IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 11, 2019, accepted March 30, 2019, date of publication April 3, 2019, date of current version April 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2909035

Feature Extraction and Simulation of EEG Signals during Exercise-Induced Fatigue

ZHONGWAN YANG[1] AND HUIJIE REN[2]
[1]School of Physical Education, Fuyang Normal University, Fuyang 236037, China
[2]Department of Sports Medicine, Dankook University, Tainan 31116, South Korea

Corresponding author: Huijie Ren (renhuijie1990@126.com)

This work was supported in part by the project of Philosophy and Social Science Planning of Anhui Province under Grant AHSKY2017D71, in part by the Project of University-level Undergraduate Teaching Engineering in 2017 under Grant 2017JYXM33, and in part by the Project of University Humanities and Social Science of Anhui Province in 2018 under Grant SK2018A0311.

ABSTRACT Accurate extraction of EEG signal characteristics during exercise fatigue can provide a scientific basis for sports fatigue detection and exercise fatigue injury treatment. In this paper, based on multivariate empirical mode decomposition (MEMD) and Hilbert-Huang (HHT) algorithm, feature extraction of EEG signals during exercise fatigue is performed. MEMD extends standard experience mode to multi-channel signal processing and solves traditional algorithms. It is not suitable for self-adaptability, modal aliasing, and scale alignment. It is suitable for analyzing multi-time sequence; multi-channel and multi-scale EEG signal decomposition. After the original EEG signal passes through the MEMD, the energy mean, median and standard deviation of the EEG bands in different levels are calculated and used to form the feature set. Then the support vector machine (SVM) classifier is used to classify the extract the extracted features. The simulation results show that the proposed method can effectively extract the features of EEG signals during exercise fatigue.

INDEX TERMS Exercise fatigue, EEG signal, multivariate empirical mode decomposition, Hilbert-Huang transform.

I. INTRODUCTION

Strenuous exercise or prolonged training can cause exercise fatigue in the body. Exercise fatigue not only causes peripheral fatigue, reduces the body's ability to exercise, but also causes central fatigue, leading to central nervous system homeostasis, damage to brain power, severely leading to perception and disturbance of consciousness, and triggering behavioral disorders. Therefore, how to effectively detect and evaluate exercise fatigue is a hot research topic in the cross-disciplinary fields such as neuroscience and rehabilitation engineering. The results have shown that with the increase of exercise fatigue, the activity of brain motor cortex neurons will be inhibited, and the strength of brain connection will also be weakened, which provides a theoretical basis for detecting sports fatigue through brain activity. During the process of exercise fatigue, EEG signals are composed of physiological and pathological information. Accurate analysis of EEG signals plays an important role in clinical brain research. However, when EEG signal feature extraction is performed by current methods, only EEG signals can be extracted. The frequency domain characteristics of EEG signals in a stationary state have great limitations. Under this circumstance, how to effectively extract the features of EEG signals during exercise fatigue has become a major problem to be solved [1].

Commonly used signal feature extraction and classification methods are generally considered to have linear and nonlinear points. Feature extraction methods generally include power spectrum method (PSM), common spatial mode (CSP), wavelet packet analysis, and Hilbert-Huang Transform (HHT) [2]. The methods used for feature selection mainly include principal component analysis (PCA), divergence analysis, sequential forward search (SFS), linear discriminate analysis (LDA), support vector machine (SVM), and artificial neural network method [3]. The power spectrum is a method that reflects the change of the energy of the signal as a function of frequency and indicates the frequency domain, but does not reflect the time domain. The common spatial mode (CSP) needs to input multiple leads and cannot reflect the frequency domain. Shortcomings such as information, but the correct rate is higher for the two

The associate editor coordinating the review of this manuscript and approving it for publication was Victor Hugo Albuquerque.

classification problems; although the wavelet packet analysis method (WPA) can adaptively select the corresponding frequency band and improve the time-frequency resolution, the basis function is relatively fixed, lacks adaptability, and is difficult to be better. The nonlinear features in the EEG signal are extracted. With the improvement of nonlinear science, nonlinear methods are used to process and analyze EEG signals. The Hilbert Huang algorithm is a classical nonlinear signal processing method. It decomposes signals into multiples through empirical mode decomposition. The intrinsic mode function (IMF) component is further analyzed by the Hilbert transform [4], [5].

In this paper, multivariate empirical mode decomposition and Hilbert-Huang algorithm are used to extract the EEG signals during exercise fatigue. MEMD extends the standard experience mode to multi-channel signal processing and solves the traditional algorithm. It is not suitable for self-adaptability and modal aliasing. It is suitable for analyzing multi-time series and can simultaneously process multi-channel multi-scale decomposition, and then compare the time-frequency characteristics of multi-time series at different scales. After the original EEG signal passes through the EMD, the energy mean, median and standard deviation of the EEG bands in different levels are calculated and used to form the feature set. Then the SVM classifier is used to classify the extract the extracted features.

II. BASIC PRINCIPLES OF ALGORITHM

A. MULTIVARIATE EMPIRICAL MODE DECOMPOSITION

Set an n-dimensional vector group sequence $\{\boldsymbol{v}(t)\}_{i=1}^{T} = \{v_1(t), v_2(t), \cdots, v_n(t)\}$, it represents an n-ary signal, the length of the signal sequence is T, $\boldsymbol{X}^{\theta_k} = \left[x_1^k, x_2^k, \cdots, x_{(n-1)}^k\right]$ indicates the set of direction vectors corresponding to the angle $\boldsymbol{\theta}^k = \left\{\theta_1^k, \theta_2^k, \cdots, \theta_{(n-1)}^k\right\}$ on $(n-1)$-dimensional spherical surface.

