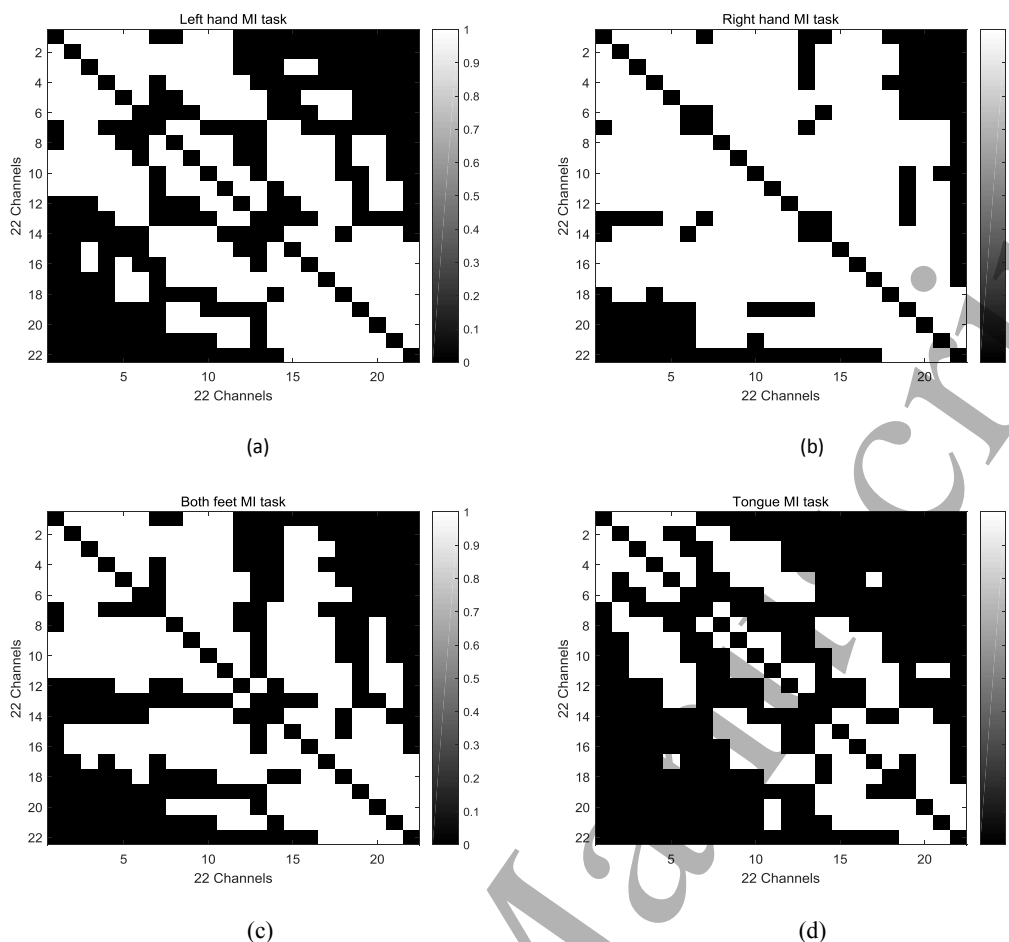


Figure 6. Adjacency matrix of binary brain networks under four kinds of MI tasks: (a) left hand, (b) right hand, (c) both feet, and (d) tongue.

Table 2. Classification accuracy (%) of five kinds of measures based on weighted network and binary network combined with CSP and LCD features.

Subject	Weighted network					Binary network				
	Degree	Betweenness centrality	Clustering coefficient	Average shortest path length	Local efficiency	Degree	Betweenness centrality	Clustering coefficient	Average shortest path length	Local efficiency
1	82.76	79.31	81.03	74.14	79.31	81.03	82.76	86.21	79.31	79.31
2	65.52	67.24	62.07	62.07	56.90	68.97	65.52	63.79	65.52	62.07
3	87.93	89.66	77.59	82.76	75.86	87.93	86.21	87.93	89.66	86.21
4	77.59	68.97	74.14	79.31	75.86	70.69	75.86	75.86	77.59	74.14
5	72.41	70.69	70.69	70.69	74.14	74.14	70.69	70.69	72.41	75.86
6	70.69	65.52	65.52	74.14	72.41	70.69	70.69	68.97	70.69	70.69
7	82.76	84.48	79.31	81.03	84.48	84.48	82.76	82.76	84.48	84.48
8	87.93	86.21	87.93	86.21	87.93	86.21	87.93	86.21	87.93	87.93
9	89.66	93.10	86.21	87.93	91.38	91.38	89.66	89.66	87.93	89.66
Mean	*79.69	78.35	76.05	77.59	77.59	*79.50	79.12	79.12	*79.50	78.93

