

Research Article

An Improved Refined Composite Multivariate Multiscale Fuzzy Entropy Method for MI-EEG Feature Extraction

Mingai Li ^{1,2}, Ruotu Wang,¹ Jinfu Yang,^{1,2} and Lijuan Duan¹

¹Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

²Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing 100124, China

Correspondence should be addressed to Mingai Li; limingai@bjut.edu.cn

Received 15 January 2019; Accepted 25 February 2019; Published 28 March 2019

Academic Editor: Amparo Alonso-Betanzos

Copyright © 2019 Mingai Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Feature extraction of motor imagery electroencephalogram (MI-EEG) has shown good application prospects in the field of medical health. Also, multivariate entropy-based feature extraction methods have been gradually applied to analyze complex multichannel biomedical signals, such as EEG and electromyography. Compared with traditional multivariate entropies, refined composite multivariate multiscale fuzzy entropy (RCmvMFE) overcomes the defect of unstable entropy values caused by the scale factor increase and is beneficial towards obtaining richer feature information. However, the coarse-grained process of RCmvMFE is mean filtered, which weakens Gaussian noise and is powerless against random impulse noise interference. This yields poor quality feature information and low accuracy classification. In this paper, RCmvMFE is improved (IRCmvMFE) by using composite filters in the coarse-grained procedure to enhance filter performance. Median filters are employed to remove the impulse noise interference from multichannel MI-EEG signals, and these filtered MI-EEGs are further smoothed by the mean filters. The multiscale IRCmvMFEs are calculated for all channels of composite filtered MI-EEGs, forming a feature vector, and a support vector machine is used for pattern classification. Based on two public datasets with different motor imagery tasks, the recognition results of 10×10 -fold cross-validation achieved 99.43% and 99.86%, respectively, and the statistical analysis of experimental results was completed, showing the effectiveness of IRCmvMFE, as well. The proposed IRCmvMFE-based feature extraction method is superior compared to entropy-based and traditional methods.

1. Introduction

Brain-computer interface (BCI) is a new type of human-computer interaction technology that enables the brain to control external devices [1, 2]. Motor imagery electroencephalogram- (MI-EEG-) based BCI has great prospects in the field of rehabilitation medical engineering. One of the key technologies of BCI is the ability to effectively extract features from complex multichannel MI-EEG signals.

Previous studies focus on time-frequency methods in MI-EEG feature extraction, including wavelet transform (WT) [3], discrete WT (DWT) [4], Hilbert-Huang transform (HHT) [3], dual-tree complex WT (DTCWT) [5], empirical mode decomposition (EMD) [6], and common spatial pattern- (CSP-) based methods, including CSP, filter bank CSP (FBCSP), discriminant FBCSP (DFBCSP), sparse

FBCSP (SFBCSP), and spectrally weighted CSP (SWCSP) [7–9]. With the development of nonlinear dynamics, entropy-based methods have been widely utilized in the analysis of biomedical signals. Specifically, the most prevalent methods are approximate entropy (ApEn) and sample entropy (SampEn) because of the power these methods have to quantify the complexity of a time series [10–12]. Nevertheless, sensitivity to selected parameters will lead to entropy mutation. To solve this weakness, fuzzy entropy (FE) was proposed for EEG analysis, where fuzzy membership functions replace Heaviside functions [13–15]. Research shows that FE alleviates the problem of entropy mutation; however, these methods analyze at a single scale, which loses useful information. Therefore, multiscale SampEn (MSE) [16–19], multiscale FE (MFE) [20, 21], and improved MFE (IMFE) [22] were put forward to explore

deeper information. Unfortunately, using larger scale factors for short time series may cause inaccurate MSEs and MFEs. To achieve more reliable results, composite MFE (CMFE) [23, 24] was developed as a rolling bearing fault detection method.

Although the above univariate methods have shown good performance, they are only suitable for single-channel recording analyses. They fail to measure multichannel data synchronously and ignore the dynamic characteristics across channels [25]. So, SampEn was extended to produce multivariate SampEn (mvSE) [26] and multivariate MSE (mvMSE) [26–29] to analyze multichannel signals more effectively. Considering the disadvantages of SampEn in mvSE and mvMSE, multivariate FE (mvFE) and multivariate MFE (mvMFE) [30, 31] were yielded by replacing SampEn with FE. Recently, as an improvement of mvMFE, a refined composite mvMFE (RCmvMFE) was proposed to analyze fault signals and biomedical signals [32, 33]. In RCmvMFE, the entropy stability is improved and the signals' length sensitivity is reduced. However, the coarse-grained process of RCmvMFE is a mean filter that smoothens signals but does not eliminate random impulse noise interference. It is inevitable to produce high-amplitude electrooculogram and electromyography interference during the acquisition of MI-EEG. This is not conducive to extracting valid feature information from multichannel MI-EEG signals. In this paper, improved RCmvMFE (IRCmvMFE) is developed by combining median [34] and mean filters in the coarse-grained process to further improve filter effect, i.e., first the median filter is applied to each channel to remove pulse interference, and then the mean filter is used for further smoothing. Subsequently, IRCmvMFE is proposed to extract features from multichannel MI-EEG signals. The experimental research shows the effectiveness of IRCmvMFE.

The rest of the paper is described as follows: Section 2 introduces the process of extracting MI-EEG features using IRCmvMFE, Section 3 describes the experiments performed, Section 4 discusses the results, and Section 5 provides the conclusions.

2. Feature Extraction with IRCmvMFE

By combining median filters and mean filters in coarse-grained processes, RCmvMFE is improved to produce IRCmvMFE, which is applied to extract features of MI-EEG. The main steps are as follows: preprocessing, optimal channel selecting, performing multivariate coarse-grained analysis of preprocessed MI-EEG data, calculating IRCmvMFE, and constructing a feature vector. The support vector machine (SVM) was used to classify the feature vector. The block diagram of the proposed method is displayed in Figure 1.

2.1. Preprocessing MI-EEG Signals. For two-class motor imagery tasks, assume that $X_{T,C_i}^0 = [X_{T,C_i}^0(1), X_{T,C_i}^0(2), \dots, X_{T,C_i}^0(e)]^T$ represents the i th channel MI-EEG sequence of the T th task, where $T \in \{1, 2\}$, $i = 1, 2, \dots, p$; e and p represent the sample points and the number of total channels, respectively. X_{T,C_i}^0 is bandpass filtered to the frequency band associated with the tasks and is expressed as $X_{T,C_i}^1 =$

