

Article



EEG Feature Extraction Based on a Bilevel Network: Minimum Spanning Tree and Regional Network

Zhizeng Luo *, Xianju Lu and Xugang Xi 💿

Institute of Intelligent Control and Robotics, Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China; luxianju@hdu.edu.cn (X.L.); xixi@hdu.edu.cn (X.X.)

* Correspondence: luo@hdu.edu.cn; Tel.: +86-0571-8691-5009

Received: 4 December 2019; Accepted: 18 January 2020; Published: 21 January 2020



Abstract: Feature extraction is essential for classifying different motor imagery (MI) tasks in a brain-computer interface (BCI). Although the methods of brain network analysis have been widely studied in the BCI field, these methods are limited by differences in network size, density, and standardization. To address this issue and improve classification accuracy, we propose a novel method, in which the hybrid features of the brain function based on the bilevel network are extracted. Minimum spanning tree (MST) based on electroencephalogram (EEG) signal nodes in different MIs is constructed as the first network layer to solve the global network connectivity problem. In addition, the regional network in different movement patterns is constructed as the second network layer to determine the network characteristics, which is consistent with the correspondence between limb movement patterns and cerebral cortex in neurophysiology. We attempt to apply MST to the classification of the MI EEG signals, and the bilevel network has better interpretability. Thereafter, a vector is formed by combining the MST fundamental features with the directional features of the regional network. Our method is validated using the BCI Competition IV Dataset I. Experimental results verify the feasibility of the bilevel network framework. Furthermore, the average classification performance of the proposed method reaches 89.50%, which is higher than that of other competing methods, thereby indicating that the bilevel network is effective for MI classification.

Keywords: motor imagery; feature extraction; bilevel network; minimum spanning tree

1. Introduction

Electroencephalogram (EEG) is a bioelectrical signal formed by the simultaneous synthesis of many postsynaptic potentials, and reflects the state of the brain and the activation of nerve cells [1]. Therefore, we can obtain substantial physiological, psychological, and pathological information through the analysis of EEG signals [2,3]. The full expression of the various levels and aspects of signal characteristics is important in the recognition process. In the past decades, numerous algorithms, such as fast Fourier transform (FFT), power spectral density (PSD), autoregressive (AR) model, band power (BP), and common spatial patterns (CSP), have been applied to extract the EEG features of different motor imageries (MIs) [4]. Wu et al. [5] combined CSP and multivariate empirical mode decomposition to classify left- and right-hand MI. Durongbhan et al. [6] extracted different features through wavelet transform (WT) and FFT to classify Alzheimer's disease.

However, the aforementioned methods are aimed at a specific motor cortex of the brain. Therefore, only the directly related features were extracted in the initial research process. As the research progresses, numerous comprehensive features will result in considerably effective expression. Recent research indicates that combining bridge neurofeedback, connectomics, and network control theory can better understand the current frontier of human cognition [7]. Functional connectivity often is referred to as the addition of quantitative measures of coordinated activity [8]. The functional connectivity

model can provide a cost-effective approach to differentiate and integrate information. Graph theory is a common tool for analyzing the topology of neural networks and provides a comprehensive framework for characterizing network topology [9]. The brain network is based on graph theory, which can visually show the dynamic interaction among neurons, neuron clusters, and brain regions [10,11]. Furthermore, the brain network can construct brain functional topology and integrate the connectivity strength between different regions. The brain network also reflects the global activation of the cerebral cortex and provides a highly reliable performance for distinguishing MIs. Many studies have demonstrated that the brain network is an effective method for describing coordination among the different brain regions. Stanley et al. [12] used global and local efficiency to analyze the relationship between the efficiency of brain information transmission and working memory performance in young and old people. Kong et al. [13] performed spectral decomposition on brain networks to classify four-class tasks. Although the current methods of brain network analysis can relatively solve the classification problems of brain function state, the construction of the network has no unified standard [14]. The construction of a brain network is mainly obtained by global threshold processing. Different thresholds often result in the same topology that generates different unweighted networks. Moreover, all valid nodes have no guarantee that they will be connected to the network, thereby resulting in insufficient brain information.

Minimum spanning tree (MST) is mathematically defined as an acyclic subgraph that connects all nodes in the original weighted network with minimal link weights [15-17]. MST is the only subnetwork in the weighted network that avoids methodological biases and solves the problem of network connectivity. Moreover, MST can simplify network characteristics and directly compare networks with the same number of nodes [18]. MST is a new and emerging method of human functional connectomics and has been effectively used in the recent research on some neuropsychiatric diseases [19]. As early as 2006, Lee et al. [20] have applied the MST analysis to the EEG signals and demonstrated its effectiveness in the study of epilepsy and in characterizing the network topology of different epilepsies. Olde Dubbelink et al. [21] found that MST of the brain network of different patients with Parkinson's disease could identify clinically relevant changes at the early stage of the disease. These studies have proven the potential of MST for network analysis. Given that MST can effectively determine the essential attributes of complex networks, this method is unbiased for studying brain networks. Methods based on MST in the classification of the MI EEG signals have been scarcely investigated. However, in [22], functional connectivity analysis and MST could successfully distinguish between imagery hand movements and silent state. Therefore, MST should be applied to the classification of the MI EEG signals in the present brain-computer interface (BCI) research, which provides an effective new idea for classifying the patterns of brain consciousness tasks.

