

Combination of Variational Mode Decomposition for Feature Extraction and Deep Belief Network for Feature Classification in Motor Imagery Electroencephalogram Recognition

Wulin Zhang, Zeyu Liang,* and Zirui Liu

College of Physics and Electronics Engineering, Mudanjiang Normal University,
Heilongjiang 157011, China

(Received July 1, 2021; accepted August 27, 2021)

Keywords: brain–computer interface, motor imagery, variational mode decomposition, deep belief network, high-dimensional feature classification, high-dimensional feature extraction

Brain–computer interface (BCI) technology based on motor imagery (MI) can establish the connection between the brain and the outside world; this has gradually become an important application of human–machine hybrid intelligence enhancement, especially for medical rehabilitation treatment. Because of the nonlinear, nonstationary, and low signal-to-noise ratio (SNR) characteristics of electroencephalogram (EEG) signals, it is a great challenge to accurately classify MI-EEG signals. Toward this end, in this study, both variational mode decomposition (VMD) and a deep belief network (DBN) were applied to MI classification based on the dataset of a previous BCI competition. Firstly, the EEG signal was decomposed by VMD to obtain the narrow-band component. Then the marginal spectrum, the instantaneous energy spectrum under the characteristic frequency band, and time-frequency joint features were extracted by Hilbert transform to achieve feature fusion. Finally, the DBN was used to reduce the dimensions of high-dimensional features and recognize MI patterns. The experimental results show that the joint VMD feature extraction and DBN feature classification method avoids information omission caused by the manual optimization of the period and the frequency band of artificial determination imagery. On the other hand, the method of using VMD and DBN to automatically extract the characteristics of the optimal period and optimal frequency band can effectively improve the recognition rate of MI.

1. Introduction

A brain–computer interface (BCI) system operates an output device by analyzing the input electrophysiological signals and decoding the user’s intention into control instructions.⁽¹⁾ In accordance with the different forms of signals, an electroencephalogram (EEG) can be divided into electroencephalograph, magnetoencephalogram (MEG), and functional magnetic resonance imaging (fMRI) signals, where EEG signals are widely used in BCI systems due to their noninvasiveness and low cost.^(2,3)

*Corresponding author: e-mail: llllzy321@126.com
<https://doi.org/10.18494/SAM.2021.3514>

The principle of motor imagery (MI) is to generate event-related desynchronization/synchronization phenomena for EEG signals on both sides of the brain during imagination-induced movement, which has become a research direction in the field of EEG signals.⁽⁴⁾ EEG signals are, for example, weak, nonlinear, nonstationary, and sensitive in time; thus, time-frequency domain analysis and spatial filtering are widely used in feature extraction.⁽⁵⁾ Time-frequency domain analysis mainly includes the short-term Fourier transform (STFT),⁽⁶⁾ wavelet transform (WT),⁽⁷⁾ wavelet package transform (WPT),⁽⁸⁾ and common spatial pattern (CSP) method.⁽⁹⁾ However, the time-frequency domain analysis methods based on STFT, WT, and WPT cannot achieve high resolution in the time and frequency domains. However, with a change in brain state, the characteristics of an EEG signal in the time-frequency domain will fluctuate. Chen *et al.* combined Shannon entropy, wavelet entropy, and sample entropy for feature extraction.⁽¹⁰⁾ Tan *et al.* proposed a multifeature extraction method based on ensemble empirical mode decomposition (EMD) and approximate entropy.⁽¹¹⁾

In summary, the above methods show good adaptability and high recognition accuracy, consider the fusion of features from a single angle, and cannot obtain a more complete description of the signal from multiple angles. However, variational mode decomposition (VMD) can decompose an EEG signal into several narrow-band components, which contain the characteristics of the signal at different time and frequency scales. Dragomiretskiy *et al.* applied this phenomenon to solve the problem of mode mixing.⁽¹²⁾ A deep belief network (DBN) can automatically learn valid features to reduce the dimensions of features.⁽¹³⁾ Movahedi *et al.* applied some of these features to EEG signal processing, such as sleep stage and emotion recognition, and epilepsy diagnosis.⁽¹⁴⁾ A DBN can reduce the dimensions of high-dimensional features to obtain the optimal features of MI and avoid reducing the recognition rate. With this background, we propose in this paper a method comprising a combination of VMD, Hilbert transform, and a DBN for the recognition of MI-EEG patterns while avoiding the reduction of the recognition rate caused by manually determining the optimal period and optimal frequency band.