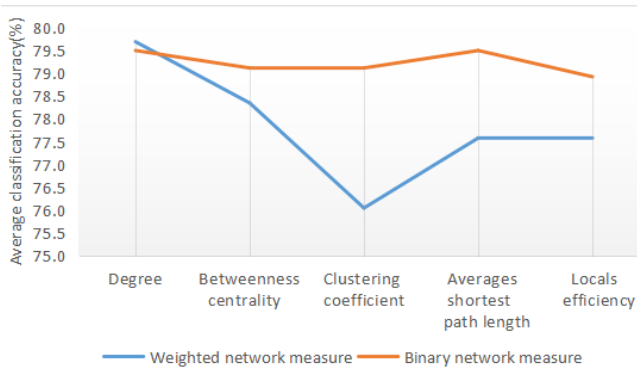


Figure 7. Comparison of average classification accuracy (%) between weighted network measures and binary network measures.

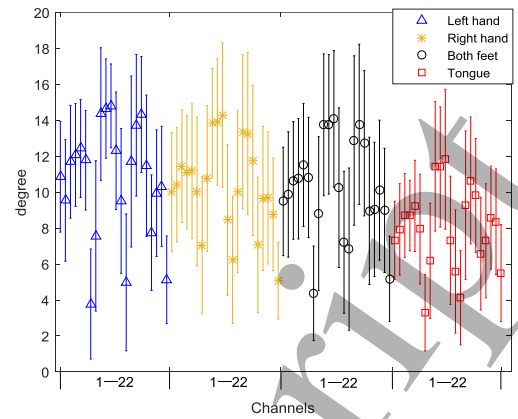


Figure 8. Measures of degree of the binary networks in four kinds of MI tasks under different channels.

Table 3. Classification accuracy (%) and kappa score of proposed three combined methods compared with three individual methods.

Subjects	CSP		LCD		Brain network		Proposed method	
	CA	K	CA	K	CA	K	CA	K
1	69.0	0.59	25.9	0.01	51.7	0.36	*82.8	0.77
2	48.3	0.31	25.9	0.01	31.0	0.08	*65.5	0.54
3	70.7	0.61	22.4	-0.03	62.1	0.49	*87.9	0.84
4	69.0	0.59	34.5	0.13	32.8	0.10	*77.6	0.7
5	70.7	0.61	32.8	0.10	39.7	0.20	*72.4	0.63
6	68.1	0.58	29.3	0.06	24.1	-0.01	*70.7	0.61
7	*84.5	0.79	32.8	0.10	41.4	0.22	82.8	0.77
8	*87.9	0.84	19.0	-0.08	65.5	0.54	*87.9	0.84
9	87.9	0.84	32.8	0.10	53.4	0.38	*89.7	0.86
Mean	72.9	0.64	28.4	0.04	44.6	0.26	79.7	0.73

Table 4. Classification accuracy (%) and kappa score of proposed method compared with three other feature extraction methods.