The network constructed by MST has the characteristics of the original network index structure and summary information and is an unbiased and effective network topology information structure. However, the brain network is a small-world network from the perspective of neurophysiology, and the MST clustering coefficient is constantly zero and has no small-world network characteristics. To overcome these limitations, we propose a novel method, which is the hybrid features of the brain function based on the bilevel network. Lastly, the experimental results demonstrate the success of the bilevel network in dealing with the MI classification. Moreover, our proposed method achieves better classification performance than using a single-layer network and other algorithms.

In summary, this study's contributions are as follows.

- We attempt to introduce MST to the MI-BCI research and improve the classification performance.
- Our proposed bilevel network framework is simple, effective, and has extensive interpretability.
- We demonstrate the superiority of our method in the BCI Competition IV Dataset I.

The remainder of this paper is organized as follows. Section 2 discusses the construction and feature analysis of the bilevel network, as well as the experimental scheme. Section 3 presents the details of the experimental results. Lastly, Section 4 provides the discussion and conclusion.

2. Materials and Methods

2.1. Overview

Figure 1 shows an overview of our proposed bilevel network method. First, we preprocess the BCI Competition IV Dataset I. Second, we constructed the MST for different MIs based on the graph theory. Third, the regional network is constructed based on MST by combining the corresponding relationship between limb movement and cerebral cortex. We compute diameter, average eccentricity, average node degree, average clustering coefficient, and average path length of the bilevel network. Lastly, we use support vector machine (SVM) to learn a classifier for the MI tasks.



Figure 1. Overview of the bilevel network framework.

2.2. Motivation

MST of each movement was taken as the first network layer to comprehensively describe the topological structure of the brain network in different MIs. On this basis, we constructed a regional network as the second layer of the network according to the corresponding relationship between limb movement and cerebral cortex in neurophysiology. The first network layer ensures that it connects all nodes in the network and maintains the network tree trend in each movement. In addition, the first network layer gives the node association and aggregation degree of the brain network. The second network layer focuses on reflecting the activation degree of the corresponding regions of different MIs. Moreover, the second network layer highlights the individualized network characteristics of each movement. Therefore, a vector with the fundamental and directional characteristics of the brain network is formed by combining the relevant features of the bilevel network. Accordingly, our method has substantial interpretability.

2.3. EEG Signal Features

This study fuses the features of MST and the regional network to characterize the EEG signals. The construction of the regional network is based on the constructed MST. Therefore, the methods of node selection and node relationship quantification of the regional network are consistent with MST. The unified description is as follows.

Step 1: We select the appropriate network nodes. Given that multichannel EEG signals are used, we define each electrode on the scalp surface as a network node.

Step 2: We quantify the functional connectivity relationship between network nodes. The commonly used methods are Pearson correlation coefficient, coherence spectrum, and mutual information. Pearson correlation coefficient is selected because of its better noise suppression and robustness compared with the other two methods [23]. The formula is as follows:

$$r_{ij} = \frac{\sum_{t=1}^{T} [x_i(t) - \overline{x_i}] [x_j(t) - \overline{x_j}]}{\sqrt{\sum_{t=1}^{T} [x_i(t) - \overline{x_i}]^2 \sum_{t=1}^{T} [x_j(t) - \overline{x_j}]^2}} \quad i, j = 1, 2, \cdots, N,$$
(1)

where $x_i(t)$ and $x_j(t)$ are the sampling values of nodes *i* and *j*, respectively, at time *t*; $x_i(t)$ and $x_j(t)$ are the average sampling values of nodes *i* and *j*, respectively; and *N* is the number of network nodes. The larger the absolute value of r_{ij} , the stronger the correlation between the two nodes. We can obtain an $N \times N$ connection coefficient symmetric matrix.

2.3.1. MST Features

In a given undirected graph G = (V, E), V is the set of nodes and E is the set of edges. Hence, (i, j) is the edge connecting nodes i and j, and w(i, j) is the weight of this edge. If T is a subset of E and a cyclic graph, so that w(T) in Equation (2) is the smallest, then T is the MST of G. Each node in T is called a connected component.

$$w(T) = \sum_{(i,j) \subset T} w(i,j).$$
⁽²⁾

The MST construction steps are listed as follows.

Step 1: Define the weight of the edge

We need to connect the nodes with high correlation because MST, as the connection edge, is the edge with the smallest weight. Therefore, we consider the reciprocal Pearson correlation coefficient r_{ij} between nodes as the weight of the edge.

Step 2: Select the algorithm to solve MST

The Kruskal and Prim algorithms are commonly used to solve MST in graph theory. We choose the Kruskal algorithm, which has a wider application [24]. The algorithm is as follows:

Algorithm1 Kruskal (Figure 2)

- 1: Initialize two sets, V and E. V is the set of nodes and E is the set of edges.
- 2: Create a graph $G \leftarrow \emptyset$.
- 3: Sort *E* according to the weight of edges.
- 4: while the num of edges in $G \neq$ the size of *E* do
- 5: Select $e_{i,i} \in E$, so that $e_{i,i}$ is the edge with the smallest weight.
- 6: **if** node *i* directly reaches node *j* **then**
- 7: Add $e_{i,j}$ and nodes *i* and *j* into the graph.
- 8: Remove $e_{i,j}$ from *E*.
- 9: end if
- 10: end while





Step 3: Calculate the MST features

We use MST of different MIs as the first network layer, which is considerably fundamental. Therefore, we select the global metrics of the MST topology, namely, diameter and average eccentricity, as features [16]. The diameter is the distance between the two nodes that are the farthest apart in the MST, and can be used to measure the similarity between MST and the small-world network. The average eccentricity can be used to measure the global distribution of MST in the brain network.