2. Related Work

Most of the existing methods are based on the use of the spatial and time-frequency characteristics of EEG signals to classify MI. The methods based on spatial features are mainly the CSP method and its modifications. The CSP method is based on matrix diagonalization to construct the optimal spatial filter for projection and then obtain the feature vectors with high discrimination. Zhang *et al.* proposed a multicore extreme learning machine based on the CSP method, in which the CSP method is used to extract spatial features for MI tasks.⁽¹⁵⁾ Zayyanu *et al.* proposed the filter bank common spatial pattern (FBCSP) to solve the problem that the effectiveness of the CSP method depends on the selection of an appropriate frequency band.⁽¹⁶⁾ Lu *et al.* proposed a deep learning model based on a restricted Boltzmann machine (RBM) that uses the fast Fourier transform to extract time-frequency features in EEG signals.⁽¹⁷⁾ In addition, Sun *et al.* proposed extraction methods based on bispectral features.⁽¹⁸⁾ However, it is necessary to acquire considerable prior knowledge to extract EEG signal characteristics. The

end-to-end deep learning framework combines multiple processing stages, such as data processing and feature extraction, into a single model, and establishes a direct projection from the input to the output. For this reason, the emergence of deep learning has provided ideas for the establishment of end-to-end models; Dai *et al.* attempted to construct end-to-end deep learning models for MI classification.⁽¹⁹⁾

However, EEG signals have the characteristics such as nonlinearity, nonstationarity, and low signal-to-noise ratio (SNR). These characteristics give rise to several major problems in the construction of an end-to-end model based on EEG signals. Schirrmeister *et al.* used the end-to-end learning advantages of ConvNets to build ConvNets models with three different architectures: the Deep-ConvNets model, Shallow-ConvNets model, and Hybrid-ConvNets model.⁽²⁰⁾ Zhao *et al.* proposed the WaSF ConvNet model as an improved ConvNets model to solve the problems of the traditional model, such as the difficulty in interpretation and the large number of parameters, to realize feature learning of space–time union.⁽²¹⁾ However, the above models generally use single-scale convolution, and the features extracted using this structure are limited. Although there are a few multiscale models for MI classification, Tang *et al.* proposed a single-scale convolution model that relies on the preprocessing of EEG signals. However, the multiscale span of these models is too small to fully extract EEG features.⁽²²⁾ In addition, the classification results of most existing models depend, to some extent, on the richness of spatial information in EEG signals. Data acquisition with fewer channels is more convenient, so when spatial information is insufficient, how EEG signals are extracted and classified is an important problem. Therefore, we use VMD to decompose EEG signals. The marginal spectrum (MS), instantaneous energy spectrum (IES), and joint time–frequency features (JT–Fs) are extracted by Hilbert transform to cover the time–frequency domain features under the complete MI time history, and the three features are integrated. A DBN is introduced to reduce the dimensions of the high-dimensional features after fusion to obtain the optimal features of MI and recognize the MI-EEG pattern, which prevents the reduction of the recognition rate that occurs upon manual determination of the optimal period and optimal frequency band.

3. High-dimensional JT–F Extraction Based on VMD

The framework of feature extraction and recognition based on VMD and DBN is shown in Fig. 1. When the MI-EEG signal is collected, several electrodes fixed on the head are used to

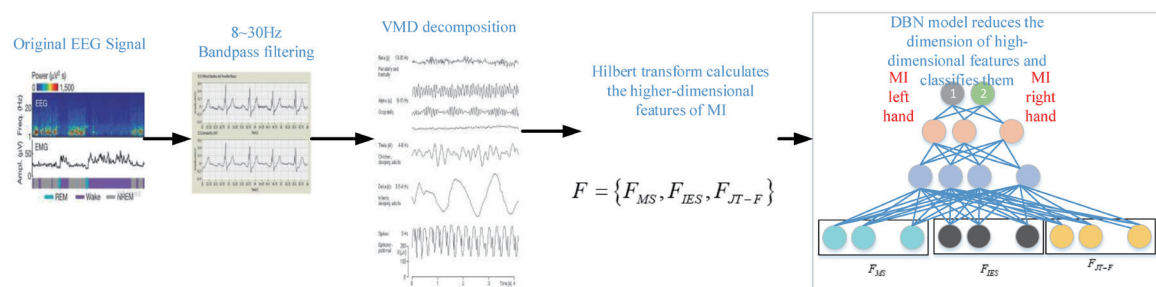


Fig. 1. (Color online) Framework of feature extraction and recognition based on DBN.

obtain multichannel signals, and all the electrodes record the signals generated by the brain at the same sampling frequency. The MI-EEG signal used by the participants of this study was typically collected between the time of the MI reminder and the end of the activity. The MI-EEG signal is defined as $X = (x_1, x_2, \dots, x_n) \in \mathfrak{R}^{n \times T \times C}$, where n is the number of samples of the MI-EEG signal, $T = t \times f$ is the number of time points of each EEG signal (t is the duration of the MI segment and f is the sampling frequency in Hz), and C is the number of channels of the EEG signal. The MI-EEG signal classification problem can be defined as

$$Y_{pre} = F(X), \quad (1)$$

where F is the mapping function, X represents the input EEG data, and Y_{pre} is the prediction result of the model output.