If K direction vectors is to be created on the spherical space, then $k = 1, 2, \cdots, K$. According to the above definition, MEMD processing steps are as follows [6]:

(1) Using the Hammersley sequence sampling method, obtain the appropriate direction vector of the n-dimensional space on the $(n-1)$-dimensional spherical surface;

(2) Calculate the mapping $\boldsymbol{P}^{\theta_k}$ of $v(t)$ on each direction vector of $\boldsymbol{X}^{\theta_k}$;

(3) Determine the instantaneous moment $\left\{t_i^{\theta_k}\right\}_{k=1}^{K}$ of the extreme value of the mapping signal $\left\{P^{\theta_k}(t)\right\}_{k=1}^{K}$ of all direction vectors, where i represents the extreme point position, $i \in [1, T]$;

(4) Interpolate the extreme point $\left[t_i^{\theta_k}, v(t_i^{\theta_k})\right]$ with a multivariate interpolation function yields K multivariate envelopes $\left\{e^{\theta_k}(t)\right\}_{k=1}^{K}$;

(5) For the K direction vector of the sphere space, the n-element mean $\boldsymbol{m}(t)$ is

$$\boldsymbol{m}(t) = \frac{1}{K} \sum_{k=1}^{K} e^{\theta_k}(t) \tag{1}$$

(6) Extract the intrinsic mode function $\boldsymbol{h}(t)$ through $\boldsymbol{h}(t) = \boldsymbol{v}(t) - \boldsymbol{m}(t)$; if $\boldsymbol{h}(t)$ satisfies the multivariate IMF judgment criterion, then use the result of $\boldsymbol{v}(t) - \boldsymbol{m}(t)$ as the input signal of step 2), continue with step 2) to step 6), and extract the new multivariate. IMF component $\boldsymbol{h}(t)$; otherwise, treat $\boldsymbol{h}(t)$ as the input signal of step 2) and continue with step 2) to step 6).

After a series of MEMD decomposition processes, similar to the EMD algorithm, the original n-ary signal $\{\boldsymbol{v}(t)\}_{i=1}^{T} = \{v_1(t), v_2(t), \cdots, v_n(t)\}$ is decomposed into a series of additions of the IMF component $\{\boldsymbol{h}(t)\}_{i=1}^{q}$ and the margin $\boldsymbol{r}(t)$:

$$\boldsymbol{v}(t) = \sum_{i=1}^{q} \boldsymbol{h}_i(t) + \boldsymbol{r}(t) \tag{2}$$

In equation (2), q represents the number of IMF. $\boldsymbol{h}_i(t) = \{\boldsymbol{h}_{i,1}(t), \boldsymbol{h}_{i,2}(t), \ldots, \boldsymbol{h}_{i,n}(t)\}$ and $\boldsymbol{r}_i(t) = \{\boldsymbol{r}_{i,1}(t), \boldsymbol{r}_{i,2}(t), \ldots, \boldsymbol{r}_{i,n}(t)\}$ correspond to n sets of IMF components and n margins of the n-ary signal, respectively. The IMF corresponding to each element variable of the n-ary signal is aligned on the frequency scale in n channels to form a multivariate IMF.

B. HILBERT-HUANG TRANSFORM

By performing Hilbert transformation on each IMF component, the instantaneous frequency and amplitude are obtained, and then the spectrum distribution of "time-frequency-amplitude" is obtained, which is called Hilbert-Huang spectrum. Hilbert marginal spectrum is obtained by integrating Hilbert-Huang spectrum over time. The algorithm is solved as follows [4]:

(1) Use the algorithm decomposition in 2.1 to get:

$$S(t) = \sum_{i=1}^{q} C_k(t) + r_n(t) \tag{3}$$

where, $C_k(t)$ is the k-th component of IMF, and $r_n(t)$ is residual function;

(2) The amplitude function $a(t)$ and the instantaneous frequency function $f(t)$ are obtained for each IMF component;

First, the Hilbert transform is performed on the IMF component $c(t)$ to obtain $H[c(t)]$;

Second, solve the parsing signal of $c(t)$, that is

$$z(t) = \boldsymbol{c}(t) + jH[\boldsymbol{c}(t)] = \boldsymbol{a}(t)\, e^{i[\varphi(t)]} \tag{4}$$

$a(t)$ is amplitude function: $a(t) = c_2(t) + H_2[c(t)]$); (t) is the phase function: $(t) = \arctan(H[c(t)]/c(t))$;

Finally, the instantaneous frequency of $c(t)$ is calculated using the phase function,

$$f(t) = \frac{1}{2\pi} \frac{dt}{d} [\varphi(t)] \tag{5}$$

(3) After Hilbert transform is performed on all IMF components, the Hilbert-Huang spectrum is obtained:

$$\begin{aligned} f(f, t) &= Re \sum_{i=1}^{n} \boldsymbol{a}_i(t)\, \boldsymbol{e}^{[j\varphi_i(t)]dt} \\ &= Re \sum_{i=1}^{n} \boldsymbol{a}_i(t)\, \boldsymbol{e}^{[j2\pi \int f_i(t)dt]} \end{aligned} \tag{6}$$

where Re represents the real part, a_i is the amplitude of the i-th IM F component in time; $f_i(t)$ is the instantaneous frequency of the IMF component;

(4) The Hilbert marginal spectrum is the integral of the Hilbert-Huang spectrum time, and the Hilbert marginal spectrum is obtained:

$$h(f) = \int_{-\infty}^{+\infty} H(f, t)dt \tag{7}$$

C. SUPPORT VECTOR MACHINE

Support vector machine is a pattern recognition method based on statistical learning theory. It has high generalization ability and global optimality of solution. It solves small sample problems, nonlinear problems and high-dimensional data. It has many unique advantages and has been widely used in many problems such as processing prediction, data fitting, comprehensive evaluation and pattern recognition [7].

Suppose training set $D = \{(X_1, y_1), (X_2, y_2), \ldots, (X_n, y_n)\}$, $X_i \in R^n$ is the i-th sample of the input mode. $y_i \in \{-1, +1\}$ is the category label corresponding to the sample (-1 means negative sample, +1 means positive sample), and the total number of training samples is n. Set the equation of the SVM's classification hyper plane L is:

$$\boldsymbol{D} \cdot \boldsymbol{X} + b = 0 \tag{8}$$

where W is the normal vector of hyper plane L and b is the constant term of hyperplane L.

$$\boldsymbol{D} \cdot \boldsymbol{X}_i + b = 1, \quad y_1 = 1 || \boldsymbol{D} \cdot \boldsymbol{X}_i + b = -1, \quad y_1 = -1 \tag{9}$$

$\boldsymbol{X}_i$ in the sample point (X_i, y_i) satisfying the equation (8) is called a support vector.