The diameter is calculated as follows:

$$D = \max\{d(i, j) | i, j \in T\}.$$
(3)

The average eccentricity is calculated as follows:

$$M_{ee} = \frac{\sum\limits_{i,j\in T} d(i,j)}{N},\tag{4}$$

where d(i, j) is the shortest path length between nodes *i* and *j* in MST, and *N* is the number of network nodes.

2.3.2. Regional Network Features

The event-related desynchronization (ERD)/event-related synchronization (ERS) phenomenon is caused by the resonance of numerous neurons on physiological electrical signals. This phenomenon indicates the interaction between neurons and local neurons in a certain frequency band of EEG, which mainly appears on the sensory motion region corresponding to the electrodes C3, C4, and Cz. For example, the ERD on the left side of the motor region (C3) is observed in the right hand MI; ERD on the right side of the motor region (C4) is observed in the left MI; and ERD in the midline central

region (Cz) is observed in foot MI [25,26]. Therefore, the regional network centered electrodes C3, C4, and Cz are constructed for the MI of the right hand, left hand, and foot, respectively. The regional network is more directional than the MST and can better reflect the brain network characteristics of each movement. The construction steps are listed as follows.

Step 1: Quantify the correlation between nodes

The Pearson correlation coefficient method is also selected for calculation. However, the regional networks constructed in different movement patterns are centered on electrodes C3, C4, and Cz. Therefore, we compute for the $1 \times N$ connection coefficient matrix.

Step 2: Matrix thresholding

By selecting an appropriate threshold value (δ) and performing threshold processing on the connection coefficient matrix, we obtain a binary matrix that can be expressed as follows:

$$a_{ij} = \begin{cases} 1 & r_{ij} \ge \delta \\ 0 & r_{ij} < \delta \end{cases}$$
(5)

 a_{ij} = 1 indicates that a connection edge exists, and the correlation between nodes *i* and *j* is strong. Otherwise, the connection does not exist and the correlation between nodes is weak.

Step 3: Calculate the regional network features

Commonly used network metrics include node degree, path length, clustering coefficient, and network density. These metrics reflect the structural characteristics of the network. We choose average node degree, average clustering coefficient, and average path length as features [10,27]. The average node degree is an important indicator in measuring the size of a network. The average clustering coefficient reflects the tightness of the local connection of the network. Lastly, the average path length can measure the connection characteristics of the network.

The average node degree (K) can be computed as follows:

$$K = \frac{1}{N} \sum_{i=1}^{N} k_i,\tag{6}$$

where k_i is the degree of node *i*, indicating the number of nodes directly connected to node *i*. The greater the degree of a node, the more important it is in the network. *N* is the number of all nodes.

The average clustering coefficient (*C*) can be computed as follows:

$$C = \frac{1}{N} \sum_{i=1}^{N} c_i = \frac{1}{N} \sum_{i=1}^{N} \frac{2E_i}{k_i(k_i - 1)},$$
(7)

where E_i is the actual number of edges directly connected to node *i*, c_i is the clustering coefficient of node *i*, and k_i is shown in Equation (6).

The average path length (L) can be computed as follows:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in V, i \neq j} d(i,j),$$
(8)

where d(i, j) is shown in Equation (3).

2.4. Feature Fusion

To characterize EEG signals more comprehensively, we would like to combine MST features and regional network features into a framework. To avoid the impact of the dimensions of two types of

features, it is necessary to perform a normalization process before feature fusion. Specifically, the feature vector (*F*) is normalized as:

$$\widetilde{F} = \frac{F\% - \mu}{\sigma},\tag{9}$$

where μ is the average vector of training feature vectors, and σ is the average of standard deviation of training feature vectors.

Serial combination is a widely implemented feature fusion method [28]. With the serial combination, the fused feature vector can be obtained, given by

$$\widetilde{F}_{s} = \left(\begin{array}{c} \widetilde{F}_{MST} \\ \widetilde{F}_{RN} \end{array}\right), \tag{10}$$

where \overline{F}_{MST} is the normalized MST feature vector, \overline{F}_{RN} is the normalized regional network feature vector.

Although serial combination can achieve satisfactory classification results, the contribution of different features to the classification is ignored. In [29], a Bayesian fusion approach based on the weighted average method was proposed. The core of this method is that the contribution of each modality can be automatically weighed to optimize performance, and it will give access to the metric that best classifies the data. This method is similar to hybrid-BCI systems and more details can be found in [29–31]. Therefore, the weight can be computed as follows:

$$\lambda_i = \frac{p_i}{p_{MST} + p_{RN}},\tag{11}$$

where p_i is the posterior probability of the modality *i*.

In this way, different features can get corresponding weights, thereby optimizing the feature selection of each individual.