Subjects	SVM [38]		TSLDA [39]		CSP-LCD [40]		Proposed method	
	CA	K	CA	K	CA	K	CA	K
1	59.3	0.46	80.5	0.74	69.0	0.59	*82.8	0.77
2	59.3	0.46	51.3	0.35	56.9	0.43	*65.5	0.54
3	57.5	0.43	87.5	0.83	84.5	0.79	*87.9	0.84
4	55.4	0.4	59.3	0.46	46.6	0.29	*77.6	0.70
5	*76.1	0.68	45.0	0.27	69.0	0.59	72.4	0.63

6	56.1	0.41	55.3	0.4	*72.4	0.63	70.7	0.61
7	84.0	0.79	82.1	0.76	*86.2	0.82	82.8	0.77
8	76.1	0.68	84.8	0.8	*87.9	0.84	*87.9	0.84
9	75.7	0.68	86.1	0.81	87.9	0.84	*89.7	0.86
Mean	66.6	0.55	70.2	0.6	73.4	0.65	*79.7	0.73

4.2 Real-time data analysis

The online BCI robot control system is used to validate the real-time performance of the proposed algorithm. In this work, we used data from eight subjects for training a classification model. Then, the training model is used to realize the synchronous control and asynchronous control of the NAO robot. In the experiment of synchronous control, each participant conducts four experiments. The execution times are listed in Table 5.

Table 5. Execution time (s) of eight subjects to complete experiment of synchronous control.

Subjects	Experim -ent 1	Experim -ent 2	Experim -ent 3	Experim -ent 4	Mean
1	40	43	39	35	39.25
2	43	46	37	41	41.75
3	44	43	38	42	41.75
4	42	37	33	34	36.5
5	39	40	38	33	37.43
6	44	53	43	40	45.05
7	42	63	44	36	46.23
8	40	51	33	49	43.2

As mentioned, several key steps are involved in producing one single motion command for the robot, including the data acquisition and processing of EEG signals, conversion of the classification results to corresponding robot control commands, and execution of the corresponding motion actions with the robot. Meanwhile, the motion status of the robot needs to be fed back to the subjects in real time so that the subjects can decide the next MI task to be performed. One experiment involves a total of 17 iterations of the above. As shown in Table 5, the time it takes each subject to complete the experiment varies greatly, and it takes about 41 s for subjects to complete one full experiment. The average time of one instruction for the robot is about 2.4 s.

Owing to the different training effects of each subject in the previous period, the classification model obtained was not the

same, which led to differences in the results of the real-time experiments for each subject. In the real-time control of the robot, the experimental effects of each subject were easily affected by subjective factors because the robot only performed the corresponding movement after the subject issued the correct instruction. The reason for the incorrect instruction may be that the subject performed the wrong MI task or there was a false classification by the algorithm.

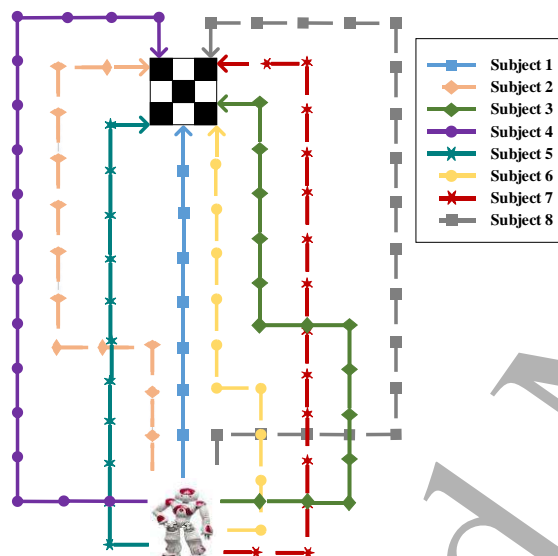
With asynchronous control, each subject plans and decides the motion actions of the robot. The mission is to control the robot to the designated position. During the experiment, the resulted trajectories of the robot controlled by the eight subjects are recorded (see Figure 9). The actual instructions received by the robot are listed in Table 6. Owing to the aforementioned error control mechanism, the instructions issued by the subjects are not all sent to the robot but need to be detected by the false alarm detector. The execution times required by the subjects to complete the experiments are also listed in Table 6.

As shown in Table 6, the subjects spent more than twice the time sending instructions. In addition to the reason of false classification by the algorithm, it should be noted that there is an additional error control mechanism that forces the subjects to perform the same MI tasks multiple times. Moreover, the subjects could unintentionally use incorrect MI mental activities that may accidentally change the trajectories of the robot. Each subject controlled the movements of the robot with different planned trajectories.

