

Research Article

Applying a Locally Linear Embedding Algorithm for Feature Extraction and Visualization of MI-EEG

Mingai Li, Xinyong Luo, Jinfu Yang, and Yanjun Sun

School of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China

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

Received 3 May 2016; Revised 18 July 2016; Accepted 31 July 2016

Academic Editor: Eugenio Martinelli

Copyright © 2016 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.

Robotic-assisted rehabilitation system based on Brain-Computer Interface (BCI) is an applicable solution for stroke survivors with a poorly functioning hemiparetic arm. The key technique for rehabilitation system is the feature extraction of Motor Imagery Electroencephalography (MI-EEG), which is a nonlinear time-varying and nonstationary signal with remarkable time-frequency characteristic. Though a few people have made efforts to explore the nonlinear nature from the perspective of manifold learning, they hardly take into full account both time-frequency feature and nonlinear nature. In this paper, a novel feature extraction method is proposed based on the Locally Linear Embedding (LLE) algorithm and DWT. The multiscale multiresolution analysis is implemented for MI-EEG by DWT. LLE is applied to the approximation components to extract the nonlinear features, and the statistics of the detail components are calculated to obtain the time-frequency features. Then, the two features are combined serially. A backpropagation neural network is optimized by genetic algorithm and employed as a classifier to evaluate the effectiveness of the proposed method. The experiment results of 10-fold cross validation on a public BCI Competition dataset show that the nonlinear features visually display obvious clustering distribution and the fused features improve the classification accuracy and stability. This paper successfully achieves application of manifold learning in BCI.

1. Introduction

In response to an external stimulus or active thinking activity, nerve cells in the brain cortex can produce Motor Imagery Electroencephalography (MI-EEG) signals with the characteristics of specificity and rhythmicity, and they not only contain a large amount of physiological or disease information but also have a close correlation with the state of consciousness. As a result, much attention has been paid to the application of MI-EEG in brain cognition; meanwhile, the correct interpretation and accurate extraction of MI-EEG features are the key to its successful applications [1, 2].

MI-EEG has the characteristics of individual differences and is nonlinear, nonstationary, and time-varying sensitive. Wavelet transform (WT) is able to take advantage of scale and shift operations to perform multiscale decomposition and time-frequency domain localization, effectively obtaining the time-frequency information of signals. Thus, the analysis of EEG signals can benefit from WT. Despite this, recent studies do not support the effective use of wavelet features for the

discrimination of EEG signals because of the redundant and irrelevant information contained in wavelet coefficients [3]. To this day, these characteristics remain a problem for the extraction of useful features from EEG signals for classification. Many approaches to the problem were proposed in the last decade. Xu et al. proposed a novel method of extracting EEG features based on discrete wavelet transform (DWT) and autoregressive (AR) model, in which the combination features of wavelet coefficient statistics and the sixth-order AR coefficients were used as input vectors for the classifier [4]. To solve the problem of a lack of spatial information and large amounts of features, a method based on the time-frequency-space domain was proposed using independent component analysis (ICA) and wavelet transform [5]. A method using only wavelet coefficients from the subband bound with specific frequency features of MI-EEG was proposed by Aldea and Fira [6]. Chaurasiya performed a linear dimensionality reduction of wavelet coefficients by principal component analysis (PCA) and then used the coordinates embedded in the low dimension space as features and obtained better

classification results [7]. Obviously, wavelet transform, as a type of classical time-frequency analysis method, played an important role in the feature extraction of the EEG signal.

However, the brain is a highly complex biological system in terms of structure and function; namely, the brain is a typical nonlinear system [8]. MI-EEG signals not only contain abundant information and rhythmic characteristics of motor imagery consciousness but also possess an obvious nonlinear structure. Taking the reason that traditional time-frequency feature extraction methods were based on linear system theory into consideration, we will inevitably find a loss of information in the original signal, and follow-up questions about the nonlinear structural characteristics cannot be excavated from high-dimensional EEG datasets [9]. Manifold learning (ML) can recover the structure of lower dimensional manifolds from high-dimensional data and can help us obtain the corresponding nonlinear embedded coordinates that are regarded as a meaningful representation of reduced data. During the dimension reduction process, ML can best preserve the local neighborhood of each object or, in other words, the potential of the manifold structure. At present, Locally Linear Embedding (LLE), as one of the typical ML algorithms, has got a certain application in feature mapping of EEG signal, for the reason that LLE as a noniterative algorithm has low computing complexity and local optimal analytical solutions. Lee et al. [10] calculated MI-EEG signal adaptive autoregression model parameters based on Kalman filter and then by using LLE performed dimension reduction to get EEG features, and the classification accuracy was between 61.9% and 74.2% with a linear discriminant analysis classifier; Pan et al. [11] extracted the features of EEG by weighted LLE for detecting epilepsy, and the classification rate was between 72.79% and 90.66% using SVM classifier; Mirsadeghi et al. [12] calculated the power, covariance, and several kinds of entropy of EEG to obtain the primitive features, and LLE was applied to reduce the feature dimension. Then, the quadratic discriminant analysis (QDA) was adopted to estimate the depth of anesthesia, and the average classification accuracy was 88.4%. From the studies, we can conclude that LLE is applicable for feature extraction and feature dimension reduction of EEG. However, there exist the following problems that need to be improved: (1) LLE algorithm is sensitive to data noise, so if it is directly used in feature extraction and feature dimension reduction of EEG signal, the low-dimensional embedding structure may be easily destroyed, which affects the quality of EEG features; (2) because EEG signal has significant time-frequency distribution and nonlinear characteristics, using only ML algorithm is difficult to get its essential characteristics comprehensively, which sometimes even will cause feature information redundancy and mismatching problem.