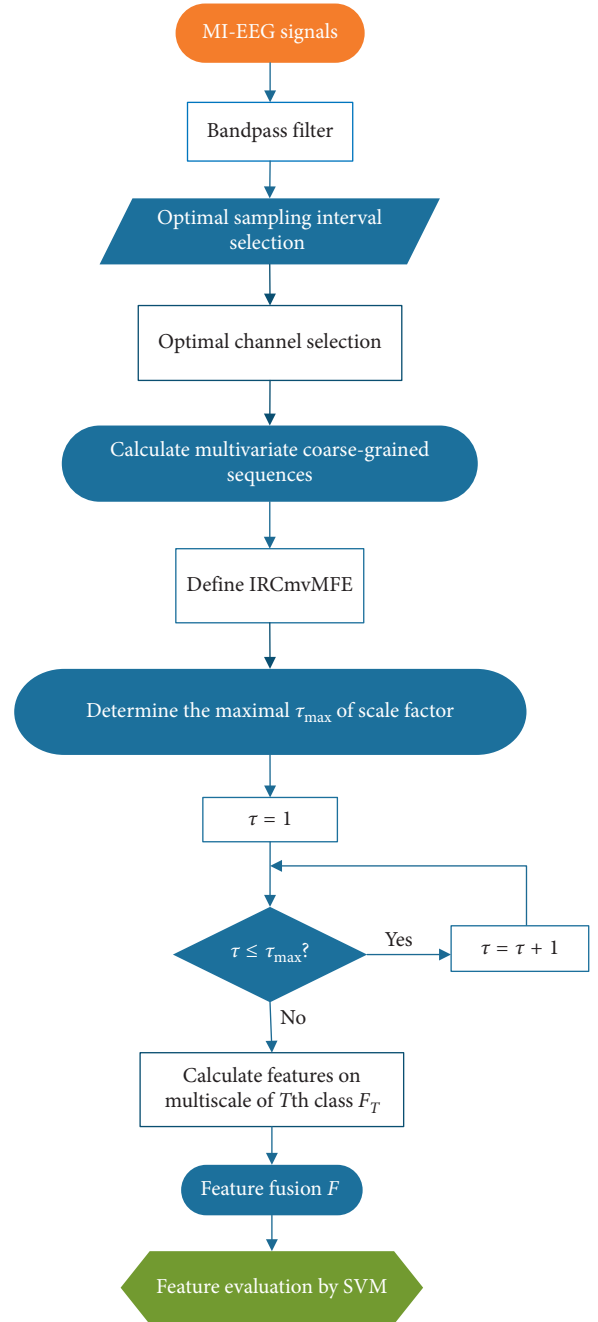


FIGURE 1: Block diagram of the proposed method.

$[X_{T,C_i}^1(1), X_{T,C_i}^1(2), \dots, X_{T,C_i}^1(e)]^T$. The motor imagery time period $[a, b]$ is taken as the optimal sampling interval, and MI-EEG signals in the segment are summarized as $X_{T,C_i}^2 = [X_{T,C_i}^2(1), \dots, X_{T,C_i}^2(N)]^T$, where $N = b - a + 1$ represents the sampled MI-EEG points within the optimal sampling interval.

2.2. Channel Selection. When the brain is engaged in motor imagery, only parts of channels are activated in the form of the power spectrum. Extracting the features of all channels not only increases the computational complexity but also increases the feature information redundancy and reduces the classification accuracy [35]. Therefore, the choice of

optimal channels is important. In this paper, the Fisher score of the average power spectrum of X_{T,C_i}^2 is calculated to select channels according to the following equation:

$$F(i) = \frac{|P_1(i) - P_2(i)|^2}{\text{var}(P_1(i)) + \text{var}(P_2(i))}, \quad (1)$$

where $P_1(i)$ and $P_2(i)$ represent the average power spectrum on the i th channel of class 1 and class 2 motor imagery tasks, respectively. $\text{var}()$ is the variance and $F(i)$ represents the i th channel Fisher score. The larger the $F(i)$, the greater contribution of the i th channel. The signals of p' channels with the top $F(i)$ are selected for subsequent research. X_{T,C_i}^2 is rewritten as $X_{T,C_i} = [X_{T,C_i}(1), \dots, X_{T,C_i}(N)]^T$, where $i = 1, 2, \dots, p'$, in which p' stands for the number of selected channels.

2.3. Coarse-Graining of IRCmvMFE

Step 1. In the coarse-grained process of IRCmvMFE, the median filter is first performed on X_{T,C_i} . Supposing the filter size is $j = 2k$ or $j = 2k + 1$, the data in the window would be sorted in ascending order with the filter output being

$$y_{T,C_i}(s) = \text{med}(X_{T,C_i}(s))$$

$$= \begin{cases} X_{T,C_i}(k+1), & j = 2k+1, \\ \frac{1}{2}[X_{T,C_i}(k) + X_{T,C_i}(k+1)], & j = 2k, \end{cases} \quad (2)$$

where $s \in \{1, \dots, N\}$ and $X_{T,C_i}(k)$ means the k th maximum value in the window.

Step 2. For the scale factor τ , the k th coarse-grained sequence on C_i channel of T th class task is

$$y_{T,C_i}^{k,\tau}(j) = \frac{1}{\tau} \sum_{s=(j-1)\tau+k}^{j*\tau+k-1} y_{T,C_i}(s), \quad 1 \leq j \leq N', 1 \leq k \leq \tau, \quad (3)$$

where $N' = \text{int}[N/\tau]$ represents the sample points of the coarse-grained sequence. Therefore, τ multivariate coarse-grained sequences are obtained and described as $Y_{T,C_i}^{1,\tau} = [y_{T,C_i}^{1,\tau}(1), y_{T,C_i}^{1,\tau}(2), \dots, y_{T,C_i}^{1,\tau}(N')]$, $Y_{T,C_i}^{2,\tau} = [y_{T,C_i}^{2,\tau}(1), y_{T,C_i}^{2,\tau}(2), \dots, y_{T,C_i}^{2,\tau}(N')]$, \dots , $Y_{T,C_i}^{\tau,\tau} = [y_{T,C_i}^{\tau,\tau}(1), y_{T,C_i}^{\tau,\tau}(2), \dots, y_{T,C_i}^{\tau,\tau}(N')]$.

2.4. IRCmvMFE Calculation

Step 1. The multivariate coarse-grained sequence $Y_{T,C_i}^{k,\tau}$ is executed for multivariate embedded reconstruction, with the multivariate composite delay vectors $Z_{T,m}^{k,\tau}(i)$ calculated as

$$Z_{T,m}^{k,\tau}(i) = \begin{bmatrix} y_{T,C_1}^{k,\tau}(i), \dots, y_{T,C_1}^{k,\tau}(i + (m_1 - 1)\lambda_1), \\ y_{T,C_2}^{k,\tau}(i), \dots, y_{T,C_2}^{k,\tau}(i + (m_2 - 1)\lambda_2), \dots, \\ y_{T,C_{p'}}^{k,\tau}(i), \dots, y_{T,C_{p'}}^{k,\tau}(i + (m_{p'} - 1)\lambda_{p'}) \end{bmatrix}, \quad (4)$$

where $i \in [1, N' - n]$ and $M = [m_1, m_2, \dots, m_{p'}]$ and $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_{p'}]$ are the embedding dimension vector and time delay vector, respectively. Additionally, $m = \sum_{i=1}^{p'} m_i$ and $n = \max\{M\} \times \max\{\lambda\}$.

Step 2. The distance of any two multivariate composite delay vectors $Z_{T,m}^{k,\tau}(i)$ and $Z_{T,m}^{k,\tau}(j)$ is computed in the following equation:

$$d[Z_{T,m}^{k,\tau}(i), Z_{T,m}^{k,\tau}(j)] = d_{T,m}^{ij,k,\tau} = \max \left\{ \left| y_{T,C_g}^{k,\tau}(i+l-1), \right. \right. \\ \left. \left. y_{T,C_h}^{k,\tau}(j+l-1) \right|, l = 1, 2, \dots, m \right\}, \quad (5)$$

where $i, j \in [1, N' - n]$, $i \neq j$, and $g, h \in [1, p']$.

Step 3. Given a threshold r , suppose the fuzzy membership function is $\mu(x, r) = e^{-d^2/r}$, the similarity $D_{T,m}^{ij,k,\tau}$ between $Z_{T,m}^{k,\tau}(i)$ and $Z_{T,m}^{k,\tau}(j)$ is

$$D_{T,m}^{ij,k,\tau} = \mu(d_{T,m}^{ij,k,\tau}, r) = \exp\left(\frac{-(d_{T,m}^{ij,k,\tau})^2}{r}\right). \quad (6)$$

Step 4. The average membership grade $\phi_{T,m}^{k,\tau}(r)$ can be obtained using the following equation:

$$\phi_{T,m}^{k,\tau}(r) = \frac{1}{N' - 1} \sum_{i=1}^{N'-n} \frac{\sum_{j=1, j \neq i}^{N'-1} D_{T,m}^{ij,k,\tau}}{N' - n - 1}. \quad (7)$$

Step 5. Repeat the above steps, extend the dimension of the multivariate composite delay vector from m to $m + 1$ and derive $\phi_{T,m+1}^{k,\tau}$. For each $Z_{T,m}^{k,\tau}(i)$, we get $\tau \phi_{T,m}^{k,\tau}(r)$ and $\tau \phi_{T,m+1}^{k,\tau}(r)$. The average $\bar{\phi}_{T,m}^{k,\tau}(r)$ and $\bar{\phi}_{T,m+1}^{k,\tau}(r)$ are calculated. The definition of IRCmvMFE is as follows:

$$\text{IRCmvMFE}(X_{T,C_i}, M, \tau, n, r) = -\ln \left[\frac{\bar{\phi}_{T,m+1}^{\tau}(r)}{\bar{\phi}_{T,m}^{\tau}(r)} \right]. \quad (8)$$

The procedure for calculating IRCmvMFE is summarized in Algorithm 1.