Subject 1 took the shortest path that only required eight instructions. The robot took the shortest time to complete the experiment at only 16.5 s. Subjects 4 and 8 decided to take the longest path that involved motions in all four directions. A total of 19 instructions were needed, and it took 41.8 s and 56.1 s to complete the experiment, respectively. On average, it took 2.6 s for each subject to send an instruction. This is slightly longer than that of the synchronous experiment (2.4 s). Because each of the subjects used different paths and different MI tasks were performed, the classification accuracy of each type of MI signal varied as well, resulting in different time spent by each subject.

Table 6. Instructions sent by eight subjects, and execution time(s) of experiment.

Subjects	Control instructions	Time(s)
1	00000000	16.5
2	0001100000022	26.4
3	222000011000001	31.9
4	111000000000022223	41.8
5	100000000002	27.5
6	20001000000	34.1
7	220000000000011	49.5
8	022220000000011113	56.1

**Figure 9.** Eight trajectories of robot controlled by eight subjects.

5. Conclusions

In this paper, a novel feature extraction method was proposed to classify four classes of MI signals by combining CSP, LCD, and functional brain networks. Features were extracted in the frequency domain and spatial domain from MI EEG signals by using the CSP and LCD algorithms. Brain networks were then constructed using the EEG signals of each subject. The measures of degree of the brain networks were extracted to characterize the subjects' brain activities. The proposed method was integrated in a real-time BCI robot control system designed for real-world experiments.

The method was validated using the BCI competition IV dataset 2a for offline study and online BCI data collected with

the real-world experiments in this work. The experimental results for the two databases showed that the proposed feature extraction method can effectively extract discriminative features of the MI signals, enhancing the classification accuracy considerably in comparison with popular state-of-the-art algorithms. In particular, it is worth noting that the method performs robustly when dealing with different subjects, meaning that the method can effectively eliminate individual differences by analyzing the functional brain network of each subject. The computation time demonstrates the capability of real-time applications and the feasibility of being applied in practical rehabilitation BCI systems.

The functional brain network constructed in this paper belongs to the category of nondirected networks, which simply omit information flows in functional brain networks. In the future, it will be interesting to construct directional functional brain networks that characterize the causal relationships of neural activities in order to further extract discriminative information. In addition, we will continue to study the channel selection algorithms in order to reduce the number of channels required to construct a compact set of more representative features extracted from the EEG signals. The ultimate goal is to improve the performance with regard to the accuracy and robustness of classification and the suitability for practical BCI rehabilitation systems.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 51675389 and 51705381), Hubei Provincial Natural Science Foundation of China (Grant No. 2018CFB370), and the Excellent Dissertation Cultivation Funds of Wuhan University of Technology (Grant No. 2017-YS-058).

References

- [1] Berger H 1929 Electroencephalogram in humans *Arch. Psychiatr. Nervenkr.* **87** 527-570
- [2] Javaid M A 2016 *Ssrn Electronic Journal* **17** 89-103
- [3] Nicolas-Alonso L F, Corralejo R, Gomez-Pilar J, Alvarez D and Hornero R 2015 *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **23** 702-712
- [4] Perdakis S, Leeb R, Williamson J, Ramsay A, Tavella M, Desideri L and d R Millán J 2014 *J. Neural. Eng.* **11** 036003
- [5] Pfurtscheller G, Muller-Putz G R, Schlogl A, Graimann B, Scherer R and Leeb R 2006 *IEEE Trans. Rehabil. Eng.* **14** 205-210
- [6] Leeb R, Lancelle M, Kaiser V, Fellner D W and Pfurtscheller G 2013 *Computational Intelligence & AI in Games* **5** 117-128
- [7] Pfurtscheller G, Leeb R, Keinrath C, Friedman D, Neuper C and Guger C 2006 *Brain Research* **1071** 145-152