In this paper, to achieve a comprehensive and accurate extraction of the essential features of MI-EEG signals, a feature extraction method denoted as DWT-LLE was proposed based on a Locally Linear Embedding algorithm, a typical manifold learning algorithm, and discrete wavelet transform. On the foundation of an analysis with a Wigner Ville distribution and power spectrum, we used DWT to obtain the time-frequency feature and then applied LLE to the approximation

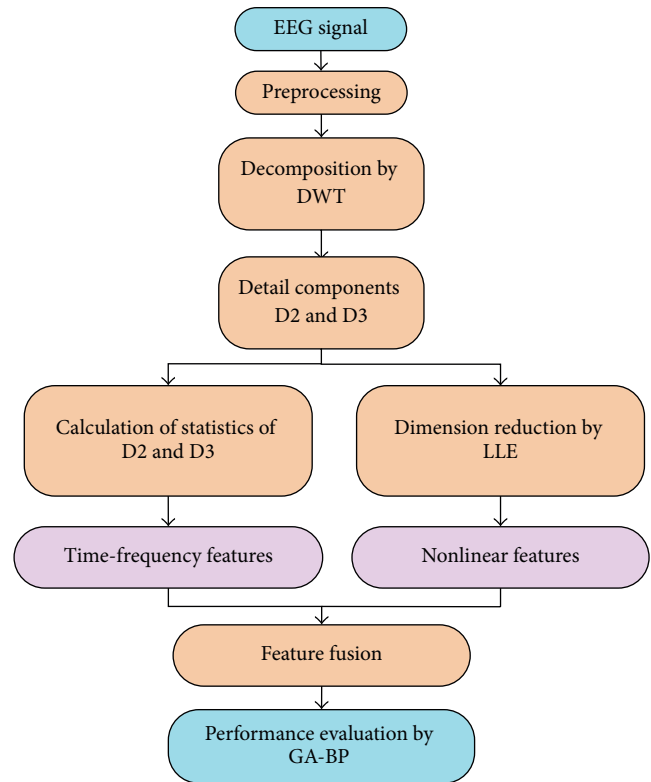


FIGURE 1: Block diagram of the proposed method.

components to obtain nonlinear features. Finally, the time-frequency features and nonlinear features were fused serially, and then a BP neural network optimized by a genetic algorithm was employed as a classifier and many experiments were conducted on a publicly available dataset with 10-fold cross validation. Compared with the traditional DWT-based feature extraction methods, the proposed method enabled the feature vector to contain more comprehensive and complete MI-EEG information because of the addition of nonlinear features. Thus, better classification results have been achieved and demonstrate the validity of this method. The overall block diagram is given to illustrate the flow of the proposed method, shown as in Figure 1.

2. Algorithm

2.1. Discrete Wavelet Transform. The time-frequency domain localization and multiresolution as characteristics of wavelet transform make it possible to obtain transient information in different scales under different frequencies that are suitable for the analysis of nonstationary signals, such as EEG. This paper uses discrete wavelet transform to analyze the Event-Related Desynchronization (ERD) and Event-Related Synchronization (ERS) of MI-EEG for right and left hand movements.

Based on the Mallat algorithm, for the L -level wavelet decomposition of the signal, the given signal $f(t)$ can be represented as

$$f(t) = A_L + \sum_{j=1}^L D_j, \quad (1)$$

where L denotes the number of the decomposition levels, A_L denotes the approximation coefficients, and D_j denotes the detail coefficients at scale j , $j = 1, 2, \dots, L$. Thus, the signal $f(t)$ can be divided into several subbands. Assuming that the sampling frequency of the signal is f_s , then for the coefficients $A_L, D_L, D_{L-1}, \dots, D_1$, the corresponding frequency band ranges are $[0, f_s/2^{L+1}]$, $[f_s/2^{L+1}, f_s/2^L]$, $[f_s/2^L, f_s/2^{L-1}]$, ..., $[f_s/2^2, f_s/2]$, respectively. In wavelet decomposition for MI-EEG, the number of decomposition layers is decided by the distribution of useful components in the signal in the frequency domain and the sampling frequency.

2.2. The Locally Linear Embedding Algorithm. The Locally Linear Embedding algorithm is a type of typical manifold learning algorithm. The main idea of LLE is to solve globally nonlinear problems using locally linear fitting, which is based on the assumption that data lying on a nonlinear manifold can be viewed as linear in local areas [13]. By computing low-dimensional, neighborhood-preserving embeddings of high-dimensional inputs, LLE maps its inputs into a single global coordinate system of lower dimensionality, and its optimizations do not involve local minima. By exploiting the local symmetries of linear reconstructions, LLE is able to learn the global structure of nonlinear manifolds [14]. The locally linear fitting is characterized by linear coefficients that reconstruct each data point from its neighbors and the linear coefficients, namely, the weight matrix; record the neighborhood information; and establish a bridge between the high-dimensional data space and the low-dimensional latent space.

The Locally Linear Embedding algorithm consists of three steps:

- (1) Select neighbors: as an input, LLE takes a set of n D -dimension vectors assembled in a matrix $X = [x_1, x_2, \dots, x_n] \in R^{D \times n}$. According to the similarities between data points measured by Euclidean distances, for each sample x_i , $i = 1, \dots, n$, its k nearest neighbors are sought, and then their indices are stored in an $n \times k$ matrix J .
- (2) Reconstruct weight matrix W with linear weights: to find the reconstruction weight matrix $W \in R^{n \times n}$, where w_{ij} represents the contribution of the data point x_j to the data point x_i , the following cost function should be minimized:

$$e(W) = \sum_{i=1}^n \left\| x_i - \sum_{j=1}^n w_{ij} x_j \right\|^2, \quad (2)$$

where the optimization should be subject to constraints

$$\sum_{j=1}^n w_{ij} = 1 \quad (3)$$

and $w_{ij} = 0$, if x_i and x_j are not neighbors. If not, w_{ij} is decided by the optimization of reconstruction errors in the way of solving a least-squares problem (2).