Assuming that there are sample points $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ as support vectors, the interval between positive and negative samples is as follows:

$$dist = \frac{\boldsymbol{W}}{\| \boldsymbol{W} \|} \cdot (\boldsymbol{X}_1 - \boldsymbol{X}_2) = \frac{2}{\| \boldsymbol{W} \|} \tag{10}$$

Equation (10) represents the projection of the vector $(\boldsymbol{X}_1 - \boldsymbol{X}_2)$ on the unit normal vector $\boldsymbol{W}$ of the hyper plane L. It can be seen that solving the optimal classification hyper plane problem is equivalent to solving the maximum interval, and transforming equation (10) into:

$$min \frac{1}{2} = \| \boldsymbol{W} \|^2, s.t. y_i (\boldsymbol{W} \cdot \boldsymbol{X}_i + b) \geq 1, \quad \forall i = 1, 2, \cdots, n \tag{11}$$

Therefore, for solving the optimal classification hyper plane of support vector machine, it is finally transformed into a solution of a quadratic optimization problem. The advantage of SVM is that it is suitable for a small number of sample data and can solve high dimensional problems. The disadvantage is that when the amount of data is large, the computing resources consumed are increased, and the calculation speed is slow. At this point, the computational performance of the logistic regression classification algorithm may be better than the classification algorithm.

III. EEG DATA COLLECTION

A. QUADRICEPS MAXIMAL VOLUNTARY CONTRACTION (MVC) EXPERIMENTAL PROTOCOL

In order to achieve exercise fatigue for the purpose of testing and analysis of related EEG signals, the maximum free contraction experiment of quadriceps was designed in this work. The specific paradigm is [8]: the subject sits on the muscle chair and the knee joint is naturally bent at 90° C, then the subject makes the knee extension exercise with maximum muscle strength, and the muscle force curve is input into the multi-channel bioelectricity through the force sensor. The recorder obtained individual MVC and then calculated the subject's 1/3 MVC value. Subjects used 1/3 MVC for static sustained isometric contraction, and 1/3 MVC of static force self-regulated muscle strength through video feedback until after the verbal encouragement, the established strength was not considered fatigue. A dynamic exercise test was conducted one day after the completion of the static exercise test. During the test, the subject performed continuous knee extension with the speed of the metronome according to the frequency of the metronome. The movement was performed by contracting the muscles at a metronome frequency and 2 seconds of continuous uninterrupted movement. Until the subject's maximum muscle strength, which was unable to maintain the video feedback for three consecutive times after oral encouragement, was fatigue and the test was terminated. The subject should wear an EEG signal acquisition device throughout the procedure and fill in the subjective fatigue scale immediately after termination.

B. EXPERIMENTAL OBJECTS

In this experiment, 20 subjects were recruited, including 10 males and 10 females, with an age range of 25±3 years. The specific recruitment targets must meet the following requirements: (1) have a good life routine; (2) no history of brain disease (brain trauma, epilepsy, etc.), no history of mental illness, no history of drug abuse; (3) two-legged bones And his muscles have no history of past injuries; (4) Do not stay up late, take medicine, drink alcohol one week before the experiment, do not smoke, drink tea or coffee 8 hours before the experiment, scalp cleansing and emotional relief 2 hours before the experiment, peace of mind; (5) All subjects need to understand the whole process of the experiment before the experiment and voluntarily participate in all aspects of the experiment.

C. EXPERIMENTAL EQUIPMENT AND DATA RECORDS

In this experiment, the EEG data of the motor fatigue process of the subjects of Neuroscan's EEG signal acquisition system were collected and analyzed. Using the international standard 10-20 lead scalp electrode system, a total of 19 channels of EEG signals were recorded (Fp2, Fp1, F8, F7, F4. F3, T4, T3, C4, C3, T6, T5, P4, P3, O2, O1, Fz, Cz and Pz), covering the frontal, parietal, central, and occipital lobe of the brain (Figure 1). The electrode of the electric cap is

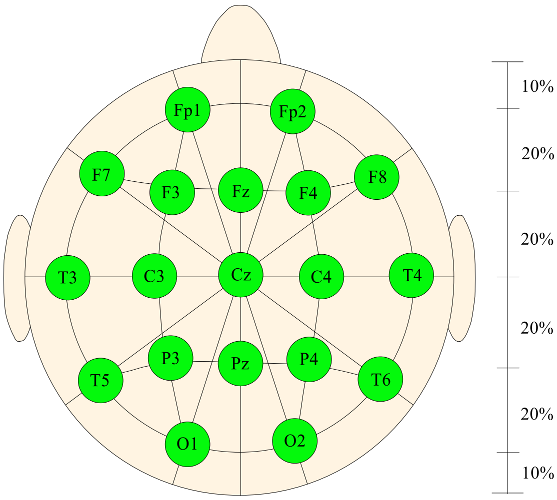

FIGURE 1. Disposition of electrodes on the scalp for EEG signal acquisition.

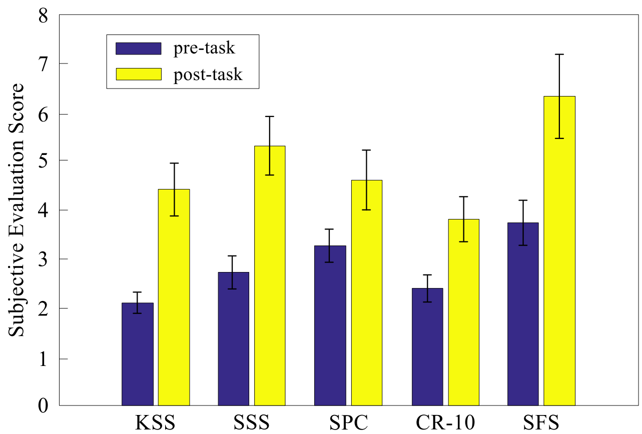

FIGURE 2. Subjective evaluation of exercise fatigue before and after the MVC task.

a silver/silver chloride (Ag/AgCl) alloy electrode which is not affected by the polarization potential, and the reference electrode is placed on the bilateral mastoid line; the contact resistance between the pole and the scalp is less than 5 kΩ, the sensor The bandwidth is 0.2-45 Hz and the sampling frequency is 128 Hz, respectively [9]. The experiment is carried out in a normal indoor lighting, quiet and comfortable environment. The experimental final data were analyzed by SPSS 13.0 software. The subjective fatigue scale scores and EEG complexity parameters of the subjects before and after the maximum voluntary contraction of the quadriceps were tested by paired test.

IV. SIMULATION RESULTS AND ANALYSIS

A. SUBJECTIVE EVALUATION OF EXERCISE-INDUCED FATIGUE

Subjective sleepiness was assessed using the Stanford Drowsiness Scale (SSS) and the Karolinska Sleepiness Scale (KSS), using the Samn-Perelli Scale (SPC), Subjective Fatigue Rating Scale (SFS), and Borg Rating Scale (CR-10) [10]. The comparison results of several subjective scale scores before and after the experimental task are shown in Figure 2.