- [8] Clerc M, Bougrain L and Lotte F (Eds) 2016 *Brain-computer Interfaces 2 Technology and Applications* John Wiley and Sons
- [9] Shi T, Wang H and Zhang C 2015 *Expert Systems with Applications* **42** 4196–4206
- [10] Müller-Putz G, Leeb R, Tangermann M, Höhne J, Kübler A, Cincotti F and Millán J D R 2015 *Proceedings of the IEEE* **103(6)** 926–943
- [11] Pfurtscheller G, Neuper C, Schlogl A and Lugger K 1998 *IEEE Trans. Rehabil. Eng.* **6** 316–25
- [12] Hsu W Y 2015 *Clinical Eeg & Neuroscience* **46** 94–99
- [13] Huang N E, Shen Z, Long S R, Wu M C, Shih H H and Zheng Q 1998 *Proceedings Mathematical Physical & Engineering Sciences* **454** 903–995
- [14] Ramoser H, Muller-Gerking J and Pfurtscheller G 2000 *IEEE Trans. Rehabil. Eng.* **8** 441–6
- [15] Aghaei A S, Mahanta M S and Plataniotis K N 2016 *IEEE Trans. Biomed. Eng.* **63** 15–29
- [16] Zheng J, Cheng J and Yang Y 2013 *Mechanism & Machine Theory* **70** 441–453
- [17] Pfurtscheller G and Aranibar A 1979 *Electroencephalography & Clinical Neurophysiology* **46** 138–146 7
- [18] Pfurtscheller G and Neuper C 1994 *Neuroscience Letters* **174** 93
- [19] Nijholt A, Tan D, Pfurtscheller G, Brunner C, Millán J D R and Allison B 2008 *IEEE Intelligent Systems* **23** 72–79
- [20] Vidaurre C and Blankertz B 2010 *Brain Topography* **23** 194–198
- [21] Qi J, Wang J, Wang J, Xiao Q and Yang H 2011 *Fluctuation and Noise Letters* **10** 371–379
- [22] Luque B, Lacasa L, Ballesteros F J and Robledo A 2011 *Plos One* **6** e22411
- [23] Huang N E, Shen Z, Long S R, Wu M C, Shih H H and Zheng Q 1998 *Proceedings A* **454(1971)** 903–995
- [24] Mcfarland D J, Miner L A, Vaughan T M and Wolpaw J R 2000 *Brain Topography* **12(3)** 177–186
- [25] Watts D J and Strogatz S H 1998 *Nature* **393** 440–442
- [26] Strogatz S H 2001 *Nature* **410** 268–276
- [27] Onnela J P, Saramäki J, Kertész J and Kaski K 2005 *Phys. Rev.* **71** 065103
- [28] Latora V and Marchiori M 2001 *Phys. Rev. Lett.* **87** 198701
- [29] Newman M E J 2003 *Social Networks* **27** 39–54
- [30] Erdős P and Rényi A 1961 *Acta Mathematica Hungarica* **12** 261–267
- [31] Cai D, Zhang C and He X 2010 *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 333–342
- [32] Cai D, He X and Han J 2008 *IEEE Transactions on Knowledge & Data Engineering* **20** 1–12
- [33] Shiman F, Eduardo LópezLarraz, Sarasolasanz A, IrastorzaLanda N, Martin Spüler and Birbaumer N 2017 *J. Neural. Eng.* **14(4)** 046018
- [34] Tangermann M, Müller K R, Aertsen A, Birbaumer N, Braun C, Brunner C, Leeb R, Mehring C, Miller K J and Nolte G 2012 *Front. Neurosci.* **6** 55
- [35] Oostenveld R and Praamstra P 2001 *Clin. Neurophysiol.* **112** 713–9
- [36] Alvarez-Meza A M, Velasquez-Martinez L F and Castellanos-Dominguez G 2015 *Neurocomputing* **151** 122–129
- [37] Supekar K, Menon V, Rubin D, Musen M and Greicius M D 2008 *Plos Computational Biology* **4(6)** e1000100
- [38] Wang D, Miao D and Blohm G 2012 *Front. Neurosci.* **6** 151
- [39] Barachant A, Bonnet S, Congedo M and Jutten C 2012 *IEEE Trans. Biomed. Eng.* **59** 920–928
- [40] Liu A, Chen K and Liu Q 2017 *Sensors* **17** 2576