- (3) Map high-dimensional data to embedded coordinates: the weight matrix W is kept fixed and the low-dimensional embedding $Y = [y_1, y_2, \dots, y_n] \in R^{d \times n}$, where d represents embedded dimension, can be obtained by minimizing the embedding cost function

$$\Phi(Y) = \sum_{i=1}^n \left\| y_i - \sum_{j=1}^n w_{ij} y_j \right\|^2, \quad (4)$$

under the constraints below:

$$\begin{aligned} \sum_{i=1}^n y_i &= 0 \\ \frac{1}{N} \sum_{i=1}^n y_i y_i^T &= I, \end{aligned} \quad (5)$$

which provide a unique solution. A new matrix M of size $n \times n$ is constructed based on the matrix W : $M = (I - W)^T (I - W)$, and then we can obtain the low-dimensional embedding Y by carrying out an eigendecomposition of the new sparse matrix M [14]. Because the first eigenvector, whose eigenvalue is close to zero, is excluded, Y can be represented as below:

$$Y = [v_2, \dots, v_{d+1}]^T, \quad (6)$$

where v_2, \dots, v_{d+1} denotes the 2nd to $(d+1)$ st smallest eigenvalues of matrix M .

3. Analysis of the Dataset

3.1. Experimental Data Description. The experimental dataset was from BCI Competition 2003 provided by BCI Lab, Graz University of Technology. The dataset was composed of 280 trials, of which 140 were for training and 140 were used for testing images of left/right hand movements. Each trial lasted for 9 seconds, and the sequence diagram of the experiment is shown schematically in Figure 2. The first 2 seconds were quiet, during which the subjects kept calm and a short beep indicated the start of the trial at the end of second 2, with a fixation cross “+” displayed on the screen simultaneously. At the third second, the visual cue (left-right arrow) was displayed as the direction of motor imagery. The data were sampled at 128 Hz. Three MI-EEG channels were measured over a C3, CZ, and C4 conductor, using AgCl as an electrode, and the placement of the electrode is shown in Figure 3.

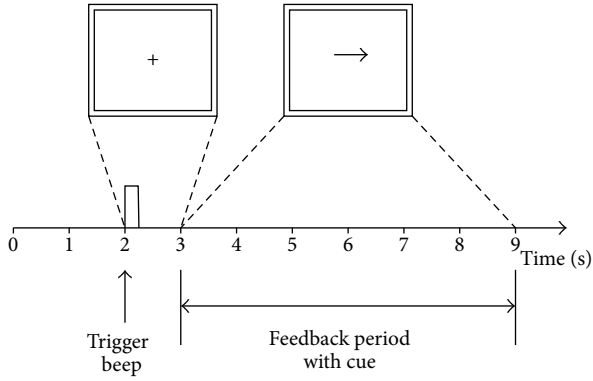


FIGURE 2: Timing scheme of the experiment.

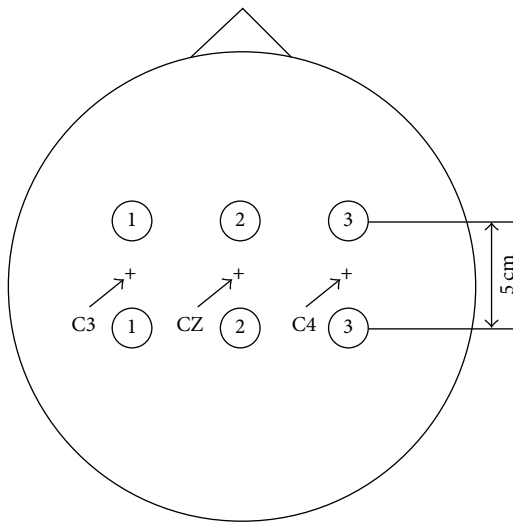


FIGURE 3: Electrode placement.

3.2. The Physiological Basis of Signal Analysis for MI-EEG.

When some region in the cerebral cortex is activated, the area of the metabolism and blood flow increase, leading to an amplitude reduction of the MI-EEG in rhythm Alpha (8~13 Hz) and rhythm Beta (14~30 Hz), which is called Event-Related Desynchronization (ERD). In resting or inert states, the amplitudes of rhythm Alpha and rhythm Beta are significantly higher, which is called Event-Related Synchronization (ERS). Therefore, when people imagine one hand moving, the amplitude of rhythms Alpha and Beta from the contralateral corresponding primary sensorimotor cortex decreases, while that from the ipsilateral corresponding primary sensorimotor cortex increases. The ERD and ERS phenomenon of MI-EEG is the most fundamental basis for the classification of left/right hand movement images [15].

3.3. The Time-Frequency Analysis Based on the Wigner Ville Distribution. Considering the generating mechanism and signal characteristics of EEG signals, on MI-EEG, the effective frequency window of feature extraction can be decided by signal time-frequency analysis based on the Wigner Ville

distribution and the energy distribution of the signal on the time-frequency domain.

The Wigner Ville distribution of MI-EEG is illustrated in Figure 4. Figure 4(a) describes the 3D relationship among the signal energy, time, and frequency, which demonstrates that the signal energy is concentrated in a frequency range of 8 to 30 Hz. Figure 4(b) shows the relationship among the MI-EEG signal energy, time, and frequency in the form of a 2D diagram and the signal energy mainly concentrated at approximately 10 Hz and 20 Hz in the way of zonal distribution. Thus, it lays a solid foundation for frequency range determination of MI-EEG, but it is difficult for the analysis above to accurately determine the scope of the valid time.

3.4. The Power Spectrum Analysis of MI-EEG. On the basis of the analysis above, the MI-EEG signal is filtered by a 8~30 Hz band-pass filter and is calculated as the signal average power. For a given $x(i, j)$, which represents the j th data in trial i , the average power $P(j)$ can be calculated as

$$P(j) = \frac{1}{N} \sum_{i=1}^N x^2(i, j), \quad (7)$$

where N denotes the number of trials. Because the phenomenon of ERD/ERS in C3 and C4 is much more obvious, the signal C3 and C4 can only be calculated in (7), and the results are presented in Figure 5.