2.5. Determination of a Maximum Scale Factor. As the number of scale factors increases, multivariate coarse-grained sequences become smoother. Scale factors that are too large omit useful information and reduce classification accuracy. Therefore, the impact on sequence smoothness and classification accuracy should be considered comprehensively to determine its maximum scale factor τ_{\max} .

2.6. Construction of a Feature Vector. For $\tau \in [1, \tau_{\max}]$, IRCmvMFE at τ scale in the T th class task, i.e., IRCmv_T^{τ} , is estimated and combined to form the feature vector F_T :

$$F_T = [\text{IRCmv}_T^1, \text{IRCmv}_T^2, \dots, \text{IRCmv}_T^{\tau_{\max}}] \in R^{\tau_{\max}}. \quad (9)$$

Input: Channels selected data X_{T,C_i}

(1) Coarse-graining of IRCmvMFE
Step 1. Calculate the output of median filter on $X_{T,C_i}; y_{T,C_i}$
Step 2. Calculate the output of mean filter on $y_{T,C_i}; y_{T,C_i}^{k,\tau}$

(2) IRCmvMFE calculation
Step 1. Set embedding dimension m and scale factor τ , calculate the multivariate reconstruction vectors: $Z_{T,m}^{k,\tau}$
Step 2. For $k=1$ to τ
 Calculate the distance and the similarity of $Z_{T,m}^{k,\tau}(i)$ and $Z_{T,m}^{k,\tau}(j)$ using equations (5) to (6)
 Calculate $\phi_{T,m}^{k,\tau}(r)$ by equation (7)
 End
Step 3. Repeat the above steps, extend the dimension from m to $m+1$, and calculate $\phi_{T,m+1}^{k,\tau}$
Step 4. Set $\vec{\phi}_{T,m}^{k,\tau}(r) = \sum_{k=1}^{\tau} \phi_{T,m}^{k,\tau}(r)/\tau$ and $\vec{\phi}_{T,m+1}^{k,\tau}(r) = \sum_{k=1}^{\tau} \phi_{T,m+1}^{k,\tau}(r)/\tau$, and calculate IRCmvMFE

Output: IRCmvMFE at τ scale

ALGORITHM 1: The procedure for calculating IRCmvMFE.

The feature vectors of the two tasks are fused in parallel to obtain the feature vector of MI-EEG:

$$F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \in R^{2 \times \tau_{\max}}. \quad (10)$$

3. Experimental Research

3.1. Data Description and Preprocessing. MI-EEG data were obtained from dataset III in the BCI Competition II [36] and dataset IVa in the BCI Competition III [37]. MI-EEG signals on channels C3, Cz, and C4 were recorded in dataset III of BCI Competition II, where the data were from a healthy subject who imagined left-right hand movement. Left- and right-motor imagery tasks were each performed 140 times for a total of 280 experimental trials. The signals were sampled at 128 Hz and filtered to 0.5–30 Hz. The MI-EEG collection timing scheme is shown in Figure 2(a). The subject was at rest for the first 2 s, and the corresponding motor imagery task was completed according to the screen prompts from 3 s to 9 s. To better distinguish the two-class tasks, this paper used the sampling interval [451, 900].

The dataset IVa of BCI Competition III recorded the MI-EEG signals of five healthy subjects using 118 channels during right-hand (RH) and right-foot (RF) motor imagery tasks. The original sampling rate was 1000 Hz, but we downsampled these data to 100 Hz. The subjects performed the corresponding imaginary movement according to the prompts in the first 3.5 s and then rested for a random epoch between 1.75 s and 2.25 s. The timing scheme of MI-EEG collection during the right-hand-foot motor imagery task is shown in Figure 2(b). Each subject performed 280 trials, with 140 each of the RH and RF motor imagery tasks. In this paper, MI-EEG related to mu rhythm (8–13 Hz) and beta rhythm (14–32 Hz) related to motor imagery tasks were selected, i.e., the original MI-EEG signals were preprocessed by a bandpass filter of 8–32 Hz. The data between 0.5 s and 3.5 s were used for subsequent experimental research.

3.2. Channel Selection. Channel selection directly affects the quality of feature information and classification accuracy. It

is essential to select the optimal channels before extracting MI-EEG features. There was a close relationship between the signals on channels C3, Cz, and C4 in the left-right-hand motor imagery task, so the data of these three channels were used for feature extraction. When RH and RF motor imagery tasks were conducted in dataset IVa from BCI Competition III, the Fisher Score of each channel was calculated by equation (1). The scores of different subjects are shown in Figure 3.

For each subject, the score of each channel is different and for different subjects, scores from the same channel are different. Thus, the optimal channels for each subject are different due to individual differences. The channels with the top three Fisher scores can be used as the optimal channels. The detailed information is shown in Table 1.

3.3. Comparison of Coarse-Grained Sequences between IRCmvMFE and RCmvMFE Methods. To confirm the effectiveness of IRCmvMFE in extracting MI-EEG features, the coarse-grained processes of RCmvMFE and IRCmvMFE were compared. The relevant parameters were selected as follows: $m_k = 2, \lambda_k = 1, r = 0.2SD$, and $\tau = 10$, where SD represents the standard deviation of X_{T,C_i} . According to Table 1, the channel k with the highest Fisher score of each subject was selected. The experimental process was as follows: when a motor imagery task was performed, at τ scale, the first $j(0 \leq j \leq N - \tau)$ points of $X_{T,k}$ were removed in turn. The RCmvMFES of the remaining points were calculated separately, and they were composed of time series recorded as $RC_{T,k,i} = [RC_{T,k,i}(1), RC_{T,k,i}(2), \dots, RC_{T,k,i}(N - \tau + 1)]$, where $T \in \{1, 2\}$, $i \in [1, n_e]$, and n_e represents the number of experiments. The $RC_{T,k,i}$ of n_e experiments were superimposed and averaged to obtain the average time series $RC_{T,k}$. The average time series of IRCmvMFE was obtained the same way as $IRC_{T,k}$. When imaging left-right-hand motor imagery, training set data were used for analysis, i.e., $n_e = 70$. Similarly, n_e was selected as 140 when the RH and RF motor imagery were performed. The amplitude of the original MI-EEG signals and coarse-grained sequences of RCmvMFE and IRCmvMFE during left-right-hand motor imagery are

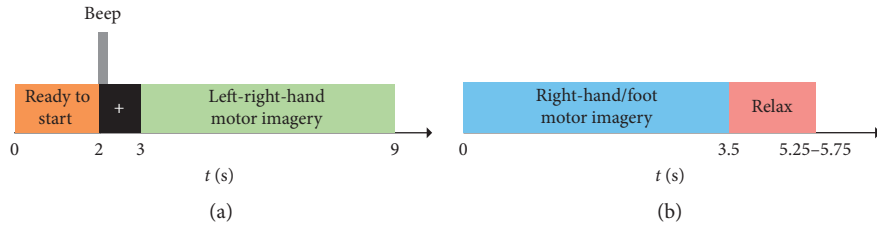


FIGURE 2: Timing scheme of MI-EEG collection. (a) Left-right-hand motor imagery task. (b) Right-hand/foot motor imagery task.

