Journal of Innovative Optical Health Sciences
Vol. 5, No. 2 (2012) 1250006 (6 pages)

DOI: 10.1142/S179354581250006X

ENERGY FEATURE EXTRACTION AND SVM CLASSIFICATION OF MOTOR IMAGERY-INDUCED ELECTROENCEPHALOGRAMS

JIANING ZHENG, LIYU HUANG* and JING ZHAO
School of Life Sciences and Technology
Xidian University, Xi-an 710071, China
**huangly@mail.xidian.edu.cn*

Received 15 September 2011
Accepted 19 February 2012
Published 27 April 2012

The precise classification for the electroencephalogram (EEG) in different mental tasks in the research on brain−computer interface (BCI) is the key for the design and clinical application of the system. In this paper, a new combination classification algorithm is presented and tested using the EEG data of right and left motor imagery experiments. First, to eliminate the low frequency noise in the original EEGs, the signals were decomposed by empirical mode decomposition (EMD) and then the optimal kernel parameters for support vector machine (SVM) were determined, the energy features of the first three intrinsic mode functions (IMFs) of every signal were extracted and used as input vectors of the employed SVM. The output of the SVM will be classification result for different mental task EEG signals. The study shows that mean identification rate of the proposed algorithm is 95%, which is much better than the present traditional algorithms.

Keywords: Electroencephalogram; empirical mode decomposition; support vector machine; motor imagery.

1. Introduction

Brain−computer interface (BCI) can provide a new access that does not depend on the surrounding nerves and muscles for the information interaction between the brain and the outside.[1] In the present BCI system, electroencephalogram (EEG) is often used to convey information produced by the human brain.[2] EEG preprocessing, feature extraction and pattern classification are the key steps in the design of a BCI system.

The traditional feature extraction algorithms include power spectrum analysis, adaptive autoregressive, wavelet transform, independent component analysis, common spatial pattern (CPS), etc. All of these algorithms are good at stationary signal analysis. But EEG is a typical nonstationary signal, which contains long-time low-frequency component and short-time high-frequency component. It is of great value to investigate a novel and effective feature extraction and classification method based on the time-frequency characteristics of the nonstationary signal.

Research shows that the energy of the specific frequency of EEG will reduce when the corresponding brain mantle is activated by mental activity, which is

*Corresponding author.

called event-related desynchronization (ERD) and that the energy of the specific frequency of EEG will rise when the corresponding brain mantle is not activated by mental activity, which is called event-related synchronization (ERS). The ERD/ERS feature of EEG becomes obvious when imagining left or right hand's movement.[3]

The experimental model of left- or right-hand motor imagery is most commonly used in designing BCI system. The EEG data recorded from three electrodes in the two mental tasks were divided into training sample and testing sample. After decomposed by EMD algorithm for denoise purpose, the EEG's first three IMF components were calculated and used as the input vectors of SVM to classify the different mental tasks. A relatively ideal result is received.

2. Method

The purpose of the proposed algorithm is to recognize the mental tasks of the subject by feature extraction and fusion classification of the EEG data. The μ rhythm can be detected by the IMFs and the ERD/ERS phenomenon during left- and right-hand movement imagination can be analyzed by the Hilbert marginal spectrum,[4] so the strategy we put forward is to combine the EMD, energy feature extraction with SVM. The details of each step are discussed as follows.

2.1. *Empirical mode decomposition*

EMD algorithm is researched from the assumption that any signal consists of different simple intrinsic modes of oscillations. The main aim of the EMD in this paper is to decompose an EEG time series into a set of components or functions, known as IMFs, each of which must satisfy the following definition[5]: (1) having the same number of extrema and zero-crossings or differ at most by one in the whole data set, and (2) symmetric with respect to local zero mean. According to EMD theory, a given EEG signal $x(t)$ can be expressed as follows:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t), \quad t = 1, 2, \ldots, N. \tag{1}$$

Here n is the number of IMFs, $r_n(t)$ denotes the final residue, and it can be interpreted as the DC component of the EEG signal, and $c_i(t)$ denotes all IMFs that have nearly zero means and are nearly orthogonal to each other. An original EEG is decomposed by EMD and shown in Fig. 1.