Figure 5 demonstrates that the average power of C3 and C4 for the MI-EEG of the left/right hand movement displays obvious ERD and ERS phenomena in the 8~30 Hz range, especially in the time range of 3.5 to 7 s, which provides the theoretical foundation for determining the time range for feature extraction, thereby reducing the complexity of the calculation.

4. A Description of Feature Extraction

4.1. Time-Frequency Feature. Based on the analysis above and the characteristics of the rhythmic distribution in the time-frequency domain of MI-EEG, the db5 wavelet function in Class Daubechies was used to make a 3-level wavelet decomposition of MI-EEG, which provides the information at different resolutions of the samples at different frequency bands [16]. After the decomposition, A_3 denotes its approximation coefficients and D_1-D_3 denote its detail components. Because the wavelet coefficients express the energy distribution of the signal in the time-frequency domain [4] and rhythm Alpha (8~13 Hz) is close to D_3 , whose frequency range is 8~16 Hz, and rhythm Beta (14~30 Hz) is similar to D_2 , whose frequency range is 16~30 Hz, the two frequency ranges are the most obvious frequency bands for the ERD and ERS phenomena [17]. Considering the frequency of our interest, the detail components D_2 and D_3 are chosen for feature extraction.

To depict the feature in the view of signal energy, the statistics of the detail components are calculated as the time-frequency feature and the statistical features, such as the

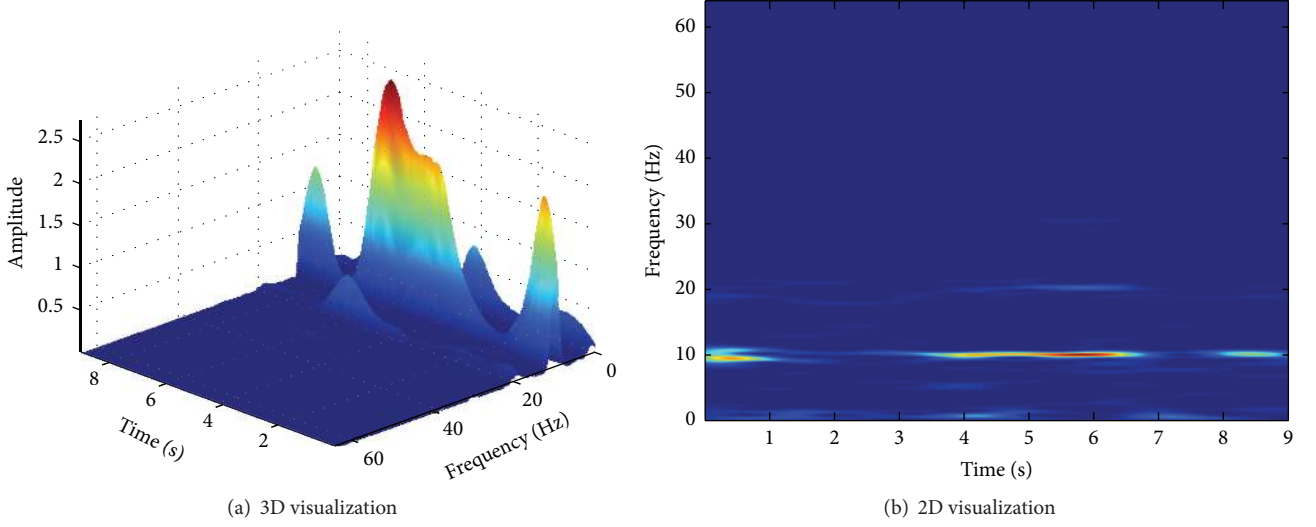


FIGURE 4: Visualization of the time-frequency analysis for MI-EEG.

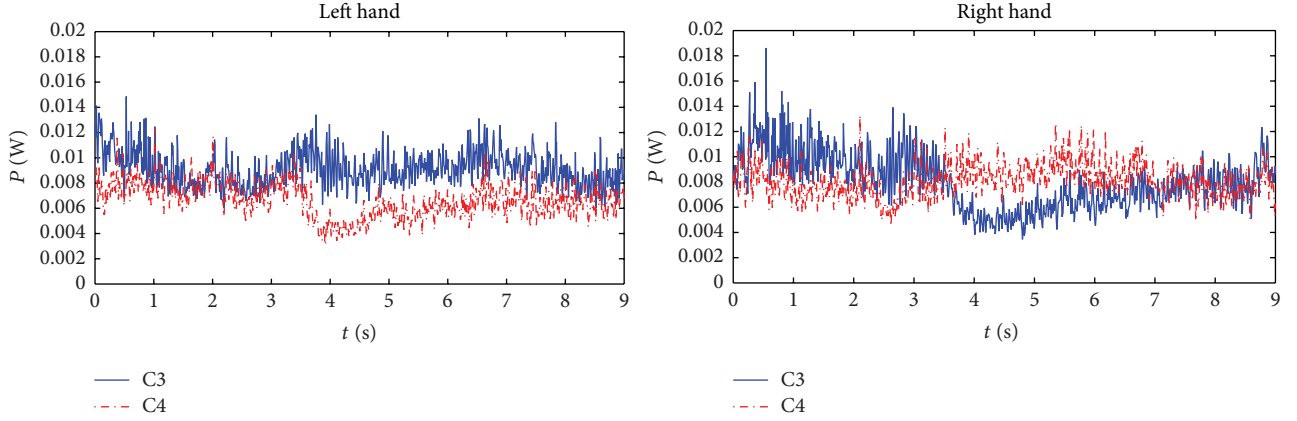


FIGURE 5: The average power of C3 and C4 for MI-EEG of left/right hand movement.