Since the WPT has the advantages of no redundancy and no omission, it can realize the local time–frequency analysis of the signal and meet the analysis requirements of an EEG signal containing many details. Therefore, after comparing various preprocessing algorithms, we decided to use the WPT to extract the characteristic frequency band. The wavelet packet function is expressed as⁽²³⁾

$$\begin{cases} u_{2n}(t) = \sqrt{2} \sum_{k \in Z} H_0 \cdot u_n(2t-l) \\ u_{2n+1}(t) = \sqrt{2} \sum_{k \in Z} H_1 \cdot u_n(2t-l) \end{cases}, \quad (2)$$

where $u_n(t)$ is the expression of the signal at time t at decomposed node n , l is the displacement, and H_0 and H_1 are a pair of conjugate orthogonal wavelet packet filters satisfying $H_1(k) = (-1)^k H_0(1-k)$. When $n = 0$, the scale function $u_0(t) = \varphi(t)$ and the wavelet basis function $u_1(t) = \Phi(t)$ can be obtained. The WPT has a similar structure to a binary tree. If the nodes in the wavelet packet decomposition are denoted as (j, p) (where j is the depth or the number of layers of wavelet packet decomposition and p is the frequency band order of the nodes), then after the wavelet packet decomposition of layer j , the frequency band range corresponding to node j is $\left[(2^j - 1)f_s / 2^{j+1}, 2^{j+1} / 2^{j+1} \right]$, where f_s is the sampling frequency. When the WPT is performed on an EEG signal, it is necessary to first select the appropriate wavelet basis function in accordance with the characteristics of the EEG signal and then to determine the characteristic frequency band of the EEG signal to determine the numbers of layers and reconstruction nodes of wavelet packet decomposition. In the VMD algorithm, the intrinsic mode function (IMF) is defined as the amplitude–frequency modulation signal, which is expressed as

$$u_k(t) = A_k(t) \cos(\varphi_k(t)), \quad (3)$$

where $A_k(t)$ ($A_k(t) \geq 0$) defines the instantaneous amplitude, and $\varphi_k(t)$ is the phase. The frequency of IMF is defined as $\omega_k(t) = \dot{\varphi}_k(t)$. A variational problem is first constructed. Assuming that the

original signal $f(t)$ is decomposed into K components, i.e., $\sum_{k=1}^K u_k(t) = f(t)$, the decomposition sequence is guaranteed to be a modal component with finite bandwidth with a central frequency, and the sum of the estimated bandwidth of each mode is minimum. The constraint condition is that the sum of all modes is equal to the original signal; then the corresponding expression for the constrained variational problem is

$$\min_{\{u_k(t), \omega_k\}} \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}, \quad (4)$$

where K is the number of modes to be decomposed, u_k and ω_k are the k th modal component and the center frequency after decomposition, respectively, $\delta(t)$ is the Dirac function, and $*$ is the convolution operator. To solve Eq. (4), we introduce the Lagrange multiplication operator λ , transform the constrained variational problem into an unconstrained variational problem, and obtain the augmented Lagrange expression as

$$L(u_k, \omega_k, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle, \quad (5)$$

where α is a quadratic penalty factor. The alternating-direction multiplier method is used to update u_k^{n+1} , ω_k^{n+1} , and λ_k^{n+1} alternately and solve the saddle point of Eq. (5). Then, the expression for u_k^{n+1} is

$$u_k^{n+1} = \arg \min_{u_k \in X} \left\{ \alpha \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_i u_i(t) + \frac{\lambda(t)}{2} \right\|_2^2 \right\}, \quad (6)$$

where X is the signal to be decomposed, $\omega_k = \omega_{k+1}$, and $\sum_i u_i(t) = \sum_{i \neq k} u_i(t)^{n+1}$. The Parseval/Plancherel Fourier isometrical transform is applied to Eq. (6) to transform it into the frequency domain, and each modal component and center frequency are optimized:

$$\hat{u}_k^{n+1}(\omega) \leftarrow \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2}, \quad (7)$$

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega \left| \hat{u}_k^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_k^{n+1}(\omega) \right|^2 d\omega}, \quad (8)$$

$$\hat{\lambda}^{n+1}(w) \leftarrow \lambda_k^n(w) + \gamma \left(\hat{f}(w) - \sum_k \hat{u}_k^{n+1}(w) \right). \quad (9)$$

In this study, the MI-EEG signal was first filtered using an 8–30 Hz bandpass filter, and then K IMF components were obtained by VMD of the filtered EEG signal. To extract the time-frequency characteristics of the MI-EEG signal, the IMF components were decomposed by Hilbert transform as follows:

$$Y_k(t) = \frac{1}{\pi} \mathfrak{R} \int_{-\infty}^{\infty} \frac{u_k(t_0)}{t-t_0} dt, \quad (10)$$

$$Z_k(t) = u_k(t) + jY_k(t) = a_k(t)e^{j\theta_k(t)}, \quad (11)$$

where \mathfrak{R} is the Cauchy principal value, t_0 is time in the time domain, $a_k(t)$ is the instantaneous amplitude, and $\theta_k(t)$ is the instantaneous phase. Therefore, the instantaneous frequency is defined as

$$\omega_k(t) = \frac{d\theta_k(t)}{dt}. \quad (12)$$

The original EEG signal can be expressed as the sum of each IMF component.