It should be noted that in the fitness function based on the classification error rate of kNN, the values are different, and the selected features are finally selected. The number is also different. Experiments show that k takes a value of about 3, and the result set fluctuates around 20% in the original feature set. The characteristics of this ratio make the output of the subsequent classifier better. Here, with the fitness function kNN algorithm parameter $k = 3$ (selected feature index 18%) as an example to illustrate the effectiveness of genetic algorithm for feature selection. After feature selection, the dimension of the feature set is significantly reduced. In order to verify whether the features selected by the genetic algorithm have obvious reparability, and verify the effect of genetic algorithm for feature selection, this paper inputs the selected features into SVM, Fisher linear discriminate classifier and Mahalanobis distance linearity. Discriminate the classifier and obtain corresponding results. The three classifiers performed 10 repeated classification experiments. The ratio of the training set to the test set was 8:2. The results are shown in Figure 2. The number of repetitions is 11, and the corresponding classification accuracy rate represents 10 classification results.

Subjective self-evaluation results showed that the subjects were neither fatigue nor sleepy before the experimental task, but moderate to extreme fatigue and drowsiness after completing the experimental task. Compared with before the experiment, the subjective scale score increased significantly after completing the experimental task ($P<0.005$), indicating that the continuous long-term quadriceps MVC task causes an increase in fatigue and drowsiness.

B. FEATURE EXTRACTION

EEG is a signal waveform that amplifies and records the electrical activity of the brain. It can be divided into four basic rhythms according to frequency: δ wave, θ wave, α wave and β wave. For normal adults, the frequency of the delta wave is 0.5-3 Hz and the amplitude is 20-200 μV, indicating that the brain is in a state of no dream deep sleep; the frequency of the θ wave is 4-7 Hz, and the amplitude is 20-100 μV, indicating the brain. In a state of relaxation (such as first nap or waking) or drowsiness and sleepiness; the alpha wave has a frequency of 8-13 Hz and a amplitude of 10-100 μV, which is a normal waveform in a quiet closed state, and is most obvious in the top and occipital regions. The largest number, disappears when deep sleep or excitement; the frequency of beta wave is 14-30 Hz, the amplitude is 5-25 μV, indicating that the brain is awake (with efforts to reach attention), in the forehead, sputum, central The area is most obvious. Therefore, by detecting the frequency information appearing in the EEG, it is possible to judge the change of the level of consciousness of the person well.

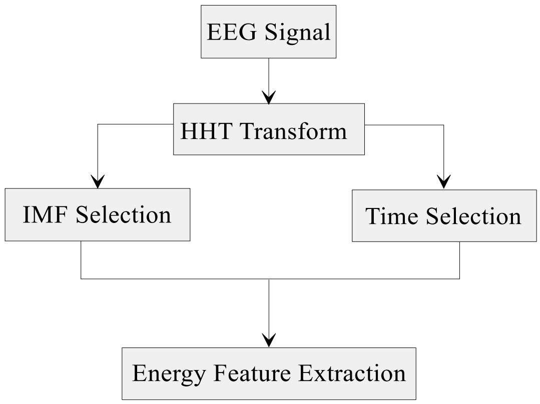

FIGURE 3. Flow chart of energy feature extraction algorithm based on HHT.

In the analysis of EEG signals, if some band energy of the signal duration is used as the characteristic of EEG signals, it does not take into account the fact that EEG signal energy changes with time, and some information about the EEG signal at time and frequency were more or less ignored. The EEG signal can be easily obtained by the Hilbert-Huang transform to obtain the distribution of time-frequency-amplitude in three dimensions. Based on this characteristic of Hilbert-Huang transform, the feature extraction method of motion imaging EEG signals is extracted by using energy feature extraction method considering time and frequency information. Figure 3 shows the flow of an energy feature extraction algorithm based on the Hilbert-Huang transform. The algorithm mainly includes Hilbert transform, IMF selection, feature book tip selection, energy feature extraction and other links.

The EEG signal can be expressed as a series of IMF sums after EMD decomposition, so IMF with most useful frequency information can be selected to improve the efficiency and classification accuracy of feature extraction. In this section, the IMF containing most of the μ rhythm and β rhythm is selected by the following method. The IMF selection steps are as follows:

(1) The EEG method is used to decompose the EEG signal into the sum of IMF, and then the Hilbert spectrum Hi(w, t) of each IMF is obtained by Hilbert transform, and the energy of 8-30 Hz band in each order IMF is calculated.

$$E_i = \int_{\omega_1}^{\omega_2} \int_{0}^{T} |H_i(\omega, t)|^2 dt d\omega \tag{12}$$

Here, ω_1 and ω_2 are the upper and lower limits of frequency, 8 Hz and 30 Hz, respectively, $i = 1, 2, \ldots, N$ is the number of IMF.

(2) Then calculate the ratio P of the energy in the 8-30 Hz frequency band in each IMF, and obtain the energy distribution of the frequency band, and select the IMF where the energy of the 8-30 Hz frequency band is located:

$$P_i = \frac{E_i}{\sum_{k=1}^{N} E_k}, \quad i = 1, 2, \cdots, N \tag{13}$$

In the study of exercise fatigue signals, the fatigue phenomenon cannot continue throughout the entire signal duration. If the motion imaging signal of the whole duration is taken as the research object, it may contain a lot of noise information, which affects the feature extraction efficiency and classification accuracy. Therefore, this paper finds the imaginary signal with obvious fatigue phenomenon by selecting the time period with the largest energy difference between the C3 and C4 lead characteristic bands (8-25 Hz):

(1) The motion imaging signals of the C3 and C4 leads are respectively decomposed by MEMD to obtain multiple sets of IMF component sums;

(2) The IMF component selection method is used to select a relatively large IMF component of the characteristic frequency band, and the instantaneous frequency is calculated;

(3) Calculate the characteristic band energy of the selected IMF component with a sliding time window length of m s and specify a sliding step size of n s:

$$E_{3i}(k) = \int_{F_1}^{F_2} \int_{(k-1)n}^{(k-1)n+m} |F_3(t,f)|^2 dt df \tag{14}$$

$$E_{4i}(k) = \int_{F_1}^{F_2} \int_{(k-1)n}^{(k-1)n+m} |F_4(t,f)|^2 dt df \tag{15}$$