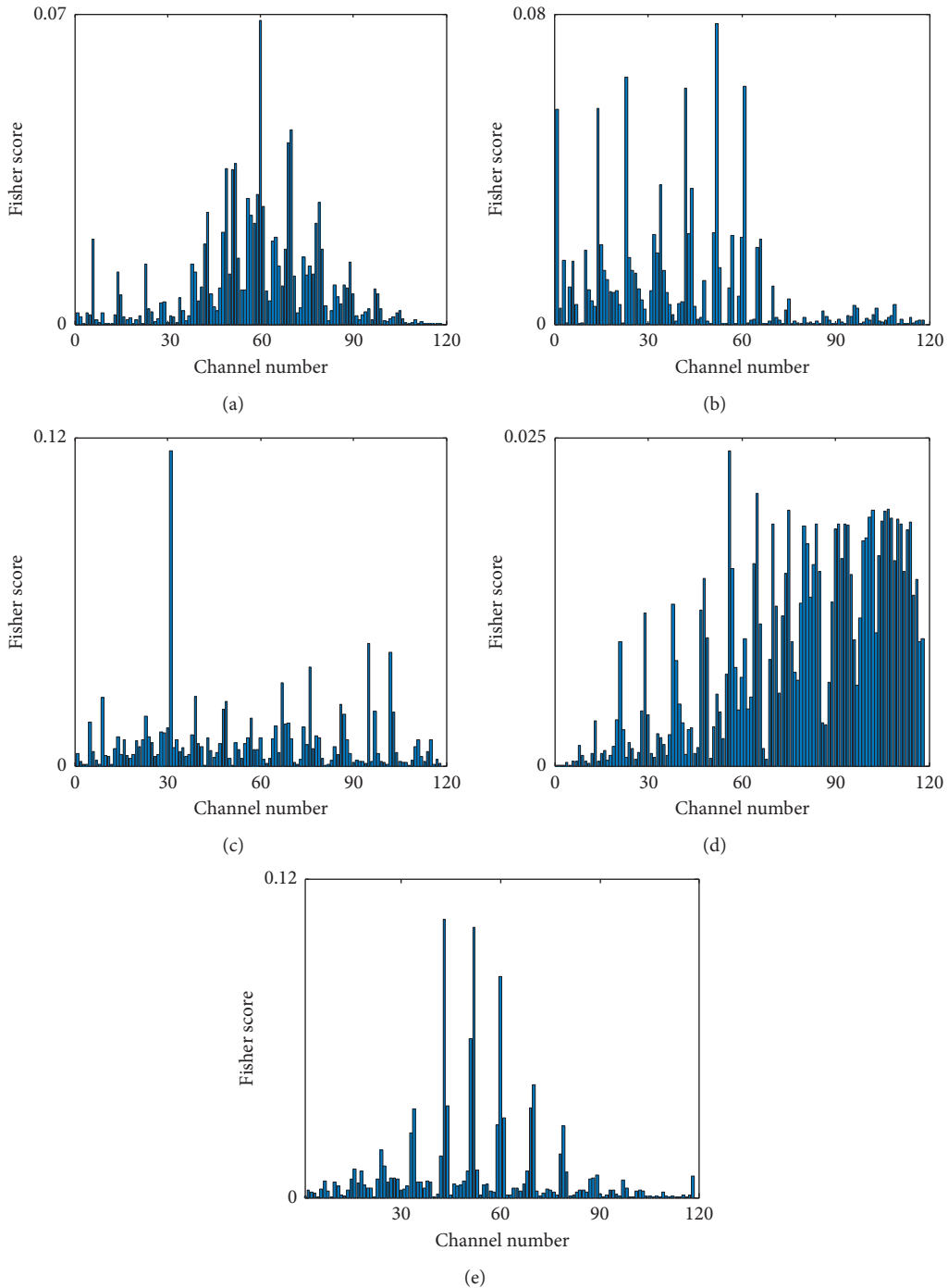


FIGURE 3: The Fisher scores of all channels with different subjects. (a) Subject "aa." (b) Subject "al." (c) Subject "av." (d) Subject "aw." (e) Subject "ay."

TABLE 1: The optimal channel combination.

Subjects	Optimal channel combination
“aa”	60, 70, 69
“al”	52, 23, 61
“av”	31, 95, 102
“aw”	56, 65, 75
“ay”	60, 52, 43

displayed in Figure 4. Similarly, the experimental results from imaging right-hand-foot movement are shown in Figure 5.

It can be seen from Figure 4 that the original MI-EEG signals had larger fluctuations, which was obviously improved after the coarse-grained process of both RCmvMFE and IRCmvMFE; and the smoothness of IRCmvMFE was better than RCmvMFE. In Figure 5, there are different intensity impulse noise interferences for different subjects. The coarse-grained sequences of both RH and RF motor imagery tasks using the RCmvMFE and IRCmvMFE of each subject changed with the fluctuations of the original MI-EEG but oscillated more smoothly. For subject “aw,” the impulse noise is not obvious, and the coarse-grained sequence of IRCmvMFE had larger fluctuations than that of RCmvMFE. But the intensity of impulse noise interference is higher for other subjects. Both $RC_{T,k}$ and $IRC_{T,k}$ showed better smoothness, and $IRC_{T,k}$ was superior to $RC_{T,k}$ for rapid MI-EEG changes. The reason is that the coarse-grained process of RCmvMFE is equivalent to a mean filter, which has the effect of low-pass filtering and smoothing and can remove some random interference. However, it is helpless against impulse noise caused by sudden factors such as eye-movements, blinks, and motion. In the coarse-grained IRCmvMFE, the median filter is assigned to remove the impulse noise interference, and then the filtered signals are smoothed by a mean filter.

3.4. Selection of Parameters in IRCmvMFE. The parameter selection will affect the estimate of IRCmvMFE. According to equation (8), the estimation of IRCmvMFE is not only related to the preprocessed MI-EEG but also involves selecting an embedding dimension vector $M = [m_1, m_2, \dots, m_{p'}]$, time delay vector $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_{p'}]$, threshold r , and scale factor τ . The selection of parameter M was similar to reference [32], i.e., $m_k = 2$. Parameter λ does not have any proven standards, so for simplicity, λ_k was selected as 1. The threshold r was determined as $r = 0.2SD$.

In addition, the selection of τ influenced the filter effect in the coarse-grained process of MI-EEG and affected the extracted features and the classification results in turn. The larger the τ , the larger the calculation and the better the recognition. In contrast, a smaller τ resulted in poor filter performance [34]. When $\tau \in [1, 75]$, the IRCmvMFEs with imaging left-right-hand movements were estimated and then classified by SVM. Gaussian kernel function was employed in this paper, and SVM optimized by grid search. When $\tau \in [1, 45]$, the same experiment was performed with right-hand-foot motor imagery tasks. The 10×10 -fold

cross-validation (CV) was used to eliminate the contingency in the feature extraction process of MI-EEG. The average classification accuracy of the 10×10 -fold CV is shown in Figure 6.

In Figure 6(a), the classification results gradually increased as the scale factor τ increased. When τ was from 55 to 75, the classification accuracy tended to be stable and close to 100%, and the highest recognition was obtained at 65 scale. Therefore, the maximum τ about left-right-hand motor imagery was selected as 65. In Figure 6(b), with the increased τ , the average recognition rate of each subject first increased and then later decreased. In this paper, the τ_{\max} values of subjects “aa”, “al”, “av”, “aw,” and “ay” during right hand-foot-motor imagery were chosen as 41, 37, 33, 38, and 39, respectively. And, τ_{\max} is related to the mathematical model of the coarse-grained process of IRCmvMFE. There is a significant difference in τ_{\max} during different types of two-class motor imagery tasks, while the difference between multiple subjects during the same type of tasks is not obvious.