2.5. Experimental Scheme

The data set in the BCI Competition IV Dataset I consists of human- and artificially generated MI data. This study, we only considered the human EEG data. The data set provided by the Berlin BCI Research Center and more details of the data set can be found in [32]. The experimental paradigm was for seven healthy subjects (i.e., "a" to "g") to perform any two classes of the left hand, right hand, and foot MI tasks in the face of the computer screen. Subjects "a" and "f" chose left hand and foot MI tasks, and the remainder chose left- and right-hand MI tasks. The EEG signals were recorded at the following 59 positions of the international 10–20 system: AF3, AF4, F5, F3, F1, Fz, F2, F4, F6, FC5, FC3, FC1, FC2, FC2, FC4, FC6, CFC7, CFC5, CFC3, CFC1, CFC2, CFC4, CFC6, CFC8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, CCP7, CCP5, CCP3, CCP1, CCP2, CCP4, CCP6, CCP8, CP5, CP3, CP1, CP2, CP2, CP4, CP6, P5, P3, P1, Pz, P2, P4, P6, PO1, PO2, O1, and O2. The sampling rate of the amplifier was 100 Hz. We selected the data set of all subjects for experimental analysis. Each subject performed a total of 200 MI tasks. The data acquisition time of each trial was 8 s. From 0 to 2 s, the screen remained blank. From 2 to 4 s, the subject was in the preparation stage of MI when the screen displayed a mark "+". At the second 4 s, an arrow to the left, right, or down was displayed as a cue randomly, and the subject had to concentrate on performing the corresponding movement according to the screen for 4 s. At the end of one trial, the subject had a few minutes break. The subject imagined only one mental task in one trial, and each of the movements was imagined 100 times. The experimental paradigm and electrode positions are shown in Figure 3.



Figure 3. (a) Experimental paradigm of one trial. (b) Electrode positions.

2.6. EEG Data Preprocessing

EEG signals were preprocessed before further analysis. Fast independent component analysis was used for denoising to obtain more pure EEG signals, ensuring the effectiveness of the extracted features in the subsequent work. The ERD/ERS phenomenon is related to the mu (8–12 Hz) and beta (13–30 Hz) rhythms of EEG signals. However, some frequencies in the beta rhythm are the harmonic waves of the mu rhythm. Thus, mu rhythm is associated with motion or MI. The ERD/ERS in the mu rhythm can be used as a direct indicator that reflects the excitation degree of neurons in the cortical region and assesses neurodevelopment [33,34]. The EEG signals of 59 channels were decomposed by wavelet packet. Thereafter, we reconstructed the wavelet coefficients to obtain the mu rhythm, which was used in the latter network construction and feature extraction. Taking subject "a" as an example, the original EEG and mu rhythm were given to the C4 channel after reconstruction by wavelet packet (Figure 4).



Figure 4. Original electroencephalogram (EEG) signal and mu rhythm of subject "a".

2.7. Feature Classification with SVM

To automatically identify subjects MI data, our extracted features used machine learning methods to learn a classifier. Widely used classifiers include SVM, decision trees, linear discriminant analysis (LDA), and neural networks. Each classifier has its own application. Meanwhile, SVM classification

algorithm has greater advantages when processing small sample data [35]. Based on the characteristics of MI data, we chose SVM to classify mental tasks. The 10-fold cross-validation technique was chosen to evaluate the classification performance of network features. For each subject, 90% of the trials were used as the training set, and the rest were used as the testing set. This process was repeated 10 times to ensure that each trial was treated as a test sample. Lastly, we set the average of the 10 accuracies as the accuracy of one subject.

3. Results

3.1. MST in Different Movements

Subjects "a" and "f" performed left hand and foot MI tasks, whereas subjects "b", "c", "d", "e", and "g" performed left- and right-hand MI tasks. Therefore, we took subjects "a" and "b" as representatives of the two groups to illustrate the MST in different movements. First, we extracted the mu rhythm of each channel by wavelet packet. Second, we calculated the Pearson correlation coefficients between the mu rhythms of any two of the channel signals. Third, we averaged the Pearson correlation coefficients for the same movement of one subject and obtained a 59 × 59 connection coefficient matrix. We took the reciprocal of the connection coefficient matrix to obtain the weight matrix and facilitate the solution of Kruskal algorithm. The connection coefficient matrix and MST of subjects "a" and "b" in different movements are shown in Figures 5 and 6, respectively.



Figure 5. Minimum spanning tree (MST) in the different motor imagery (MI) tasks of subject "a". (a) Left hand; (b) foot.



Figure 6. MST in different MI tasks of subject "b". (a) Left hand; (b) right hand.

Figures 5 and 6 intuitively show that MST considerably simplifies the network structure. Thereafter, all nodes in the network are connected, addressing the issue of node loss during brain network construction. A significant difference in MST, particularly in the network tree trend, node association, and aggregation, occurs when the subject imagines different movements. The reason is that different movements activate different regions of the cerebral cortex, thereby verifying the functions of different brain regions that are specific [36]. In addition, MST also differs when different subjects imagine the same movement. The reason is that when different subjects imagine the same movement, the corresponding activation degree of the cerebral cortex region is different, which is related to the environment and psychological state of the subjects.

The differences in MST structures can be visually presented by the tree representation. By comparing the tree representation of the subjects "a" and "b" in Figure 7, we can observe that the root of each MST starts from the channel AF3. However, the subsequent trend of the tree is significantly different. The number of branches in Figure 7a is the largest, and that in Figure 7d is the least. In addition, the tree representation shows the changes in node association and aggregation of the same node in different movements and reflects the trend of the MST network nodes.



Figure 7. Cont.