average, mean value of energy, and mean value of the standard deviation of the detail components. Assume that $x_{j,k}^i \in R^{1 \times n}$ represents the detail components D_j ($i = 3, 4$ $j = 2, 3$) from channel C_i of MI-EEG and $k = 1, 2, \dots, n$, where n represents the number of detail components. Thus, the time-frequency feature can be obtained as shown below, and the average is computed as follows:

$$\bar{\mu}_j^i = \frac{1}{n} \sum_{k=1}^n x_{j,k}^i. \quad (8)$$

The mean value of energy is determined by

$$\bar{E}_j^i = \frac{1}{n} \sum_{k=1}^n (x_{j,k}^i)^2. \quad (9)$$

And the mean value of the standard deviation is calculated by

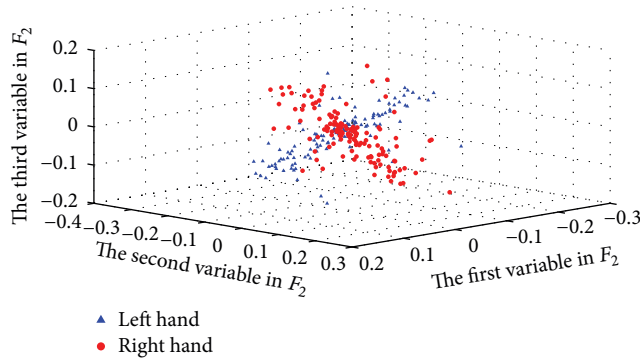
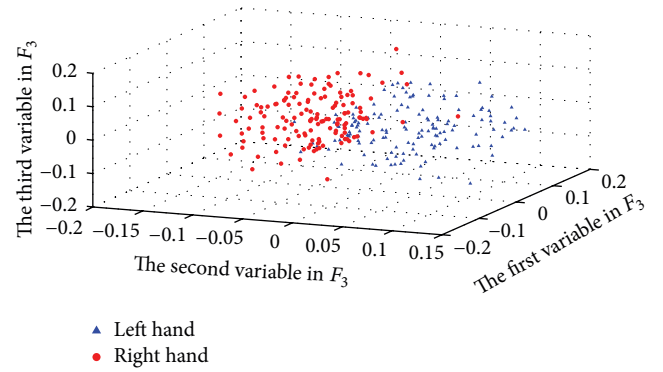
$$S_j^i = \sqrt{\frac{1}{n} \sum_{k=1}^n (x_{j,k}^i - \bar{\mu}_j^i)^2}. \quad (10)$$

Considering the performance phenomena ERD and ERS in channels C3 and C4, to make the time-frequency feature show a sharp distinction, the time-frequency feature $F_1 \in R^{6 \times 1}$ is defined as

$$F_1 = \begin{bmatrix} \frac{\bar{\mu}_2^3 - \bar{\mu}_2^4}{\|\bar{\mu}_2^3 - \bar{\mu}_2^4\|}, \frac{\bar{\mu}_3^3 - \bar{\mu}_3^4}{\|\bar{\mu}_3^3 - \bar{\mu}_3^4\|}, \frac{\bar{E}_2^3 - \bar{E}_2^4}{\|\bar{E}_2^3 - \bar{E}_2^4\|}, \frac{\bar{E}_3^3 - \bar{E}_3^4}{\|\bar{E}_3^3 - \bar{E}_3^4\|}, \\ \frac{S_2^3 - S_2^4}{\|S_2^3 - S_2^4\|}, \frac{S_3^3 - S_3^4}{\|S_3^3 - S_3^4\|} \end{bmatrix}^T, \quad (11)$$

where $\|\cdot\|$ represents the second-order norm.

4.2. Nonlinear Feature. To fully acquire the accurate nonlinear information of the distribution of MI-EEG, LLE is used to extract features, specifically, features that can loyally represent the nonlinear information in a high-dimensional observation space and simultaneously reduce nonlinear

FIGURE 6: 3D visualization of feature F_2 .FIGURE 7: 3D visualization of feature F_3 .

dimensionality and visualization in the view of manifold learning.

4.2.1. The Nonlinear Feature Based on Original MI-EEG. First, the MI-EEG from channels C3 and C4 during the 3.5~7 s period, in which the ERD and ERS phenomena are remarkable, are concatenated together; then, a nonlinear dimensionality reduction is performed based on the LLE and the low-dimensional embedding is Y_2 . Thus, the maximum differentiation of three coordinates was chosen as the nonlinear feature F_2 and is defined as

$$F_2 = Y_2^* \in R^{3 \times 1}. \quad (12)$$

The 3D visualization of F_2 is shown in Figure 6.

4.2.2. The Nonlinear Feature Based on Wavelet Coefficients. First, we performed 3-level wavelet decomposition of MI-EEG; then, because the approximation components reflect the time-frequency feature distribution of a low-frequency signal in the original MI-EEG and the high-frequency details contain more noise, in the process of nonlinear dimensionality reduction from the embedded high-dimensional observation space to the low dimension space, the noninteracting far points are mapped as the local neighbor points [18], which destroys the results of low-dimensional embedding. In addition, the LLE algorithm requires that a data set is sampled from a low-dimensional manifold in uniform density [19] and that it is vulnerable to noise. Thus, DWT was combined with LLE for feature extraction of MI-EEG, and the proposed method was called DWT-LLE. Considering the analysis above, the approximation of A_3 coefficients was chosen to reduce the dimensionality by LLE, and the nonlinear features of F_3 were defined as

$$F_3 = Y_3^* \in R^{3 \times 1}. \quad (13)$$

A 3D visualization of F_3 is shown in Figure 7.

Figure 6 shows that the low-dimensional embedding of F_2 presents a clustering distribution in the 3D space, and the distributions of two classes of nonlinear features cross each other. Further, we can know from Figure 7 that, compared with F_2 , the distribution of F_3 in the 3D space shows better separability, from which the classification

significantly benefits. We can thus conclude that there is an intimate connection between the nonlinear information contained in the MI-EEG and wavelet coefficients, which reflects the energy distribution of the signal in the time-frequency domain.