3.5. Comparison of Multiple Entropy-Based Feature Extraction Methods. In this section, the comparative study of IRCmvMFE and various entropy-based feature extraction methods was conducted. To make the comparison process more objective, the same dataset was selected as reference [13, 22], i.e., dataset III from BCI Competition II, and SVM was used for classification. The classification result of IMFE was derived from [22], and the related parameters of other entropy-based methods were selected as references [13]. The average recognition results of 10×10 -fold CV and standard deviations are displayed in Figure 7.

In Figure 7, the classification result of MFE was higher than SampEn, FE, and MSE. Because the fuzzy membership function was used to enhance the stability of MFE, richer feature information from the multiscale was collected. At the same scale, the information of multiple coarse-grained sequences was integrated by CMFE, yielding a slightly better result. Based on the parameters’ independent optimization strategy, the preferred parameters were used by IMFE to extract features from the MI-EEG, and the recognition accuracy was further improved. Despite the results of mvSE, mvFE, and mvMSE being poor, mvMFE, RCmvMFE, and IRCmvMFE showed the advantages of multivariate entropy methods over traditional univariate entropies, both in terms of classification accuracy and standard deviation. This was mainly because these feature extraction methods evaluated the multivariate complexity of multichannel data and expressed the dynamic relationships and synchronizations across channels.

IRCmvMFE, RCmvMFE, and mvMFE methods displayed superiority on dataset III from BCI Competition II. To further illustrate the improvement of IRCmvMFE, a comparative study of these three methods was performed based on dataset IVa and using SVM for classification. The classification results with 10×10 -fold CV are shown in Table 2. For each subject, the recognition rates obtained by using RCmvMFE to extract features of MI-EEG were higher

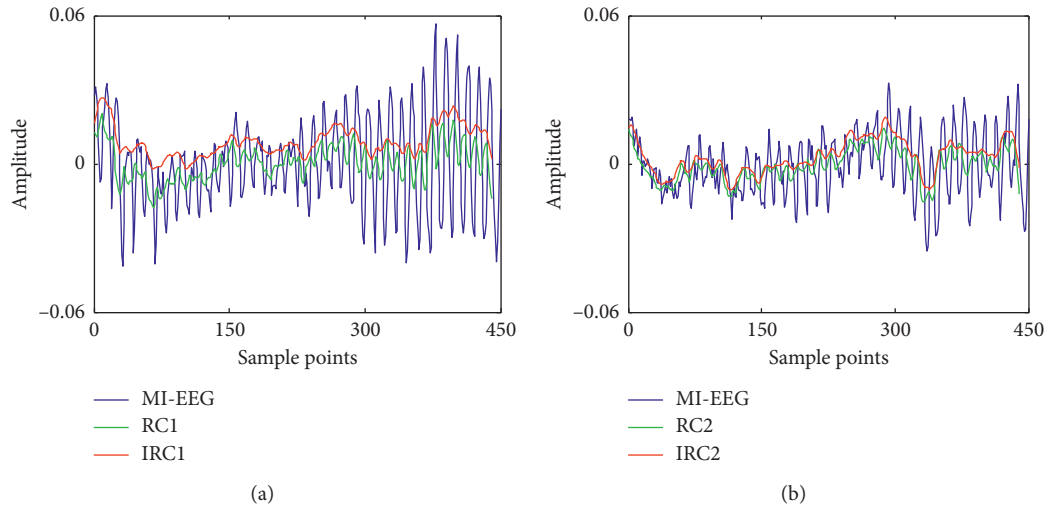


FIGURE 4: Comparison of MI-EEG and the coarse-grained sequences using RCmvMFE and IRCmvMFE during (a) left-hand motor imagery task and (b) right-hand motor imagery task.

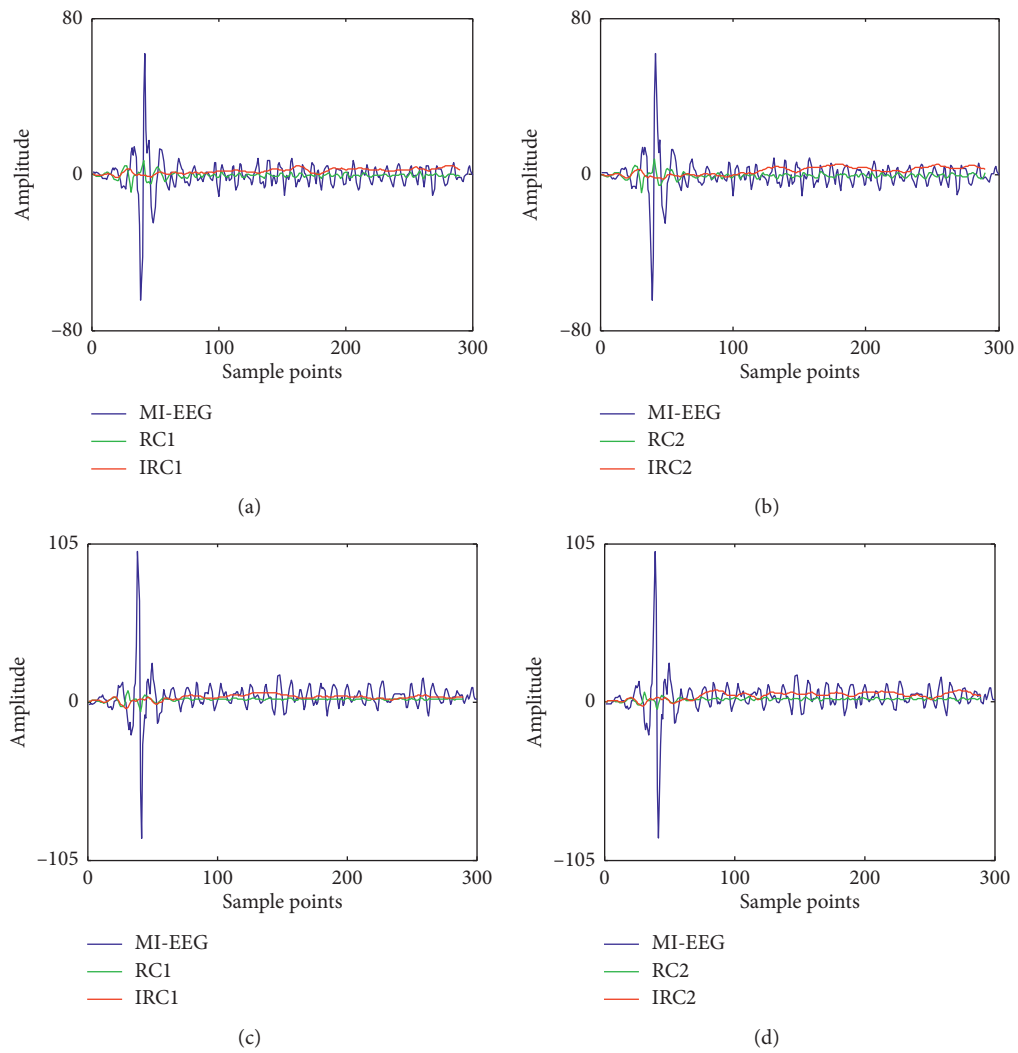


FIGURE 5: Continued.