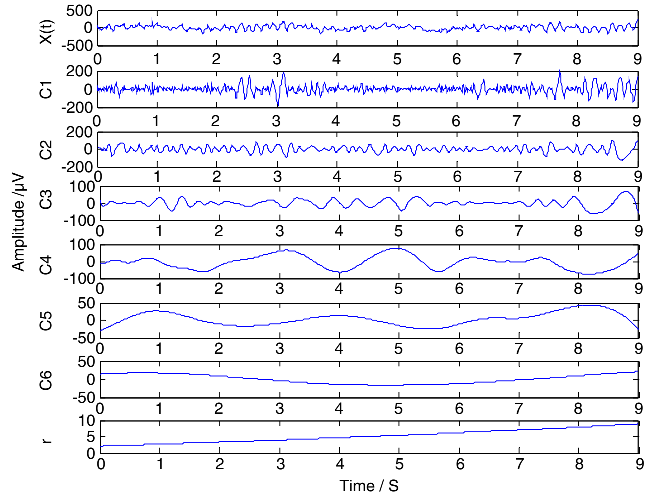


Fig. 1. Empirical mode decomposition of an original EEG.

2.2. *The energy calculation of IMF*

The energy of a time series $c(n)$ with a length of N is defined as:

$$E = \sum_{n=1}^{N} |c(n)|^2. \tag{2}$$

After decomposing an EEG signal with EMD method, the energy values of the first three IMFs were calculated and shown in Table 1. Obviously, the energy distribution of the IMFs is in a regular pattern for both of the mental tasks.

2.3. *Support vector machine*

The invention of SVM was driven by statistical learning theory.[5] The output of a binary SVM classifier can be computed by the following expression:

$$f(x) = \operatorname{sgn}\left[\sum_{i=1}^{N} \alpha_i y_i k(x_i, x) + b\right]. \tag{3}$$

Here $x_i \in R^d$, $y_i \in \{-1, 1\}$ are the class labels, $\{x_i, y_i\}_{i=1}^{N}$ is a separable training sample set, b denotes a bias and $\alpha_i \geq 0$ are Lagrangian multipliers, $k(x_i, x_j)$ is called kernel function in SVM. The commonly used kernel functions include linear function, polynomials function, radial basis function (RBF) and sigmoid function.

Table 1. The energy value of the first three IMF of the EEG signal in different mental tasks.

	C3 lead			Cz lead			C4 lead		
Motor imagery	IMF1	IMF1	IMF1	IMF1	IMF2	IMF3	IMF1	IMF2	IMF3
Left hand	2705.03	2331.49	2299.36	162.11	16.77	18.76	2299.36	162.11	16.77
Right hand	2459.88	1994.25	2940.30	518.91	91.76	12.50	2940.30	518.91	91.76
Left hand	1927.33	1328.23	1006.67	118.97	6.59	31.83	1006.67	118.97	6.59
Right hand	1046.20	806.93	1582.49	147.07	69.40	10.27	1582.49	147.07	69.40
Left hand	451.96	130.04	129.97	30.18	4.46	11.78	129.97	30.18	4.46
Right hand	398.43	287.89	599.15	197.87	20.31	24.85	599.15	197.87	20.31

The function that satisfied the Mercer theorem can all serve as kernel function theoretically. Vapnik, the founder of SVM, believed that kernel function is not the main factor that affects the performance of the classifier as long as its parameters are appropriately determined.[6] In this paper, the RBF is used as kernel function of the SVM in our algorithm.

2.4. *Optimization of the parameters in SVM using RBF*

In RBF of the SVM, there are two parameters: penalty factor c and g (g equals to $1/\sigma^2$, and σ is variance) must be chosen. The default values of these two factors are both one in most cases.[7] Nevertheless, if the optimization parameters are used in RBF, the classification accuracy rate of SVM could be improved. A common method for parameter optimization is "cross validation".[8] The result of the optimization by cross-validation method is shown in Fig. 2. There are multi-group values of c and g that correspond to the best accuracy rate.