5. Pattern Classification

For the objective assessment evaluation of the quality of the extracted features, a backpropagation (BP) neural network was chosen as a classifier for the pattern recognition of the MI-EEG signal. Although the BP neural network has good self-learning abilities and strong adaptive capabilities, the random selection of its initial weights and thresholds results in its vulnerability to the local optimum problem, thus affecting the generalization ability of the classifier. Therefore, a three-layer BP neural network optimized by genetic algorithm (GA) was used as a classifier and is denoted by GA-BP.

5.1. The Structure of GA-BP. The neurons number of input-layer was decided by the feature dimension of the EEG, and the neurons of output-layer were 1. The neurons number of hidden-layer was determined by experiments, and the optimal value was selected as 28. The transfer functions for the hidden-layer and output-layer were “tansig” and “pruelin,” respectively, and the training function was “trainlm.” Mean Square Error (MSE) was chosen as the performance function.

5.2. The Optimization Procedure of GA-BP. The optimization procedure for the weights and threshold for the input-layer and hidden-layer in the BP neural network was as follows.

Step 1. Make a serial connection between the weight/bias of the input-layer and hidden-layer and the hidden-layer and output-layer of the candidate solution, called individual, creature, or phenotype.

Step 2. Transform it into an array of bits in binary encoding and set the array digits to 20.

Step 3. The population of candidate solutions is set to 40, and the maximum generation of evolution is 100. Additionally, the

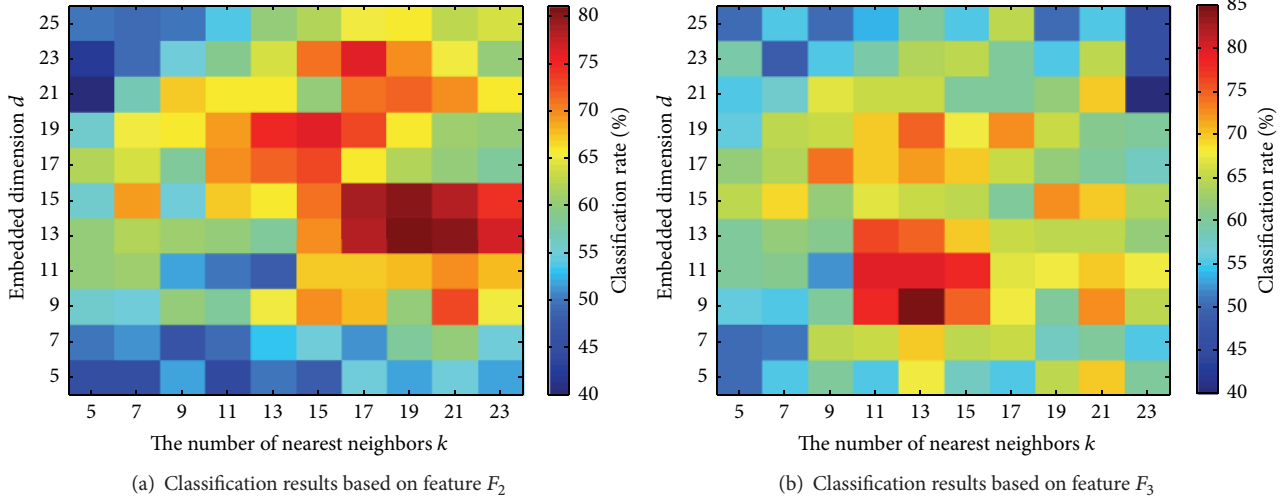


FIGURE 8: The average classification accuracy under different combination of parameters k and d .

probabilities of crossover and mutation are set to 0.7 and 0.05, respectively.

Step 4. Use training samples and testing samples for network training and testing, respectively, and test that the error is regarded as an individual fitness value for each individual. The smaller the value is, the more outstanding the individual is.

Step 5. Perform a heuristic search in the solution domain of genetic representation and use the fitness value of the individual to evaluate the solution by techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. As a result, we can obtain GA-BP neural network with a better set of weight/bias facilitating classification.

6. The Experimental Results

To evaluate the proposed feature extraction method, a three-layer BP neural network and a three-layer GA-BP neural network were, respectively, employed to classify the MI-EEG signal, and the 10-fold cross validation method was used to eliminate the contingency in the feature extraction process.

6.1. The Optimal Selection of Parameters in the LLE Algorithm.

LLE is an unsupervised, noniterative method, without the local minima problems plaguing many comparative methods. However, the number of nearest neighbors k and embedded dimension d have a great influence on its nonlinear structural features and the classification performance, the parameters will be optimally determined by experiments, in which the GA-BP is used as the classifier. Figure 8 shows the average classification accuracy of two nonlinear features under different combination of parameters k and d . Figure 8(a) was obtained with nonlinear feature F_2 and Figure 8(b) with F_3 .

Figure 8 demonstrates that the best average classification accuracy of feature F_2 is different from that of feature F_3 and

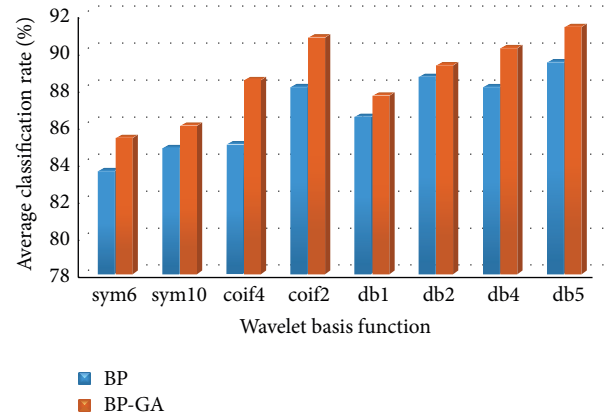
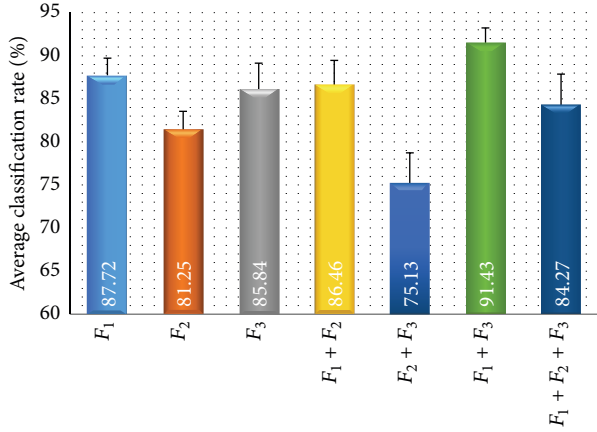


FIGURE 9: Impact of different wavelet basis functions on classification accuracy.

the corresponding combination of parameters k and d are $k = 19$, $d = 13$ and $k = 13$, $d = 9$, respectively.