$$x(t) = \text{Re} \left\{ \sum_{k=1}^K a_k(t) e^{j\theta_k(t)} \right\} = \text{Re} \left\{ \sum_{k=1}^K a_k(t) e^{j \int \omega_k(t) dt} \right\} \quad (13)$$

In addition, the Hilbert spectrum is defined as

$$H(\omega, t) = \text{Re} \left\{ \sum_{k=1}^K a_k(t) e^{j \int \omega_k(t) dt} \right\}. \quad (14)$$

The Hilbert MS, IES, and JT-F under the characteristic frequency band are calculated as follows.

$$MS(\omega) = \int_0^T H(\omega, t) dt \quad (15)$$

$$IES(t) = \int_{\omega_1}^{\omega_2} H^2(\omega, t) d\omega \quad (16)$$

$$JT - F = \int_{t_1}^{t_2} \int_{\omega_1}^{\omega_2} H(\omega, t) d\omega dt \quad (17)$$

For Eq. (15), to avoid the omission of time–frequency domain information caused by manual selection, T is set as the end moment of MI. Because channels C_3 and C_4 are distributed on the left and right sides of the brain, respectively, whereas channel C_Z is at the center of the brain, in the left-handed and right-handed MI experiments on the subjects, only the EEG signals of channels C_3 and C_4 can reflect the event-related desynchronization/event-related synchronization (ERS/ERD) phenomenon. Therefore, in feature extraction, we do not process the channel C_Z EEG signal, and the MS feature F_{MS} of channels C_3 and C_4 can be obtained under the complete MI time course,

$$F_{MS} = \{MS_{C_3}, MS_{C_4}\}. \quad (18)$$

In Eq. (16), ω_1 and ω_2 respectively correspond to the μ and β characteristic frequency bands of MI, and the corresponding values of the two groups are taken. When ω_1^μ and ω_2^μ in the Mu frequency band are 8 and 13 Hz and ω_1^β and ω_2^β in the Beta frequency band are 16 and 28 Hz, respectively, the instantaneous energy characteristics IES_μ and IES_β in the two characteristic frequency bands are obtained. Thus, the instantaneous energy characteristics F_{IES} and F_{JT-F} of channels C_3 and C_4 in the characteristic frequency band are obtained.

$$F_{IES} = \{IES_{C_3}^{\mu}, IES_{C_4}^{\mu}, IES_{C_3}^{\beta}, IES_{C_4}^{\beta}\} \quad (19)$$

In addition, the $JT-F$ can be expressed as

$$J = \{JT-F_{1,1}, JT-F_{1,2}, \dots, JT-F_{1,11}, JT-F_{2,1}, JT-F_{2,2}, \dots, JT-F_{2,11}, \dots\}. \quad (20)$$

Thus, the time-frequency joint feature F_{JT-F} of channels C_3 and C_4 is

$$F_{JT-F} = \{J_{C_3}, J_{C_4}\}. \quad (21)$$

4. High-dimensional Feature Classification Based on DBN

A DBN consists of several RBMs in the bottom layer and a classifier in the top layer. The two adjacent layers make up one RBM, and each layer is independent of the other. However, data can be transformed among layers in accordance with the learning rules of the RBM through specific activation functions. Figure 2 shows the structure of the five-layer DBN. In Fig. 2, each RBM is composed of a visible layer (v) and a hidden layer (h). The layers are interconnected with each other through weight w , but there is no connection within the layers. It is assumed that layer v includes n visible units and layer h includes m hidden units.

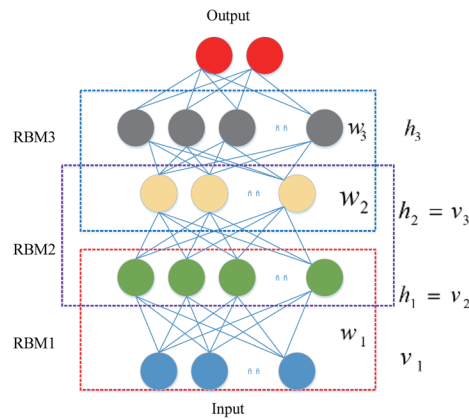


Fig. 2. (Color online) Structure of DBN.