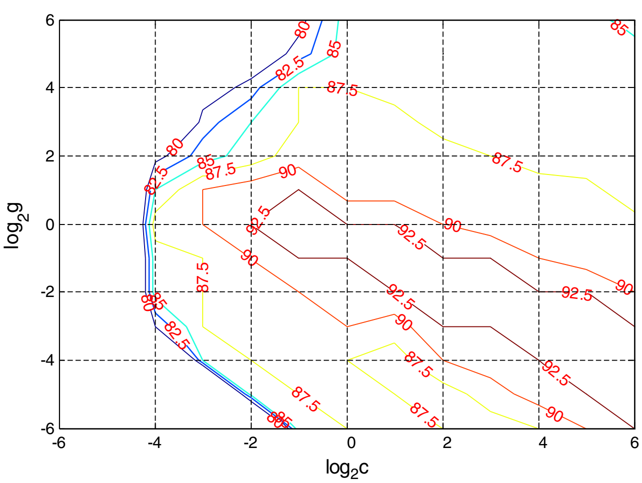


Fig. 2. The contour map of the optimization result for parameters in RBF.

Tests show that a too much higher c will cause an overfitting phenomenon in the classification process, that is to say the classification has a good accuracy rate on training set but a bad accuracy rate on testing set. In our study, the minimal c and corresponding g were chosen in all the combination of c and g under which can result in the best accuracy rate. According to our study, the best values of these two parameters are $c = 0.25$ and $g = 0.75$, respectively.

3. Data Acquisition and Preprocessing

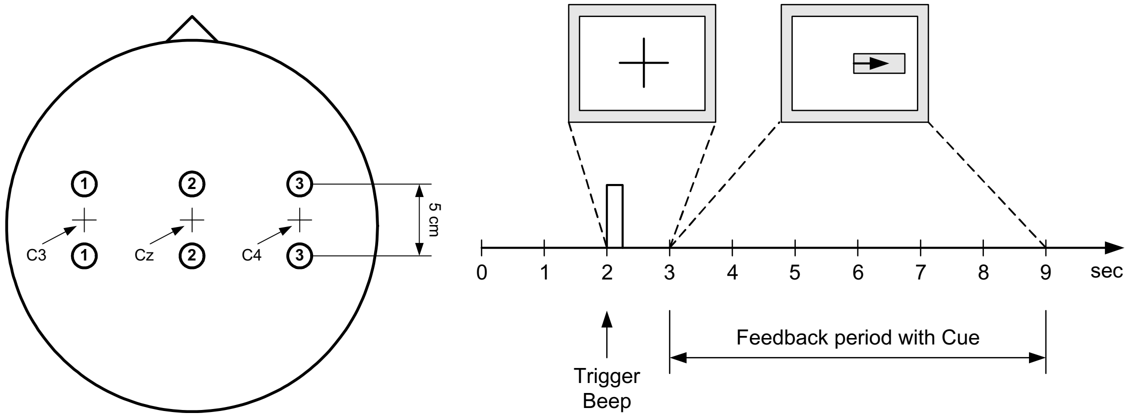
3.1. *Experiment description*

A benchmark EEG data set that was used in the BCI competition 2008 was used to evaluate the performance of the proposed feature extraction and classification algorithm.[9] The experiment was achieved and data set was provided by the Department of Medical Informatics, Institute for Biomedical Engineering, Graz University of Technology, Austria.

There are 280 trials in the available data set, each trial lasts 9 s, as shown in Fig. 3, and the subject was quiet in the first 2 s. When $t = 2$ s, an acoustic stimulus indicates the start of the trial. When $t = 3$ s, an arrow (right or left) was showed as a cue, and the subject was asked to do motor imagery along the cue direction. Three bipolar EEG channels were measured over C3, Cz and C4 lead. The EEG was sampled with 128 Hz and the frequency range of the filter in the recording system is from 0.5 to 30 Hz.

3.2. *Data preprocessing*

The low-frequency noise will be induced in the EEG recording system because of the amplitude of the



(a) (b)

Fig. 3. Electrode positions (a) and timing scheme (b) in the experiment.