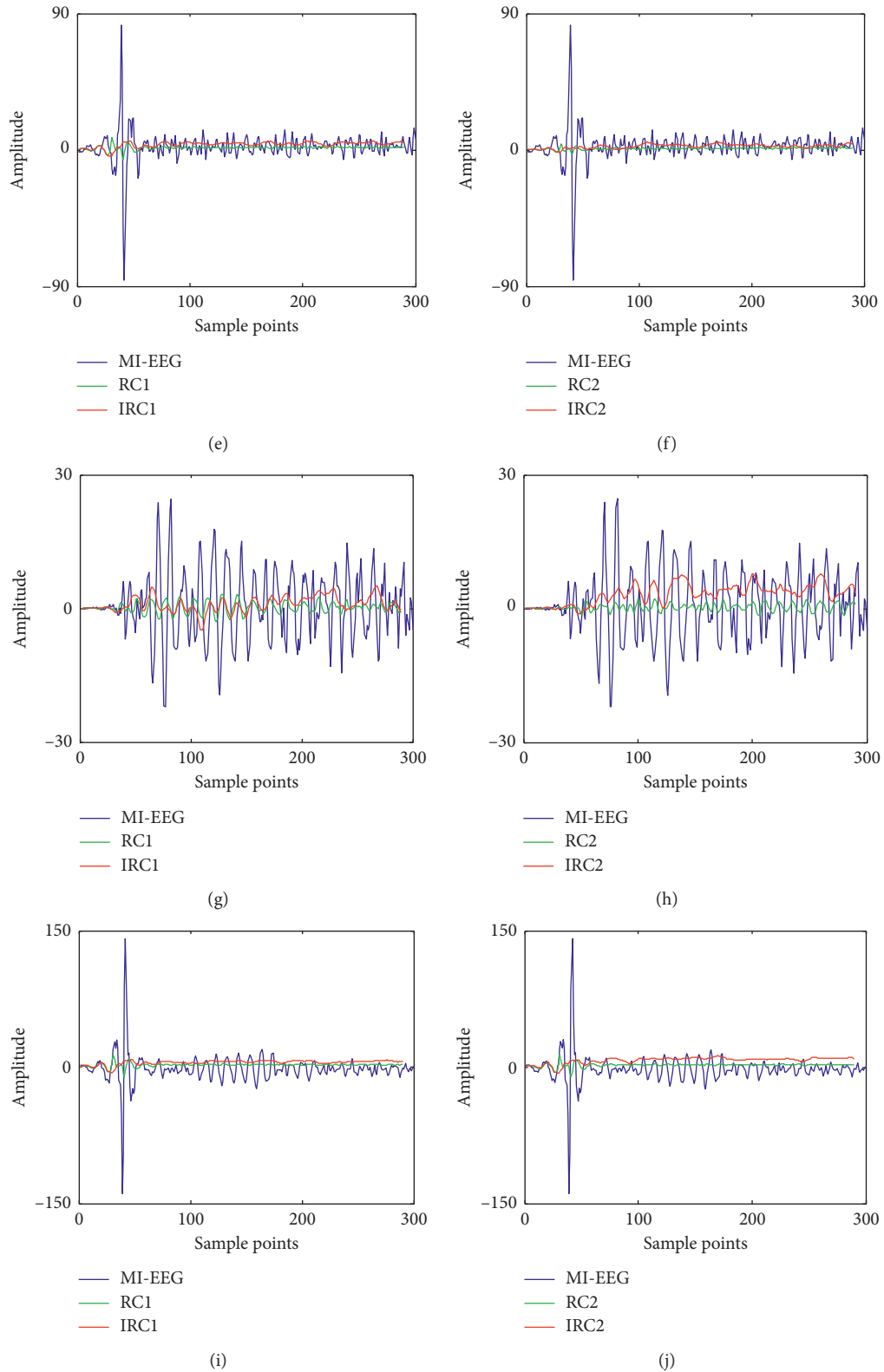


FIGURE 5: Comparison of MI-EEG signals and the coarse-grained sequences by RCmvmFE and IRCmvmFE during RH and RF motor imagery tasks. (a) Subject “aa” with RH motor imagery. (b) Subject “aa” with RF motor imagery. (c) Subject “al” with RH motor imagery. (d) Subject “al” with RF motor imagery. (e) Subject “av” with RH motor imagery. (f) Subject “av” with RF motor imagery. (g) Subject “aw” with RH motor imagery. (h) Subject “aw” with RF motor imagery. (i) Subject “ay” with RH motor imagery. (j) Subject “ay” with RF motor imagery.

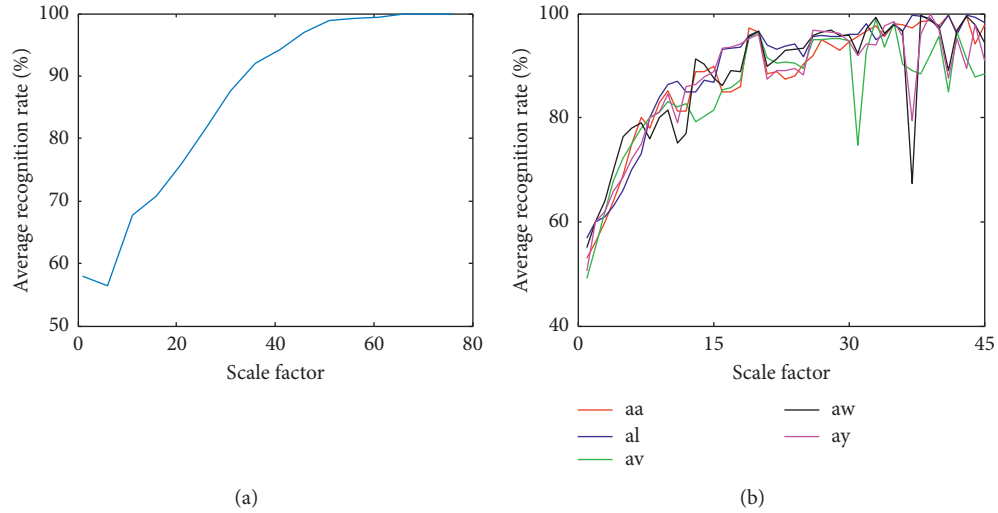


FIGURE 6: Effects of changing τ on the classification results during (a) left-right-hand motor imagery and (b) right-hand-foot motor imagery.

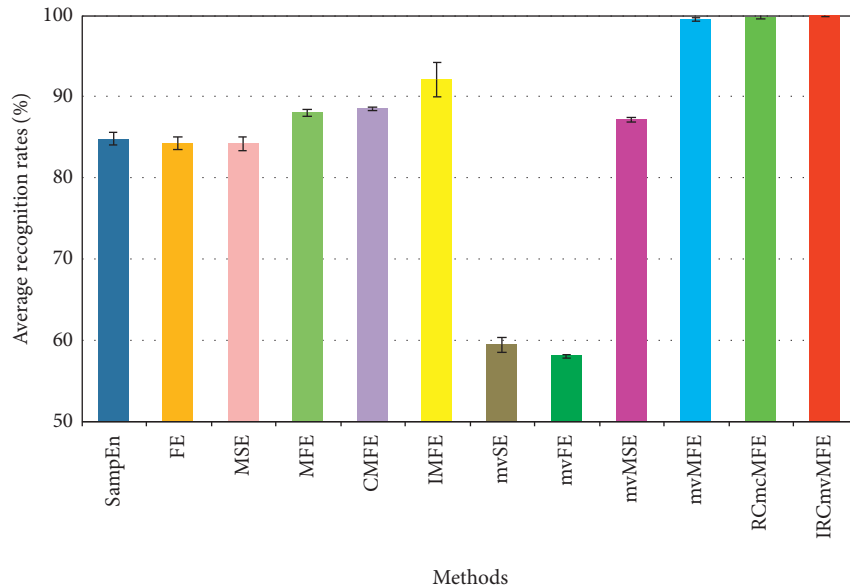


FIGURE 7: Comparison of multiple entropy-based feature extraction methods.

than those of mvMFE because the multivariate feature of RCmvMFE was considered at the same scale, and the defect of unstable entropy values, i.e., coarse-grained time series shortening with scale factor increases, was overcome. Moreover, a composite filter technique was applied in the coarse-grained process of IRCmvMFE to eliminate burst-like impulse noise and the Gaussian noise of the MI-EEG, which produced better quality information. For different subjects, IRCmvMFE achieved better recognition accuracy and a smaller standard deviation than RCmvMFE, illustrating the stability and superiority of IRCmvMFE. Further, according to Figure 5, the impulse noise interference was not obvious for subject “aw,” and the recognition result by IRCmvMFE was slightly better than RCmvMFE. However, there was greater impulse noise interference for most subjects (“aa,” “al,” “av,” “ay”), after using IRCmvMFE to enhance the filter effect, the recognition results were obviously improved.