Figure 7. Tree representation of MST of subjects "a" and "b". (a) Left hand of subject "a"; (b) foot of subject "a"; (c) left hand of subject "b"; (d) right hand of subject "b".

Although MST is not looped and the clustering coefficient is zero, the ability of MST to capture the original network index structure and summary information is invariant. Therefore, the MST features can be used to characterize the EEG signals in different MIs. Figure 8 shows the distribution of the diameter and average eccentricity of subjects "a" and "b" in the different MIs.



Figure 8. Distribution of diameter and average eccentricity of subjects "a" and "b" in different MIs. (**a**) Subject "a"; (**b**) subject "b".

As shown in Figure 8, the degree of distinction is slightly evident although overlaps among the points of movements of data are observed. Hence, the MST analysis is at least as sensitive as or more sensitive than the traditional graph theory analysis. This finding strongly confirms that the diameter and average eccentricity of the MST can be used as criteria for distinguishing MIs. This result further reveals that our assumption of introducing MST into the BCI research is promising. However, MST does not reflect attributes, such as clustering, that depend on the cycle. Thus, the MST clustering coefficient is constantly zero and has no small-world network properties. The brain

network is a small-world network. We constructed regional networks to characterize EEG signals more comprehensively because the MST does not have small-world properties.

3.2. Regional Network in Different Movements

In neurophysiological studies, the ERS/ERD phenomenon mainly appears on the sensory motion region corresponding to electrodes C3, C4, and Cz when subjects imagine right hand, left hand, and foot movements, respectively. The construction of the regional network is based on MST, and then threshold processing is carried out for the functional connection between the electrodes (C3, C4, and Cz) corresponding to the movements (right hand, left hand, and foot) and other electrodes. Therefore, we constructed regional networks centered on electrodes C3, C4, and Cz based on MST in Section 2.3.2. The selection of connection threshold in the process of regional network construction has no uniform standard. Therefore, we selected the threshold $\delta = 0.7$ according to the method in [37]. Figures 9 and 10 show the regional networks of subjects "a" and "b," respectively.



Figure 9. Regional network in the different MI tasks of subject "a". (a) Left hand; (b) foot.



Figure 10. Regional network in the different MI tasks of subject "b". (a) Left hand; (b) right hand.

Evidently, regional networks have different loops, which make them more compatible with small-world network. The regional network under different movements can reflect the changes in network location and connection well. After threshold processing, nodes with high correlation are clustered and retained. Indeed, when subject "a" imagined left hand movement, a regional network centered on electrode C4 was constructed on the basis of MST, which made the network in the vicinity of electrode C4 in the motor control cortex more agglomerated. In addition, when subject "a" imagined foot movement, the network in the vicinity of electrode C2 in the motor control cortex was highly agglomerated. Similarly, for the left- or right-hand MIs of subject "b," the network in the vicinity

based on the MST can more effectively characterize each movement pattern with better interpretability. Figure 11 shows the features distribution of subjects "a" and "b" in different MIs. The average node degree, average clustering coefficient, and average path length as features to distinguish MIs are still slightly evident (Figure 11).



Figure 11. Feature distribution of subjects "a" and "b" in the different MIs. K represents the average node degree, L represents the average path length, and C represents the average clustering coefficient. (a) Subject "a"; (b) subject "b".

3.3. Classification

To evaluate the significant difference between the two mental tasks of each subject, the one-way analysis of variance (ANOVA) was performed. As shown in Table 1, the *p*-value of MST features, regional network features, and bilevel network features is less than 0.05. This result suggests that there are significant differences in these mental tasks.

Subject	MI -	<i>p</i> -Value			
		MST	RN	SC	BF
а	Left hand/Foot	0.031	0.026	< 0.01	< 0.01
b	Left/Right hand	0.024	0.018	< 0.01	< 0.01
с	Left/Right hand	0.020	0.019	< 0.01	< 0.01
d	Left/Right hand	0.019	0.015	< 0.01	< 0.01
e	Left/Right hand	0.027	0.011	< 0.01	< 0.01
f	Left hand/Foot	0.038	0.033	< 0.01	< 0.01
g	Left/Right hand	0.023	0.012	< 0.01	< 0.01

Table 1. *p*-values of the different features.

Notes: regional network is abbreviated as RN, serial combination is abbreviated as SC, and Bayesian fusion approach based on the weighted average is abbreviated as BF.

To evaluate our method, we chose the SVM method to classify mental tasks. We respectively input the MST features, regional network features, and bilevel network features as vectors into SVM. In this study, there are two ways of fusing the bilevel network features. One method is based on serial combination and the other is Bayesian fusion approach based on the weighted average. Table 2 shows the classification accuracies of the seven subjects.

Subject	МІ	Features			
		MST	RN	SC	BF
а	Left hand/Foot	69.32	71.86	83.96	86.24
b	Left/Right hand	72.34	77.48	87.24	88.31
с	Left/Right hand	77.86	80.16	91.34	92.89
d	Left/Right hand	73.24	77.66	88.64	89.51
e	Left/Right hand	71.62	79.37	89.17	90.92
f	Left hand/Foot	65.82	75.93	86.27	88.46
g	Left/Right hand	72.19	76.87	88.37	90.16
Average accuracy of left hand/foot		67.57 ± 2.47	73.90 ± 2.88	85.12 ± 1.63	87.35 ± 1.57
Average accuracy of left/right hand		73.45 ± 2.53	78.31 ± 1.39	88.95 ± 1.51	90.36 ± 1.71
Average accuracy of all subjects		71.77 ± 3.68	77.05 ± 2.70	87.85 ± 2.34	89.50 ± 2.12

Table 2. Classification accuracies of the different features (%).