As shown in Fig. 2, the energy of one group of RBM systems in a certain state is

$$E(v, h; \theta) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij}, \quad (22)$$

where v_i defines the state of the i th visible unit and h_j is the j th hidden unitary state. $\forall i, j, v_i \in \{0, 1\}, h_j \in \{0, 1\}$, and $\theta = \{a_i, b_j, w_{ij}\}$ are the parameters of the RBM model, w_{ij} is the connection weight between visible element i and hidden element j , and a_i and b_j are the offsets of visible element i and hidden element j , respectively. The joint probability density of the visible layer and hidden layer is defined as

$$P(v, h; \theta) = \frac{e^{-E(v, h; \theta)}}{\sum_{v, h} e^{-E(v, h; \theta)}}. \quad (23)$$

From Eq. (23), the conditional probabilities of the hidden layer and visible layer are derived as follows.

$$\begin{cases} P(h | v; \theta) = \prod_j P(h_j | v; \theta) \\ P(v | h; \theta) = \prod_i P(v_i | h; \theta) \end{cases} \quad (24)$$

According to Eq. (24), σ is the activation function, and the probabilities of the hidden node and visible node being activated are

$$\begin{cases} p(h_j | v; \theta) = \sigma \left(b_j + \sum_i v_i w_{ij} \right) \\ p(v_i | h; \theta) = \sigma \left(a_i + \sum_j h_j w_{ij} \right) \end{cases} \quad (25)$$

To make the RBM fit the training samples well, the training objective is set as the maximum likelihood function,

$$L(\theta) = \prod_v L(\theta | v) = \prod_v P(v). \quad (26)$$

Contrastive divergence is used to obtain a better estimate of the visible layer with only a few state transitions. More specially, the algorithm first calculates the conditional probabilities of all the hidden layer elements, then uses Gibbs sampling to determine the states of the hidden elements. Following this, it calculates the state of the visual layer to complete the reconstruction of the visual layer and finally solves Eq. (26). Therefore, the updating process of weight and bias can be defined as

$$\begin{cases} \Delta w_{ij} = \varepsilon \left(\langle v_i h_j \rangle_{orig} - \langle v_i h_j \rangle_{recon} \right) \\ \Delta a_i = \varepsilon \left(\langle v_i \rangle_{orig} - \langle v_i \rangle_{recon} \right) \\ \Delta b_j = \varepsilon \left(\langle h_j \rangle_{orig} - \langle h_j \rangle_{recon} \right) \end{cases}, \quad (27)$$

where $\langle \cdot \rangle_{orig}$ denotes the expectation of the original data, $\langle \cdot \rangle_{recon}$ denotes the expectation of the reconstructed data, and ε is the learning rate. For the parameter ε , Liu *et al.* proposed the use of the parameter *Momentum* of the momentum term to make the convergence more accurate.⁽²⁴⁾ *Momentum* is defined as

$$Momentum = \rho \theta + \varepsilon \frac{\partial \ln L}{\partial \theta}, \quad (28)$$

where ρ is the learning rate of *Momentum*. Then the update rule of the weight bias becomes

$$\begin{cases} \Delta w = \rho w + \varepsilon \frac{\partial \ln L}{\partial w} \\ \Delta a = \rho a + \varepsilon \frac{\partial \ln L}{\partial a} \\ \Delta b = \rho b + \varepsilon \frac{\partial \ln L}{\partial b} \end{cases}. \quad (29)$$

5. Results and Discussion

5.1 Datasets and processes

To verify the universality and effectiveness of the proposed method, BCI 2003 Competition Dataset III (S1), BCI 2005 Competition Dataset III b (S2–S4), and BCI 2008 Competition Dataset II b (S5–S7) were used in the experiment. The information of each dataset is shown in Table 1.

During the signal collection process, for a reason related to the acquisition equipment, some of the time block signals were not successfully collected, and these data points were represented as nonnumeric points in the signals. To improve the SNR, the experimental data were preprocessed and nonnumeric points in the data were set to zero. To reduce the workload of data processing, a downsampling process was performed on the response data of subjects S5–S7, and the sampling frequency was reduced to 125 Hz. Since the ERD/ERS phenomenon of MI mainly occurred in the Mu rhythm (8–13 Hz) and Beta rhythm (16–28 Hz), the original signal was passed through a bandpass filter of 8–30 Hz, which is a Butterworth filter of order six, and the stopband cutoff frequency of the Butterworth filter was set to 6 and 32 Hz, respectively. The learning rate, number of pretrainings, and number of fine-tunings of each subject are shown in Table 2.

5.2 Results and analysis

To compare the effect of different feature extraction methods on the recognition rate, three reference groups, i.e., F_{MS} , F_{IES} , and F_{JT-F} , were designed and compared with the proposed method. Figure 3 shows the maximum recognition rate of each subject for the three reference groups and the proposed method. As shown in Fig. 3, the average maximum recognition rates obtained using MS features, IES features, time-frequency joint features, and fusion features are 60.5, 73.1, 77.5, and 78.5%, respectively. The classification effect of MS features is relatively

Table 1
Information of datasets.