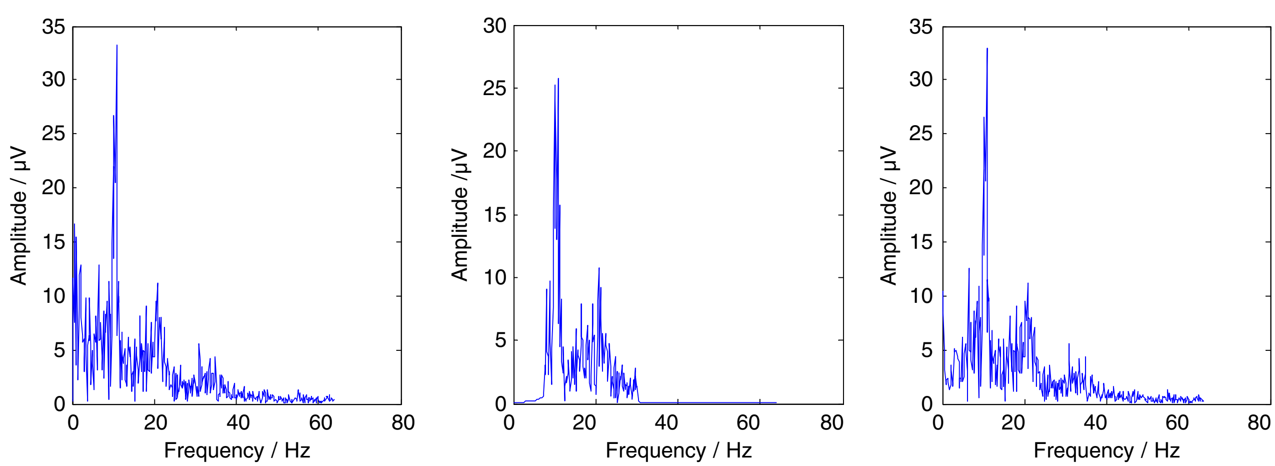
EEG signal is as small as microvolt, as shown in Fig. 4(a). The FFT filtering will cause frequency leakage and signal distortion which lowers the accuracy rate of the mental task classification, as show in Fig. 4(b). We used EMD to filter the low-frequency component, which will perfectly keep the useful frequency components.[10]

The research shows that the ERD/ERS phenomenon mainly present in the μ and β frequency ranges in the offside and onside of the primary perception motor cortical area in the human brain. The β frequency range is between 8 Hz and 12 Hz, while the μ frequency range is between 18 Hz and 25 Hz. Thus it can be seen that the frequency range between 8 Hz and 30 Hz in the signal contain the main features.

Based on the local feature time scale of the signal, EMD can decompose a signal into several IMFs with their frequency from high to low, as shown in Fig. 5. Using the first three IMFs to reconstruct the original signal, we can filter the low-frequency noise and keep the useful feature we need in classification for the motor imagery mental task. The frequency spectrum of the reconstructive signal is shown in Fig. 4(c).

4. Results

Randomly selected 240 trail data from all the 280 trails were used as the training set and the other 40 trails as testing set. Then the sample or the number of the training set and the testing set were changed



(a) Original signal (b) Filtered by FFT (c) Filtered by EMD

Fig. 4. The frequency spectrogram of the EEG signal in different conditions.

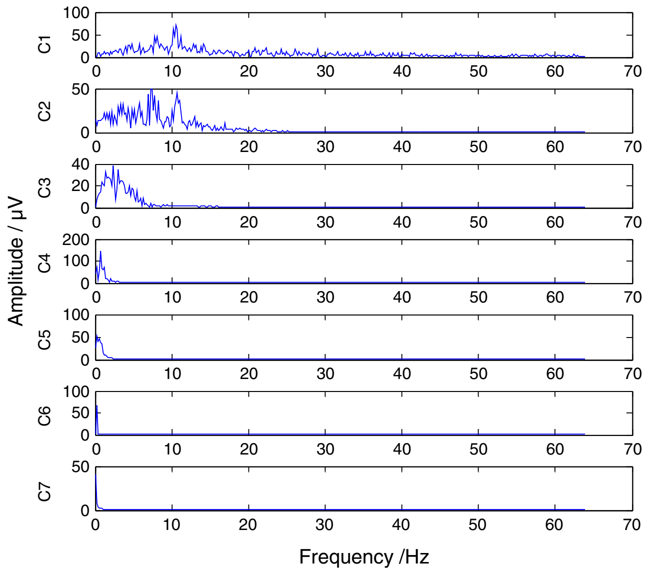


Fig. 5. The frequency spectrum of the each IMF.

Table 2. The accuracy rate in two different RBF parameters.

	Default parameters of SVM	Optimum parameters of SVM
The number of training sets	240	240
The number of testing sets	40	40
The maximal accuracy rate (%)	92.50	95.00

Table 3. The accuracy of different methods.