6.2. The Optimal Selection of the Wavelet Basis Function in DWT.

In the analysis of a given signal based on DWT, the wavelet basis function is hard to choose and relies on whether we can acquire an accurate time-frequency feature or not for the reason that the parts contained in the signal will be magnified by the waveform as it approximates the shape of the selected wavelet basis function and vice versa. Given that there is no mature and reliable analytical method to guide the selection of the wavelet basis function, this paper selects some wavelet basis functions with a certain similarity to the MI-EEG signal waveform, including sym6 and sym10, as listed in Figure 9. To evaluate the fitness of the wavelet basis function, the BP neural network and GA-BP neural network were used as classifiers after the decomposition of MI-EEG into subbands by DWT utilizing the corresponding wavelet basis



The combination of features

FIGURE 10: Experimental results of the average classification accuracy and variation range in different feature combinations.

function. The results of the average classification accuracy of the experiments are shown in Figure 9.

Obviously, we can find that GA-BP outperforms BP regardless of which wavelet basis function is used. Additionally, the results tell us that db5 is the best wavelet basis function among those listed in Figure 9 for MI-EEG classification, and with it, we obtained the best classification results.

6.3. Feature Fusion. The features extracted by different methods can provide different information contained in MI-EEG, while there might be a problem of information redundancy and feature matching among different types of features. To obtain a feature vector containing comprehensive essential information, it is necessary that consideration should be given to the nonlinear nature and the ERD and ERS phenomena of MI-EEG. To find the optimal combinations between the time-frequency feature F_1 and nonlinear features F_2 and F_3 , various feature combinations are conducted using the GA-BP neural network as a classifier in a test with 10-fold cross validation. The average classification accuracy and variance are shown in Figure 10, where “+” represents feature fusion in serial order.

Figure 10 illustrates that we can obtain the best average classification accuracy (91.43%) and the smallest standard deviation when feature F_1 is concatenated with feature F_3 , meaning that the two types of features show perfect complementarity and stability and the least redundant information.

6.4. The Comparison of Multirecognition Methods. For MI-EEG, a comparative study was performed, and it mainly included the DWT-based feature extraction methods, which have recently been proposed, the corresponding classifier, the feature dimension, and the optimal classification accuracy, as shown in Table 1.

Table 1 shows that the proposed method, called DWT-LLE, represents a better classification with fewer feature dimensions compared with recent ones.

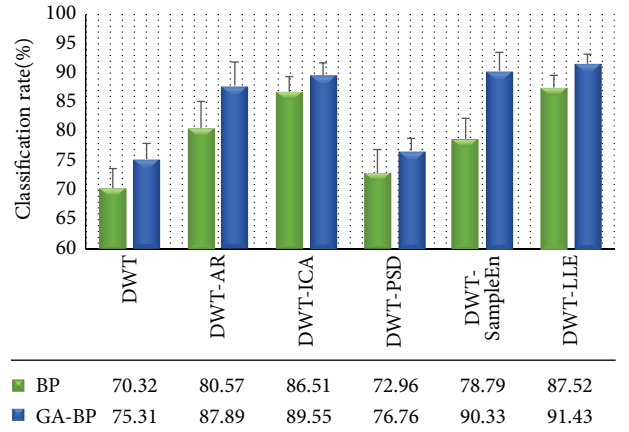


FIGURE 11: Comparison of the average classification accuracy for multi-DWT-based feature extraction methods.

TABLE 1

Number of reference	Feature extraction	Classifier	Feature dimension	Top classification rate (%)
4	DWT-AR	LDA	24	92.14
6	DWT	LDA	20	91.00
20	DWT-ICA	BP	10	90.50
5	DWT-ICA	GA-SVM	14	90.71
8	DWT-SampleEn	SVM	5	91.43
21	DWT-PSD	KNN	1	84.29
This paper	DWT-LLE	GA-BP	9	93.20

Note: “-” represents the combination or optimization of methods for feature extraction or classification.

6.5. The Comparison of Multi-DWT-Based Feature Extraction Methods. To further verify the effectiveness of the proposed method for feature extraction and to evaluate the matching degree between the features and classifier, some comparative experiment studies were conducted based on the proposed method and traditional DWT-based methods were conducted on the same dataset. The classification performance was evaluated by the BP neural network and GA-BP neural network. The evaluation results of the experiments are listed in Figure 11, where “-” represents the combination of methods for feature extraction and the test was 10-fold cross validation.

In the six experiments shown in Figure 11, the optimal numbers of hidden-layer neurons were 26, 30, 25, 7, 18, and 28. Figure 11 shows that the method proposed in this paper had the best performance with a top average classification accuracy of 91.43% and a smaller variance of approximately $\pm 2\%$ around the average classification rate. The results demonstrate that the validity of this method for feature extraction of MI-EEG is encouraging and the proposed method demonstrates a high degree of matching with the GA-BP neural network.