As shown in Table 2, the classification performance of feature fusion is significantly better than that of an independent feature. Notably, among the two ways of feature fusion, the Bayesian fusion approach based on the weighted average has better performance in each subject. The reason is that it can combine the most significant features of each network layer, thereby reducing the subjects' misclassification. This result further reveals that the Bayesian fusion approach based on the weighted average can progressively capture the characteristics of the neural plasticity phenomena. In addition, Table 2 shows that the average accuracy of all subjects based on the features of bilevel network is 89.50%, which has 17.73% and 12.45% gains compared with using only MST or regional network features, respectively. Furthermore, the average accuracy of the bilevel network for left-/right-hand MI is 90.36%, which has 3.01% gain over the left hand/foot MI. This result reveals that the bilevel network is better at handling the left-/right-hand MI. In particular, the highest accuracy of subject "c" reached 92.89%. These results demonstrated the efficacy of the proposed feature extraction method based on a bilevel network and verified that the combination of the fundamental and directional features can achieve higher accuracy than independent features.

At present, extensive studies have been conducted on the classification of MI based on BCI. Table 3 shows the comparison of different feature extraction methods with our proposed method. A valid comparison between different methods was ensured by sourcing all of their experimental data from BCI Competition IV Dataset I.

MI	Method	Accuracy (%)	Database
	Our method	87.35	BCI Competition IV Dataset I
Left hand/Foot	SBLFB [38]	79.85	BCI Competition IV Dataset I
	FBCSP [39]	76.85	BCI Competition IV Dataset I
	Our method	90.36	BCI Competition IV Dataset I
Left/Right hand	COL [40]	86.25	BCI Competition IV Dataset I
-	CSP [39]	75.60	BCI Competition IV Dataset I

Table 3. Studies with regard to the MI classification using different methods.

Notes: sparse Bayesian learning is abbreviated as SBLFB, filter bank common spatial pattern is abbreviated as FBLFB, and channel optimization based on l_1 -norm is abbreviated as COL.

Table 3 shows that our method is superior to that of references [38–40] for different types of MI. Hence, our method has a feasible application and strong robustness, and we can conclude that the bilevel network is an effective feature extraction method for the BCI research.

4. Discussion and Conclusions

Significant breakthrough has been achieved when brain network theory was applied to neuroscience. However, the construction of the brain network often causes nodes to be incompletely

connected. In this study, we applied the MST to the classification of MI EEG signals. MST is the only subnetwork in the weighted network and effectively solves the problem of network disconnection (Figures 5 and 6). In addition, MST can determine the original network index structure and summary information. Regional network is constructed on the basis of MST with the functional cortex region corresponding to different movements as the center. The most prominent role of the regional network is to include more functional connectivity information in the cerebral cortex, highlight the network characteristics of each movement, and compensate the shortcomings of the MST, which does not have small-world characteristics. Figures 9 and 10 show that the regional network is more directional than the MST. In summary, we combine MST with the regional network into a bilevel network. This bilevel network alone. Lastly, we combined the fundamental features of the MST with the directional features of the regional network to characterize the brain network and apply them to the different MI classifications. The results on real world data sets suggest that our method achieves better classification performance than using a single layer network and other algorithms regardless of left/right hand or left hand/foot MI.

This study further reveals that our assumption of introducing MST into MIs classification may be promising for the BCI research. The classification performance of the regional network is better than that of MST because the regional network contains more information of the cerebral cortex. Therefore, our proposed regional network under different movements can reflect the changes in network location and connection well. This study also demonstrates the superiority of the bilevel network in BCI Competition IV Dataset I.

However, our method continues to have some drawbacks and deserves further study. (1) Given that MST substantially simplifies the network structure, it may help find a generation model of the brain network. This research only chose the diameter and average eccentricity to characterize the fundamental features of MST. Therefore, our challenge in the future is to find a more comprehensive measurement framework for MST. (2) Our method effectively solves the two-class problem. Thus, extending our method to a multi-class framework may achieve generalization in practical applications. (3) Although our proposed feature extraction method has achieved promising results, it has strong a priori both in terms of frequency bands and EEG electrodes used to perform the classification. Given that volume conduction effects typically affect scalp EEG activity, further research is needed in source space. (4) Computational complexity is a concern for online BCI. Although the bilevel network achieves better classification performance than using a single layer network, the calculation time is higher than independent features. Hence, a topic worth considering is to reduce the computational complexity of our method, which will hopefully promote the applications of our method in BCI systems.

Author Contributions: All authors contributed to this paper. Conceptualization and methodology, Z.L. and X.L.; Software, X.L.; Validation, Z.L.; Formal analysis and investigation, Z.L., X.L. and X.X.; Data curation, X.L.; Writing—Original draft preparation, X.L.; Writing—Review and editing, Z.L. and X.X.; Funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: The work is supported by National Natural Science Foundation of China (Nos. 61671197 and 61971169).