3.6. Statistical Analysis. In this section, statistical analysis was performed to further describe the development of IRCmvMFE. The kappa coefficient, which was designed to measure the classification precision and the comparison of performance in multiclass tasks, was made fairer. This method is a common indicator for evaluating the performance of BCI systems [38, 39]. The calculation of κ coefficient was expressed as

$$\kappa = \frac{p_0 - p_e}{1 - p_e}, \quad (11)$$

where p_0 represents the classification accuracy and p_e means the probability of opportunity consistency. For a two-class task, if the number of samples across classes was equal, then the value of p_e was 0.5. Using equation (11), the mean kappa coefficients of IRCmvMFE, RCmvMFE, and mvMFE with 10×10 -fold CV were calculated. The results are shown in Table 3.

TABLE 2: Comparison of average recognition rates with 10×10 -fold CV (%) of multivariate entropy-based feature extraction methods.

Methods	Subjects					Average results (%)
	“aa”	“al”	“av”	“aw”	“ay”	
mvMFE	75.54 ± 1.70	86.00 ± 0.98	74.46 ± 1.54	84.07 ± 1.52	79.32 ± 1.24	79.88 ± 1.40
RCmvMFE	93.82 ± 1.48	98.46 ± 0.61	95.86 ± 1.01	98.79 ± 0.72	97.04 ± 0.79	96.97 ± 0.92
IRCmvMFE	99.39 ± 0.38	99.71 ± 0.15	98.61 ± 0.62	99.61 ± 0.26	99.82 ± 0.35	99.43 ± 0.35

TABLE 3: Kappa coefficients of multivariate entropy-based methods.

Methods	Subjects					Mean
	“aa”	“al”	“av”	“aw”	“ay”	
mvMFE	0.5108	0.7200	0.4928	0.6814	0.5864	0.5983
RCmvMFE	0.8764	0.9692	0.9172	0.9794	0.9408	0.9366
IRCmvMFE	0.9928	0.9942	0.9722	0.9942	0.9986	0.9904

TABLE 4: Comparison of multiple traditional feature extraction methods on BCI competition II.

Reference number	Methods	Top classification rates (%)	Average classification rates (%)
[3]	WT	83.57	—
[3]	HHT	87.86	—
[4]	DWT	96.06	—
[5]	DTCWT	91.07	—
[6]	EMD	99.48	—
[7]	CSP	82.86	—
This paper	IRCmvMFE	100	99.86

Note: “—” represents that average recognition rate of 10×10 -fold CV is not given in the reference.

TABLE 5: Comparison with multiple CSP-based feature extraction methods in BCI competition III.

Reference number	Methods	Subjects					Average results (%)
		“aa”	“al”	“av”	“aw”	“ay”	
[8]	CSP	79.89	97.89	70.39	92.14	92.14	86.67
[8]	FBCSP	90.39	97.82	72.54	97.21	94.54	90.50
[8]	DFBCSP	90.32	98.46	75.14	97.82	95.29	91.40
[8]	SFBCSP	91.54	98.57	77.43	97.03	94.69	92.05
[9]	SWCSP	94.2	99.2	78	97.7	95.6	93.0
This paper	IRCmvMFE	99.39	99.71	98.61	99.61	99.82	99.43

Comparing the mean kappa values, the results of MI-EEG feature extraction from each subject was highest when using IRCmvMFE; this result revealed that IRCmvMFE had better consistency than those of RCmvMFE and mvMFE.

3.7. Comparison of Multiple Traditional Feature Extraction Methods. A variety of traditional feature extraction methods [3–9] were compared with the method presented in this paper, using SVM as a classifier. In Table 4, the top classification results and average classification of 10×10 -fold CV of referenced feature extraction methods [3–7] on BCI competition II are displayed. IRCmvMFE achieved the highest classification accuracy over the referenced methods, and its 10×10 -fold CV results were also better; it also showed the ability of IRCmvMFE to quantify the complexity of multichannel signals and implied its superiority in extracting features from MI-EEG signals.

The CSP-based feature extraction methods have been extensively studied on BCI competition III. The experimental

results of 10×10 -fold CV with CSP, filter bank CSP (FBCSP), discriminant FBCSP (DFBCSP), sparse FBCSP (SFBCSP), and spectrally weighted CSP (SWCSP) methods were from references [8, 9]. The method presented in this paper was compared with these methods, and the recognition rates are shown in Table 5. The results of CSP-based feature extraction were lower than those of IRCmvMFE. CSP-based methods only considered the spatial characteristics of MI-EEG signals, ignoring the features in other domains. IRCmvMFE effectively extracted nonlinear dynamic features of MI-EEG, correctly analyzed multichannel signals, and had good applicability in multiple subjects.

4. Discussion

In this paper, IRCmvMFE was proposed as a feature extraction method for MI-EEG signals. In IRCmvMFE, a composite filter technique was applied to improve the coarse-grained process of RCmvMFE, which eliminated

impulse noise interference due to random factors, produced smoother MI-EEG time series, and enhanced the filter results. The optimal channels and the optimal parameters were selected to calculate IRCmvMFE for each subject when imaging left-right-hand or right-hand-foot movement. Multiscale IRCmvMFEs were constructed as a feature vector. Entropy-based and traditionally referenced feature extraction methods were compared on two public datasets. The kappa coefficients of IRCmvMFE, RCmvMFE, and mvMFE were calculated for statistical analysis. The results implied the superiority and applicability of IRCmvMFE for the analysis of two-class motor imagery tasks. In the future, we will continue to focus on the research of multiclass motor imagery tasks.

5. Conclusions

A novel nonlinear dynamics method based on RCmvMFE, called IRCmvMFE, was introduced in this study. This method provides a potential tool for the nonlinear dynamic analysis of multichannel MI-EEG signals. RCmvMFE was developed using a composite filter technique in the coarse-grained process, which effectively removes impulse noise interference, better reflects the dynamic correlations both within and across channels, and is more closely matched the nonlinear and time-varying characteristics of MI-EEG and produced better features and classification accuracy. IRCmvMFE was applied to the analysis of multichannel MI-EEG signals and was compared to other commonly used feature extraction methods. IRCmvMFE yielded the highest classification results and improved stability; it also displayed the applicability of IRCmvMFE for MI-EEG feature extraction and provided a useful tool for the analysis of other complex, two-class biological signals.

Data Availability

Two previously reported datasets were used to support this study and are available at <http://bbci.de/competition/ii/> and <http://www.bbci.de/competition/iii>. These datasets are cited at relevant places within the text as references [36, 37].

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This research was financially supported by the National Natural Science Foundation of China (Nos. 81471770, 11882003, and 61672070) and the Natural Science Foundation of Beijing (No. 7132021).