7. Conclusions

Regarding MI-EEG, the proposed feature extraction method based on LLE and DWT outperforms the existing methods in classification accuracy with fewer feature dimensions. The experiment results show the following. (1) The time-frequency features with DWT and the nonlinear features F_3 from approximation components with LLE algorithm are highly complementary for classification. The combined features contain less redundant information compared to ones extracted by other methods, and the combined ones are advantageous to the comprehensive representation of MI-EEG and have better compatibility with GA-BP classifier. (2) The application of manifold learning as a nonlinear dimension reduction method enables us to find that high-dimensional MI-EEG signals show a significant clustering distribution in the low-dimensional space, proving the existence of a nonlinear structure in MI-EEG. (3) The nonlinear information contained in the MI-EEG signals is mainly located in specific subbands related to certain motor imagery tasks, and the information can be represented by wavelet coefficients that depict the time-frequency energy distributions of the original signal. The nonlinear characteristics extracted by the method proposed in this paper can promote the further analysis of MI-EEG, paving the way for expanded applications of manifold learning to brain-computer research and posing a novel perspective of the nature of MI-EEG.

Competing Interests


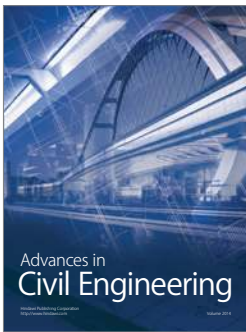
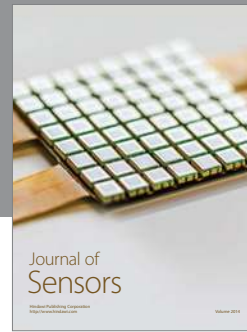
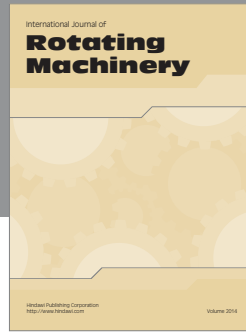
The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is financially supported by the National Natural Science Foundation of China (no. 81471770 and no. 61201362), the Natural Science Foundation of Beijing (no. 7132021), and Promotion Project on Intelligent Robot in Beijing University of Technology (Key Technologies on Modularization and Practicalization of Intelligent Service Robot).

References

- [1] T. M. Vaughan and J. R. Wolpaw, "The third international meeting on brain-computer interface technology: making a difference," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 126–127, 2006.
- [2] X. Baoguo, "Feature extraction and classification of single trial motor imagery EEG," *Journal of Southeast University*, vol. 37, no. 4, pp. 629–633, 2007.
- [3] N. Boonnak, S. Kamonsantiroj, and L. Pipanmaekaporn, "Wavelet transform enhancement for drowsiness classification in EEG records using energy coefficient distribution and neural network," *International Journal of Machine Learning and Computing*, vol. 5, no. 4, pp. 288–293, 2015.
- [4] B.-G. Xu, A.-G. Song, and S.-M. Fei, "Feature extraction and classification of EEG in online brain-computer interface," *Tien Tzu Hsueh Pao/Acta Electronica Sinica*, vol. 39, no. 5, pp. 1025–1030, 2011.
- [5] Y. Wang, X. Li, H. Li, C. Shao, L. Ying, and S. Wu, "Feature extraction of motor imagery electroencephalography based on time-frequency-space domains," *Journal of Biomedical Engineering*, vol. 31, no. 5, pp. 955–961, 2014.
- [6] R. Aldea and M. Fira, "Classifications of motor imagery tasks in brain computer interface using linear discriminant analysis," *International Journal of Advanced Research in Artificial Intelligence*, vol. 3, no. 7, pp. 5–9, 2014.
- [7] R. K. Chaurasiya, "Statistical wavelet features, PCA, and SVM based approach for EEG signals classification," *International Journal of Electrical, Computer, Electronics and Communication Engineering*, vol. 9, no. 2, pp. 182–186, 2015.
- [8] Z. Yi, "EEG feature extraction method based on wavelet transform and sample entropy," *CAAI Transactions on Intelligent Systems*, vol. 7, no. 4, pp. 339–344, 2012.
- [9] X. Hao, W. Yang, J. Li et al., "Correlation analysis of nonlinear characteristics in EEG with CRS-R score and visual characterization of rehabilitation process in DOC patients," *Chinese Journal of Biomedical Engineering*, vol. 34, no. 2, pp. 153–159, 2015.
- [10] F. Lee, R. Scherer, and R. Leeb, "Feature mapping using PCA, locally linear embedding and isometric feature mapping for EEG-based brain computer interface," in *Digital Imaging in Media and Education, Proceedings of the 28th AAPR Workshop*, pp. 189–196, Austrian Computer Society, 2004.
- [11] Y. Pan, S. S. Ge, A. Al Mamun, and F. R. Tang, "Detection of seizures in EEG signal using weighted locally linear embedding and SVM classifier," in *Proceedings of the IEEE International Conference on Cybernetics and Intelligent Systems (CIS '08)*, pp. 358–363, Chengdu, China, September 2008.
- [12] M. Mirsadeghi, H. Behnam, R. Shalhaf, and H. J. Moghadam, "Characterizing awake and anesthetized states using a dimensionality reduction method," *Journal of Medical Systems*, vol. 40, no. 1, article 13, 2016.
- [13] X. Geng, D.-C. Zhan, and Z.-H. Zhou, "Supervised nonlinear dimensionality reduction for visualization and classification," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 35, no. 6, pp. 1098–1107, 2005.
- [14] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [15] X. Baoguo, "Feature extraction and classification of single trial motor imagery EEG," *Journal of Southeast University*, vol. 37, no. 4, pp. 629–633, 2007.
- [16] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform," *Journal of Neuroscience Methods*, vol. 123, no. 1, pp. 69–87, 2003.
- [17] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [18] L. Yingke, *The Study of Manifold Learning Algorithms and Their Applications*, University of Science and Technology of China, Hefei, China, 2011.
- [19] L. K. Saul and S. T. Roweis, "Think globally, fit locally: unsupervised learning of low dimensional manifolds," *Journal of Machine Learning Research*, vol. 4, no. 2, pp. 119–155, 2004.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