Calculate the energy difference between C3 and C4 corresponding to the time window, and use the corresponding time period when the energy difference is maximum as the best time period:

$$E_{34i}(k) = |E_{3i}(k) - E_{4i}(k)|, \quad k = 1, 2, \cdots, K \tag{16}$$

C. SVM CLASSIFICATION

In order to further distinguish the normal state and the fatigue state, after extracting the Hilbert EEG feature, the SVM model is used to classify the EEG data segments before and after the experiment, and the kernel function selects the Gaussian kernel function. Due to the limited number of EEG data segments in this experiment, 5 double cross-validations were applied in the analysis of this paper. For each subject, 80% EEG data segment was randomly selected as the training part, and the remaining 20% was used for the classifier test. The original training data set is equally divided into k disjoint subsets, each subset is taken as a test data set in turn, and the remaining k-1 group subsets are used as training data sets. After taking turns k times, the classification correctness rate of k models and k test data sets under the model will be obtained, and the average accuracy of all classifications will be averaged as the combined performance index of the features and classifiers under the k-segment cross-validation. This indicator can verify the performance of the classifier and verify that the combination of different feature states and classifiers will result in a better classification effect for the same classifier. This article uses cross-validation. A feature selects different feature states. The average training data was used to classify the test data. For the original SVM classifier, the parameter σ ranges from 0.5 to 3, and the classification

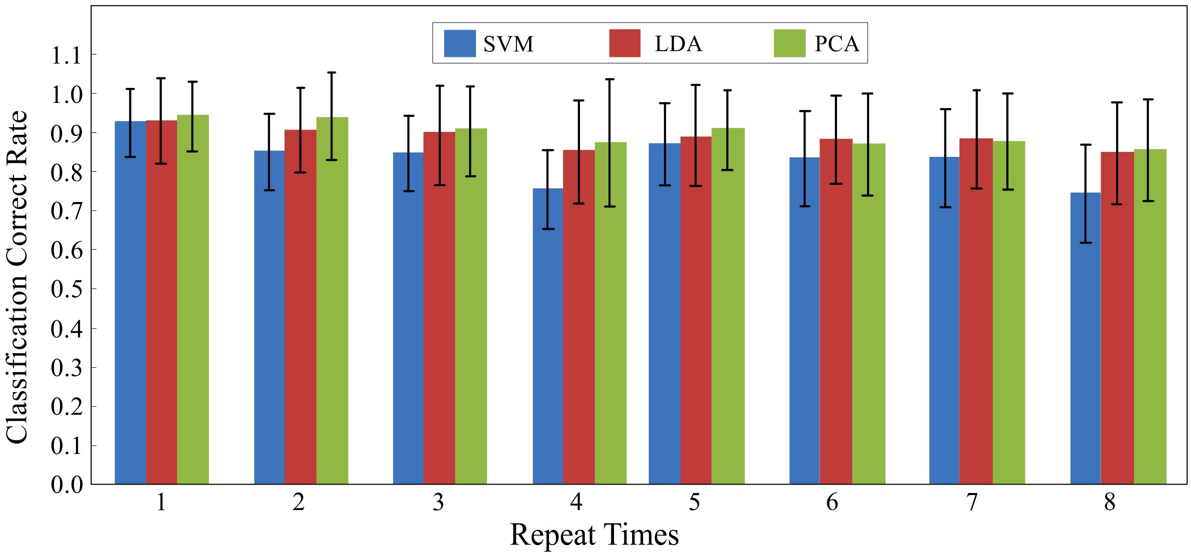

FIGURE 4. Classification results after future extraction using MEMD and HHT.

accuracy rate is calculated for each σ value. The specific steps are as follows [11]:

(1) Selected training sets and test sets for EEG samples. The processed fatigue EEG feature data randomly divides the training set and the test set according to a ratio of 3:2;

(2) Normalization. There are some abnormal data points in the EEG data. In order to make these abnormal data points not affect our classification results, normalization is used;

(3) Train the SVM classifier with the training set to obtain the classification model. In the training process, 10 cross-validation methods are used to obtain the optimal penalty factor. The penalty factor C is 0.5 in this paper. Import the test set to test the trained model. The resulting classification accuracy is the average of 5 test results. The classification result of the SVM is shown in Figure 4. The figure shows that the SVM algorithm can effectively distinguish the two brain fatigue states. When the kernel parameter σ is equal to 2.7, the maximum classification accuracy rate is obtained.

D. HILBERT-HUANG PARAMETER ANALYSIS

After MEMD decomposition, modal components are obtained. For example, the decomposition results of EEG signals before and after exercise in a single subject are shown in Figure 5. The center frequency of each IMF component reflects the distribution of the energy of the modal function, characterizing the frequency of the signal component with higher energy in the modal function, and the center frequency of the EEG signal after reconstruction of all modal functions is characterized by different conditions migration of the energy center of gravity of the entire EEG. When awake, the cortex is in an excited state, and the EEG is mainly composed of high frequency and low amplitude waves. When the degree of central fatigue increases, the degree of inhibition of the cerebral cortex increases; the high frequency components in the EEG decrease and the low frequency components increase. Experiments show that as the degree of fatigue deepens, the IMF component and the energy center of gravity of the entire EEG migrate toward the low frequency, and the frequency of the signal component with a larger energy component in the signal decreases.

The main advantage of the period gram as the power spectrum estimation is the convenience of calculation. The main disadvantage is that the frequency resolution is low. This is because the period gram method considers the observed finite length of N data to be zero in the calculation. This is obviously the fact. Considering the unobserved value as zero is equivalent to multiplying x(n) by a rectangular window function in the time domain, which is equivalent to expanding a sin c function convolved with it in the frequency domain, due to the sin c function and There are two differences between the δ function. First, the main lobe is not infinitely narrow and has side lobes, so the convolution result must be distorted.

As a nonlinear feature extraction method, Kernel Principal Component Analysis (KPCA) has proved to be a very necessary step in data preprocessing of classification algorithms. This method first maps the input vector to a feature space F through the nonlinear mapping function φ. The maximum entropy power spectrum estimation method is to extrapolate data other than the known finite-length autocorrelation sequences, instead of treating them as zero. There are many extrapolation methods to ensure that the autocorrelation function matrix is positive. Burg believes that the extrapolated autocorrelation function should make the time series exhibit the maximum entropy, so it is called the maximum entropy power spectrum. And then perform linear principal component analysis (LPCA) in the feature space. KPCA uses the kernel function that satisfies the Mercer condition instead of the inner product of the feature space, so that it is not necessary to know the exact form of the nonlinear transformation accurately; since the mapping function can be nonlinear