Dataset	Subjects	Sampling channel	Number of training sets	Number of testing sets	Sampling frequency
BCI 2003 Competition Dataset III	S1	C ₃ , C ₄ , C _Z	140	140	135
BCI 2005 Competition Dataset III b	S2	C ₃ , C ₄ , C _Z	160	160	128
	S3, S4	C ₃ , C ₄ , C _Z	540	540	128
BCI 2008 Competition Dataset II b	S5, S7	C ₃ , C ₄ , C _Z	400	320	250
	S6	C ₃ , C ₄ , C _Z	400	300	250

Table 2
Parameter settings of the seven subjects.

Parameter	S1	S2	S3	S4	S5	S6	S7
Learning rate	0.01	0.008	0.001	0.1	0.001	0.001	0.001
Number of pretrainings	540	500	1000	500	1000	1100	1000
Number of fine-tunings	33	30	360	300	71	335	310

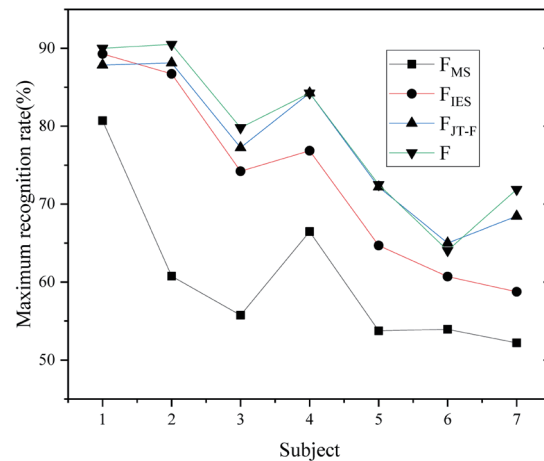


Fig. 3. (Color online) Maximum recognition rates of different feature extraction methods.

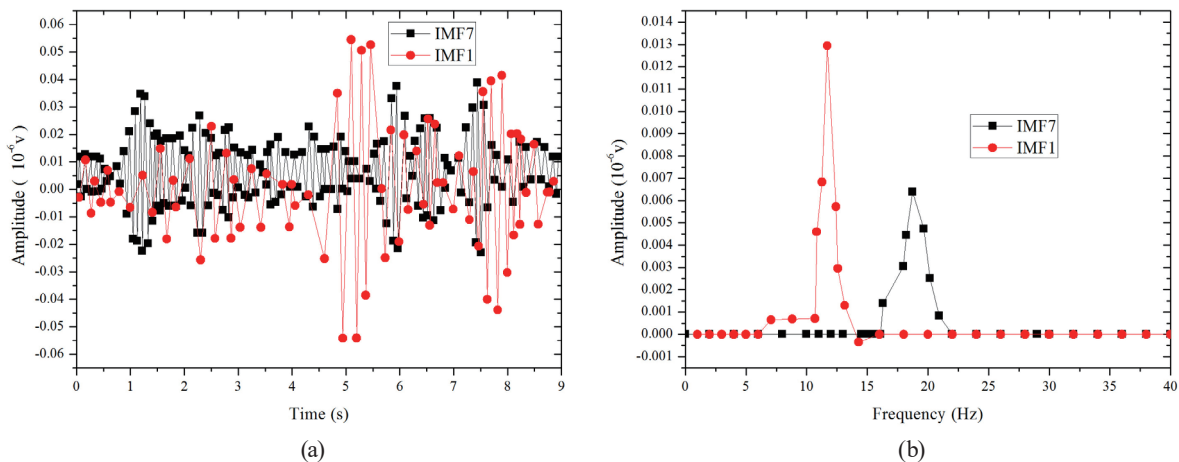


Fig. 4. (Color online) (a) Results of VMD and (b) spectral diagram of components.

poor because the experimental settings cause the strength of the MI signal to change with time. The time-domain resolution of MS features under the complete time history of MI is low; thus, the MS features cannot accurately characterize MI.

The preprocessed EEG signal was decomposed by VMD in the experiment. The results of VMD of the channel C_3 EEG signal of the first subject and its component spectrum are shown in Fig. 4. In the experiment, the EEG signal was decomposed by EMD, which served as the benchmark algorithm of VMD. The EMD results and component spectrum of the channel C_3 EEG signal of the first subject are shown in Fig. 5. The frequencies of the IMF1 and IMF7 components of the adaptive decomposition are concentrated in the range of 8–30 Hz as a result of prefiltering of the EEG signal. In addition, the frequency band aliasing of components obtained by EMD is serious, while the results of VMD are all narrow-band components, so the same component information does not appear in different components. In the experiment, it was found that the characteristic frequency band is mainly distributed in the first three order components.