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

- 1. Belwafi, K.; Gannouni, S.; Aboalsamh, H.; Mathkour, H.; Belghith, A. A dynamic and self-adaptive classification algorithm for motor imagery EEG signals. *J. Neurosci. Methods* **2019**, 327. [CrossRef]
- 2. Allen, E.A.; Damaraju, E.; Eichele, T.; Wu, L.; Calhoun, V.D. EEG Signatures of Dynamic Functional Network Connectivity States. *Brain Topogr.* **2018**, *31*, 101–116. [CrossRef] [PubMed]
- Zhou, P.Y.; Chan, K.C.C. Fuzzy Feature Extraction for Multichannel EEG Classification. *IEEE Trans. Cogn. Dev. Syst.* 2018, 10, 267–279. [CrossRef]

- Hamada, M.; Zaidan, B.B.; Zaidan, A.A. A Systematic Review for Human EEG Brain Signals Based Emotion Classification, Feature Extraction, Brain Condition, Group Comparison. J. Med. Syst. 2018, 42, 25. [CrossRef] [PubMed]
- Wu, L.Y.; Wang, T.; Wang, Q.; Zhu, Q.; Chen, J.H. EEG Signal Processing Based on Multivariate Empirical Mode Decomposition and Common Spatial Pattern Hybrid Algorithm. *Int. J. Pattern Recognit. Artif. Intell.* 2019, 33, 16. [CrossRef]
- Durongbhan, P.; Zhao, Y.F.; Chen, L.Y.; Zis, P.; De Marco, M.; Unwin, Z.C.; Sarrigiannis, P.G. A Dementia Classification Framework Using Frequency and Time-Frequency Features Based on EEG Signals. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2019, 27, 826–835. [CrossRef]
- 7. Bassett, D.S.; Khambhati, A.N. A network engineering perspective on probing and perturbing cognition with neurofeedback. *Ann. N. Y. Acad. Sci.* **2017**, *1396*, 126–143. [CrossRef]
- Buch, V.P.; Richardson, A.G.; Brandon, C.; Stiso, J.; Khattak, M.N.; Bassett, D.S.; Lucas, T.H. Network Brain-Computer Interface (nBCI): An Alternative Approach for Cognitive Prosthetics. *Front. Neurosci.* 2018, 12, 790. [CrossRef]
- Guo, H.; Yan, P.P.; Cheng, C.; Li, Y.; Chen, J.J.; Xu, Y.; Xiang, J. fMRI classification method with multiple feature fusion based on minimum spanning tree analysis. *Psychiatry Res. Neuroimaging* 2018, 277, 14–27. [CrossRef]
- 10. Sporns, O.; Tononi, G.; Kotter, R. The human connectome: A structural description of the human brain. *PLoS Comput. Biol.* **2005**, *1*, 245–251. [CrossRef]
- 11. Zhang, X.; Lei, X.; Wu, T.; Jiang, T.Z. A review of EEG and MEG for brainnetome research. *Cogn. Neurodyn.* **2014**, *8*, 87–98. [CrossRef]
- 12. Stanley, M.L.; Simpson, S.L.; Dagenbach, D.; Lyday, R.G.; Burdette, J.H.; Laurienti, P.J. Changes in Brain Network Efficiency and Working Memory Performance in Aging. *PLoS ONE* **2015**, *10*, e0123950. [CrossRef]
- Kong, W.; Guo, X.; Zhao, X.; Wei, D.; Hu, S.; Dai, G. Spectral analysis of brain function network for the classification of motor imagery tasks. In Proceedings of the International Conference on Biomedical Engineering & Informatics, Datong, China, 15–17 October 2016; IEEE: Piscataway, NJ, USA, 2016. [CrossRef]
- 14. Van Straaten, E.C.W.; Stam, C.J. Structure out of chaos: Functional brain network analysis with EEG, MEG, and functional MRI. *Eur. Neuropsychopharmacol.* **2013**, *23*, 7–18. [CrossRef]
- Wang, B.; Miao, L.W.; Niu, Y.; Cao, R.; Li, D.D.; Yan, P.F.; Xiang, J. Abnormal Functional Brain Networks in Mild Cognitive Impairment and Alzheimer's Disease: A Minimum Spanning Tree Analysis. *J. Alzheimers Dis.* 2018, 65, 1093–1107. [CrossRef]
- 16. Tewarie, P.; Van Dellen, E.; Hillebrand, A.; Stam, C.J. The minimum spanning tree: An unbiased method for brain network analysis. *Neuroimage* **2015**, *104*, 177–188. [CrossRef]
- 17. Kruskal, J.B. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Am. Math. Soc.* **1956**, *7*, 48. [CrossRef]
- Li, X.W.; Jing, Z.A.; Bin, H.; Zhu, J.; Zhong, N.; Li, M.; Majoe, D. A Resting-State Brain Functional Network Study in MDD Based on Minimum Spanning Tree Analysis and the Hierarchical Clustering. *Complexity* 2017, 2017, 1–11. [CrossRef]
- Tewarie, P.; Hillebrand, A.; Schoonheim, M.M.; Van Dijk, B.W.; Geurts, J.J.G.; Barkhof, F.; Stam, C.J. Functional brain network analysis using minimum spanning trees in Multiple Sclerosis: An MEG source-space study. *Neuroimage* 2014, *88*, 308–318. [CrossRef]
- 20. Lee, U.V.; Kim, S.; Jung, K.Y. Classification of epilepsy types through global network analysis of scalp electroencephalograms. *Phys. Rev. E* 2006, *73*, 9. [CrossRef]
- 21. Dubbelink, K.T.E.O.; Stoffers, D.; Deijen, J.B.; Twisk, J.W.R.; Stam, C.J.; Hillebrand, A.; Berendse, H.W. Resting-state functional connectivity as a marker of disease progression in Parkinson's disease: A longitudinal MEG study. *Neuroimage Clin.* **2013**, *2*, 612–619. [CrossRef]
- 22. Demuru, M.; Fara, F.; Fraschini, M. Brain network analysis of EEG functional connectivity during imagery hand movements. *J. Integr. Neurosci.* 2013, *12*, 441–447. [CrossRef]
- 23. Wang, F.W.; Zhang, X.L.; Fu, R.R.; Sun, G.B. EEG characteristic analysis of coach bus drivers based on brain connectivity as revealed via a graph theoretical network. *RSC Adv.* **2018**, *8*, 29745–29755. [CrossRef]
- Stam, C.J.; Tewarie, P.; Van Dellen, E.; Van Straaten, E.C.W.; Hillebrand, A.; Van Mieghem, P. The trees and the forest: Characterization of complex brain networks with minimum spanning trees. *Int. J. Psychophysiol.* 2014, 92, 129–138. [CrossRef]