References

- [1] J. H. Shi, J. Z. Shen, and P. Wang, "Feature extraction and classification of EEG signals in four kinds of motion imagination," *Journal of Zhejiang University (Engineering Science)*, vol. 46, no. 2, pp. 338–344, 2012.
- [2] S. K. Bashar, A. R. Hassan, and M. I. H. Bhuiyan, "Motor imagery movements classification using multivariate EMD and short time fourier transform," in *Proceedings of the Annual IEEE India Conference 2015*, New Delhi, India, December 2015.
- [3] H. W. Sun, Y. F. Fu, X. Xiong et al., "Identification of EEG induced by motor imagery based on hilbert-huang transform," *Acta Automatica Sinica*, vol. 41, no. 9, pp. 1686–1692, 2015.
- [4] K. R. He, X. G. Wang, L. Zou et al., "Brain-computer interface design using signal powers extracted during motor imagery tasks," *Chinese Journal of Biomedical Engineering*, vol. 20, no. 4, pp. 139–149, 2011.
- [5] S. K. Bashar, A. R. Hassan, and M. I. H. Bhuiyan, "Identification of motor imagery movements from EEG signals using dual tree complex wavelet transform," in *Proceedings of the 2015 International Conference on Advances in Computing, Communications and Informatics*, pp. 290–296, IEEE, Kochi, India, August 2015.
- [6] S. B. Medina and M. L. Duque, "Fuzzy entropy relevance analysis in DWT and EMD for BCI motor imagery applications," *Ingeniería*, vol. 20, no. 1, pp. 9–19, 2015.
- [7] C. Liu, H. B. Zhao, C. S. Li, and H. Wang, "CSP/SVM-based EEG classification of imagined hand movements," *Journal of Northeastern University (Natural Science)*, vol. 31, no. 8, pp. 1098–1011, 2010.
- [8] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface," *Journal of Neuroscience Methods*, vol. 255, no. 1, pp. 85–91, 2015.
- [9] J. Meng, G. Huang, D. Zhang, and X. Zhu, "Optimizing spatial spectral patterns jointly with channel configuration for brain-computer interface," *Neurocomputing*, vol. 104, pp. 115–126, 2013.
- [10] K. K. Kale and J. P. Gawande, "Automated feature extraction of epileptic EEG using approximate entropy," in *Proceedings of the 12th International Conference on Hybrid Intelligent Systems (HIS)*, pp. 474–477, IEEE, Adelaide, Australia, December 2012.
- [11] L. Wang, G. Z. Xu, S. Yang, J. Wang, M. Guo, and W. Yan, "Motor imagery BCI research based on sample entropy and SVM," in *Proceedings of the International Conference on Electromagnetic Field Problems & Applications*, pp. 1–4, IEEE, Dalian, Liaoning, China, June 2012.
- [12] L. Li, R. Cao, and J. Xiang, "Comparative study of approximate entropy and sample entropy based on characterization of EEG," *Computer Engineering and Design*, vol. 35, no. 3, pp. 1021–1026, 2014.
- [13] J. Tian and Z. Z. Luo, "Motor imagery EEG feature extraction based on fuzzy entropy," *Journal of HuaZhong University of Science and Technology*, vol. 41, no. 1, pp. 92–95, 2013.
- [14] Y. Cao, L. Cai, J. Wang et al., "Characterization of complexity in the electroencephalograph activity of Alzheimer's disease based on fuzzy entropy," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 8, article 083116, 2015.
- [15] J. Xiang, C. Li, H. Li et al., "The detection of epileptic seizure signals based on fuzzy entropy," *Journal of Neuroscience Methods*, vol. 243, pp. 18–25, 2015.
- [16] Q. Liu, Y. F. Chen, S. Z. Fan, M. F. Abbod, and J. S. Shieh, "EEG signals analysis using multiscale entropy for depth of anesthesia monitoring during surgery through artificial neural networks," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 232381, 16 pages, 2015.

- [17] T. Mizuno, T. Takahashi, R. Y. Cho et al., "Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy," *Clinical Neurophysiology*, vol. 121, no. 9, pp. 1438–1446, 2010.
- [18] P. H. Tsai, S. C. Chang, F. C. Liu, Jenho Tsao, Y. H. Wang, and M. T. Lo, "A novel application of multiscale entropy in electroencephalography to predict the efficacy of acetylcholinesterase inhibitor in alzheimer's disease," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 953868, 8 pages, 2015.
- [19] A. Humeau-Heurtier, "The multiscale entropy algorithm and its variants: a review," *Entropy*, vol. 17, no. 5, pp. 3110–3123, 2015.
- [20] X. Zou and M. Lei, "Pattern recognition of surface electromyography signal based on multi-scale fuzzy entropy," *Journal of Biomedical Engineering*, vol. 29, no. 6, pp. 1184–1188, 2012.
- [21] J. Zheng, J. Cheng, Y. Yang, and S. Luo, "A rolling bearing fault diagnosis method based on multi-scale fuzzy entropy and variable predictive model-based class discrimination," *Mechanism and Machine Theory*, vol. 78, pp. 187–200, 2014.
- [22] M. Li, H. Liu, W. Zhu, and J. Yang, "Applying improved multiscale fuzzy entropy for feature extraction of MI-EEG," *Applied Sciences*, vol. 7, no. 1, p. 92, 2017.
- [23] J. Zheng, H. Pan, and J. Cheng, "Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines," *Mechanical Systems and Signal Processing*, vol. 85, pp. 746–759, 2017.
- [24] J. D. Zheng, H. Y. Pan, S. J. Cheng et al., "Composite multiscale fuzzy entropy based rolling bearing fault diagnosis method," *Journal of Vibration and Shock*, vol. 35, no. 8, pp. 116–123, 2016.
- [25] P. Li, C. Y. Liu, L. P. Li et al., "Multiscale multivariate fuzzy entropy analysis," *Acta Physica Sinica*, vol. 62, no. 12, article 120512, 2013.
- [26] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy: a tool for complexity analysis of multichannel data," *Physica A*, vol. 465, pp. 261–276, 2017.
- [27] L. Ni, J. Cao, and R. Wang, "Time-dependent multivariate multiscale entropy based analysis on brain consciousness diagnosis," in *Proceedings of the BICS: International Conference on Brain Inspired Cognitive Systems*, Beijing, China, June 2013.
- [28] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy analysis," *IEEE Signal Processing Letters*, vol. 19, no. 2, pp. 91–94, 2012.
- [29] L. Ni, J. Cao, and R. Wang, "Dynamic multivariate multiscale entropy based analysis on brain death diagnosis," *Science China Technological Sciences*, vol. 58, no. 3, pp. 425–433, 2015.
- [30] X. Y. Wang and Y. H. Xu, "Signal classification method for Parkinson gait based on multivariate multiscale fuzzy entropy," *Beijing Biomedical Engineering*, vol. 35, no. 5, pp. 489–496, 2016.
- [31] M. U. Ahmed, T. Chanwimalueang, S. Thayyil, and D. Mandic, "A multivariate multiscale fuzzy entropy algorithm with application to uterine EMG complexity analysis," *Entropy*, vol. 19, no. 1, p. 2, 2016.
- [32] J. Zheng, D. Tu, H. Pan, X. Hu, T. Liu, and Q. Liu, "A refined composite multivariate multiscale fuzzy entropy and laplacian score-based fault diagnosis method for rolling bearings," *Entropy*, vol. 19, no. 11, p. 585, 2017.
- [33] H. Azami and J. Escudero, "Refined composite multivariate generalized multiscale fuzzy entropy: a tool for complexity analysis of multichannel signals," *Physica A: Statistical Mechanics and its Applications*, vol. 465, pp. 261–276, 2017.
- [34] L. Zhang and Z. M. Pan, "Application of multi-filter in processing of data received from three-dimension digital compass," *Modern Electronics Technique*, vol. 22, pp. 148–150, 2008.
- [35] M. A. Li, S. D. Guo, J. F. Yang, and Y. J. Sun, "A novel EEG feature extraction method based on OEMD and CSP algorithm," *Journal of Intelligent & Fuzzy Systems*, vol. 30, no. 5, pp. 2971–2983, 2016.
- [36] B. Blankertz, K. R. Müller, G. Curio et al., "The BCI Competition 2003: progress and perspectives in detection and discrimination of EEG single trials," *IEEE Transactions on Bio-Medical Engineering*, vol. 51, no. 6, pp. 100–106, 2004.
- [37] BCI competition III, <http://www.bbci.de/competition/iii>.
- [38] J. J. Meng, X. J. Sheng, L. Yao et al., "Common spatial spectral pattern for motor imagery tasks in small channel configuration," *Chinese Journal of Biomedical Engineering*, vol. 32, no. 5, pp. 553–561, 2013.
- [39] Q. S. She, Y. L. Ma, M. Meng, and Z. Luo, "Multiclass posterior probability twin SVM for motor imagery EEG classification," *Computational Intelligence and Neuroscience*, vol. 2015, Article ID 251945, 9 pages, 2015.