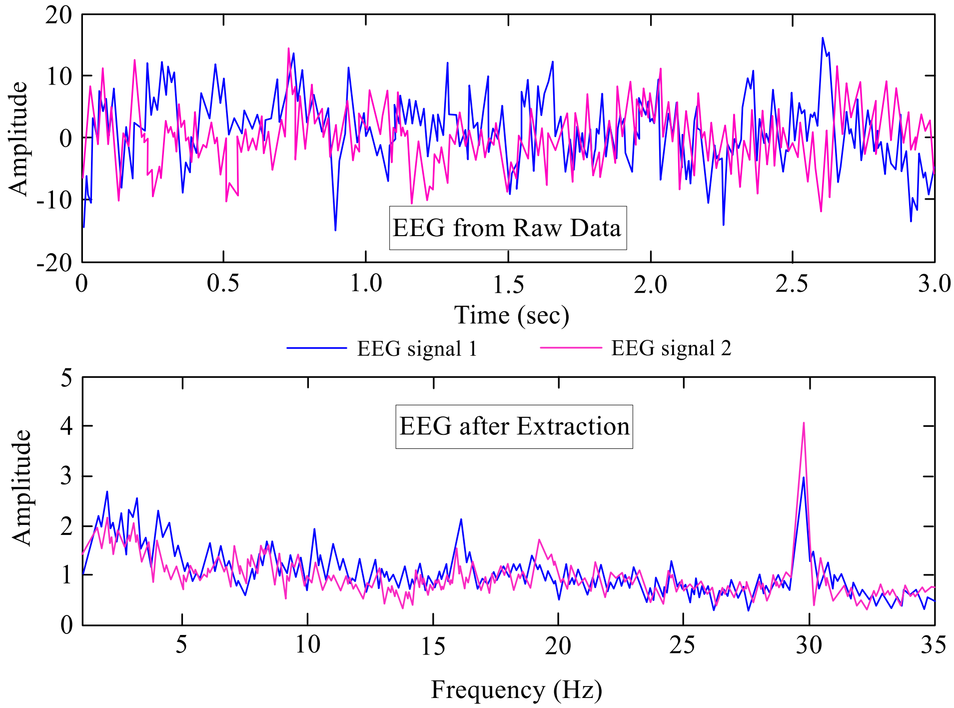

FIGURE 5. Schematic diagram of the decomposition of EEG signals before and after exercise.

Therefore, by selecting the appropriate kernel function and kernel parameters, the effective main feature components can be obtained. KPCA can not only reduce the dimension of the input vector, but also improve the generalization ability of the classifier and speed up the classification.

In both states, the Hilbert marginal spectrum of the EEG signal is significantly different. In the normal movement, the α wave of 8-13 Hz and the β wave of 14-30 Hz appear mainly, while the frequency of occurrence of fatigue motion is mainly 5-7 Hz θ wave and 8-13 Hz α wave. The Hilbert yellow algorithm is used to extract the Hilbert marginal energy values of the EEG parameters θ, α, β, $\beta \cdot \alpha$, $(\alpha + \theta)/\beta$ in the EEG signals of the subjects. The disappearance of the beta wave in the EEG rhythm indicates that the subject is not conscious, and the appearance of the slow wave (θ wave) is a manifestation of sleepiness and mental inhibition, consistent with the subjective fatigue scale recorded during the trial, that is, Hilbert marginal spectrum is a good indicator of the state of the subject.

From the center frequency of the IMF component, the energy center of gravity of IMF1 and IMF2 is in the frequency range of β and α brain electrical rhythm, while the energy centers of IMF4 and IMF5 are in the frequency range of the delta brain electrical rhythm. The energy ratio analysis results of each IMF component reflect the changes of brain electrical rhythm in different central fatigue states from another aspect (Figure 6). It is generally believed that β high-frequency fast waves are the main waveforms of the cerebral cortex when excited and δ high-amplitude slow waves are the main manifestations of electrical activity when the cerebral cortex is in a state of inhibition [12], [13]. In the awake state, the brain is highly excited, EEG is characterized by low amplitude fast waves, and β wave accounts for a large proportion; as the central fatigue deepens, the brain's alertness and excitability decrease, and thus the δ in the EEG signal the slow wave specific gravity increases, while the beta wave specific gravity decreases, and the EEG exhibits a high amplitude.

The EEG signal is generated by the potential change of the brain nerve cells after stimulation, and can be detected at the scalp by a dedicated sensor. The EEG signal is a low signal-to-noise ratio non-stationary random signal, and the components involved are very complex. It needs to be processed and analyzed in order to meet the requirements of diagnosis or research. When the traditional signal processing method analyzes the thinking EEG, it needs to be framed and windowed first, that is, a finite-length window sequence is used to intercept an EEG signal for analysis. The purpose of framing is to ensure that the data is as smooth as possible, and the choice of frame length and window function directly affects the results of the analysis. Since the Herbert-Huang T method is suitable for the analysis and processing of nonlinear non-stationary data, it does not make short-term stationary assumptions for EEG signals, so there is no need to frame and window the EEG signals in the application.

The IMF component of the C4 leads to EEG signal after MEMD in both states. The number of layers of the IMF component is different in the two states, which is the result of the MEMD adaptive decomposition, indicating that the EEG signal contains different frequency components in different

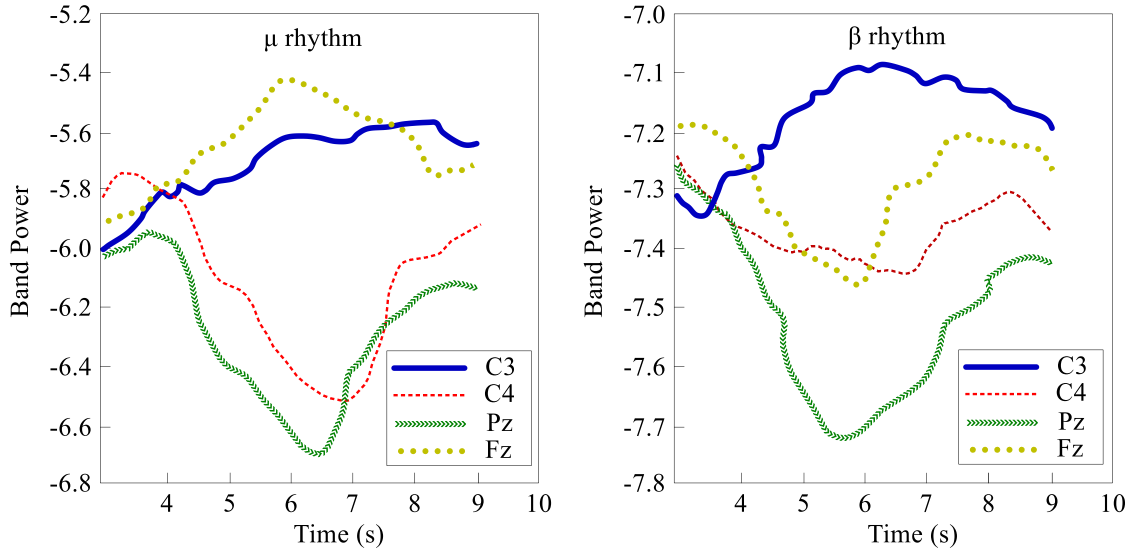

FIGURE 6. Average band energy at C3 and C4 under μ and β rhythm.