- 25. Yi, W.B.; Qiu, S.; Qi, H.Z.; Zhang, L.X.; Wan, B.K.; Ming, D. EEG feature comparison and classification of simple and compound limb motor imagery. *J. Neuroeng. Rehabil.* **2013**, *10*, 12. [CrossRef]
- 26. Nagamori, S.; Tanaka, H. Analysis Method for ERD in Mu-rhythm Detection in Motor Imagery Brain-Computer Interface. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics, Budapest, Hungary, 9–12 October 2016; IEEE: New York, NY, USA, 2016; pp. 867–870.
- 27. Sporns, O. Structure and function of complex brain networks. *Dialogues Clin. Neurosci.* 2013, 15, 247–262.
- 28. Li, X.; Fan, H.; Wang, H.; Wang, L. Common spatial patterns combined with phase synchronization information for classification of EEG signals. *Biomed. Signal Process. Control* **2019**, *52*, 248–256. [CrossRef]
- 29. Corsi, M.C.; Chavez, M.; Schwartz, D.; Hugueville, L.; Khambhati, A.N.; Bassett, D.S.; Fallani, F.D. Integrating EEG and MEG Signals to Improve Motor Imagery Classification in Brain-Computer Interface. *Int. J. Neural Syst.* **2019**, *29*, 12. [CrossRef]
- Leeb, R.; Sagha, H.; Chavarriaga, R.; Millan, J.D. Multimodal Fusion of Muscle and Brain Signals for a Hybrid-BCI. In Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Buenos Aire, Argentina, 30 August–4 September 2010; IEEE: New York, NY, USA, 2010; pp. 4343–4346.
- 31. Muller-Putz, G.; Leeb, R.; Tangermann, M.; Hohne, J.; Kubler, A.; Cincotti, F.; Millan, J.D. Towards Noninvasive Hybrid Brain-Computer Interfaces: Framework, Practice, Clinical Application, and Beyond. *Proc. IEEE* **2015**, *103*, 926–943. [CrossRef]
- 32. BCI Competition IV. Available online: http://www.bbci.de/competition/ (accessed on 18 January 2020).
- 33. Pineda, J.A. The functional significance of mu rhythms: Translating "seeing" and "hearing" into "doing". *Brain Res. Rev.* 2005, *50*, 57–68. [CrossRef]
- Lapenta, O.M.; Boggio, P.S. Motor network activation during human action observation and imagery: Mu rhythm EEG evidence on typical and atypical neurodevelopment. *Res. Autism Spectr. Disord.* 2014, *8*, 759–766. [CrossRef]
- 35. Burges, C.J.C. A tutorial on support vector machines for pattern recognition. *Datamining Knowl. Discov.* **1998**, 2, 121–167. [CrossRef]
- 36. Ehrsson, H.H.; Fagergren, A.; Jonsson, T.; Westling, G.; Johansson, R.S.; Forssberg, H. Cortical activity in precision- versus power-grip tasks: An fMRI study. *J. Neurophysiol.* **2000**, *83*, 528–536. [CrossRef]
- Zhang, L.H.; Wang, B.Z.; Hong, J.; Li, T. The Analysis of Decoding Parameter Selection of Hand Movements Based on Brain Function Network. In Proceedings of the 2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence, Goyang, Korea, 28–30 October 2015; IEEE: New York, NY, USA, 2015; pp. 587–592.
- 38. Kumar, S.; Sharma, A.; Tsunoda, T. An improved discriminative filter bank selection approach for motor imagery EEG signal classification using mutual information. *BMC Bioinform.* **2017**, *18*, 13. [CrossRef]
- 39. Zhang, H.H.; Chin, Z.Y.; Ang, K.K.; Guan, C.T.; Wang, C.C. Optimum Spatio-Spectral Filtering Network for Brain-Computer Interface. *IEEE Trans. Neural Netw.* **2011**, *22*, 52–63. [CrossRef]
- 40. Zhao, Y.W.; Han, J.Q.; Chen, Y.S.; Sun, H.J.; Chen, J.Y.; Ke, A.; Wang, C.Y. Improving Generalization Based on l(1)-Norm Regularization for EEG-Based Motor Imagery Classification. *Front. Neurosci.* **2018**, *12*, 16. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).