First author	Feature extraction	Classification tool	Accuracy (%)
V. Jeyabalan[11]	AAR parameters	LDA	81.4
	ARBF + AAR parameters	LDA	88.6
	AAR parameters	SVM	82.9
	ARBF + AAR parameters	SVM	87.9
X. Baoguo[12]	AR, power spectrum	BP	85
H. Jianfeng[13]	SFFT, energy entropy	BP	83.3
		SVM	91.6
V. Jeyabalan[14]	Coefficient of adaptive recursive bandpass filter	SFAM	92.0 (C3)
			94.0 (C4)
Our scheme	EMD, energy	SVM	95.0

and the above experiment program was repeated for enough times to test the accuracy of the classification model.

The classification accuracy rates in different RBF parameters are shown in Table 2. From the table, we can see that the accuracy is improved to 95% by using the optimization RBF parameters in the SVM. Table 3 shows several results from the references using different method with the same dataset. Obviously, the accuracy of our method is much better than the traditional algorithms.

5. Conclusion

According to the ERD/ERS phenomenon of the EEG recorded from different positions in the scalp of human brain in imagining the left-hand and right-hand movement, a new classification method is presented and used to classify the different EEGs recorded in different mental tasks in BCI system, with the average accuracy rate 95.00%.

The selection of the kernel function is a highlight in the research on SVM. Even though the pros and cons of application of different kernel functions are not verified theoretically, RBF is a common choice with the advantages of few parameters must be determined and that it can map the sample into a higher-dimensional space which is good at dealing with the nonlinear situation.

For the typical nonlinear and nonstationary signal like EEG, EMD can decompose the signal into several IMFs with different frequency range, which represent the μ and β frequency ranges in ERD/ERS of the EEG signal, these frequency ranges keep the useful information maximally in the classification of the motor mental tasks.[15] The preferable results we acquired in this paper by the above method can be interpreted reasonable in theory.

Acknowledgments

This work is supported by National Natural Science Foundation of China under Grant No. 81071221.

References

1. J. R. Wolpaw, N. Birbaumer, D. J. McFarland *et al.*, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.* **113**, 767 (2002).
2. J. Kroneg, G. Chanel, S. Voloshynovskiy *et al.*, "EEG-based synchronized brain-computer interfaces: A model for optimizing the number of mental tasks," *IEEE Trans. Neural Syst. Rehab. Eng.* **15**, 50 (2007).
3. G. Pfurtscheller, "EEG event-related desynchronization and event-related synchronization," *Electroenceph. Clin. Neurophysiol.* **103**, 26 (1997).
4. G. Xiaojing, W. Xiaopei, Z. Dexiang, "Motor imagery EEG detection by empirical mode decomposition," *IEEE Int. Joint Conf. Neural Networks*, *IEEE World Congress on Computational Intelligence* (2008).
5. N. E. Huang, Z. Shen, S. R. Long, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Royal Soc. London Series* **454**, 903 (1998).
6. V. N. Vapnik, *Statistical Learning Theory* (John Wiley & Sons, New York, 1998).
7. Available from: http://www.csie.ntu.edu.tw/∼cjlin/libsvm/.
8. V. Cherkassky, Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks* **17**, 113 (2004).
9. Available from: http://ida.first.fraunhofer.de/projects/bci/competition_iii/.
10. C.-L. Yeh, H.-C. Chang, C.-H. Wu, P.-L. Lee, "Extraction of single-trial cortical beta oscillatory activities in EEG signals using empirical mode decomposition," *BioMedical Engineering* [Online] (2010).
11. V. Jeyabalan, A. Samraj, L. C. Kiong, "Motor imaginary signal classification using adaptive recursive bandpass filter and adaptive autoregressive models for brain machine interface designs," *Int. J. Biol. Life Sci.* **3**, 2 (2007).
12. X. Baoguo, S. Aiguo, "Feature extraction and classification of single trial motor imagery EEG," *J. Southeast Univ.* **37**(4), (2007).
13. H. Jianfeng, M. Zhendong, X. Dan, "Classification of motor imagery EEG signals based on energy entropy," *Computer Eng. Appl.* **44**(33), 235–238 (2008).
14. V. Jeyabalan, A. Samraj, L. C. Kiong, "Classification of motor imaginary signals for machine communication," *Int. Conf. Signal Acquisition and Processing* (2009).
15. G. Pfurtscheller, C. Brunner, A. Schlögl, "Mu rhythm (de)synchronization and EEG singletrial classification of different motor imagery tasks," *NeuroImage* **31**, 153 (2006).