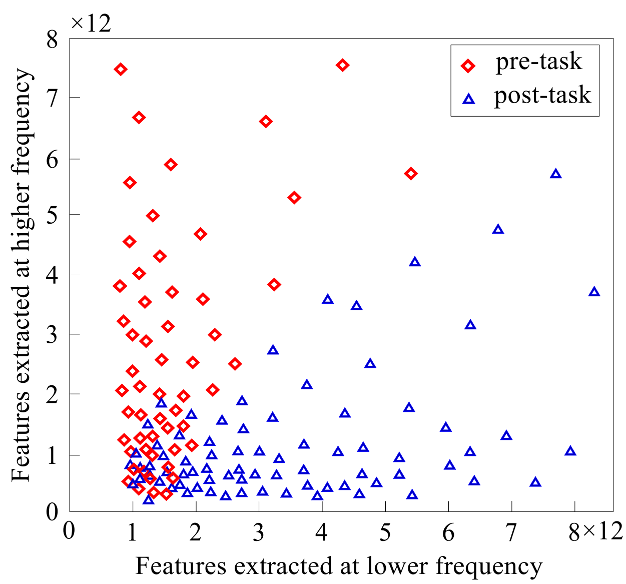

FIGURE 7. Feature extracted before and after exercise at different frequency.

states [14], [15]. In both states, the Hilbert marginal spectrum of the EEG signal is significantly different (Figure 7). The main occurrence of normal driving is the α wave of 8-13 Hz and the β wave of 14-30 Hz, while the frequency of the main occurrence of fatigue driving is 5-7 Hz θ wave and 8-13 Hz α wave Compared with the normal driving state, the θ wave increases, and the β wave decreases (almost approaches 0). The disappearance of the beta wave in the EEG rhythm indicates that the driver's consciousness is not clear. The appearance of the slow wave (θ wave) is a manifestation of sleepiness and mental inhibition, which is consistent with the subjective fatigue scale recorded during the test, that is, the Hilbert margin. The spectrum is a good indicator of the state the driver is in.

In general, the first few IMF components decomposed by the EMD method tend to concentrate the most significant and important information in the original signal. By spectral analysis of the IMF component obtained by decomposition, the first five IMF components in this experiment concentrate the main energy of the EEG signal [16], [17]. In order to further study the relationship between the frequency of each IMF component and the energy characteristics and the central fatigue, the center frequency and the energy ratio parameter of each modal function of each EEG data segment are calculated. In order to eliminate the influence of accidental fluctuations of each parameter, the average value of each parameter within 1 min was calculated and statistical analysis was performed. The central frequency of the EEG signals reconstructed by all modal functions under different central fatigue states was statistically analyzed. The analysis results are shown in Figure 8.

The EEG signal is generated by the potential change of the brain nerve cells after stimulation, and can be detected at the scalp by a dedicated sensor. The EEG signal is a low signal-to-noise ratio non-stationary random signal, and the components involved are very complex. It needs to be processed and analyzed in order to meet the requirements of diagnosis or research. The first order IMF of the signal contains a large amount of local high frequency components of the signal. As the order increases, the high-frequency components of the contained signal gradually decrease, so the EMD decomposition has the characteristics of frequency modulation and amplitude modulation. The IMF function obtained by EMD transform gives a description of the characteristics of EEG signals, and also provides a basis for establishing a time-varying model of nonlinear and non-stationary features of EEG signals. Different data segments are decomposed to different IMF orders, which mean that the EMD decomposition

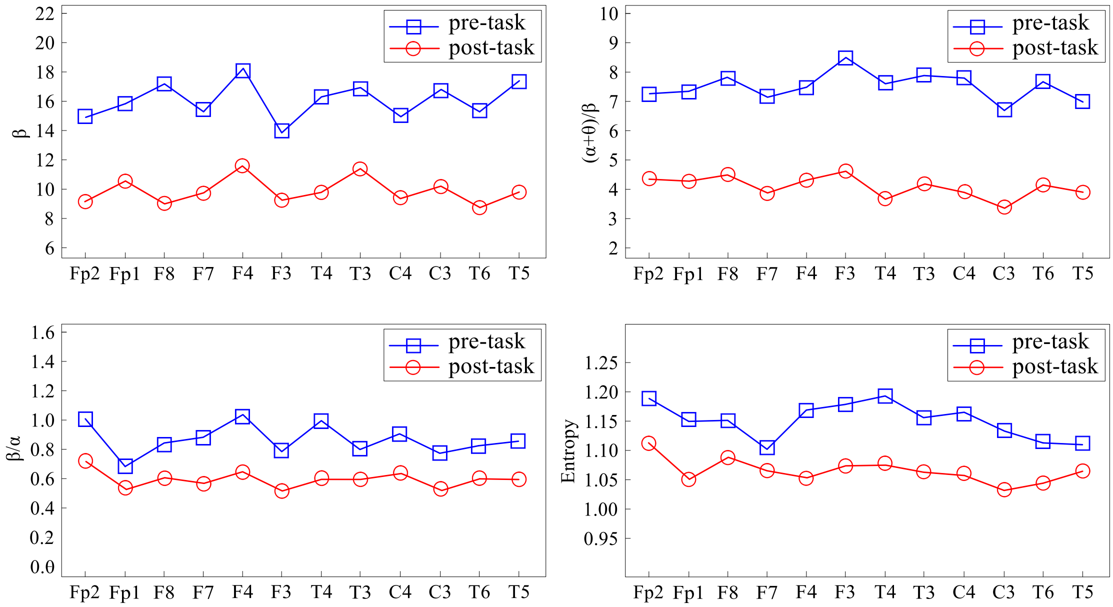

FIGURE 8. Marginal spectral energy curve before and after exercise fatigue.

is adaptive. For different signals, the number of IMF components decomposed and the frequency range corresponding to each IM F component are unknown before decomposition. But because of the high frequency components, it reflects the intrinsic characteristics of the signal.