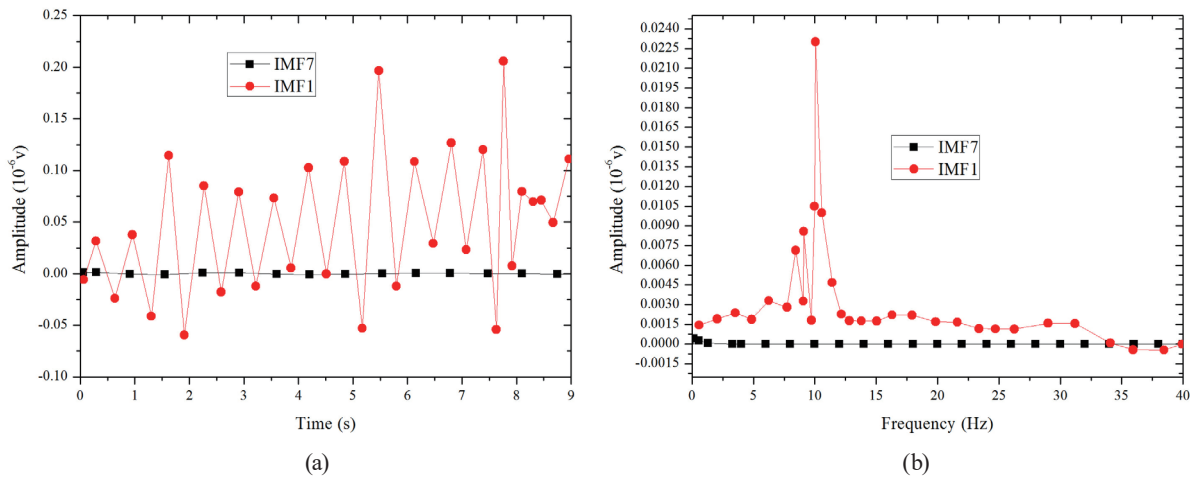


Fig. 5. (Color online) (a) Results of EMD and (b) spectral diagram of components.

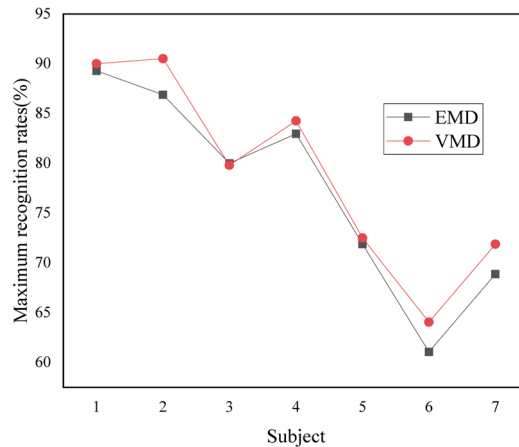


Fig. 6. (Color online) Recognition rates of EMD and VMD.

The first three order components of EMD were selected to extract the same high-dimensional features, and the DBN was used to reduce the feature dimensions and classify the features obtained by the two methods. The recognition rates obtained by the two decomposition methods were compared, and the recognition results are shown in Fig. 6. Because of the better decomposition effect, the features extracted by VMD have a better recognition rate than those extracted by EMD.

6. Conclusions

BCIs, as an important application of human–computer hybrid intelligence enhancement, can control external devices through brain activities and thus establish the connection between the brain and the outside world. In early studies, BCIs were mainly used for the rehabilitation of stroke patients, and later, they began to be used in a wide range of fields, such as wheelchair

control, spelling machines, and emotion recognition. BCIs can make full use of brain activities such as electrophysiological and hemodynamic activities to enable interaction between the brain and the outside world. In this study, the VMD method was applied to decompose the EEG signal using the Hilbert transform to extract the MI of MS, IES, and JT–F. In addition, the four-layer DBN network was introduced to achieve dimensionality reduction of the integrated high-dimensional feature and to obtain the optimal feature and classification. The recognition rate of the MI-EEG signal was improved. The proposed method was used to conduct classification experiments with BCI competition datasets, and the average accuracy was 79%. The proposed method effectively improves the recognition rate of MI-EEG signals and lays a foundation for the application of MI BCIs.

Acknowledgments

This work was supported by the Joint Project of the Natural Science Foundation of Heilongjiang (LH2019F036).