When the traditional signal processing method analyzes the thinking EEG, it needs to be framed and windowed first, that is, a finite-length window sequence is used to intercept an EEG signal for analysis. The purpose of framing is to ensure that the data is as smooth as possible, and the choice of frame length and window function directly affects the results of the analysis. Since the HHT method is suitable for the analysis and processing of nonlinear non-stationary data, it does not make short-term stationary assumptions for EEG signals, so there is no need to frame and window the EEG signals in the application.

V. CONCLUSIONS

This paper proposes a method based on multivariate empirical mode decomposition for analyzing EEG signals. The multivariate empirical mode decomposition is used to extract the corresponding marginal spectral features from the motor fatigue EEG signals, and the SVM classifier is used to classify the motor image EEG signals. MEMD not only can effectively solve the modal aliasing and scale alignment problems of EMD, but also has better adaptability and time-frequency localization ability. The simulation results confirm that the correct recognition rate of motion imaging EEG signals based on MEMD feature extraction can reach 90%, which is significantly higher than the correct recognition rate using traditional feature extraction algorithm. Therefore, this method has research application value in BCI and other biomedical signal processing.

MEMD extends the standard experience mode to multichannel signal processing and solves the traditional algorithm. It is not suitable for self-adaptability and modal aliasing. It is suitable for analyzing multi-time series and can simultaneously process multi-channel multi-scale decomposition. An important part of HHT analysis is empirical mode decomposition. EMD is an adaptive decomposition method, which is not affected by human factors. Therefore, the decomposition is more objective, and the function (IMF) component obtained from it is also it can reflect the characteristics of the EEG signal very well. As a time-frequency analysis method, it is not bound by the Fourier transform and has good time-frequency resolution. In view of the above advantages, this paper applies it to the state detection of sports fatigue, and has achieved good results.

REFERENCES

[1] M. X. Gao, "Sports fatigue in the electrical signal feature extraction simulation," *Comput. Simul.*, vol. 34, no. 5, pp. 277–280, 2017.
[2] S. L. Duan, Y. K. Shang, and L. Z. Pan, "Feature extraction and classification of multi-class motor imagery EEG data," *Comput. Meas. Control*, vol. 24, no. 2, pp. 283–287, 2016.
[3] X. Q. Niu, Q. W. Ye, Y. Zhou, and X. D. Wang, "Autoregressive model electroencephalogram signal identification based on feature selection of genetic algorithm," *Comput. Eng.*, vol. 42, no. 3, pp. 283–288, 2016.
[4] J. Nan, L. Ai, and J. Shen, "Application of HHT to driving fatigue in EEG analysis," *J. Biomed. Eng.*, vol. 28, no. 4, pp. 653–657, 2011.
[5] J. Z. Liu, Y. Xie. X. Q. Chen, and Y. Wu, "Feature extraction of visual fatigue EEG signals based on HHT," *Chin. J. Med. Phys.*, vol. 35, no. 12, pp. 1473–1478, 2018.
[6] Y. Zhang, C. Y. Zhou, and Y. Luo, "Feature extraction and analysis of imaginary movements in EEG based on MEMD," *J. Chongqing Univ. Posts Telecommun. (Natural Sci. Ed.)*, vol. 27, no. 3, pp. 386–391, 2015.
[7] L. M. Zhao and X. J. Zhu, "Feature extraction and classification of electroencephalogram signals based on local mean decomposition and sample entropy," *Comput. Eng.*, vol. 43, no. 2, pp. 299–303, 2017.
[8] L. Zhang and B. Zhou, "Characteristics on change of EMG parameters and EEG α index before and after muscle fatigue at different types of muscle exercise," *J. Wuhan Inst. Phys. Educ.*, vol. 44, no. 4, pp. 65–69, 2010.

[9] N. N. Zhang, H. Wang, and R. R. Fu, "Feature extraction of fatigued driver's electroencephalogram signals based on wavelet entropy," *Automot. Eng.*, vol. 35, no. 12, pp. 1139–1142, 2013.
[10] C. Zhang, C. X. Zheng, and X. L. Yu, "Brain fatigue estimation based on multiple psychophysiological parameters and nuclear learning algorithms," *Sci. China*, vol. 53, no. 12, pp. 1403–1413, 2008.
[11] T. Cao, F. Wan, C. M. Wong, J. N. da Cruz, and Y. Hu, "Objective evaluation of fatigue by EEG spectral analysis in steady-state visual evoked potential-based brain-computer interfaces," *Biomed. Eng. Online*, vol. 13, no. 28, pp. 1–13, 2014.
[12] C. Liu, H. B. Zhao, C. S. Li, and H. Wang, "Research on band power extraction and classification of EEG signal," *J. Syst. Simul.*, vol. 24, no. 12, pp. 2496–2499, 2012.
[13] H. L. Jin and Z. H. Zhang, "Research of movement imagery EEG based on Hilbert-Huang Transform and BP neural network," *J. Biomed. Eng.*, vol. 30, no. 2, pp. 249–253, 2013.
[14] S. N. Carvalho *et al.*, "Comparative analysis of strategies for feature extraction and classification in SSVEP BCIs," *Biomed. Signal Process. Control*, vol. 21, pp. 34–42, Aug. 2015.
[15] H. U. Amin *et al.*, "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," *Australas. Phys. Eng. Sci. Med.*, vol. 38, no. 1, pp. 139–149, 2015.
[16] A. K. Jaiswal and H. Banka, "Local pattern transformation based feature extraction techniques for classification of epileptic EEG signals," *Biomed. Signal Process. Control*, vol. 34, pp. 81–92, Apr. 2017.
[17] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," *Biomed. Signal Process. Control*, vol. 31, pp. 398–406, Jan. 2017.

ZHONGWAN YANG was born in Anhui, China, in 1975. He received the bachelor's degree from Anhui Normal University, in 1999, and the master's degree from Fujian Normal University, in 2009. Since 2010, he has been an Associate Professor with the School of Physical Education, Fuyang Normal University. He has published nine papers, one book, and three textbooks. His research focuses on the physical education and sports training.

HUIJIE REN was born in Shanxi, China, in 1990. He received the bachelor's degree from Shanxi Normal University Sports Institute, in 2012, and the master's degree from the Postgraduate College of Shanxi Normal University, in 2015. He is currently pursuing the Ph.D. degree with the General College, University of Dankook, South Korea. He has published a total of five papers in China and South Korea.