References

- 1 J. R. Wolpaw and N. Birbaumer: *Nat. Rev. Neurol.* **12** (2017) 513. <https://doi.org/10.1017/CBO9780511995583.042>
- 2 L. Carmona, P. F. Diez, E. Laciari, and V. Mut: *IEEE Trans. Neural Syst. Rehabil. Eng.* **28** (2020) 825. <https://doi.org/10.1109/TNSRE.2020.2979684>
- 3 M. Kim, B. H. Kim, and S. Jo: *IEEE Trans. Neural Syst. Rehabil. Eng.* **23** (2015) 159. <https://doi.org/10.1109/TNSRE.2014.2365834>
- 4 R. Djemal, A. G. Bazyed, K. Belwafi, S. Gannouni, and W. Kaaniche: *Brain Sci.* **6** (2016) 36. <https://doi.org/10.3390/brainsci6030036>
- 5 K. S. Biju and M. G. Jibukumar: *Procedia Comput. Sci.* **125** (2018) 468. <https://doi.org/10.1016/j.procs.2017.12.061>
- 6 C. Yücelbaş, Ş. Yücelbaş, S. Özşen, G. Tezel, S. Küçüktürk, and Ş. Yosunkaya: *Neural Comput. Appl.* **29** (2018) 17. <https://doi.org/10.1007/s00521-016-2445-y>
- 7 S. Taran and V. Bajaj: *Neural Comput. Appl.* **31** (2018) 6925. <https://doi.org/10.1007/s00521-018-3531-0>
- 8 Y. Cao, Y. Sun, G. Xie, and T. Wen: *IEEE Trans. Veh. Technol.* **68** (2019) 7544. <https://doi.org/10.1109/TVT.2019.2925903>
- 9 X. Ioannis, A. Alkinoos, P. Niki, K. Dimitris, and P. D. Bamidis: *Comput. Intell. Neurosci.* **2018** (2018) 1. <https://doi.org/10.1155/2018/7957408>
- 10 S. Chen, Z. Luo, and H. Gan: *Neural Comput. Appl.* **29** (2018) 857. <https://doi.org/10.1007/s00521-016-2594-z>
- 11 M.-G. Tan, J.-H. Ho, H.-T. Goh, H. K. Ng, L. A. Latif, and M. Mazlan: *Biomed. Signal Process. Control* **61** (2020) 102045. <https://doi.org/10.1016/j.bspc.2020.102045>
- 12 K. Dragomiretskiy and D. Zosso: *IEEE Trans. Signal Process.* **62** (2014) 531. <https://doi.org/10.1109/TSP.2013.2288675>
- 13 G. E. Hinton and R. Salakhutdinov: *Science* **313** (2006) 504. <https://doi.org/10.1126/science.1127647>
- 14 F. Movahedi, J. L. Coyle, and E. Sejdic: *IEEE J. Biomed. Health. Inf.* **22** (2017) 642. <https://doi.org/10.1109/JBHI.2017.2727218>
- 15 Y. Zhang, Y. Wang, G. Zhou, J. Jin, B. Wang, X. Wang, and A. Cichocki: *Expert Syst. Appl.* **96** (2018) 302. <https://doi.org/10.1016/j.eswa.2017.12.015>
- 16 S. Zayyanu and Q. Li: *Int. J. Comput. Appl. Technol.* **175** (2020) 975. <https://doi.org/10.5120/ijca2020920646>
- 17 N. Lu, T. Li, X. Ren, and H. Miao: *IEEE Trans. Neural Syst. Rehabil. Eng.* **25** (2017) 566. <https://doi.org/10.1109/TNSRE.2016.2601240>
- 18 L. Sun, Z. Feng, N. Lu, B. Wang, and W. Zhang: *Expert Syst. Appl.* **131** (2019) 9. <https://doi.org/10.1016/j.eswa.2019.04.021>
- 19 G. Dai, J. Zhou, J. Huang, and N. Wang: *J. Neural Eng.* **17** (2020) 016025. <https://doi.org/10.1088/1741-2552/ab405f>

- 20 R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggenberger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball: *Hum. Brain Mapp.* **38** (2017) 5391. <https://doi.org/10.1002/hbm.23730>
- 21 D. Zhao, F. Tang, B. Si, and X. Feng: *Neural Networks* **114** (2019) 67. <https://doi.org/10.1016/j.neunet.2019.02.009>
- 22 X. Tang, W. Li, X. Li, W. Ma, and X. Dang: *Expert Syst. Appl.* **149** (2020) 113285. <https://doi.org/10.1016/j.eswa.2020.113285>
- 23 M. J. Meyer, E. J. Malloy, and B. A. Coull: *Stat. Comput.* **31** (2021) 14. <https://doi.org/10.1007/s11222-020-09981-3>
- 24 W. Liu, L. Chen, Y. Chen, and W. Zhang: *IEEE Trans. Parallel Distrib. Syst.* **31** (2020) 1754. <https://doi.org/10.1109/TPDS.2020.2975189>

About the Authors



Wulin Zhang is a research scholar at Mudanjiang Normal University. He received his doctorate degree in signal and information processing from Harbin Engineering University in 2015. His research interests include computer vision, image processing, and deep learning. (hechaoctgu@163.com)



Zeyu Liang received his bachelor's degree from Mudanjiang Normal University, P.R. China. He is currently studying at the College of Physics and Electronics Engineering, Mudanjiang Normal University. His research interests include variational mode decomposition and deep learning. (lllzy321@126.com)



Zirui Liu was born in Yilan, Harbin, P.R. China, in 1998. He received his bachelor's degree from Mudanjiang Normal University, P.R. China. He is currently studying at the College of Physics and Electronics Engineering, Mudanjiang Normal University. His research interests include computer vision, image processing, and deep learning. (msylzr@163.com)